JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

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

Terri S. Armstrong, Jeffrey S. Wefel, Meihua Wang, Mark R. Gilbert, Minhee Won, Andrew Bottomley, Tito R. Mendoza, Corneel Coens, Maria Werner-Wasik, David G. Brachman, Ali K. Choucair, and Minesh Mehta

Terri S. Armstrong, University of Texas Health Science Center-School of Nursing; Terri S. Armstrong, Jeffrey S. Wefel, Mark R. Gilbert, Tito R. Mendoza, MD Anderson Cancer Center, Houston, TX; Meihua Wang, Minhee Won, Radiation Therapy Oncology Group Statistical Center; Maria Werner-Wasik, Thomas Jefferson University Hospital, Philadelphia, PA; Andrew Bottomley, Corneel Coens, European Organisation for Research and Treatment of Cancer, Brussels, Belgium; David G. Brachman, Arizona Oncology Services Foundation and Barrow Neurological Institute, Phoenix, AZ; Ali K. Choucair, Mayo Clinic, Jacksonville, FL; Minesh Mehta, University of Maryland, Baltimore, MD.

ABSTRACT

Purpose

Radiation Therapy Oncology Group trial 0525 tested whether dose-intensifying temozolomide versus standard chemoradiotherapy improves overall survival (OS) or progression-free survival (PFS) in newly diagnosed glioblastoma. Tests of neurocognitive function (NCF) and symptoms (using the MD Anderson Symptom Inventory–Brain Tumor module; MDASI-BT) and of quality of life (European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire [EORTC QLQ] –C30/BN20) examined the net clinical benefit (NCB) of therapy.

Patients and Methods

NCF tests (Hopkins Verbal Learning Test–Revised, Trail Making Test, and Controlled Oral Word Association), MDASI-BT, and EORTC QLQ-C30/BN20 were completed in a subset of patients. Multivariate Cox proportional hazard regression modeling determined the prognostic value of baseline and early change from baseline to cycle 1 for OS and PFS. Two-sample proportional test statistic was used to evaluate differences between treatments (dose-dense v standard-dose) on NCB measures from baseline to cycle 4 in stable patients.

Results

Overall, 182 patients participated in the study. Baseline NCF tests and the physical functioning quality of life scale were associated with OS and PFS. Baseline to cycle 1 in all NCB components were associated with OS and PFS. There was greater deterioration in the dose-dense arm from baseline to cycle 4 in the Global Health and Motor Function subscales (EORTC QLQ-C30/BN20) as well as in overall symptom burden, overall symptom interference, and activity-related symptom interference subscales (MDASI-BT). There were no between-arm differences in NCF.

Conclusion

Longitudinal collection of NCB measures is feasible in cooperative group studies and provides an added dimension to standard outcome measures. Greater adverse symptom burden and functional interference, as well as decreased global health and motor function were observed in patients randomly assigned to the dose-dense arm. Baseline and early change in NCB measures were associated with decreased rates of survival.

J Clin Oncol 31:4076-4084. © 2013 by American Society of Clinical Oncology

INTRODUCTION

Glioblastoma (GBM) is one of the most malignant primary brain tumors. Most studies show a median survival rate of approximately 14 months, with a treatment course often marked by a decline in overall functional status. Most patients succumb to their illness within 2 years with modern treatments having produced modest survival gains.[1] Because of the impact of the tumor on neurologic and cognitive function, improvements in survival may not be associated with improvements in patient function. As a consequence, traditional end points may not measure the overall impact on the patient. Precedence exists in oncology for measurement of nontherapeutic end points, but no standard approach has

been used to measure this effect in neurooncology. To date, approaches include subjective measures such as quality of life (QOL) and symptom burden, as well as objective neurocognitive testing. Using these end points coupled with the primary impact on disease control will allow the investigator to more fully interrogate the net clinical benefit (NCB) of a treatment on the patient.[2] Radiation Therapy Oncology Group (RTOG) trial 0525 hypothesized that using a dose-dense temozolomide regimen would modulate methylguanine methyltransferase (MGMT) activity, a presumed mechanism of chemotherapy resistance, resulting in improved survival. However, the experimental dose-dense regimen, administered for 21 of 28 days of each cycle, had the potential to either cause more treatment-related toxicity or conversely improve clinical functioning and outcome consequential to improved tumor control. Therefore, this companion analysis was designed to determine the impact of standard-dose versus dose-dense temozolomide on QOL, symptom burden, and neurocognitive function (NCF), providing information to evaluate the NCB of these treatment approaches. Specifically, the study was designed to determine if dose-dense therapy was associated with better QOL, symptoms, and NCF because of better tumor control; if dose-dense therapy slowed expected NCF deterioration owing to tumor progression, or if dose-dense therapy was associated with more or less short- and long-term neurotoxicity. The goal was that this information, coupled with traditional outcome data, could then be used in important risk-benefit considerations for patients and their health care providers.

PATIENTS AND METHODS

Sample

A subset of RTOG 0525 participants who could speak and read English were offered participation in the NCB portion of the study (Fig 1). All patients had histopathologically confirmed GBM (WHO grade 4) with a supratentorial component. Participation in the sub-study was voluntary. All patients provided written informed consent to participate in this institutional review board–approved trial.

