

Clinical outcome assessment in malignant glioma trials: measuring signs, symptoms, and functional limitations

Jaishri O. Blakeley, Stephen Joel Coons, John R. Corboy, Nancy Kline Leidy, Tito R. Mendoza, and Jeffrey S. Wefel

Johns Hopkins University, Departments of Neurology, Neurosurgery and Oncology, Baltimore, Maryland (J.O.B.); Critical Path Institute, 1730 E. River Road, Tucson, Arizona (S.J.C.); University of Colorado School of Medicine, Denver Veterans Affairs Medical Center, 12631 East 17th Avenue B185, Aurora, Colorado (J.R.C.); Evidera, 7101 Wisconsin Avenue, Suite 1400, Bethesda, Maryland (N.K.L.); University of Texas MD Anderson Cancer Center, 1515 Holcombe Bld., Houston, Texas (T.R.M., J.S.W.)

Corresponding Author: Jaishri O. Blakeley, MD, Johns Hopkins University, Departments of Neurology, Neurosurgery and Oncology, Baltimore, MD 21287 (jblakel3@jhmi.edu).

The shared goal of all parties developing therapeutics against malignant gliomas is to positively impact the lives of people affected by these cancers. Clinical outcome assessment (COA) tools, including measures of patient-reported outcome, performance outcome, clinician-reported outcome, and observer-reported outcome, allow patient-focused assessments to complement traditional efficacy measures such as overall survival and radiographic endpoints. This review examines the properties of various COA measures used in malignant glioma clinical trials to date and cross references their content to the priority signs, symptoms, and functional limitations defined through a community survey conducted by the National Brain Tumor Society. The overarching goal of this initiative is to identify COA measures that are feasible and have appropriate psychometric properties for use in this patient population as well as highlight where further development is needed.

Keywords: clinical outcome assessment, clinical trials, endpoints, malignant glioma.

People with malignant gliomas suffer a wide range of neurologic signs, symptoms, and functional limitations. These are variable in timing, severity, and nature but occur at a high frequency across the majority of people with these brain cancers. The vast majority of clinical trials for people with malignant gliomas assess therapeutic benefit via survival endpoints such as overall survival (OS) and surrogates of biologic activity such as time to progression, progression-free survival (PFS), and radiographic response. However, it is also increasingly recognized that disciplined study and integration of clinical outcome assessments (COAs) provide another approach to assess the impact of anticancer therapy on patient benefit.¹ For malignant gliomas specifically, as new agents with broad mechanisms of action (including anti-angiogenesis and immunomodulatory agents) enter the clinical trial arena, there is increasing evidence that radiographic assessments are limited in predicting clinical benefit and that measurements of signs, symptoms, and functional limitations provide important data about treatment efficacy.^{2–5} Hence, there is growing interest in identifying COA tools that provide reliable and meaningful data about a given drug's impact on patient-focused clinical outcomes in people with malignant gliomas.

The FDA has invested extensively in resources to guide the development and incorporation of COA-based endpoints in clinical trials, including a COA staff within the Center for Drug

Evaluation and Research's (CDER) Office of New Drugs specifically tasked with promoting the development and implementation of patient-focused endpoint measures in medical product development. Together with members of the Center for Biologics Evaluation and Research and the Center for Devices and Radiological Health, this team created the "Guidance for Industry—Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims" to describe the FDA's expectations regarding patient-reported outcome (PRO)-based labeling claims for new drugs.¹

With the FDA's recommendations in mind and with the goal of inspiring a community-wide conversation about "ideal" COAs for each priority sign, symptom, and functional limitation in people with malignant glioma (See the article by Armstrong et al., on pages 1–12), the authors of this review (a group of investigators with a wide range of expertise in neuro-oncology, COA tool development, and clinical trial design) identified and described COA measures that have been used in large therapeutic trials for malignant gliomas; summarized evidence of reliability, validity, responsiveness, and feasibility in people with malignant gliomas; and presented the results of the group's analysis at the Jumpstarting Brain Tumor Drug Development Coalition and FDA Clinical Trials Clinical Outcome Assessment Endpoints Workshop, October 15, 2014. The intention of this work is to inform stakeholders in glioma therapeutic discovery

Received 3 September 2015; accepted 25 October 2015

Published by Oxford University Press on behalf of the Society for Neuro-Oncology 2016. This work is written by (a) US Government employee(s) and is in the public domain in the US.

of the core principles of COA endpoint measures and review the strengths and limitations of the available COA measures to inform their use (or the development and use of new measures) in future glioma clinical trials.

Clinical Outcome Assessment Measures and Application for Malignant Glioma

There are several types of COAs, including PROs, performance outcomes (PerfOs), clinician-reported outcomes (ClinRO), and observer-reported outcomes (ObsROs) (Table 1 and Fig. 1). In order to choose the best type of measure for a given study, it is important to identify: the context of use, the primary construct to be measured (eg, pain, cognitive function), and the approach that best assesses that construct. In regard to

