AJNR

Measurement of Tumor "Size" in Recurrent Malignant Glioma: 1D, 2D, or 3D?

Mary F. Dempsey, Barrie R. Condon and Donald M. Hadley

This information is current as of July 6, 2025.

Measurement of Tumor "Size" in Recurrent Malignant Glioma: 1D, 2D, or 3D?

Mary F. Dempsey, Barrie R. Condon, and Donald M. Hadley

BACKGROUND AND PURPOSE: Tumor "size" is used internationally as a surrogate marker for overall survival when following current response assessment protocols (World Health Organization and Response Evaluation Criteria in Solid Tumors). With little evidence of a relationship between tumor "size" and survival in intrinsic brain tumors, this study was undertaken to investigate the predictive value of MR imaging–defined tumor size for survival in patients with recurrent malignant glioma and to compare the different measures of tumor size used in these current response assessment protocols.

METHODS: Volumetric, bidimensional, and unidimensional measurements of tumor size were made using baseline contrast-enhanced T1-weighted images of 70 patients with recurrent malignant glioma receiving intravenous chemotherapy. Cox's proportional hazards model was used to investigate the prognostic importance of tumor size using survival as the end point. Further statistical analysis was undertaken to investigate the relationship between the different measurement techniques.

RESULTS: Only the volumetric measurement of tumor size was found to be predictive of survival in recurrent malignant glioma on both univariate and multivariate analysis. Furthermore, analysis demonstrated that the unidimensional and bidimensional measures of tumor were not comparable with the more accurate and direct volumetric measurement.

CONCLUSION: Indirect unidimensional and bidimensional measurement techniques do not have a significant association with overall survival or adequately assess tumor size in recurrent malignant glioma. These findings have serious implications about the validity of using current response assessment protocols in therapy trials for recurrent malignant glioma.

Glioblastoma multiforme (GBM) and anaplastic astrocytoma (AA; high-grade malignant glioma) are the most common primary brain tumors occurring in adulthood, and they are among the most lethal and difficult to treat forms of cancer. Identifying prognostic indicators is a crucial component of the ongoing search for more-effective tumor therapies. MR imaging of patients with primary brain tumors provides a sensitive tool for assessing tumor size, which is thought to have a bearing on patient outcome. There is, however, a lack of validation of this belief within the literature, as few studies have undertaken a clinical correlation of tumor size with outcome, and, of these few studies, the findings were inconsistent (1–3). Despite this lack of validation, MR-defined tumor size is heavily relied upon when assessing response to

therapy in accordance with accepted response assessment criteria (4–6). Internationally accepted protocols devised to standardize response assessment in solid tumors—World Health Organization (WHO) criteria (4) and Response Evaluation Criteria in Solid Tumors (RECIST; 6)—and in supratentorial malignant glioma—MacDonald's criteria (5)—advise use of unidimensional or bidimensional measurement of tumor size to assess response.

The primary aim of this report is to assess the predictive value of tumor size in recurrent malignant glioma using the methods of tumor-size measurement advised by the above 1D and 2D response assessment protocols and the more accurate volumetric measurement (3D) and the "best" end point of survival. For this study, 1D and 2D measurement techniques have been optimized and automated by the development of in-house software to minimize measurement error. To our knowledge, such an analysis has not been performed for the three main measurement techniques in recurrent malignant glioma. The secondary aim of this study is to undertake the basic, yet important, comparison of these different measurement techniques to investigate the comparability between

these nondirect measures (1D and 2D) and the direct measure (3D) of actual tumor size.

Methods

We retrospectively reviewed the MR imaging data of 70 patients enrolled in a multicenter phase II clinical trial of intravenous RMP-7 (Cereport) and Carboplatin for the treatment of recurrent glioma (7). Patients were enrolled in the study if they had a radiographically measurable recurrence of a pathologically documented AA or GBM in accordance with WHO criteria. Recurrent disease was defined as an increase in tumor size (on MR imaging or CT) with deterioration in neurological status relapsing after a progression free interval of greater than 3 months. Patients were included in the trial if they were older than 18 years with a Karnofsky index >60 (this is a performance measure for rating the ability of a person to perform usual activities) with an expected lifespan of at least 8 weeks. RMP-7 is a drug that, by a receptor-mediated mechanism, modifies the permeability of the blood-brain barrier (BBB) throughout the tumor and around the peripheral infiltrations for a short time to allow entry of the chemotherapeutic agent, Carboplatin. All patients were treated at centers in France, Sweden, and the United Kingdom between 1994 and 1996.