Study Design

The NCB assessment schedule is provided in Appendix Figure A1 (online-only). The primary end point of the clinical study was overall
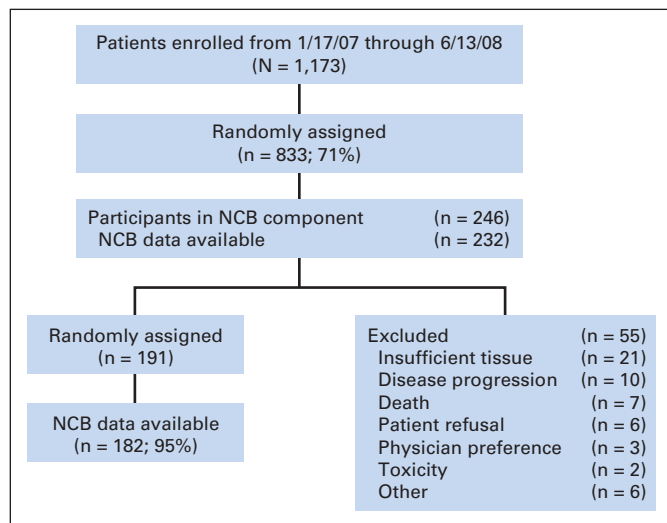

Fig 1. CONSORT diagram. NCB, net clinical benefit.

survival (OS). Symptom and QOL measures were collected at the following timepoints: baseline; before the initiation of cycles 1, 2, 3, 4, 5, 6, and 10 (if administered, or 3 months after completing cycle 6); and 1 month after cycle 12 (if administered, or 6 months after completing cycle 6). NCF data were collected at baseline; before the initiation of cycles 1, 4, 7 (if administered), and 10 (if administered, or 3 months after completing cycle 6); and 1 month after completing cycle 12 (if administered, or 6 months after completing cycle 6).

The primary study objective was to determine whether dose-intensifying adjuvant temozolomide (increasing the dose density) enhances treatment efficacy as measured by overall survival (Gilbert et al[3]). The primary objective of the NCB portion of the study was to compare changes in QOL, symptom burden, and NCF between the two treatment arms in patients whose disease did not progress. In addition, we evaluated the ability of baseline NCB measures and deterioration in NCB measures after concurrent chemoradiotherapy to predict progression-free and overall survival (see text box for the prespecified NCB study objectives).

Outcome Measures

Patient-reported outcome measures were used to measure QOL and symptoms, and objective NCF testing was used to assess cognition in our study.

QOL. Health-related quality of life was assessed using the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire C30/BN20 (EORTC-QLQ-C30/BN20),[4,5] developed and validated for use in a cancer patient population; the BN20, in particular, was developed as a module for patients with brain cancer. QLQ-C30 is a 30-item self-report questionnaire in which patients rate the items on a four-point scale, with 1 as "not at all" and 4 as "very much." The measure produces five domain scores including physical, role, emotional, cognitive, and social functioning, as well as nine single-item individual symptom scales (fatigue, pain, nausea, vomiting, dyspnea, insomnia, anorexia, constipation, and diarrhea) and financial impact. The BN20 consists of four domain scores, including future uncertainty, visual disorder, motor dysfunction, and communication deficit, as well as seven individual symptom items (headache, seizures, drowsiness, hair loss, itching, difficulty with bladder control, and weakness of both legs). Reference and normative data exist for this measure and it was scored using the standardized recommended approach of the EORTC.[6,7]

Symptom assessment. Symptom burden and interference was assessed using the MD Anderson Symptom Inventory Brain Tumor (MDASI-BT),[8] which was developed and validated for use in this patient population. It consists of 23 symptoms rated on an 11-point scale (0 to 10) to indicate the presence and severity of the symptom, with 0 as "not present" and 10 "as bad as you can imagine." Each symptom is rated at its worst level in the last 24 hours. Symptoms included on the instrument include those commonly associated with cancer therapies, those associated with increased intracranial pressure, and those related to focal deficits. The symptom items on the MDASI-BT were averaged to generate a mean symptom severity score (ie, symptom burden), and they were grouped into six previously identified factors: affective, cognitive, neurologic, treatment-related, generalized/disease, and gastrointestinal-related.[8] Seizure was removed from the neurologic factor score for this analysis because of the low frequency of reporting by patients in this study. The MDASI-BT also includes ratings of how symptoms interfered with different aspects of a patient's life in the preceding 24 hours. The interference items are measured on a 0 to 10 scale and include: general activity, mood, work (includes both work outside the home and housework), relationships with other people, walking, and enjoyment of life. The interference items are averaged to generate a mean symptom interference score (ie, symptom interference) and are grouped into two types: activity-related (general activity, work, and walking) and mood (mood, relations with other people, and enjoyment of life).