Table 1. Types of COA tools

COA	Definition
PRO measures	Measurement based on a report that comes from the patient (ie, study subject) about the status of his or her health condition without amendment or interpretation of the patient’s report by a clinician or anyone else. A PRO measure can be self- or interview administered, provided that the interviewer records only the patient’s response. Symptoms or other unobservable concepts known only to the patient (eg, pain severity, nausea) can only be assessed by PRO measures. PRO measures can also assess the patient perspective on functioning or activities that may also be observable by others.
PerfO measures	Based on a task performed by a patient according to standardized instructions and administered by a health care professional. These include measures of gait speed (eg, timed 25-foot walk test), memory recall, and other cognitive testing (eg, digit symbol substitution test).
ClinRO measures	Measurement based on a report that comes from a trained health care professional after observation of a patient’s health condition. A ClinRO involves a clinical judgment or interpretation of the observable signs, behaviors, or other physical manifestations thought to be related to a disease or condition.
ObsRO measures	Measurement based on an observation by someone other than the patient or a health professional. This may be a parent, spouse, or other nonclinical caregiver who is in a position to regularly observe and report on a specific aspect of the patient’s health. An ObsRO measure does not include medical judgment or interpretation. Examples of ObsROs include a parent report of a child’s vomiting episodes or a report of wincing thought to be the result of pain in people who are unable to report for themselves.

therapeutic efficacy studies for malignant gliomas, the central questions that must be addressed to choose the “best” COA tool include:

- (1) What component(s) is the treatment of interest expected to impact (eg, memory, attention, gait, mood, energy, motor strength)? This defines the domain(s) of interest to be measured.
- (2) What are the strengths and limitations of the various COA types (PRO, PerfO, ClinRO, ObsRO measures) for the domain(s) of interest?
- (3) What are the strengths and limitations of specific tools purported to assess the domain(s) of interest in the population of interest?
- (4) What are the psychometric properties of those tools, including evidence of construct validity and the amount of change that can be considered clinically meaningful?

Carefully addressing these four questions early in study development greatly facilitates selection of a specific COA tool based on its fit for purpose. This in turn vastly improves the likelihood of generating meaningful data about the impact of treatment on a patient-focused endpoint.

An in-depth examination of existing tools is necessary to determine if there is sufficient evidence for a given tool applied to assess a specific domain of interest in clinical trials for people with malignant gliomas. For each COA measure, we examined: the concepts or domains it is assessing, the strength of the evidence supporting its psychometric properties for assessment of those concepts or domains in people with malignant gliomas, its feasibility in a multicenter clinical trial setting, and, if possible, its relationship to other COA measures and assessments of disease response in people with malignant gliomas. The goal of this evaluation was to determine if there is evidence that a given COA tool validly and reliably measures priority concepts or domains that reflect treatment benefit and are proximally related to biologic effects on tumor.

Patient-Reported Outcome Measures

PRO measures assess symptoms or other concepts known only to the patient (eg, pain severity, nausea) via patient self-report (Table 1). Malignant glioma therapeutic efficacy trials have long incorporated PRO measures as exploratory endpoints, but more recently some PRO measures have been required study endpoints, resulting in much higher compliance (Supplementary Table S1).⁶ Such studies have provided extensive experience with commonly used PRO measures in the context of multicenter efficacy studies for glioma and serve as a valuable resource to investigate their performance, including: (i) the European Organisation for Research and Treatment of Cancer’s (EORTC) health-related quality of life (HRQoL) instruments (the 30-item Quality of Life Questionnaire [QLQ-C30] and its 20-item Brain Neoplasm version [QLQ-BN20]), (ii) the MD Anderson Symptom Inventory–Brain Tumor (MDASI-BT), (iii) the Functional Assessment of Cancer Therapy–Brain (FACT-Br), and (iv) the EuroQol Group’s 5-dimension health questionnaire (EQ-5D). Collectively, these PRO measures assess multiple domains of HRQoL, symptoms, and functional limitations.

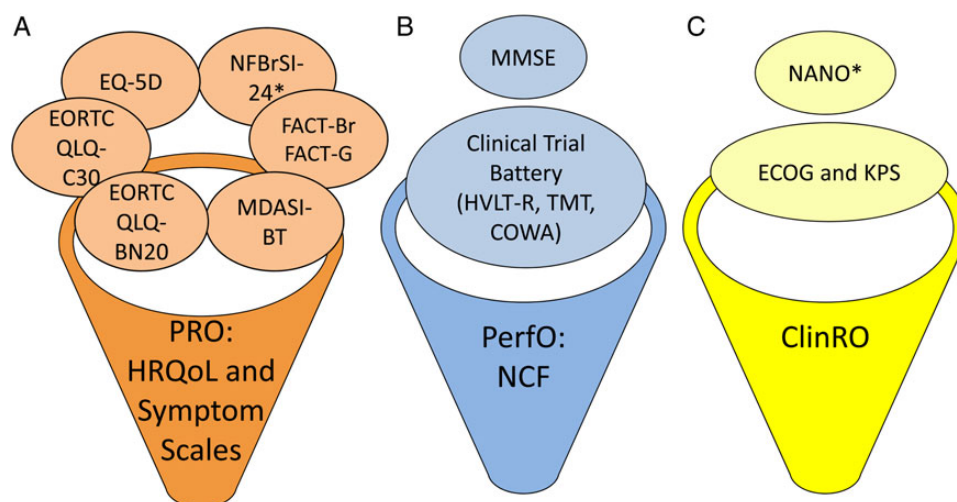


Fig. 1. Categories of COA commonly used in malignant glioma clinical trials. The COA measures commonly used in glioma clinical trials can be divided into (A) PRO measures, (B) Perfo measures, and (C) ClinRO measures. PRO measures can include items assessing symptoms, functional limitations, and/or HRQoL. *Newly developed or in development. NCF, neurocognitive function; TMT, Trail Making Test; COWA, Controlled Oral Word Association.