A 3D T1-weighted imaging protocol with contrast enhancement (0.1 mmol/kg gadopentetate dimeglumine) was followed at each center resulting in images of similar resolution and section thickness (voxel dimensions, $0.98 \times 0.98 \times 2.8$ mm). Because this was a multicenter trial, the images were acquired in different MR systems but all operating at a 1.5T field strength. Because of variations between manufacturers, individual sites were unable to duplicate sequences exactly, and images were acquired using one of two 3D sequences selected to produce optimal enhancement on T1-weighted images (MPRAGE [TR, 10 ms; TE, 4 ms] or 3D FLASH [TR, 30 ms; TE, 6 ms]) appropriate to the MR system. The baseline MR images were anonymized and transferred electronically to a centralized processing unit at our institution—which was not involved in treatment—to allow analysis by a single consulting neuroradiologist, blinded to the clinical data and outcome (D.M.H.). Analysis involved manual outlining of contrast enhancement on T1-weighted images using a standardized windowing protocol, maximizing gray and white matter contrast while ensuring areas of contrast enhancement could be delineated from white matter, to obtain a 3D measure of tumor volume (mm³) as illustrated in Fig 1.

In-house software was developed, using Microsoft Excel and Visual Basic, to convert these regions to coordinate data and objectively calculate the maximum diameter for each region to obtain a 1D measurement of tumor size (Appendix). When more than one region per section was observed, the maximum diameters were summed in accordance with RECIST (6). Standardizing measurement in the case of multiple foci was one of the main motivations for producing this revised set of response assessment guidelines (RECIST), because previous guidelines did not offer clear advice on their inclusion/exclusion. The overall largest diameter (or summed diameters) was then recorded as the 1D measurement of tumor size. This measurement was converted to a bidimensional measurement by multiplication with tumor length in the perpendicular sagittal plane. Validation of this technique was undertaken using simulated lesions with known maximum diameters. Although these calculations were computationally demanding (eg, >1 hour per image volume), this method eliminated further error involved in the visual estimation of maximum diameters.

Survival was calculated in days from the time of diagnosis of recurrence to time of death (with patients having intravenous chemotherapy between these times) regardless of cause or
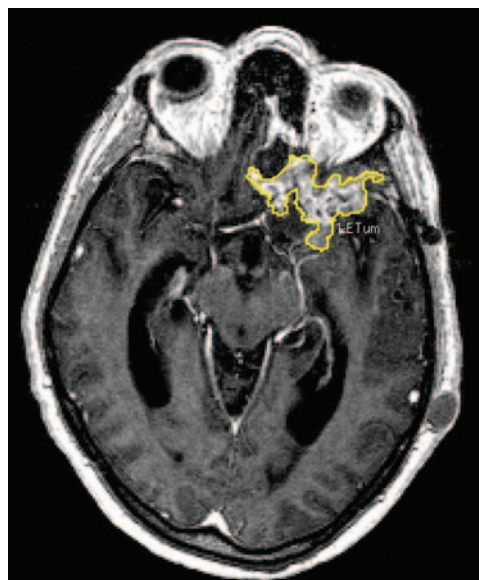

FIG 1.   Image showing software used (developed for Alkermes, Inc., by Evergreen Technologies, Inc., 00000, ME) by neuroradiologist to outline areas of enhancement manually on each transverse section to determine tumor volume.