NCF. Neurocognitive function was assessed using the Hopkins Verbal Learning Test–Revised (HVLT-R; tests for Total Recall, Delayed Recall, and Delayed Recognition),[9] the Trail Making Test (Part A and Part B),[10] and the Controlled Oral Word Association (COWA)[11] The NCF tests were administered by a health care professional (eg, nurse, psychologist) who was trained and certified by the study neuropsychologist (J.S.W.). The tests have published normative data that take into account age and, when appropriate, education and gender. The tests in the neurocognitive battery were selected because they

are widely used standardized psychometric instruments that have been shown to be sensitive to the impact of cancer and the neurotoxic effects of cancer treatment in other clinical trials.[12-14]

Statistical Methods

Participation in the NCB sub-study was not mandatory. Therefore to determine whether the results would be generalizable to the entire study population, patient characteristics were compared between NCB sub-study participants and nonparticipants via χ^2 tests. NCB sub-study participant characteristics were also compared between two treatment arms using χ^2 tests. Only patients without disease progression were required to finish the NCB measures at protocol-specified assessment time points. A two-sample z test was used to compare the percentage of received forms between the two treatment arms at baseline, and before cycles 1, 4, and 10. QOL, symptom, and NCF data collected at baseline and before cycle 4 were used for the primary analysis.

Clinically meaningful change. For the MDASI-BT, a change in symptom severity of one point was classified as the minimum clinically meaningful change. A decline for an individual patient was calculated as an increase of one point or more in the score from the baseline measurement to before cycle 4. For the EORTC QLQ-C30/BN20, differences of 10 points from the baseline measurement to before cycle 4 based on the standardized score (range, 0 to 100) were classified as the minimum clinically meaningful change in QOL.[15] At cycle 4, change in raw test scores relative to baseline was calculated for each NCF test. Neurocognitive status was categorized as improved, stable, or declined, using the Reliable Change Index (RCI)[16] for each test. The RCI is derived from the SE of measurement of each test and represents the 90% CI for the difference in raw score from baseline to the next assessment that would be expected if no real change occurred:

$$RCI = 1.64(SEdiff), \text{ where } SEdiff = [2(SEM^2)]^{1/2}$$

$$\text{and } SEM = SD_1[(1 - r_{xy})^{1/2}] \quad (1)$$

SEdiff is the SE of difference, SEM is the SE of measurement, SD_1 is the standard deviation, and r_{xy} is the test-retest reliability statistic. All RCI thresholds were rounded to the nearest whole number. Changes that did not meet the RCI criterion for decline were categorized as stable/improved. The RCI values were determined from published studies[11,17,18] and have been previously reported.[14]

Between-arm differences. For the analysis related to between-arm differences, two-sample proportion tests were used to compare proportions of deterioration (the percentages of patients with clinically meaningful worsening before cycle 4).

Prediction of OS and progression-free survival. Individual baseline scores and early change scores (ie, baseline to before cycle 1) were evaluated for association with progression-free survival (PFS) and OS using Cox proportional hazards models. The standardized scores for the neurocognitive tests were used at baseline.[10,11,20] OS was defined as the interval from random assignment to patient death as a result of any cause, and PFS was defined as the interval from random assignment to disease progression or death, whichever occurred first. Those items with a P value less than .1 in the univariate analysis were then included in the multivariate analysis with step-wise selection. recursive partitioning analysis (RPA) class and MGMT were forced to remain in the multivariate model. NCB items with a P value less than .05 on multivariate analysis were deemed to be statistically significant and are reported in this article.

All testing was done at the overall significance level of .05 and there were no multiple comparison adjustments owing to the exploratory purpose of this analysis. Statistical Analysis System (SAS Institute, Cary, NC) was used to perform these analyses.

RESULTS

The primary study opened for accrual on January 17, 2006, and closed to accrual on June 13, 2008, with 1,173 patients enrolled. Among registered patients, 1,125 patients (96%) were eligible for the concur-

Table 1. Pretreatment Characteristics Since NCB Component Available by Randomized Versus Nonparticipating Patients

Characteristic	Randomized (n = 182) No. of Patients	%	Nonparticipants (n = 337) No. of Patients	%	P
Age, years					.81
Median	58		57		
Min-max	21-84		22-84		
Q1-Q3	49-66		50-64		
< 50	46	25.3	82	24.3	
≥ 50	136	74.7	255	75.7	
Sex					.09
Male	94	51.6	200	59.3	
Female	88	48.4	137	40.7	
Race					.20*†
American Indian/Alaska Native	1	0.5	0	0.0	
Asian	2	1.1	3	0.9	
Black or African American	7	3.8	1	0.3	
Native Hawaiian or Other Pacific Islander	0	0.0	1	0.3	
White	170	93.4	205	60.8	
More than one race	0	0.0	2	0.6	
Unknown or not reported	2	1.1	125	37.1	
Ethnicity					.25*
Hispanic or Latino	5	2.7	11	3.3	
Not Hispanic or Latino	159	87.4	188	55.8	
Unknown (individuals not reporting ethnicity)	18	9.9	138	40.9	
Education					.26*
≤ 12 years	61	33.5	80	23.7	
> 12 years	107	58.8	117	34.7	
Other	9	4.9	5	1.5	
Prefers not to answer/unknown	5	2.7	135	40.1	
KPS					.04
90-100	128	70.3	207	61.4	
Surgery					.001
Biopsy	3	1.6	9	2.7	
Partial resection	64	35.2	172	51.0	
Total resection	115	63.2	156	46.3	
Neurologic function					.28
No symptoms	62	34.1	111	32.9	
Minor symptoms	86	47.3	143	42.4	
Moderate symptoms	33	18.1	80	23.7	
Severe symptoms	1	0.5	3	0.9	
RPA class					.24
III	37	20.3	57	16.9	
IV	116	63.7	207	61.4	
V	29	15.9	73	21.7	
Lateralization of tumor					.23
Right side only	104	57.1	172	51.0	
Left side only	77	42.3	156	46.3	
Bilateral	1	0.5	5	1.5	
Unknown	0	0.0	4	1.2	

