

Neuro-Oncology Practice

Neuro-Oncology Practice 1(2), 55–63, 2014
doi:10.1093/nop/npu006
Advance Access date 5 May 2014

Impact of recall period on primary brain tumor patient's self-report of symptoms

Terri S. Armstrong, Elizabeth Vera-Bolanos, Alvina Acquaye, Mark R. Gilbert, and Tito R. Mendoza

Department of Family Health, University of Texas Health Science Center School of Nursing, Houston, Texas (T.S.); Department of Neuro-Oncology, The University of Texas: MD Anderson Cancer Center, Houston, Texas (T.S., E.V., A.A., M.R.G.); Department of Symptom Research, The University of Texas: MD Anderson Cancer Center, Houston, Texas (T.R.M.)

Corresponding Author: Terri S. Armstrong, PhD, RN, ANP-BC, FAANP, 6901 Bertner Ave, Houston, TX 77030 (terri.s.armstrong@uth.tmc.edu).

Background. Understanding the severity of symptoms is an integral part of patient care. The MD Anderson Symptom Inventory-Brain Tumor (MDASI-BT) was developed using a 24-hour recall period. The choice of recall period is dependent on the treatment and disease of interest. The aim of this study was to evaluate the congruence and equivalency of 24-hour and 7-day symptom reporting using the MDASI-BT.

Method. Adult brain tumor patients completed the MDASI-BT using 24-hour and 7-day recall periods and a tablet format. Equivalence and congruence were determined using equivalency testing and Bland-Altman analysis. Reliability and known group's validity were then assessed by use of Cronbach's alpha and evaluating differences based on performance status.

Results. One hundred patients (mean age, 48 y; range 19 y–77 y), who were primarily white (86%) males (62%) with a variety of brain tumors, most commonly glioblastoma (69%), participated. KPS scores ranged from 50%–100%, with 28% of participants scoring 80% or lower. Overall severity reporting using the 7-day recall was congruent and equivalent with the 24-hour rating, with difference scores of one point or less on the overall instrument and individual symptoms. The 7-day recall period instrument demonstrated psychometric properties similar to the established 24-hour recall instrument.

Conclusion. This study supports the use of the 7-day recall period in addition to the 24-hour recall period for symptom reports of patients with primary brain tumors. Future studies should continue to explore the reliability and validity of this recall period and its utility in other central nervous system tumor populations.

Keywords: patient reported outcomes (PRO), primary brain tumor, symptom burden.

Primary brain tumor patients experience symptoms related to both the tumor and the effects of therapy. These symptoms can impact functional status and can vary over the disease epoch in relation to disease status and therapy received. Patient-reported outcomes (PROs) are questionnaires that allow patients to report the occurrence and severity of their symptoms. PRO use is important because it allows reporting of symptoms that may only be known by the patient and avoids the filtering of this information by a clinician.¹

Most PROs ask patients to rate their symptoms over either a 7-day or a 24-hour recall period. The choice of recall period is dependent on the disease and treatment characteristics to be tested.^{1,2} In research, the clinical trial design and frequency of testing are important considerations for determining the best recall period to use. Whereas a 24-hour recall period allows for

more frequent testing, PROs may be administered less frequently in the clinical trial setting as a consequence of issues such as administrative burden.²

The MD Anderson Symptom Inventory-Brain Tumor (MDASI-BT) is a PRO multisymptom assessment tool that was developed specifically for symptoms that occur in patients with primary brain tumors. The use and validation of the MDASI-BT have been established using a 24-hour recall period.^{3,4} The purpose of this study was to evaluate the congruence and equivalency of 24-hour and 7-day symptom reporting using the MDASI-BT and to establish psychometric properties of the 7-day recall period. Whereas congruency evaluates whether the ratings of symptom severity between the 2 recall periods are in agreement, equivalency testing evaluates whether the difference in ratings is small enough to be clinically unmeaningful and for the results to be essentially similar.

Received 6 January 2014

© The Author(s) 2014. Published by Oxford University Press on behalf of the Society for Neuro-Oncology. All rights reserved.
For permissions, please e-mail: journals.permissions@oup.com.