censoring at last follow-up. Survival data were analyzed using Cox's proportional hazards model (8). The 5% significance level was chosen for univariate and multivariate analysis. Multivariate techniques were applied to determine the joint effects of other potential prognostic variables acting simultaneously to select out the variables that are independently most closely related to survival. Although there are various established prognostic factors for patients with newly diagnosed malignant glioma—including age, histology, extent of surgery, and performance status—prognostic factors for patients with recurrent disease are not so well established. One group has performed an analysis of potential prognostic variables (histology, age, performance status, and salvage therapy) in recurrent high-grade glioma and found histology (whether AA or GBM) to be the dominant factor in determining outcome (9). Because several studies with similar findings are required before we achieve "established" prognostic factors, for this study we have included in the model several potentially prognostic variables: age, sex, whether the patient had previous chemotherapy, and histologic grade (AA or GBM). Unfortunately, because of the retrospective and multicentric nature of this analysis, pathological information of actual tumor grade (AA or GBM) at recurrence was available only in the chemotherapy-naive group ($n = 35$; 37% AA and 63% GBM). The remaining patients ($n = 35$), were known to be 40% AA and 60% GBM. Information on the extent of any previous surgery at initial diagnosis was not available and could not be included in this analysis (surgery was not performed when recurrence was diagnosed); performance status was not assessed, as inclusion criteria required all patients to have a Karnofsky index >60.

The age of patients included in this study ranged from 24 to 68 years (mean age, 45.6 ± 10.7 years). Fifty-one patients were male, and 19 were female. Thirty-five of the 70 patients had previously undergone chemotherapy, whereas 35 were chemotherapy-naive. Actual tumor grade (AA or GBM) was known only in the chemotherapy-naive group (13 AA and 22 GBM). Age and tumor size were entered in the analysis as continuous covariates, whereas the other covariates (sex, previous chemotherapy, and tumor grade) were entered as binary coded factors. The median survival of the 70 patients was 196 days.

Measurement techniques were compared using standard statistical methods for comparing measures (10). Because the

TABLE 1: Summary of 1D, 2D and 3D measures of tumor size

	1D (mm)	2D (mm^2)	3D (mm^3)	1D→2D (mm^2)	1D→3D (mm^3)	2D→3D (mm^3)
Minimum	2.8	22.5	53.5	6.2	11.6	80.3
Maximum	206.6*	22645.4*	131387.4	33501.7*	4613963.1*	2564157.0*
Mean	78.6	4240.2	34416.5	5922.4	454253.8	253035.7

*Such large maximum values originate from images with several regions of contrast enhancement outlined separately and consequent summation of the appropriate diameters resulting in a large 1D measurement and subsequent conversions.

TABLE 2: Cox model applied to assess the effect of 1D measured tumor size on survival

	Univariate Analysis				Multivariate Analysis			
Covariate	Coefficient (b$_i$)	HR [exp(b$_i$)]	95% CI	P-value	Coefficient (b$_i$)	HR [exp(b$_i$)]	95% CI	P-value
Age	0.036	1.039	(1.014–1.064)	.002	0.036	1.037	(1.013–1.062)	.002
Sex				.015				.045
Female	(0.00)	(1.00)			(0.00)	(1.00)		
Male	0.725	2.070	(1.153–3.715)		0.625	1.868	(1.015–3.436)	
Previous				.096				.545
Chemotherapy								
No	(0.00)	(1.00)			(0.00)	(1.00)		
Yes	0.433	1.542	(0.926–2.566)		0.164	1.178	(0.684–2.000)	
1D Size	0.067	1.069	(0.997–1.147)	.062	0.052	1.053	(0.182–1.129)	.146

HR = hazard ratio, CI = confidence interval

measurement techniques (1D, 2D, and 3D) produce results of different dimensions, these measurements were converted into area or volume measurement as appropriate by assuming circular or spherical geometry to eliminate bias during measurement comparison.

Results

Predictive Value of Tumor Size

Cox's proportional hazards model was used to investigate the dependence of survival on the measurements of tumor size at baseline for 70 patients with recurrent malignant glioma taking into account the effect of age, sex, tumor grade, and previous chemotherapy. Table 1 summarizes the range of tumor-size measurements within this patient group.