Abbreviations: KPS, Karnofsky performance score; max, maximum; min, minimum; NCB, net clinical benefit; RPA, recursive partitioning analysis.
*P values are based on the χ^2 test.
†White versus other race.

rent radiation and temozolomide stage and 48 patients (4%) were ineligible. A total of 833 patients (74%) were randomly assigned for the adjuvant temozolomide stage. The NCB component became available on July 12, 2007. Of the 569 patients subsequently enrolled to

the primary study, 246 patients (43%) agreed to participate in the NCB component and 232 patients (94%) have at least one NCB assessment available. Among the randomly assigned patients, 182 patients have NCB data available (Fig 1).

Sample Characteristics

Patient characteristics for NCB sub-study participants and non-participants included in this analysis are listed in Table 1. The patient characteristics were balanced except for sex ($P = .09$), the surgery type ($P = .008$), and Karnofsky performance score ($P = .04$). There was no statistically significant difference on important stratification factors including RPA class ($P = .76$) and MGMT ($P = .27$). Patient characteristics between two treatment arms for all of the randomly assigned patients with NCB data are listed in Table 2. The patient characteristics were balanced between these two treatment arms.

Table 3 lists the compliance to NCF and QOL assessments at baseline and before cycles 1, 4, and 10. At least 85% of patients completed a baseline assessment on QOL, MDASI-BT, and NCF measures. There were no statistically significant differences in compliance on these instruments between the two arms, except for QOL and MDASI-BT at the assessment performed before cycle 4 ($P = .014$ and $P = .02$, respectively). The compliance rates were lower in the dose-dense arm than in the standard-dose arm for the QOL and MDASI-BT instruments at the assessment performed before cycle 4.

Differences in Deterioration Status

Detecting between-arm differences is an important component of assessing the impact of a new therapy versus a standard approach. We were interested in exploring differences in deterioration status from baseline to the assessment before cycle 4 for each component of the NCB assessment separately. Complete results are provided in Table 4.

QOL. There were no statistically significant differences in the rates of deterioration between the two treatment arms, except for the EORTC QLQ-C30 global health status factor ($P = .03$) and for the BN-20 motor dysfunction item ($P = .014$), although a trend for greater decline in patient-reported QOL in the dose-dense arm was noted.

Symptom burden. There were statistically significant differences in the rates of deterioration between the two treatment arms in both mean symptom burden and symptom interference. For symptom interference, there was a statistically significant difference in the frequency of deterioration between the two treatment arms for the activity-related and work and walking factors, but no statistically significant difference for the mood-related factor.

NCF. There were no statistically significant differences in the frequency of deterioration on any of the NCF tests between the two treatment arms.

Prediction of Overall Survival

We examined baseline QOL, symptoms, and NCF as well as changes in these measures after concurrent chemoradiotherapy as predictors of OS. Measures that were significant at the $P \leq .1$ level on univariate analyses were retained for inclusion in multivariate analysis. At baseline, QOL and NCF measures yielded prognostic information regarding overall survival time. Specifically, the EORTC physical functioning scale and nausea/vomiting items, the HVLT-R Delayed

Table 2. Pretreatment Characteristics for Randomly Assigned Patients Participating in NCB Component

Characteristic	Arm 1 (n = 92)		Arm 2 (n = 90)		P
	No. of Patients	%	No. of Patients	%	
Age, years					.80
Median	57		59		
Min-Max	23-84		21-79		
Q1-Q3	49-67.5		50-66		
< 50	24	26.1	22	24.4	
≥ 50	68	73.9	68	75.6	
Sex					.10
Male	53	57.6	41	45.6	
Female	39	42.4	49	54.4	
Race					.25*
American Indian/Alaska Native	1	1.1	0	0.0	
Asian	1	1.1	1	1.1	
Black or African American	5	5.4	2	2.2	
White	84	91.3	86	95.6	
Unknown or not reported	1	1.1	1	1.1	
Ethnicity					.24
Hispanic or Latino	4	4.3	1	1.1	
Not Hispanic or Latino	77	83.7	82	91.1	
Unknown (individuals not reporting ethnicity)	11	12.0	7	7.8	
Education					.40
≤ 12 years	32	34.8	29	32.2	
> 12 years	53	57.6	54	60.0	
Other	6	6.5	3	3.3	
Prefers not to answer/unknown	1	1.1	4	4.4	
KPS					.58
60-80	29	31.5	25	27.8	
90-100	63	68.5	65	72.2	
Surgery					.20
Biopsy	0	0.0	3	3.3	
Partial resection	32	34.8	32	35.6	
Total resection	60	65.2	55	61.1	
Neurologic function					.32
No symptoms	31	33.7	31	34.4	
Minor symptoms	40	43.5	46	51.1	
Moderate symptoms	21	22.8	12	13.3	
Severe symptoms	0	0.0	1	1.1	
MGMT status					.65
Methylated	28	30.4	30	33.3	
Unmethylated	52	56.5	52	57.8	
Unknown (indeterminate, invalid)	12	13.0	8	8.9	
RPA class					.85
III	18	19.6	19	21.1	
IV	58	63.0	58	64.4	
V	16	17.4	13	14.4	
Lateralization of tumor					.41†
Right side only	55	59.8	49	54.4	
Left side only	36	39.1	41	45.6	
Bilateral	1	1.1	0	0.0	