Methods

Patients and Procedures

Patients diagnosed with a primary brain tumor and being seen for initial consultation or ongoing clinical management were approached in the Brain Tumor Clinic at The University of Texas MD Anderson Cancer Center in Houston and invited to participate. Patients were eligible for the study if they (i) were at least 18 years old, (ii) had a pathologically proven diagnosis of a primary brain tumor, (iii) spoke English, and (iv) had signed a study-specific informed consent prior to study entry. Patients were not eligible to participate if they were (i) unable to understand the symptom assessment questionnaire in English due to language or cognitive deficits, (ii) unwilling to participate, or (iii) unable to comply with protocol requirements. The MD Anderson Cancer Center Institutional Review Board approved the study. Participants were asked to rate their symptoms during one of their regularly scheduled clinic visits. All participants first completed a symptom tool with a 24-hour recall and then immediately completed the same tool with a 7-day recall. Both instruments were presented on an electronic tablet for data entry.

Assessment Measures

To assess symptoms, we used the brain tumor module of the MD Anderson Symptom Inventory (MDASI-BT). The MDASI-BT is a brief, psychometrically validated multisymptom assessment tool that assesses 22 symptoms commonly associated with brain tumors or its treatment.⁴ There are also 6 interference items, designed to measure the impact of symptom severity on the patient. It provides clinicians with pertinent, easily attainable information to help guide patient-specific evaluations and interventions. All MDASI-BT symptom items are rated on numeric 0-to-10 scales from “not present” (0) to “as bad as you can imagine,” (10), and the MDASI-BT interference items are rated on numeric 0-to-10 scales from “did not interfere” (0) to “interfered completely” (10). Each MDASI-BT took an average of 4 minutes to complete.

Overall symptom burden and component scores, namely gastrointestinal (GI), neurological, and cognitive scores, were calculated for comparison between the 2 reporting periods. These scores correspond to the items comprising component scores identified in the MDASI-BT validation (See Armstrong et al., 2006).⁴ For example, the GI component score consists of the average of nausea and vomiting. We also computed the overall interference score and its 2 symptom interference component scores (activity and mood-related interference).

Participant sex, ethnicity, race, age, level of education, marital status, and employment status were collected using a demographic information sheet that was created and used in all prior psychometric studies of the MDASI-BT. Information on tumor type and characteristics and Karnofsky performance status (KPS) were recorded on the clinical assessment tool that was used in all prior psychometric studies of the MDASI-BT.

Statistical Analysis

Descriptive statistics were used to summarize participant demographic data, clinical characteristics, and MDASI-BT ratings by

Table 1. Patient characteristics

	n	%
Age (years)		
Mean, median, SD	48, 48, 13	
Range	19–77	
Sex		
Male	57	62
Female	35	38
Marital status		
Married	74	80
Single	11	12
Divorced/separated/widowed	7	8
Employment		
Full-time	45	49
Part-time	7	8
Retired	9	10
Unemployed due to diagnosis of tumor	11	12
Other	17	19
Race		
White non Hispanic	79	86
Hispanic	6	7
Asian	4	4
Black or African American	2	2
Other	1	1
Highest level of education		
High school graduate	7	8
Some college	22	24
College graduate	23	25
Graduate degree	39	42
Household income		
\$100 000 or more	42	46
\$50 000–\$99 999	23	27
\$30 000–\$49 999	7	8
Less than \$30 000	9	10
Treatment status		
Newly diagnosed	22	24
On treatment	42	46
In follow-up without active treatment	28	30
Tumor grade		
Grade II	12	13
Grade III	16	17
Grade IV	63	69
Extent of initial surgery		
Biopsy	24	26
Partial resection	29	32
Gross total resection	38	41
Recurrence status		
None	53	58
First recurrence	26	28
Repeat recurrence	13	14
Concomitant medications		
Corticosteroid	24	26
Anticonvulsant	70	76
Antidepressant	16	17
Stimulant	1	1
Analgesic	17	19

Abbreviation: SD, standard deviation.