Preliminary Cox proportional hazard modeling was performed on the chemotherapy naive group (n = 35) with known histological information to assess the effect of tumor grade on survival. Median survival was 317 days for AA patients (n = 13) and 185 days for GBM patients (n = 22). The effect of tumor grade on survival, however, was not found to be statistically significant (P > .05) on both univariate and multivariate analysis. It is important to note that, because tumor grade was not shown to significantly affect survival, this group could be joined with the previous chemotherapy group to increase statistical power for further Cox proportional hazard modeling.

Considering the effect of tumor size on survival, all measures (1D, 2D, and 3D) resulted in a hazard ratio of >1, which indicates that larger tumor size was associated with shorter survival. For the 1D measurement of tumor size, however, this effect was not found to be statistically significant (P > .05) on both uni-

variate and multivariate analysis as shown in Table 2. The effect of the 2D measurement of tumor size on survival was statistically significant on univariate analysis (P < .05) but was not statistically significant on multivariate analysis, as shown in Table 3. Nevertheless, the effect of the 3D measurement of tumor size and survival was found to be statistically significant on both univariate and multivariate analysis as shown in Table 4.

Overall, age, sex, and 3D volumetric measurement of tumor size were the most important predictors of survival in the univariate and multivariate Cox models. This implies that for patients of the same age with the same volume of tumor, males have poorer survival than females, a finding that has not previously been reported and raising the question of what differs between the male and female brain to result in this effect? Longer survival was observed in younger patients with smaller tumor volumes.

Having previously shown that tumor grade did not significantly affect survival, for completeness, the above analysis was repeated to ensure that adding tumor grade information (when possible) to the model did not alter the results. Adding tumor grade to the above model did not alter the results, and age, sex, and volumetric measurement remained the most important predictors of survival.

Comparison of Tumor Size Measurements

Analysis was also performed to investigate the comparability between the different measurements (direct and indirect) of tumor size. The mean and range of measurements and conversions are shown in Table 1.

TABLE 3: Cox model applied to assess the effect of 2D measured tumor size on survival

	Univariate Analysis				Multivariate Analysis			
Covariate	Coefficient (b$_i$)	HR [exp(b$_i$)]	95% CI	P-value	Coefficient (b$_i$)	HR [exp(b$_i$)]	95% CI	P-value
Age	0.038	1.039	(1.014–1.064)	.002	0.036	1.037	(1.013–1.062)	.003
Sex				.015				.039
Female	(0.00)	(1.00)			(0.00)	(1.00)		
Male	0.725	2.070	(1.153–3.715)		0.648	1.912	(1.033–3.539)	
Previous				.096				.601
Chemotherapy								
No	(0.00)	(1.00)			(0.00)	(1.00)		
Yes	0.433	1.542	(0.926–2.566)		0.144	1.155	(0.673–1.082)	
2D Size	0.006	1.006	(1.000–1.012)	.046	0.004	1.004	(0.999–1.010)	.140

HR = hazard ratio, CI = confidence interval

TABLE 4: Cox model applied to assess the effect of 3D measured tumor size on survival

	Univariate Analysis				Multivariate Analysis			
Covariate	Coefficient (b$_i$)	HR [exp(b$_i$)]	95% CI	P-value	Coefficient (b$_i$)	HR [exp(b$_i$)]	95% CI	P-value
Age	0.038	1.039	(1.014–1.064)	.002	0.034	1.034	(1.010–1.058)	.005
Sex				.015				.003
Female	(0.00)	(1.00)			(0.00)	(1.00)		
Male	0.725	2.070	(1.153–3.715)		0.936	2.551	(1.361–4.781)	
Previous				.096				.115
Chemotherapy								
No	(0.00)	(1.00)			(0.00)	(1.00)		
Yes	0.433	1.542	(0.926–2.566)		0.430	1.537	(0.900–2.626)	
3D Size	0.019	1.019	(1.009–1.029)	.000	0.025	1.026	(0.015–1.036)	.000