Abbreviations: KPS, Karnofsky performance score; max, maximum; MGMT, methylguanine methyltransferase; min, minimum; NCB, net clinical benefit; RPA, recursive partitioning analysis.
*White versus other race.
†Right versus left.

Table 3. Symptom Burden, QOL, and Neurocognitive Assessment Compliance

Evaluation Status by Measure and Treatment Arm	Baseline		Cycle 1		Cycle 4	
	No. of Patients	%	No. of Patients	%	No. of Patients	%
QLQ-C30/BN20						
Standard dose						
Expected	96		96		77	
Received	85	89	73	76	52	68
Dose dense						
Expected	95		94		80	
Received	87	92	74	79	39	49
Difference		3		3		19
P*		.48		0.62		.014
MDASI-BT						
Standard dose						
Expected	96		96		77	
Received	87	91	73	76	52	68
Dose dense						
Expected	95		94		80	
Received	86	91	73	78	40	50
Difference		0		2		12
P*		1.0		.74		.02
HVLT-R						
Standard dose						
Expected	96		96		77	
Received	83	86	68	71	41	53
Dose dense						
Expected	95		94		80	
Received	86	91	64	68	38	48
Difference		5		3		5
P*		.28		.65		.53
TMT part A						
Standard dose						
Expected	96		96		77	
Received	84	88	71	74	41	53
Dose dense						
Expected	95		94		80	
Received	86	91	63	67	38	48
Difference		3		7		5
P*		.50		.29		.53
TMT part B						
Standard dose						
Expected	96		96		77	
Received	82	85	67	70	39	51
Dose dense						
Expected	95		94		80	
Received	85	89	61	65	37	46
Difference		4		5		5
P*		.41		.46		.53
COWA						
Standard dose						
Expected	96		96		77	
Received	83	86	68	71	41	53
Dose dense						
Expected	95		94		80	
Received	83	87	63	67	39	49
Difference		1		4		4
P*		.84		.55		.62

NOTE. Patients who discontinued therapy and did not complete assessments were considered noncompliant.
Abbreviations: COWA, Controlled Oral Word Association; HVLT-R, Hopkins Verbal Learning Test–Revised; TMT, Trail Making Test; QOL, quality of life; QLQ-C30/BN20, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30/BN20.
*P value is from the two-sample z test to compare the percentage of received forms between the two arms.

Table 4. Between Arm Differences in Deterioration of QOL, Symptoms, and NCF

Test/Measure and Component	Standard-Dose Deterioration		Dose-Dense Deterioration		
	No. of Patients	%	No. of Patients	%	P^*
EORTC QOL C30					
Scale					
Global health status/QOL	11	23	17	45	.03*
Physical	12	24	14	36	.25
Role	12	24	14	36	.25
Emotional	6	13	11	28	.07
Cognitive	10	21	15	38	.07
Symptom items					
Social	11	23	14	36	.18
Fatigue	12	24	17	44	.06
Nausea/vomiting	8	16	12	31	.11
Pain	8	16	9	23	.43
Dyspnea	8	16	11	28	.18
Insomnia	5	10	7	18	.32
Appetite loss	16	33	13	33	.95
Constipation	8	16	10	26	.29
Diarrhea	8	17	5	13	.58
Financial	9	19	6	17	.77
BN20					
Scale					
Future uncertainty	3	6	5	14	.26
Visual disorder	15	31	6	16	.09
Motor dysfunction	10	20	17	45	.014*
Communication deficit	15	31	12	32	.92
Symptom items					
Headaches	5	10	9	24	.09
Seizures	3	6	2	5	.86
Drowsiness	15	31	13	34	.77
Hair loss	7	15	11	31	.09
Itchy skin	13	27	12	34	.45
Weak legs	7	14	9	24	.27
Bladder	2	4	5	13	.14
MDASI-BT					
Summary scores					
Symptom burden	5	10	11	27	.03*
Symptom interference	7	14	13	32	.03*
Work and walking	8	16	15	39	.01*
Activity-related	8	16	15	38	.02*
Mood-related	12	24	12	30	.49
Factor groups					
Affective	7	14	12	30	.06
Cognitive	9	18	11	28	.27
Neurologic	7	14	9	23	.28
Treatment-related	13	25	16	40	.14
Generalized/disease	6	12	8	20	.27
GI-related	4	8	7	18	.19
NCF					
HVLT-R					
Total Recall	11	27	11	30	.78
Delayed Recall	8	21	8	24	.76
Delayed Recognition	10	24	7	19	.60
TMT					
Part A	8	20	11	29	.35
Part B	10	26	11	30	.69
COWA	3	7	6	16	.25
CTB composite†	11	29	10	29	.99