Table 2. Symptom severity by recall method

MDASI-BT	Group	Mean	SD	Range	% ≥ 5 ^b	% = 0 ^c
Symptom (rank order) ^a						
Fatigue	24 h	3.43	2.74	0-10	34	19
	7 day	3.54	2.86	0-10	35	20
Drowsiness	24 h	3.15	2.86	0-10	32	22
	7 day	3.27	2.81	0-10	34	23
Difficulty remembering	24 h	2.42	2.71	0-10	27	39
	7 day	2.27	2.73	0-10	22	38
Distress	24 h	2.23	2.65	0-10	21	45
	7 day	2.49	2.72	0-10	23	35
Disturbed sleep	24 h	2.10	2.56	0-10	21	42
	7 day	2.61	2.93	0-10	27	39
Sadness	24 h	1.92	2.77	0-10	17	14
	7 day	2.34	2.86	0-10	22	41
Difficulty speaking	24 h	1.78	2.44	0-10	17	47
	7 day	1.98	2.59	0-10	20	47
Difficulty concentrating	24 h	1.63	2.18	0-8	13	47
	7 day	2.08	2.41	0-9	20	41
Dry mouth	24 h	1.40	2.34	0-10	13	60
	7 day	1.39	2.12	0-10	10	58
Weakness	24 h	1.37	2.35	0-10	12	65
	7 day	1.43	2.44	0-10	17	65
Irritability	24 h	1.33	1.83	0-10	7	48
	7 day	1.75	2.20	0-10	12	45
Difficulty understanding	24 h	1.27	2.04	0-10	8	60
	7 day	1.61	2.17	0-8	15	51
Numbness	24 h	1.04	2.06	0-8	11	74
	7 day	1.10	1.97	0-8	9	66
Vision	24 h	0.98	2.00	0-8	10	73
	7 day	1.13	2.08	0-8	11	65
Change in bowel pattern	24 h	0.95	2.01	0-8	9	73
	7 day	1.10	1.96	0-9	10	64
Pain	24 h	0.93	1.70	0-7	5	65
	7 day	1.61	2.33	0-10	15	54
Lack of appetite	24 h	0.87	1.69	0-8	7	71
	7 day	0.77	1.75	0-10	8	76
Shortness of breath	24 h	0.75	1.54	0-7	4	73
	7 day	0.68	1.49	0-7	5	75
Change in appearance	24 h	0.48	1.46	0-7	5	87
	7 day	0.72	1.81	0-8	7	80
Nausea	24 h	0.42	1.09	0-7	2	79
	7 day	0.59	1.51	0-7	5	82
Seizure	24 h	0.38	1.33	0-7	3	90
	7 day	0.45	1.42	0-7	7	89
Vomiting	24 h	0.09	0.74	0-7	1	98
	7 day	0.12	0.84	0-7	1	98
Interference Items (rank order) ^a						
Work	24 h	2.13	2.73	0-10	18	45
	7 day	2.38	2.68	0-10	21	37
General activity	24 h	1.98	2.39	0-10	17	42
	7 day	2.28	2.52	0-10	20	36
Enjoyment of life	24 h	1.98	2.73	0-10	16	45
	7 day	2.29	2.84	0-10	21	40

Continued

Table 2. Continued

MDASI-BT	Group	Mean	SD	Range	% ≥ 5 ^b	% = 0 ^c
Mood	24 h	1.84	2.35	0-8	16	42
	7 day	2.33	2.49	0-10	20	34
Walking	24 h	1.65	2.54	0-10	18	57
	7 day	1.74	2.45	0-10	16	54
Relations with others	24 h	1.47	2.25	0-10	14	57
	7 day	1.67	2.26	0-10	12	48
Subscale Scores						
All Symptoms	24 h	1.41	1.15	0-6	3	3
	7 day	1.59	1.32	0-5	6	15
Affective	24 h	2.20	1.84	0-8	14	10
	7 day	2.55	2.18	0-8	22	14
Cognitive	24 h	1.78	2.08	0-10	15	28
	7 day	1.98	2.28	0-9	20	30
Neurologic	24 h	0.93	1.35	0-6	4	41
	7 day	1.15	1.56	0-7	5	35
Treatment-related	24 h	1.81	1.69	0-8	9	12
	7 day	1.81	1.63	0-7	9	13
Generalized disease	24 h	0.79	1.26	0-6	3	47
	7 day	0.91	1.32	0-7	2	41
GI (composite)	24 h	0.49	0.90	0-5	1	60
	7 day	0.60	1.02	0-5	1	59
Interference	24 h	1.84	2.05	0-8	12	20
	7 day	2.12	2.18	0-8	14	22
WAW	24 h	1.92	2.20	0-9	15	26
	7 day	2.13	2.32	0-8	18	25
REM	24 h	1.76	2.19	0-9	15	29
	7 day	2.10	2.28	0-10	14	33

^aRanking based on 24 hour recall mean scores.

^bPercent moderate to severe.