European Organisation for Research and Treatment of Cancer's HRQoL Instruments

EORTC's HRQoL instruments are among the most widely used PRO measures across all cancer trials around the world. The QLQ-C30 is an HRQoL questionnaire for people with any cancer and the QLQ-BN20 was developed specifically for people with brain cancer (Supplemental Table S1). The QLQ-C30 (version 3.0) produces 15 scores based on 9 multi-item scales (global health/QoL, physical functioning, role functioning, emotional functioning, cognitive functioning, social functioning, fatigue, nausea/vomiting, and pain) and 6 single items (dyspnea, insomnia, anorexia, constipation, diarrhea, and financial impact).⁷ All items and scale scores for the QLQ-C30 and QLQ-BN20 are linearly transformed to a 0–100 scale (higher scores reflect higher levels of the concept being measured). Internal consistency coefficients for the multi-item scales of the QLQ-C30 ranged from 0.52 to 0.89 in a study of 305 people with lung cancer.⁷ There were variable interscale correlations, with weak correlations between the emotional functioning scale and physical functioning scale, but stronger correlations for physical functioning, role functioning, and fatigue scales.⁷ There was also a significant change in physical and role functioning, fatigue, and global health/QoL in people who had improved versus worsening performance status on treatment. However, there was no significant difference across scales from pre- to on-treatment without stratification by performance status.⁷ Finally, the cognitive functioning scale had poor internal consistency, suggesting that this measure is not ideal for assessing cognitive domains.^{8–11}

The brain-specific QLQ-BN20 produces 11 scores based on 4 multi-item scales (future uncertainty, visual disorder, motor dysfunction, and communication deficit) and 7 single items. The QLQ-BN20's psychometric properties were assessed using data from EORTC protocols 26951 and 26891 and showed high internal consistency, with Cronbach's coefficient alpha ranging from 0.71 to 0.90.¹² However, similar to the QLQ-C30,

the QLQ-BN20 did not appear to have high sensitivity to change based on data from the phase II study of bevacizumab versus lomustine versus the combination in people with recurrent glioblastoma.^{13,14} Several of the QLQ-BN20's single items (eg, headaches, hair loss, itchy skin, leg weakness) have shown poor test-retest reliability.¹⁵ Although concerns have been raised about the time required to complete all 50 items that make up the QLQ-C30 and QLQ-BN20, completion rates in both the Radiation Therapy Oncology Group (RTOG) 0825 and Avastin in Glioblastoma (AVAglio) trials were 78%–86% at the latest timepoint in 0825¹⁶ and 86%–88% at a comparable timepoint in AVAglio.¹⁷ Overall, the QLQ-C30 and QLQ-BN20 have been important PRO tools in trials for malignant glioma. Their major limiting factor is that several of the items and scales reflect outcomes distal from the impact of therapy on the disease. Within the context of pivotal trials used for regulatory approval of drugs in the US, endpoints used to support labeling claims are expected to reflect proximal markers of treatment benefit (survival, symptom relief) rather than more distal outcomes (QoL, financial well-being). It is possible, however, to select a subset of scales or items from the QLQ-C30 and QLQ-BN20 to measure target constructs in malignant glioma trials (eg, QLQ-C30's physical function scale).

MD Anderson Symptom Inventory–Brain Tumor

The MDASI-BT has a general cancer module as well as a disease-specific module designed to assess symptom severity and degree of interference across a range of domains pertinent to people with malignant gliomas. It provides a single composite score for symptom severity and a second composite score for symptom interference. The 21 symptom items and 7 interference items in the MDASI-BT have content validity established through concept elicitation from people with brain cancer, clinicians, and caregivers.¹⁸ Internal consistency reliability (Cronbach's alpha of 0.91) and test-retest reliability ($r = 0.95$) for

the symptom severity composite score were observed in a study of people with heterogeneous grades of primary brain tumor.^{19,20} A factor analysis revealed the following domains measured by the MDASI-BT: focal neurologic deficit and cognitive, affective, treatment-related, and generalized disease status. The sensitivity of the MDASI-BT has been demonstrated in a randomized study of temozolomide in people with newly diagnosed glioblastoma multiforme (GBM) (RTOG 0525), in which a worsening neurologic symptom score on the MDASI-BT from baseline to cycle 1 in people receiving concurrent chemoradiation was predictive of lower PFS, while worsening cognitive score was predictive of OS.²¹ Baseline neurologic symptom scores and early change in cognitive symptom scores were associated with both PFS and OS in the RTOG 0825 study of bevacizumab added to standard chemoradiation for people with newly diagnosed GBM.¹⁶ In addition, in RTOG 0825 the treatment and placebo groups significantly differed in global MDASI-BT symptom severity score ($P = .02$) and symptom interference score over the progression-free period ($P = .004$).¹⁶ Finally, differences between groups were observed in mood-related symptom interference ($P < .001$), affective factors ($P = .04$), cognitive factors ($P = .01$), treatment factors ($P = .03$), and generalized or disease factors ($P = .01$), and the MDASI-BT discriminated among known patient groups based on performance status, disease status, and acuity status.¹⁹ For example, people with malignant glioma with significant activity-related interference (ratings of >5) were 3.8 times more likely to have recurrence on imaging.^{19,22} Hence, the MDASI-BT has been demonstrated to have validity, reliability, and sensitivity as a symptom assessment instrument in people with malignant gliomas within the context of therapeutic trials.