Abbreviations: COWA, Controlled Oral Word Association; CTB, clinical trial battery; EORTC QOL C30, European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire C30; HR, hazard ratio; HVLT-R, Hopkins Verbal Learning Test–Revised; MDASI-BT, MD Anderson Symptom Inventory–Brain Tumor module; NCF, neurocognitive function; TMT, Trail Making Test; QOL, quality of life.
*From two-sample proportional test statistic comparing the percentage of people who deteriorated since baseline.
†CTB composite score is the mean of standardized scores (HVLT-R–Total Recall, HVLT-R–Delayed Recall, HVLT-R–Delayed Recognition, TMT parts A and B, and COWA).

Table 5. Prediction of Overall Survival

	Multivariate Analysis		
Variable	P	HR	95% CI
Baseline, continuous			
EORTC, physical functioning scale	.029	0.99	0.98 to 1.00
EORTC, nausea/vomiting item	.030	1.02	1.00 to 1.04
HVLT-R, delayed recognition	.022	0.87	0.77 to 0.98
COWA	.010	0.81	0.69 to 0.95
Early change, deterioration			
EORTC, cognitive functioning scale	.004	1.95	1.23 to 3.09
EORTC, motor dysfunction item	.041	1.59	1.02 to 2.47
EORTC, hair loss item	.039	0.57	0.33 to 0.97
MDASI-BT, cognitive factor	.012	1.82	1.14 to 2.89
HVLT-R, Total Recall	.013	1.90	1.14 to 3.15
TMT, part B	.003	2.11	1.29 to 3.44

Abbreviations: COWA, Controlled Oral Word Association; EORTC, European Organisation for the Research and Treatment of Cancer; HR, hazard ratio; HVLT-R, Hopkins Verbal Learning Test–Revised; MDASI-BT, MD Anderson Symptom Inventory–Brain Tumor module; TMT, Trail Making Test.

Recognition, and COWA were prognostic for OS. No baseline MDASI-BT factors were associated with overall survival time.

After concurrent chemoradiotherapy, deterioration on QOL, symptom, and NCF measures were associated with inferior OS. Specifically, deterioration in the MDASI-BT cognitive factor, EORTC cognitive functioning scale, EORTC motor dysfunction scale, EORTC bothered-by-hair-loss item, HVLT-R Total Recall, and Trail Making Test Part B were associated with shorter overall survival time (Table 5).

Prediction of Progression-Free Survival

We examined baseline QOL, symptoms, and NCF as well as changes in these measures after concurrent chemoradiotherapy as predictors of progression-free survival (Table 6). At baseline, QOL and NCF measures yielded prognostic information regarding progression-free survival time. Specifically, the EORTC physical functioning scale, EORTC constipation item, and the COWA were prognostic for progression-free survival time. No baseline MDASI-BT factors were associated with PFS.

After concurrent chemoradiotherapy, deterioration on QOL, symptom, and NCF measures were associated with inferior PFS.

Table 6. Prediction of Progression-Free Survival

	Multivariate Analysis		
Variable	P	HR	95% CI
Baseline, continuous			
EORTC, physical functioning scale	.009	0.987	0.977 to 0.997
EORTC, constipation item	.012	0.990	0.981 to 0.998
COWA	.007	0.824	0.716 to 0.949
Early change, deterioration			
EORTC, motor dysfunction item	.005	1.74	1.18 to 2.55
MDASI-BT neurologic factor	.007	1.87	1.19 to 2.94
HVLT-R Delayed Recall	.030	1.71	1.06 to 2.77

Abbreviations: COWA, Controlled Oral Word Association; EORTC, European Organisation for the Research and Treatment of Cancer; HR, hazard ratio; HVLT-R, Hopkins Verbal Learning Test–Revised; MDASI-BT, MD Anderson Symptom Inventory–Brain Tumor module.

Net Clinical Benefit Objectives

- Primary

- To compare between the two treatment arms the symptom burden, neurocognitive function (NCF), and health-related quality of life (HRQOL) in patients whose disease has not progressed 6 months after adjuvant therapy (6-month progression-free survival).

- To evaluate midcycle differences in symptom burden and HRQOL in patients on the two arms at day 14 of course 1 and course 4.

- Secondary

- To evaluate longitudinal changes in HRQOL measures and determine the impact of dose-intense chemotherapy on these parameters.

- To measure symptom burden and degree of interference over the course of therapy to evaluate differences between patients' individual symptom severity, overall mean symptom severity, and difference in scores on the interference items between the two treatment arms.