^cPercent of patients scoring at the floor (score =0 on the 0-10 scale). Abbreviations: GI, gastrointestinal; REM, relate-enjoy-mood; SD, standard deviation; WAW, walk-activity-work.

recall period. Mean global and individual symptom severity, symptom component, interference, and interference subscale scores were calculated for both the 24-hour and 7-day recall periods. Mean difference scores between the 2 recall periods (with confidence intervals) were calculated, as was the percentage of participants who rated symptoms as moderate to severe (defined as a symptom score ≥5).⁵

To evaluate if symptom severity, as reported by the 7-day recall period, was similar to the symptom severity reported by the established 24-hour recall period, we examined both the congruency and equivalency of the symptom ratings for both recall periods. With congruency, we are interested in determining whether the ratings are in agreement or concordant between the 2 recall periods. The Bland-Altman approach⁶ was used to visually inspect the congruency of the component scores across the 2 recall periods. The Bland-Altman approach is a 2-step analysis. First, a scatterplot was generated for symptom and interference severity for both recall periods and inspected visually for deviation from a 45° line of equality. Then, difference scores between the 2 recall periods were calculated and examined to see if they fell within

Table 3. Mean differences between recall methods in scale items and subscales

	Mean Difference ^e
Symptoms	-0.19 ^a
Affective	-0.34 ^c
Fatigue	-0.11
Disturbed sleep	-0.51
Distress	-0.26
Sadness	-0.41
Irritability	-0.42
Cognitive	-0.21
Difficulty remembering	0.15
Difficulty understanding	-0.34
Difficulty speaking	-0.20
Difficulty concentrating	-0.45
Neurologic	-0.21 ^c
Pain	-0.67 ^d
Numbness	-0.05
Weakness	-0.07
Seizure	-0.07
Treatment related	0.00
Lack of appetite	0.10
Drowsiness	-0.12
Dry mouth	0.01
Generalized disease	-0.12
Shortness of breath	0.07
Vision	-0.15
Change in appearance	-0.24
Change in bowel pattern	-0.15
GI (composite)	-0.12
Nausea	-0.16
Vomiting	-0.03
Interference	-0.28 ^a
WAW	-0.21
General activity	-0.30
Work	-0.25
Walking	-0.09
REM	-0.34 ^b
Mood	-0.49 ^d
Relations with others	-0.21
Enjoyment of life	-0.32

^a*P* < .05.^b*P* < .025.^c*P* < .01.^d*P* < .0025.^eBetween 24-hour and 7 day recall.

Abbreviations: CI, confidence interval; REM, relate-enjoy-mood; WAW, walk-activity-work.

the 95% limits of agreement (as determined by their standard deviation [SD]).

Once congruency between the 2 recall periods was established, we then evaluated the equivalency of the ratings using confidence intervals (CIs). With equivalency, we were interested in determining whether the mean difference in ratings was small

enough that we could consider the ratings to be essentially similar. Based on the work of Sloan with PRO instruments, a difference of approximately half SD is considered clinically significant.⁷ Therefore, any rating differences that fell within 0.5 SD were considered equivalent. A reference CI (-0.5SD to +0.5SD) was constructed based on the SD of the mean ratings of the established 24-hour recall period. Symptom and interference severity, as reported by the 7-day recall period, were deemed equivalent to the symptom and interference severity reported by the 24-hour recall period if a 90% CI of the differences in recall rating was inside the reference CI. A 90% CI was used instead of the more commonly used 95% CI because two 1-sided *t* tests are performed in equivalency testing, and both 1-sided null hypotheses must be rejected. However, it is sufficient to perform only 1 test, allowing for a test without halving the classic α of 0.05.⁸

Finally, we reported psychometric testing of the 7-day recall period: in particular, reliability and known-group validity.⁹ As in our previous validation, we powered our study to detect a half SD difference in symptom severity between good and poor performance status participants as evidence of known group validity. Our power analysis indicated that we would require 50 participants per group to detect such a difference with 80% power using a 1-tailed test at 5% significance level. Cronbach's alphas for the MDASI-BT using the 7-day recall period were calculated to determine if they met the minimum standard requirement for an internally reliable instrument. Coefficients of at least 0.70 were considered acceptable.¹⁰ The Cronbach's alphas for the established 24-hour recall period were also reported for this sample as reference. To establish known group validity of the 7-day recall instrument, we investigated the ability of the instrument to distinguish between participants with high daily function and those with low daily function. In several prior studies, a significant difference in symptom burden and interference was demonstrated in participants with KPS of 80% or less compared with those having KPS of 90%–100%.^{3,4,10} Therefore, we categorized participants into 2 groups of good versus poor performance status based on their KPS score (90%–100% vs 80% or less, respectively). Independent sample tests of the 7-day recall scores were used to compare participants with poor KPS to those with good KPS. All statistical analyses were performed using IBM SPSS 21.