The MDASI-BT has a short completion time (~4 min) and can be administered on paper or tablet or via a phone interview.¹⁹ Both the “past 24 hour” and “past week” recall periods have been validated.²³ The MDASI-BT was completed by 85%–95% of participants at baseline in both RTOG 0525 and RTOG 0825.^{16,21} Comparable to the EORTC QLQ-C30/BN20, the MDASI-BT had $>75\%$ completion rates at all study timepoints in RTOG 0825.¹⁶

A limitation of the MDASI-BT is that it asks only about the severity of each symptom and does not ask about other symptom attributes such as frequency or distress associated with each symptom. In addition, the composite symptom scale includes both disease- and treatment-related symptoms, requiring the investigator to isolate one from the other.

Functional Assessment of Cancer Therapy–Brain and National Comprehensive Cancer Network/Functional Assessment of Cancer Therapy–Brain Symptom Index

The FACT-Br includes 27 items measuring general (FACT-G) cancer-related physical, social, emotional, and functional well-being, and a 23-item scale for symptoms and problems specific to brain tumors. There are multiple scoring options for the measure, including subscale and total scores. Recently, a tool derived from the FACT-Br, the 24-item National Comprehensive Cancer Network/Functional Assessment of Cancer Therapy–Brain Symptom Index (NFBrsI-24), was designed for people with advanced brain cancer. The NFBrsI-24 evaluates

symptoms and concerns in people with primary brain tumors with 3 subscales: symptoms (12 physical items, 5 emotional items), side effects (5 items), and function/well-being (2 items), as well as a summary score.²⁴ Although the NFBrsI-24 is derived from the FACT-Br, because it includes 4 new symptoms not assessed by the FACT-Br, one cannot derive NFBrsI-24 scores from the FACT-Br, although some prorating can be performed. Given that NFBrsI-24 evaluates advanced brain cancer symptoms, it may play a role in future studies requiring a symptom focus; however, it is the newest PRO measure for this population and does not yet have a performance record in clinical trials.

EQ-5D

The EQ-5D is a PRO measure that has been used in several general cancer and brain tumor trials.^{25,26} This 6-item measure is a generic, preference-weighted measure of health status that has been used across a wide range of conditions and treatments. It combines responses on 5 items (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) into a single, preference-weighted, interval-level score where 0 is dead, 1 is perfect health, and negative values represent states worse than death. The EQ-5D is used most commonly for informing economic models evaluating the cost-effectiveness of treatment (cost per quality-adjusted life year).

Performance Outcome Measures

PerfO measures are defined by CDER as “based on a task(s) performed by a patient according to instructions . . . administered by a health care professional. Performance outcomes require patient cooperation and motivation. These include measures of gait speed (eg, timed 25-foot walk test), memory recall, or other cognitive testing (eg, digit symbol substitution test).” The domains that are best measured by PerfO in the setting of malignant glioma studies include cognitive and motor function.

Impairments in cognitive function are ubiquitous in people with malignant glioma and are one of the most frequent signs of a brain tumor.^{27–30} People with malignant glioma also report cognitive dysfunction as a common presenting symptom.^{21,27} Cognitive dysfunction has been identified across many studies as a priority area for patients, caregivers, and clinicians (See the article by Armstrong et al., on pages 1–12). Although cognitive dysfunction can impact HRQoL, result in symptoms, and contribute to functional limitations, the most proximal and objective measures of cognitive function are PerfO measures.

PerfO measures of cognitive function range from screening tests such as the Mini-Mental Status Examination (MMSE) to tests assessing specific cognitive domains. In order to make application of cognitive tests feasible in the setting of wide-scale clinical trials, it is necessary to choose a limited set of domains and measures. There are associations between OS and the domains of memory, executive function, and processing speed in people with malignant gliomas.^{21,28–32} Memory (60%), executive function (54%), and psychomotor processing speed (32%) are the most commonly impaired cognitive domains in people with newly diagnosed GBM.³³ Given that these 3 domains are important for daily functional activities, are commonly

impacted in people with malignant gliomas, and can be assessed with standardized PerFO tests, including the Hopkins Verbal Learning Test–Revised (HVLTR) for memory; the Trail Making Test Part A for processing speed; and the Trail Making Test Part B and Multilingual Aphasia Examination–Controlled Oral Word Association for executive function, together making up the Clinical Trial Battery (CTB), they have been applied in recent large clinical trials for malignant glioma.^{16,21}

The instruments within the CTB have published normative data that take into account relevant demographic variables, have adequate psychometric properties, have known cutoff values for clinically significant change, and can assess change over time without frequent floor or ceiling effects in people with malignant glioma.^{21,32,34}

Moreover, they are feasible in that they can be performed in roughly 30 minutes and have standardized administration procedures to ensure similar test administration across sites, although a limitation is that dedicated training of study personnel is required. People with newly diagnosed GBM, recurrent GBM, and multiple brain metastases have demonstrated good compliance with test completion at baseline (up to 98%) and serially during the trial (>90% at 6 mo).^{16,35,36} Test performance at baseline and changes during treatment predict PFS and OS,^{16,21,33,37} worsen in advance of imaging progression,^{16,21,38} and discriminate between treatment groups.^{16,39–43}