- To describe the association between quality of life as measured by the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire C30/BN20 (EORTC-QL30/BCM20) and mean symptom severity as measured by the MD Anderson Symptom Inventory–Brain Tumor module (MDASI-BT) in patients enrolled onto this study.

- To describe the variability of symptom severity across the epoch and follow-up period to compare differences between the two treatment arms.

- To evaluate these instruments as a useful composite measurement of the impact of treatment and disease response in analysis of efficacy.

- To evaluate differences in longitudinal changes on measures of NCF associated with dose-intense chemotherapy.

- To evaluate the relationship between self-reported cognitive dysfunction and NCF testing.

Specifically, deterioration on the MDASI-BT neurologic factor, EORTC motor dysfunction score, and on the HVLT-R Delayed Recall measure were associated with shorter progression-free survival time.

DISCUSSION

RTOG 0525 is the first large, randomized phase III trial in GBM to incorporate a multidimensional battery of measures to prospectively assess patient-oriented outcomes, including cognition, symptoms, and quality of life. The net clinical benefit sub-study was initiated 18 months after the primary trial was activated and yet still succeeded in reaching its accrual goals. Baseline data were collected on more than 80% of patients, supporting the feasibility of this approach.

Patients randomly assigned to the dose-dense arm reported experiencing more severe symptoms and overall and activity-related interference on the MDASI-BT and worse global quality of life and motor dysfunction on the EORTC QLQ-C30/BN20 from baseline to before cycle 4, which is likely a reflection of increased toxicities reported with the dose-dense treatment. There were no between-arm differences in cognitive function from baseline to before cycle 4, potentially commensurate with the absence of between-arm differences in treatment efficacy and tumor control.

Analysis of the baseline NCB measures as independent predictors of survival after accounting for traditional prognostic factors including RPA class and MGMT status demonstrated that worse memory, executive function, physical functioning, and nausea/vomiting were associated with shorter overall survival. Our results are similar to other articles recently published that support the prognostic importance of select symptoms assessed by the EORTC measure in brain cancer.[5,21,22]

Decline after concurrent chemoradiotherapy in memory, executive function, subjective cognitive function, motor function, and increased bother with hair loss was associated with shorter overall survival. Similarly, worse baseline executive function, physical function, and constipation were independently associated with shorter progression-free survival. Decline after concurrent chemoradiotherapy in memory and neurologic function was associated with shorter progression-free survival.

Employing multidimensional outcome measures as in RTOG 0525 is feasible in large cooperative group clinical trials and provides complimentary information to standard outcome measures. Our study demonstrated that overall and progression-free survival outcomes are predicted by early deterioration of patient function, supporting the utility of these measures and suggesting that this test battery may have profound implications for clinical trial design. These assessments are not without their limitations. Whereas select components of baseline QOL were associated with both PFS and OS, there were no significant between-arm differences in QOL, despite differences in symptom burden and interference. This may represent the effect of a response shift of the patient report, leading to an absence of a reported decline in QOL despite more severe symptoms. Though there were global differences in symptom burden and interference, baseline symptom reporting was not associated with PFS and OS. Neurocognitive testing was associated with the impact of tumor on cognitive function, prognostic for both PFS and OS, but differential treatment effects were not seen. Finally, completion rates declined over time, primarily a consequence site error. This limitation affects the assessment of longitudinal treatment-related changes in patient outcomes. Strategies to encourage data collection and completion similar to traditional measures are warranted.

Although future trials incorporating the NCB measures are needed to validate these findings, our results strongly support making the NCB measures a component of brain tumor clinical trials. These NCB measures can potentially be used for stratification and also provide a dimension of treatment efficacy that is not adequately determined by traditional measures.

AUTHOR CONTRIBUTIONS

Conception and design: Terri S. Armstrong, Jeffrey S. Wefel, Meihua Wang, Mark R. Gilbert, Andrew Bottomley, Tito R. Mendoza, Corneel Coens, Ali K. Choucair, Minesh Mehta
Provision of study materials or patients: David G. Brachman
Collection and assembly of data: Terri S. Armstrong, Jeffrey S. Wefel, Meihua Wang, Minhee Won, Maria Werner-Wasik, David G. Brachman
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Stupp R, Mason WP, van den Bent MJ, et al: Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 352:987-996, 2005

2. Taphoorn MJ, Sizoo EM, Bottomley A: Review on quality of life issues in patients with primary brain tumors. Oncologist 15:618-626, 2010

3. Gilbert MR, Wang M, Aldape KD, et al: Dose-dense temozolomide for newly diagnosed glioblastoma: A randomized phase III clinical trial. J Clin Oncol 31:4085-4091, 2013

4. Efficace F, Bottomley A: Health related quality of life assessment methodology and reported outcomes in randomised controlled trials of primary brain cancer patients. Eur J Cancer 38:1824-1831, 2002

5. 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 46:1033-1040, 2010

6. Scott NW, Fayers PM, Aaronson NK, et al: Reference Values Manual. Brussels, Belgium, European Organisation for Research and Treatment of Cancer, 2008

7. Fayers PM, Aaronson NK, Bjordal K, et al: The EORTC QLQ-C30 Scoring Manual (3rd ed). Brussels, Belgium, European Organisation for Research and Treatment of Cancer, 2001