Results

Only 2 of the 102 patients who were approached declined to participate, resulting in 100 patients who were consented and participated in the study. Ninety-two participants completed both the 24-hour and 7-day recall instruments, with none being unable to complete the questionnaires due to cognitive or neurologic deficits. Median age was 48 years (range 19y–77y), and 62% were male. Participants were primarily white non-Hispanic (86%); 25% completed a college degree, and 46% had a household income of >\$100 000. Seventy percent of participants had a KPS classified as good to excellent (90%–100%). The majority of participants were diagnosed with a grade IV primary brain tumor (69%), and a relatively high percentage (41%) had undergone gross total tumor resection at their initial surgery. Many were on active treatment (46%) when they were tested. Forty-two percent had at least 1 recurrence by the time of their assessment,

and 76% were on an anticonvulsant regimen. See Table 1 for a complete list of demographic and clinical characteristics.

Symptom Severity

Table 2 presents the symptoms and interference severity, with all items having reports through the full range of severity grouping

(mild, moderate, or severe). The most severe symptoms reported were the same in both recall periods and included fatigue, drowsiness, difficulty remembering, distress, sadness, and disturbed sleep. The mean difference between recall methods in scale items and subscales were <1 point for overall symptom scores, interference scores, and subgroupings, which is less than the clinically meaningful difference^{4,11} (see Table 3).