HR = hazard ratio, CI = confidence interval

Figure 2 demonstrates that, although a relationship exists between the 1D and 2D measurement of tumor size ($r = 0.85$ $P < .05$), the measurements are not in agreement with the 1D measurement, giving higher values than the 2D measurement. This is expected, as by definition, the perpendicular diameter used in the 2D measurement will be less than the largest 1D diameter, introducing a systematic error in the conversion of the 1D measurement to 2D proportions. For these measurements, the mean difference was 1681.91 mm^2, which is not close to zero, thus implying there is an overall bias. This is supported by the second plot, in which we see a clear relationship between difference and mean ($r = 0.73$; $P < .05$). The 95% confidence interval for the bias is 834–2528 mm^2.

Similar findings are shown in Fig 3 comparing the 1D with the 3D measurement ($r = 0.44$; $P < .05$). This demonstrates that (as expected) recurrent gliomas are not spherical in shape and that the 1D measurement overestimates tumor size. For these measurements, the mean difference was 419,837.3 mm^3, which, again, is not close to zero, implying there is bias. This is supported by the second plot, in which we see a clear relationship between difference and mean ($r = 0.998$; $P < .05$). The 95% confidence interval for the bias is 230,985.2–608,689.4 mm^3.

Again, Fig 4 shows that, although a relationship is present between the 2D and 3D measures ($r = 0.47$; $P < .05$), the measurements are not in agreement with the
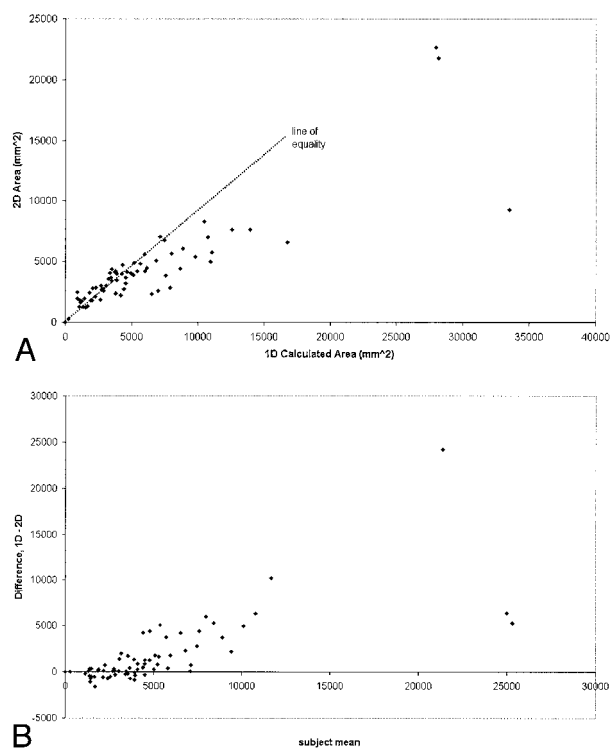

FIG 2. Tumor size measured using 1D and 2D techniques. *A,* 1D versus 2D technique plotted showing line of equality ($r = 0.85$; $P < .05$), *B,* Difference versus mean of tumor size measured by the two methods.
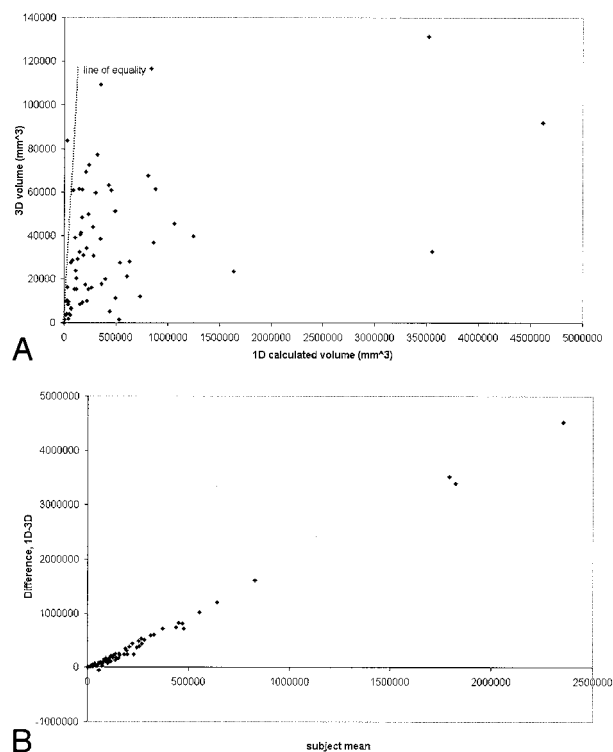
Fig 3.   Tumor size measured using 1D and 3D techniques. *A,* 1D versus 3D technique plotted showing line of equality ($r =$ 0.44; $P < .05$). *B,* Difference versus mean of tumor size measured by the two methods.