Notably, there is a relationship among PRO measures such as the MDASI-BT, FACT-Br, and CTB. In people with brain tumors, cognitive function performance declined in advance of changes in the FACT-Br total scores or the Functional Independence Measure and predicted future worsening on these measures.³⁸ There were significant correlations between the CTB and the FACT-Br total score as well as between the CTB and the Functional and Social well-being factors of the FACT-G in people with newly diagnosed malignant glioma before initial surgery ($r = 0.31–0.38$).⁴⁴ In the same patient population, Bradshaw et al⁴⁵ demonstrated that greater neurocognitive impairment was associated with lower global functional status as measured by both KPS and Functional Independence Measure. Resendiz et al⁴⁶ also found a relationship between the MDASI-BT item “Problems with Remembering” and HVLTR memory impairment in 115 people with primary brain tumors. Taken together, these data demonstrate that cognitive dysfunction assessed with PerFO measures such as the CTB correlates with PRO measures of HRQoL, functional activities, and cognitive dysfunction as well as PFS and OS.

There has been far less experience with PerFO measures of neurologic function outside of cognition. There are a handful of examples of noncognitive PerFO measures in clinical trials for malignant gliomas, but none used to assess efficacy of an anticancer drug. Trial NCT01169415, “Effects of Steroid Tapering on Functional Capacity and Neurocognition,” assessed performance via skeletal muscle strength, a 6-minute walk test, and cognitive measures. The 6-minute walk test was shown to be feasible in people with recurrent malignant gliomas and correlated with KPS in one study but was not associated with survival endpoints.⁴⁷

Mobility is a priority symptom, but PerFO evaluations of mobility and the functions that contribute to it (vision, motor, and sensory function) have not been incorporated into clinical trials for drugs targeting malignant gliomas to date. Such measures

exist and have been incorporated into clinical trials for other neurologic diseases. For example, for multiple sclerosis, the Multiple Sclerosis Functional Composite (MSFC) is a combination of 3 PerFO tools, including the Paced Auditory Serial Addition Test, the timed 25-foot walk, and a bilateral 9-hole peg test, incorporating cognitive and physical (arm and leg) assessments. The MSFC is commonly used in combination with the Expanded Disability Status Scale, which is a ClinRO measure.

Clinician-Reported and Observer-Reported Outcome Measures

ClinRO measures, including KPS and the Eastern Cooperative Oncology Group Performance Status (ECOG, or World Health Organization [WHO] performance scale), are extensively used across clinical trials for malignant gliomas (Supplementary Table S1). These scales are clinician rated and are designed to broadly quantify a given patient’s functional capacity for activities of daily living from the clinician’s perspective. KPS is the ClinRO used most commonly across clinical trials for malignant gliomas. It also has been validated as a prognostic factor for OS in this population.⁴⁸ The KPS score ranges from 0 (death) to 100 (no evidence of disease) in 11 intervals, whereas the ECOG/WHO scale has 6 points from 0 (no evidence of disease) to 5 (death). Overall, they have similar psychometric properties.⁴⁹ Across a variety of cancers, KPS has been found to have inter-rater correlations (Pearson product-moment) ranging from 0.66 to 0.89.^{50–52} The limitations of KPS as a clinical trial outcome include: (i) it is not specific for either the impact of the cancer or the treatment, and (ii) it is not clear how sensitive it is to change with treatment. Hence, KPS has prognostic significance but is not a sensitive or specific measure of response to treatment.

Additional ClinRO measures have been assessed in malignant glioma trials, including the Glasgow Outcome Score (initially designed for people with brain trauma), modified Rankin Scale (initially designed for stroke assessment), and the Medical Research Council brain prognostic index (designed expressly for people with brain tumors). Interestingly, all measures performed reasonably similarly in people with GBM.⁵³

There is work under way to develop a specific ClinRO for malignant gliomas termed the Neurologic Assessment in Neuro-Oncology (NANO) to complement the imaging criteria in use for assessing tumor response in malignant gliomas (Response Assessment in Neuro-Oncology, RANO).^{54,55} The NANO is composed of 8 clinician-assessed domains (neurologic signs). It provides a single score as well as domain-specific scores and, importantly, accounts for preexisting (and possibly irreversible) signs by rating the associated domain as nonevaluable. Current work is assessing inter- and intrarater reliability.

Selecting or Developing Clinical Outcome Assessment Tools for Malignant Glioma Clinical Trials

Addressing the core questions about the domain(s) of interest, the context of use, and the psychometric properties of the

tool(s) in the target population allows one to make an informed decision about whether existing tools, scales within tools, or individual items within tools are appropriate for assessing a target domain within the context of an efficacy trial for malignant glioma. The design of glioma treatment trials incorporating COA tools and strategies for statistical analysis of COA-based endpoints in conjunction with traditional efficacy endpoints such as OS and PFS are addressed by panel 3 (See the article by Gilbert et al., on pages 21–25).