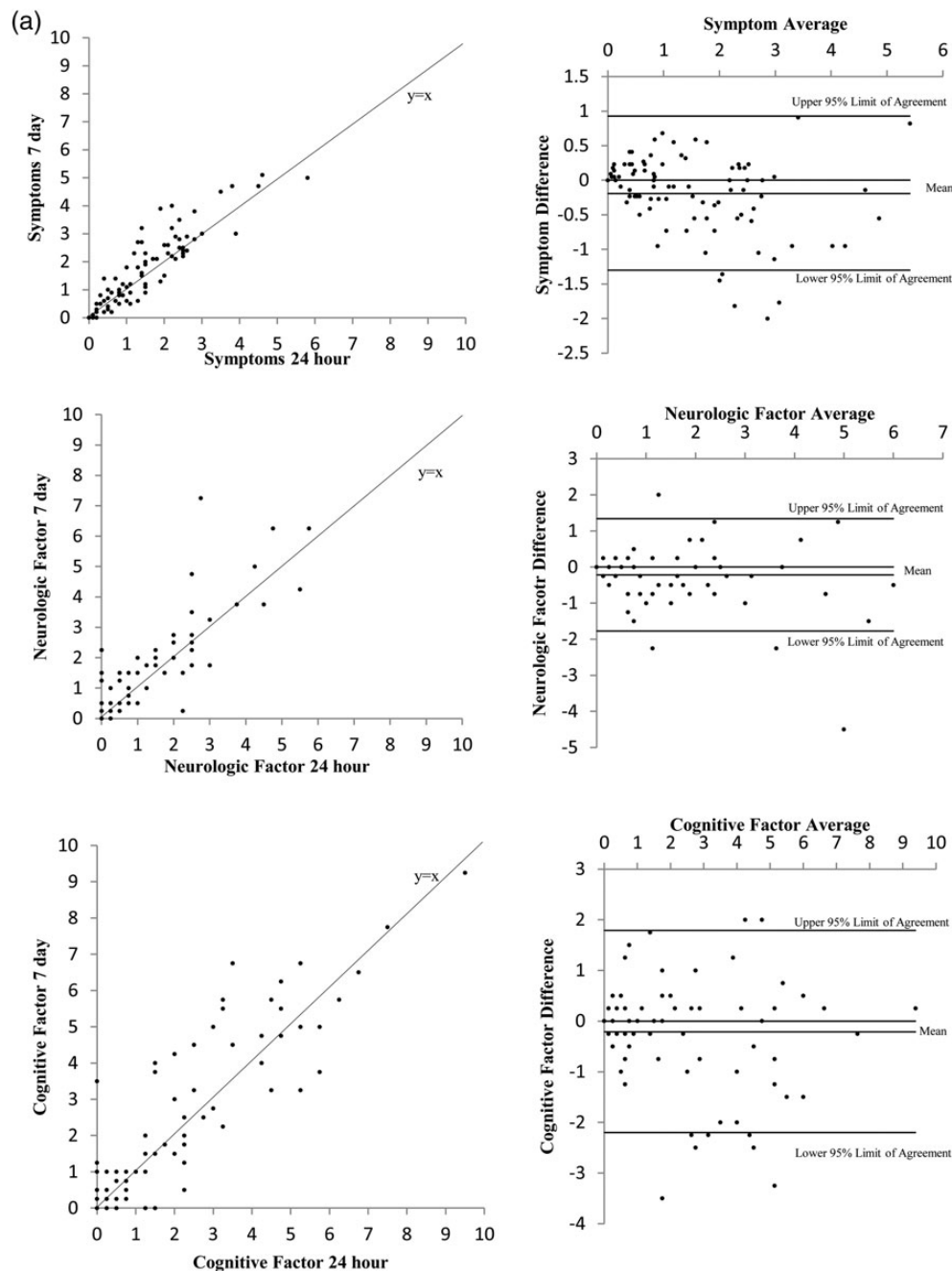


Fig. 1. (a) Plots of 7-day versus 24-hour recall symptom and symptom factor scores. Comparisons in these figures include overall symptom burden and the factor groupings of neurologic and cognitive symptoms as representative examples of the equivalency between the 7-day and 24-hour recall periods. (b) Plots of 7-day versus 24-hour recall interference and interference subscale scores. Comparisons in these figures include overall interference and the subscales of interference items related to activity and mood.

Test of Congruency

Figure 1(a) and (b) present the Bland-Altman plots for the component scores of interest. Visual inspections of the left column suggest that most data points lie close to the 45° line of equality. Although the right column of the figure shows some data points outside the lines of agreement, more than 90% of data points still lie within the limits (see Table 4). Because of the subjectivity of inspecting the Bland-Altman plots, we formalized the inspection by using equivalency testing.

Test of Equivalency

As previously stated, if our computed 90% CIs (of the differences between recall periods) lay completely within the prespecified equivalency interval, we could declare equivalency and conclude that the ratings based on a -7-day recall were similar to those based on a 24-hour recall. Table 4 illustrates that the ratings for all component scores were similar regardless of the recall period, indicating equivalency.