8. Armstrong TS, Mendoza T, Gning I, et al: Validation of the MD Anderson Symptom Inventory Brain Tumor Module (MDASI-BT). J Neurooncol 80:27-35, 2006

9. Benedict RH, Schretlen D, Groninger L, et al: Hopkins Verbal Learning Test-Revised: Normative data and analysis of inter-form and test-retest

reliability. Clin Neuropsychol 12, 1998

10. Tombaugh TN: Trail Making Test A and B: Normative data stratified by age and education. Arch Clin Neuropsychol 19:203-214, 2004

11. Ruff RM, Light RH, Parker SB, et al: Benton Controlled Oral Word Association Test: Reliability and updated norms. Arch Clin Neuropsychol 11:329-338, 1996

12. Wefel JS, Saleeba AK, Buzdar AU, et al: Acute and late onset cognitive dysfunction associated with chemotherapy in women with breast cancer. Cancer 116:3348-3356, 2010

13. Meyers CA, Smith JA, Bezjak A, et al: Neurocognitive function and progression in patients with brain metastases treated with whole-brain radiation and motexafin gadolinium: Results of a randomized phase III trial. J Clin Oncol 22:157-165, 2004

14. Wefel JS, Cloughesy T, Zazzali JL, et al: Neurocognitive function in patients with recurrent glioblastoma treated with bevacizumab. Neuro Oncol 13:660-668, 2011

15. Maringwa J, Quinten C, King M, et al: Minimal clinically meaningful differences for the EORTC QLQ-C30 and EORTC QLQ-BN20 scales in brain cancer patients. Ann Oncol 22:2107-2112, 2011

16. Jacobson NS, Truax P: Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 59:12-19, 1991

17. Woodard JL, Benedict RH, Salthouse TA, et al: Normative data for equivalent, parallel forms of the Judgment of Line Orientation Test. J Clin Exp Neuropsychol 20:457-462, 1998

18. Levine AJ, Miller EN, Becker JT, et al: Normative data for determining significance of test-retest differences on eight common neuropsychological instruments. Clin Neuropsychol 18:373-384, 2004

19. Reference deleted

20. Shapiro AM, Benedict RH, Schretlen D, et al: Construct and concurrent validity of the Hopkins Verbal Learning Test-revised. Clin Neuropsychol 13:348-358, 1999

21. Quinten C, Maringwa J, Gotay CC, et al: Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival. J Natl Cancer Inst 103:1851-1858, 2011

22. Gotay CC, Kawamoto CT, Bottomley A, et al: The prognostic significance of patient-reported outcomes in cancer clinical trials. J Clin Oncol 26:1355-1363, 2008

■ ■ ■

Resource for Fellows

Oncology fellows can interact with colleagues and peers by subscribing to ASCO's fellows listserve. The fellows listserve is an unmoderated online discussion, which means that all replies to an e-mail are posted immediately to the entire group of subscribers. There is no better way to keep up with your peers and ask those tough questions than ASCO's fellows listserve. To subscribe, please email **ListserveAdmin@asco.org.** For more information, visit **asco.org** and click the Education & Training tab, Resources for Fellows, ASCO Fellows Listserve.

ASCO
American Society of Clinical Oncology
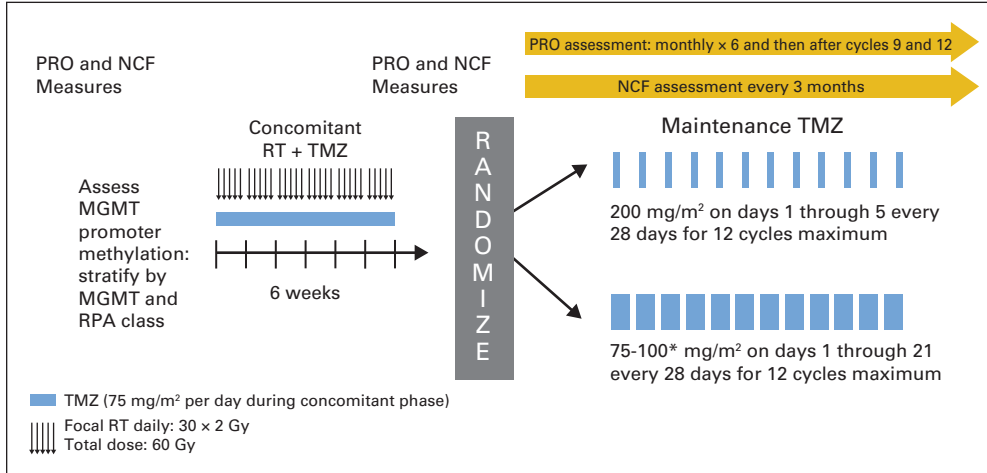
Acknowledgment

Appendix

Figure A1. Study schema. (*) increased after cycle 1 if no myelotoxicity. MGMT, methylguanine methyltransferase; NCB, net clinical benefit; NCF, neurocognitive function; PRO, patient-reported outcomes; RPA, recursive partitioning analysis; RT, radiation therapy; TMZ, temozolomide.