There is a substantial library of analytic data demonstrating the psychometric properties of COA tools, including the EORTC QLQ-C30/BN20, MDASI-BT, FACT-Br, and the CTB, that are commonly used in malignant glioma trials. For many other COA tools, there is considerable experience, and therefore a wealth of data available for analysis, but their psychometric properties are insufficiently examined in the context of malignant glioma. It is important that COA tools used in glioma clinical trials have prespecified responder definitions, that is, score changes that have been determined to be clinically meaningful. Unfortunately, responder definitions have not been estimated for many of the COA measures applied to malignant gliomas historically. In order to incorporate COA measures into decisions about efficacy with the greatest degree of accuracy, the neuro-oncology field will need to invest in understanding what the clinically meaningful change (or “response”) is for each of the measures, much like the field has dedicated resources to defining the context and guidelines for interpretation of radiographic response in clinical trials for malignant glioma.⁵⁵ In a similar vein, the use of glucocorticoids and other medications (eg, anti-epileptics, mood altering medications, analgesics) may influence the results of some of the COA measures. Recognition of this fact requires that efforts are made to accurately collect concomitant medication information and consider these medications during the analysis of COA measures.

There is an opportunity to dramatically advance the assessment of patient-centered outcomes in malignant glioma trials by performing secondary analyses of COA data from phase III studies completed in this patient population and aggregating data from trials with similar eligibility criteria using the same COA measure. This analysis would allow: (i) evaluation of items, scales, and whole measures that address priority signs, symptoms, and functional limitations; (ii) assessment of sensitivity to change of the measures relative to pertinent external criteria such as OS and radiographic endpoints; (iii) examination of feasibility based on completion rates over time; and (iv) determination of responder definitions and clinically meaningful change.

Conclusion

Although no single COA measure is sufficient to address all the signs, symptoms, and functional limitations identified as priorities for people with malignant glioma, there are several COA tools that address many of these key domains and have data available from their use in prior clinical trials. Analysis of these data will be immensely valuable for prioritizing select COA tools for specific domains, identifying where modification is desirable, or determining where new COA tool development is required.

For immediate use, the MDASI-BT as a PRO measure of symptoms and the CTB as a PerFO tool for assessing cognition are thought to have attributes (eg, assessing signs, symptoms, and function proximal to the disease pathology) and measurement properties (eg, demonstrated reliability and validity) that support their use as key efficacy endpoints in clinical trials for malignant glioma. However, it is critical to ensure that these COA measures address the specific domains of interest in a given trial, and in fact, it may be desirable to select isolated domains of interest and apply only the items that address those domains to minimize respondent burden. Finally, there are several new COA tools, including the NANO, the NFBRSI-24, and a brain tumor-specific EORTC measure of instrumental activities of daily living that require further development and testing but may prove useful in assessing the outcomes of malignant glioma treatment trials.

Acknowledgments

The authors would like to thank Drs Terri Armstrong, Mark Gilbert, and Patrick Wen and Jennifer Helfer and David Arons for their efforts to coordinate and lead the Jumpstarting Brain Tumor Drug Development Coalition and FDA Clinical Trials Clinical Outcome Assessment Endpoints Workshop on October 15, 2014; Drs Elektra Papadopoulos and Ruthann Giusti for their expert advice during the preparation of the Clinical Outcome Assessment Endpoints Workshop, and Rhonda Jackson for her administrative support of the manuscript.

Conflict of interest statement. None declared.

References

1. FDA Center for Drug Evaluation and Research. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. <http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf>. 2009.
2. Batchelor TT, Sorensen AG, di Tomaso E, et al. AZD2171, a pan-VEGF receptor tyrosine kinase inhibitor, normalizes tumor vasculature and alleviates edema in glioblastoma patients. *Cancer Cell*. 2007;11(1):83–95.
3. Batchelor TT, Duda DG, di Tomaso E, et al. Phase II study of cediranib, an oral pan-vascular endothelial growth factor receptor tyrosine kinase inhibitor, in patients with recurrent glioblastoma. *J Clin Oncol*. 2010;28(17):2817–2823.
4. Hygino da Cruz LC Jr, Rodriguez I, Domingues RC, et al. Pseudoprogression and pseudoresponse: imaging challenges in the assessment of posttreatment glioma. *AJNR Am J Neuroradiol*. 2011;32(11):1978–1985.
5. Avery RA, Hwang EI, Jakacki RI, et al. Marked recovery of vision in children with optic pathway gliomas treated with bevacizumab. *JAMA Ophthalmol*. 2014;132(1):111–114.
6. Armstrong TS, Gilbert MR. Patient reported endpoints for measuring clinical benefit in (high grade glioma) primary brain tumor patients. *Curr Treat Options Oncol*. 2014;15(4):519–528.
7. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85(5):365–376.