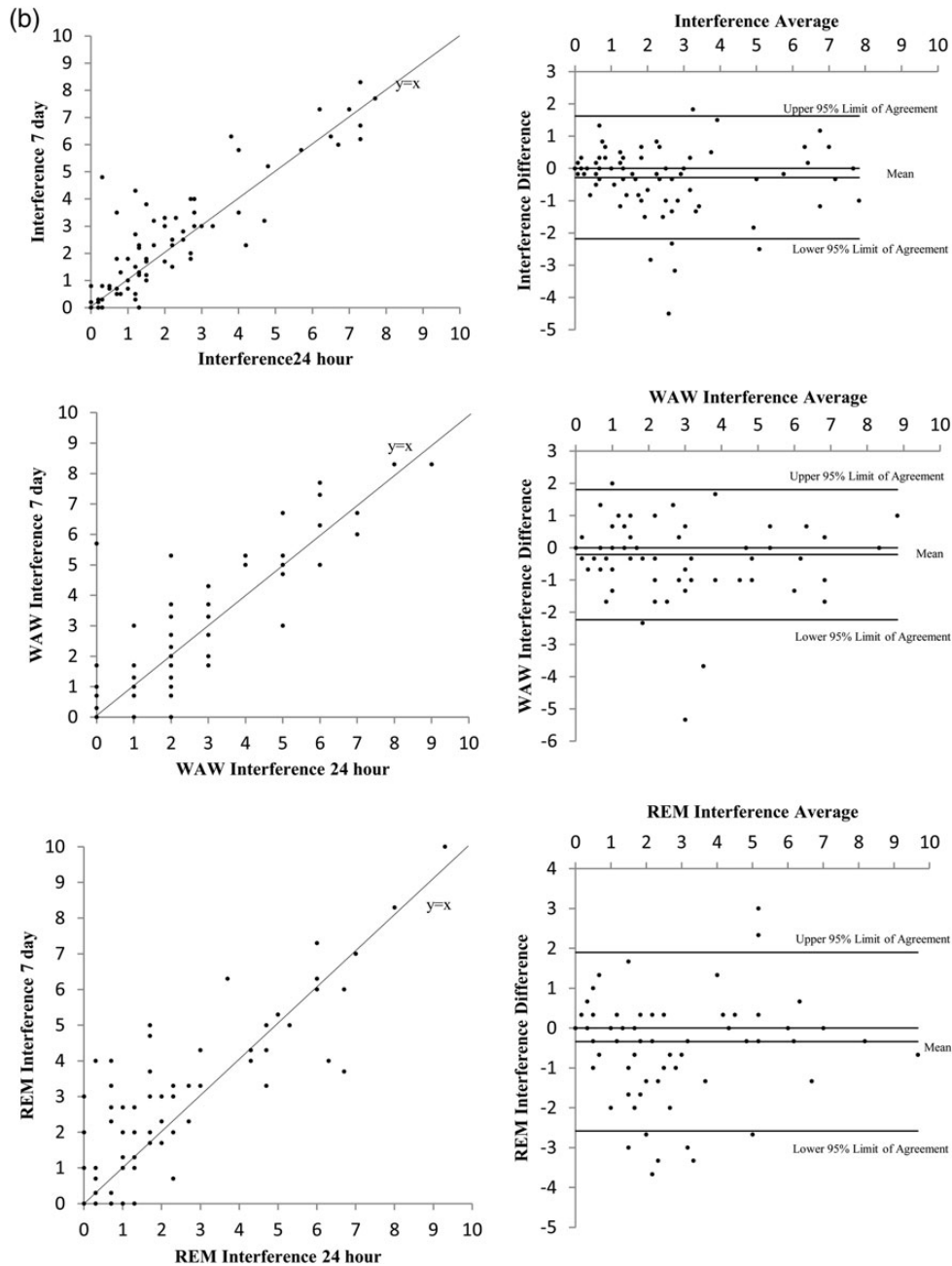


Fig. 1. Continued.

Table 4. Equivalency testing confidence intervals

	24 mean ^a	24 h SD	Reference CI ±0.5SD	90% CI of Differences ^b		Equivalent?
				Lower Bound	Upper Bound	
All symptoms	1.41	1.15	±0.57	-0.28	-0.09	Yes
Neurologic factor	0.93	1.35	±0.68	-0.35	-0.08	Yes
Cognitive factor	1.787	2.08	±1.04	-0.38	-0.03	Yes
Interference	1.84	2.05	±1.03	-0.44	-0.11	Yes
WAW	1.92	2.20	±1.10	-0.39	-0.04	Yes
REM	1.76	2.18	±1.09	-0.53	-0.14	Yes

^aThe 24-hour recall is the reference recall to evaluate equivalency.

^bBetween 24-hour and 7-day recall.

Abbreviations: CI, confidence interval; REM, relate-enjoy-mood; SD, standard deviation; WAW, walk-activity-work.

Psychometric Testing

Reliability of the 7-day Versus 24-hour Recall MDASI-BT

Table 5 shows internal consistency results of both recall periods. Both the overall symptom burden and interference scales revealed high and similar internal consistency (symptom Cronbach alpha of 0.88–0.91 and interference scores of 0.9–0.93). When evaluating factor scores, with the exception of the treatment-related and GI factors, all subscales of the MDASI-BT using the 7-day recall met the minimum standard requirement for a reliable instrument with Cronbach’s alphas ranging from 0.68 to 0.94. The findings of these 2 factor groupings were similar with each recall period and were impacted by the lower mean ratings of these items and the participants in the sample who were not on active treatment.

Validity of the 7-day versus 24-hour Recall MDASI-BT

The impact of KPS on the report using the 7-day recall period was then assessed. The 7-day recall tool showed differences in KPS. Participants with low KPS reported higher symptom severity ($t = 4.0, P < .001, r = 0.53$) and interference severity ($t = 4.2, P < .001, r = 0.40$) in the previous 7 days than those with high KPS, as shown in Table 6. Similar results were found using the 24-hour recall tool, with participants having low KPS reporting higher symptom and interference severity ($P < .01$) (reported in Table 6 as reference).

Discussion

Use of the 7-day versus 24-hour recall period for reporting symptom and interference severity was found to be congruent and equivalent in this study. These results mirror findings with other solid tumor patients.^{2,12} Although it is reasonable to expect that the symptom report for a 7-day recall may be greater than a 24-hour recall because the 24-hour recall is subsumed within the longer time period, our results suggest that these 2 time periods result in severity ratings that are quite similar.

The design and goal of a study typically dictates the choice of recall period to use. For daily symptom assessment, the previous

Table 5. Internal consistency reliability of the 24-hour and 7-day recall MDASI-BT

Scale	Cronbach’s Alpha	
	24 hours	7 days
Symptoms	0.88	0.91
Affective	0.78	0.86
Cognitive	0.91	0.94
Neurologic	0.67	0.74
Treatment related	0.54	0.53
Generalized disease	0.68	0.68
GI (composite)	0.31	0.42
Interference	0.90	0.93
WAW	0.83	0.89
REM	0.87	0.88

Abbreviations: REM, relate-enjoy-mood; WAW, walk-activity-work.