Fig 4.   Tumor size measured by 2D and 3D techniques. *A,* 2D versus 3D technique plotted showing line of equality ($r = 0.47$; $P < .05$). *B,* Difference versus mean of tumor size measured by the two methods.
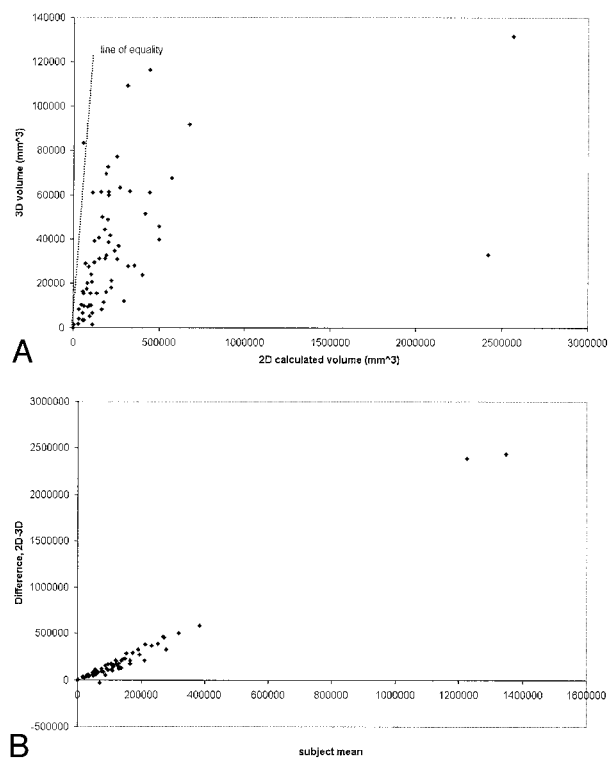
2D measurement, giving higher values than the more accurate 3D measurement. For these measurements, the mean difference was 218,619.2 mm^3, implying there is bias. This is supported by the above plot in which we see a clear relationship between difference and mean ($r = 0.99; P < .05$). The 95% confidence interval for the bias is 123,507.9–313,730.6 mm^3.

Discussion

Measurement of tumor size is heavily relied upon to assess response to therapy in recurrent malignant glioma. It is not known whether the use of tumor size for these purposes is clinically valid, and consequently it is not clear which measurement, if any, should be used. We have addressed this issue by clinically evaluating the three most-common techniques employed and have shown the 3D measurement of contrast-enhancing tumor volume to be prognostic (with age and sex) for survival whereas corresponding 1D and 2D measurements are not.

Few studies have assessed the relationship between tumor size and outcome in this form of cancer, and of those that have only one measurement technique was studied per patient group. One group showed 3D contrast-enhancing volume to be prognostic (1), and two groups showed 2D area measurement to be non-prognostic (2, 3). The effect of measurement technique on actual response has been studied previously and showed significant differences (11). Ours, how-

ever, is the first study to investigate the three measurement techniques on the same patient group and assess the clinical validity of the measurement. Our findings, supported by the above previous studies, must cast serious doubt on the use of 1D and 2D measures to assess response to therapy as currently advised by international protocols (4–6).