8. Wan C, Meng Q, Yang Z, et al. Validation of the simplified Chinese version of EORTC QLQ-C30 from the measurements of five types of inpatients with cancer. *Ann Oncol*. 2008;19(12):2053–2060.
9. Gundy CM, Aaronson NK. Effects of mode of administration (MOA) on the measurement properties of the EORTC QLQ-C30: a randomized study. *Health Qual Life Outcomes*. 2010;8:35, 1–7.
10. Cheng JX, Liu BL, Zhang X, et al. The validation of the standard Chinese version of the European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire 30 (EORTC QLQ-C30) in pre-operative patients with brain tumor in China. *BMC Med Res Methodol*. 2011;11:56, 1–9.
11. Shin YS, Kim JH. Validation of the Korean version of the European Organization for Research and Treatment of Cancer brain cancer module (EORTC QLQ-BN20) in patients with brain tumors. *Health Qual Life Outcomes*. 2013;11:145, 1–5.
12. Taphoorn MJ, Claassens L, Aaronson NK, et al. An international validation study of the EORTC brain cancer module (EORTC QLQ-BN20) for assessing health-related quality of life and symptoms in brain cancer patients. *Eur J Cancer*. 2010;46(6):1033–1040.
13. Taal W, Oosterkamp HM, Walenkamp AM, et al. Single-agent bevacizumab or lomustine versus a combination of bevacizumab plus lomustine in patients with recurrent glioblastoma (BELOB trial): a randomised controlled phase 2 trial. *Lancet Oncol*. 2014;15(9):943–953.
14. Dirven L, van den Bent MJ, Bottomley A, et al. The impact of bevacizumab on health-related quality of life in patients treated for recurrent glioblastoma: results of the randomised controlled phase 2 BELOB trial. *Eur J Cancer*. 2015;51(10):1321–1330.
15. Osoba D, Aaronson NK, Muller M, et al. The development and psychometric validation of a brain cancer quality-of-life questionnaire for use in combination with general cancer-specific questionnaires. *Qual Life Res*. 1996;5(1):139–150.
16. Gilbert MR, Dignam JJ, Armstrong TS, et al. A randomized trial of bevacizumab for newly diagnosed glioblastoma. *N Engl J Med*. 2014;370(8):699–708.
17. Taphoorn MJ, Henriksson R, Bottomley A, et al. Health-related quality of life in a randomized phase III study of bevacizumab, temozolomide, and radiotherapy in newly diagnosed glioblastoma. *J Clin Oncol*. 2015;33(19):2166–2175.
18. Armstrong TS, Cohen MZ, Eriksen L, et al. Content validity of self-report measurement instruments: an illustration from the development of the brain tumor module of the M.D. Anderson Symptom Inventory. *Oncol Nurs Forum*. 2005;32(3):669–676.
19. Armstrong TS, Mendoza T, Gning I, et al. Validation of the M.D. Anderson symptom inventory brain tumor module (MDASI-BT). *J Neurooncol*. 2006;80(1):27–35.
20. Armstrong TS, Wefel JS, Gning I, et al. Congruence of primary brain tumor patient and caregiver symptom report. *Cancer*. 2012;118(20):5026–5037.
21. Armstrong TS, Wefel JS, Wang M, et al. Net clinical benefit analysis of radiation therapy oncology group 0525: a phase III trial comparing conventional adjuvant temozolomide with dose-intensive temozolomide in patients with newly diagnosed glioblastoma. *J Clin Oncol*. 2013;31(32):4076–4084.
22. Armstrong TS, Vera-Bolanos E, Gning I, et al. The impact of symptom interference using the MD Anderson Symptom Inventory-Brain Tumor Module (MDASI-BT) on prediction of recurrence in primary brain tumor patients. *Cancer*. 2011;117(14):3222–3228.
23. Armstrong TS, Vera-Bolanos E, Acquaye A, et al. Impact of recall period on primary brain tumor patient's self-report of symptoms. *Neuro-Oncol Pract*. 2014;1:55.
24. Lai JS, Jensen SE, Beaumont JL, et al. Development of a symptom index for patients with primary brain tumors. *Value Health*. 2014;17(1):62–69.
25. Jakola AS, Gulati S, Weber C, et al. Postoperative deterioration in health related quality of life as predictor for survival in patients with glioblastoma: a prospective study. *PLoS One*. 2011;6(12):e28592.
26. Jakola AS, Gulati M, Gulati S, et al. The influence of surgery on quality of life in patients with intracranial meningiomas: a prospective study. *J Neurooncol*. 2012;110(1):137–144.
27. Tucha O, Smely C, Preier M, et al. Cognitive deficits before treatment among patients with brain tumors. *Neurosurgery*. 2000;47(2):324–333; discussion 333–334.
28. Meyers CA, Hess KR, Yung WK, et al. Cognitive function as a predictor of survival in patients with recurrent malignant glioma. *J Clin Oncol*. 2000;18(3):646–650.
29. Brown PD, Ballman KV, Rummans TA, et al. Prospective study of quality of life in adults with newly diagnosed high-grade gliomas. *J Neurooncol*. 2006;76(3):283–291.
30. Johnson DR, Wefel JS. Relationship between cognitive function and prognosis in glioblastoma. *CNS Oncol*. 2013;2(2):195–201.
31. Gorlia T, van den Bent MJ, Hegi ME, et al. Nomograms for predicting survival of patients with newly diagnosed glioblastoma: prognostic factor analysis of EORTC and NCIC trial 26981-22981/CE.3. *Lancet Oncol*. 2008;9(1):29–38.
32. Wefel JS, Patwardhan S, Strange C. Concurrent and Criterion Validity Related Evidence for the Neurocognitive Function Clinical Trial Battery in Brain Tumor patients. *Paper presented at: the 15th Annual Meeting of the Society for Neuro-Oncology; November 18–21; 2010; Montreal, Quebec*.
33. Johnson DR, Sawyer AM, Meyers CA, et al. Early measures of cognitive function predict survival in patients with newly diagnosed glioblastoma. *Neuro Oncol*. 2012;14(6):808–816.
34. Wefel JS, Cloughesy T, Zazzali JL, et al. Neurocognitive function in patients with recurrent glioblastoma treated with bevacizumab. *Neuro Oncol*. 2011;13(6):660–668.
35. Meyers CA, Smith JA, Bezjak A, et al. Neurocognitive function and progression in patients with brain metastases treated with whole-brain radiation and metexafin gadolinium: results of a randomized phase III trial. *J Clin Oncol*. 2004;22(1):157–165.
36. Wefel JS, Vardy J, Ahles T, et al. International Cognition and Cancer Task Force recommendations to harmonise studies of cognitive function in patients with cancer. *Lancet Oncol*. 2011;12(7):703–708.
37. Lee ST, Park CK, Kim JW, et al. Early cognitive function tests predict early progression in glioblastoma. *Neuro-Oncol Pract*. 2015;2(3):137–143.
38. Meyers CA, Hess KR. Multifaceted end points in brain tumor clinical trials: cognitive deterioration precedes MRI progression. *Neuro Oncol*. 2003;5(2):89–95.
39. Groves MD, Maor MH, Meyers C, et al. A phase II trial of high-dose bromodeoxyuridine with accelerated fractionation radiotherapy followed by procarbazine, lomustine, and vincristine for glioblastoma multiforme. *Int J Radiat Oncol Biol Phys*. 1999;45(1):127–135.
40. Sun A, Bae K, Gore EM, et al. Phase III trial of prophylactic cranial irradiation compared with observation in patients with locally