24 hours is the obvious choice. In addition, evaluating symptoms more frequently may be beneficial with treatments associated with intense but varied symptom profiles. In studies where symptom assessment coincides with clinic visits that are weeks apart, the use of the 7-day recall period may be more appropriate. The decision to use the 7-day recall is no longer hampered by the lack of data on its properties. As stated below, we now have evidence demonstrating the desirable psychometric properties of the 7-day recall. Depending on the goal or design of the study, either the 24-hour or 7-day window can be used. Our findings indicate that both recall period versions of the MDASI-BT are reliable and valid. Hence, our results offer more flexibility in designing studies.

Psychometric testing of the 7-day recall period yielded similar results to the well-established 24-hour recall period version of the MDASI-BT. Overall, select psychometric testing of the 7-day recall period demonstrated that it had adequate reliability as evaluated by internal consistency, with the exception of treatment-related and GI-related symptoms. These results, however, may be impacted by the small number of items in the symptom grouping. Cronbach’s alpha is directly proportional to the number of items

Table 6. Performance status^a effect on symptom and interference severity

	<i>n</i>	Mean	Mean Difference	95% CI	Effect Size
<i>7-day recall</i>					
All symptoms					
Low (KPS ≤80%)	28	2.4	1.2*	0.6–1.8	0.53
High (KPS ≥90%)	64	1.2			
Interference					
Low (KPS ≤80%)	28	3.4	1.9*	0.8–3.0	0.40
High (KPS ≥90%)	64	1.5			
<i>24-hour recall</i>					
All symptoms					
Low (KPS ≤80%)	28	2.2	1.0*	0.5–1.6	0.52
High (KPS ≥90%)	64	1.1			
Interference					
Low (KPS ≤80%)	28	3.3	2.1*	1.0–3.2	0.48
High (KPS ≥90%)	64	1.2			

**P* value <.01.

^aas evaluated with Karnofsky performance scale.

Abbreviations: CI, confidence interval; KPS, Karnofsky performance scale.

in the subscale, and these 2 subscales have the fewest number of items at 3 and 2, respectively. In addition, there was a floor effect for the GI-related items (nausea and vomiting). Evaluation of validity of the 7-day recall period showed that it was also able to detect differences in symptom severity in those with good versus poor performance status, as does the 24-hour recall period. Psychometric validation of an instrument is an iterative process. In general, validation is never finalized with one study. Once an instrument is adapted and its use become widespread, new evidence of validity and reliability accumulates and becomes part of the validation dossier for that instrument.

Only select measures of validity and reliability are reported here. Ability to detect change over time and test-retest reliability among other psychometric testing should also be evaluated for this instrument.

There are several limitations to this study. We employed a cross-sectional design; thus, we did not evaluate any differences in severity associated with recall period over time. However, the cross-sectional sample, which included participants who were newly diagnosed, those on active treatment, and those who were not on active treatment but were being followed with imaging, allowed for generalizability across the disease epoch. In addition, 30% of the sample had a poor KPS, which allowed evaluation of the sensitivity of the 7-day recall period to known differences based on patient functional status.

This study can also be limited by order effects, in which the order the instruments are given may impact the severity rating. In this study, participants first answered for the previous 42 hours and then for the previous week. Overall, the 7-day recall period was associated with more severe symptom reporting, although the difference was <1 point for all symptom and symptom groupings that were assessed.

This study supports the use of either the 24-hour or the 7-day recall period as based on study design and expected symptom outcomes. Future studies should continue to explore the reliability and validity of this recall period and its utility in other CNS tumor populations, including brain metastases.

Funding

Support for this study provided by the Collaborative Ependymoma Research Network (CERN).

Conflict of interest statement. None declared.

References

- Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:79.
- Shi Q, Trask PC, Wang XS, et al. Does recall period have an effect on cancer patients' ratings of the severity of multiple symptoms?. *J Pain Symptom Manage*. 2010;40(2):191–199.
- Armstrong TS, Gning I, Mendoza TR, et al. Clinical utility of the MDASI-BT in patients with brain metastases. *J Pain Symptom Manage*. 2009;37(3):331–340.
- 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.
- Serlin RC, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. *Pain*. 1995;61(2):277–284.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310.
- Sloan JA, Symonds T, Vargas-Chanes D, Fridley B. Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Drug Inf J*. 2003;37:23–31.
- Rogers JL, Howard KI, Vessey JT. Using significance tests to evaluate equivalence between two experimental groups. *Psychol Bull*. 1993; 113(3):553–565.
- Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd edn. McGraw-Hill, 1994.
- 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.

11. 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.

12. Broderick JE, Schneider S, Schwartz JE, Stone AA. Interference with activities due to pain and fatigue: accuracy of ratings across different reporting periods. *Qual Life Res*. 2010;19(8):1163–1170.