- advanced non-small-cell lung cancer: neurocognitive and quality-of-life analysis. *J Clin Oncol*. 2011;29(3):279–286.
41. Wolfson AH, Bae K, Komaki R, et al. Primary analysis of a phase II randomized trial Radiation Therapy Oncology Group (RTOG) 0212: impact of different total doses and schedules of prophylactic cranial irradiation on chronic neurotoxicity and quality of life for patients with limited-disease small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2011;81(1):77–84.
 42. Gehring K, Patwardhan SY, Collins R, et al. A randomized trial on the efficacy of methylphenidate and modafinil for improving cognitive functioning and symptoms in patients with a primary brain tumor. *J Neurooncol*. 2012;107(1):165–174.
 43. Brown PD, Pugh S, Laack NN, et al. Memantine for the prevention of cognitive dysfunction in patients receiving whole-brain radiotherapy: a randomized, double-blind, placebo-controlled trial. *Neuro Oncol*. 2013;15(10):1429–1437.
 44. Turner C, Strange C, Bradshaw M, et al. Quality of Life, Mood, and Neurocognitive Functioning in Patients with Temporal Lobe Glioma. *Poster session presented at the 4th Quadrennial Meeting of the World Federation of Neuro-Oncology held in conjunction with the Scientific Meeting and Education Day of the Society for Neuro-Oncology; November 21; 2013; San Francisco, CA*.
 45. Bradshaw ME, Noll K, Ziu M, et al. Neurocognitive Functioning and Functional Status in Patients with Glioma of the Temporal Lobes. *Poster session presented at the 4th Quadrennial Meeting of the World Federation of Neuro-Oncology held in conjunction with the Scientific Meeting and Education Day of the Society for Neuro-Oncology; November 21; 2013; San Francisco, CA*.
 46. Resendiz CV, Armstrong TS, Acquaye A, et al. When Do Patients/ Caregiver Report of Cognitive Symptoms Predict Neurocognitive Dysfunction? *Poster session presented at the 4th Quadrennial Meeting of the World Federation of Neuro-Oncology held in conjunction with the Scientific Meeting and Education Day of the Society for Neuro-Oncology; November 21; 2013; San Francisco, CA*.
 47. Jones LW, Cohen RR, Mabe SK, et al. Assessment of physical functioning in recurrent glioma: preliminary comparison of performance status to functional capacity testing. *J Neurooncol*. 2009;94(1):79–85.
 48. Jeremic B, Milicic B, Grujicic D, et al. Clinical prognostic factors in patients with malignant glioma treated with combined modality approach. *Am J Clin Oncol*. 2004;27(2):195–204.
 49. Peus D, Newcomb N, Hofer S. Appraisal of the Karnofsky Performance Status and proposal of a simple algorithmic system for its evaluation. *BMC Med Inform Decis Mak*. 2013;13:72.
 50. Yates JW, Chalmer B, McKeegney FP. Evaluation of patients with advanced cancer using the Karnofsky performance status. *Cancer*. 1980;45(8):2220–2224.
 51. Mor V, Laliberte L, Morris JN, et al. The Karnofsky Performance Status Scale. An examination of its reliability and validity in a research setting. *Cancer*. 1984;53(9):2002–2007.
 52. Schag CC, Heinrich RL, Ganz PA. Karnofsky performance status revisited: reliability, validity, and guidelines. *J Clin Oncol*. 1984; 2(3):187–193.
 53. Stark AM, Stepper W, Mehdorn HM. Outcome evaluation in glioblastoma patients using different ranking scores: KPS, GOS, mRS and MRC. *Eur J Cancer Care (Engl)*. 2010;19(1):39–44.
 54. Reardon D, Nayak L, DeAngelis L, et al. The Neurologic Assessment in Neuro-Oncology (NANO) scale: a tool to assess neurologic function for integration in the Radiologic Assessment in Neuro-Oncology (RANO) criteria. *Neuro Oncol*. 2014;16(suppl 2): ii76.
 55. Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol*. 2010;28(11):1963–1972.