Although we have shown the 3D measure to be superior to the 1D and 2D measures, we recognize that manual outlining of contrast enhancement is far from ideal. This measurement is based on the assumption that enhancing regions represent tumor and not some other disease process. This can be problematic, because there can be significant necrotic and cystic components apparent when high-grade malignant glioma recurs, although these areas will be excluded from the measurement when possible. Radiation necrosis often appears as an enhancing mass and can be indistinguishable from recurrent tumor on MR imaging. Functional imaging is required to identify and exclude radiation necrosis, although this imaging is rarely included in clinical trial imaging protocols. Therefore, while the potential of inclusion of radiation necrosis in our measurement of tumor size is a limitation, our measurements represent the clinical circumstances. Furthermore, there is the argument that we are directly measuring BBB dysfunction and therefore only indirectly measuring tumor size. Another concern is that cells will inevitably have spread

beyond the neoplasm's identifiable enhancing mass into surrounding tissue (12).

The second aim of this study was to compare the different tumor size measurement techniques. Although it is well recognized that 1D and 2D tumor measurements are not as representative of tumor size as true volume calculations, especially in irregularly shaped tumors, they are widely employed, and indeed recently published response assessment criteria will increase the use of the 1D measurement (6). Despite "best-case" circumstances, our study has shown that 1D, 2D, and 3D techniques commonly used to measure tumor size are not comparable and both the 1D and 2D techniques overestimate tumor size. The 95% confidence intervals for the bias between the different measures are of an extent that we must assume to be clinically significant.

A slight limitation of this study is that, whereas the 2D WHO criteria (4) state that when there are multiple lesions all the products should be summed, this study differs from this as the 1D diameters have been summed and then multiplied by the perpendicular diameter to gain the product. It is unlikely that this difference would affect the overall result.

There are various sources of error present in all quantitative analysis of MR imaging that can be difficult to assess: partial volume, head tilt, plane of view, noncontiguous sections, contrast and intensity manipulation, and radio-frequency field inhomogeneities. Lesion-size measurement error can be assessed and has been shown in multiple studies to be significant with various influencing factors: lesion size and intra- and interobserver variabilities (13–15). Obviously, the 1D and 2D measurements require subjective selection of section showing maximum diameter and maximum area and so on, introducing further sources of error into the measurement (16, 17). It should be noted that most of these studies assessed measurement error with spherically shaped lesions, and, therefore, these errors could be accentuated greatly when investigating malignant glioma.

For our study, all measurements are based on the manual outlining of regions of enhancement by a single, experienced neuroradiologist, thus eliminating interobserver variability. Nevertheless, inherent intraobserver variability remains.

Conclusion

Consideration of the results presented here—showing that, in best-case circumstances, the 1D or 2D measurement techniques, as advocated by current response criteria (4–6, 18), are neither prognostic of overall survival nor compare with the more accurate volumetric measurement—must cast serious doubt on the use and validity of these response protocols in recurrent malignant glioma.

Appendix

Objective Calculation of 1D and 2D Measurement of Tumor Size Using Manually Drawn Regions of Interest

Regions of interest defined by the single observer were imported into Microsoft Excel using Text Import Wizard.

Header and footer information was deciphered and characters converted into x, y coordinate data for calculation of maximum diameter.

Validation of conversion was undertaken by plotting converted data using Microsoft Excel Chart Wizard and comparing with original region of interest.

For each region, the length between each x, y coordinate and every other x, y coordinate within region was calculated using Pythagorean theorem, to find the maximum diameter.

If more than one region per section, these maximum diameters were summed.

The overall largest diameter (or summed diameters) was recorded as 1D measurement of tumor size.

The number of sections with a region of interest multiplied by section thickness in the orthogonal orientation was multiplied by this overall largest diameter to obtain an objective 2D measurement of tumor size.

Acknowledgments

References

1. Chow KL, Gobin YP, Cloughesy T, et al. **Prognostic factors in recurrent glioblastoma multiforme and anaplastic astrocytoma treated with selective intra-arterial chemotherapy.** *AJNR Am J Neuroradiol* 2000;21:471–478
2. Reeves GI, Marks JE. **Prognostic significance of lesion size for glioblastoma multiforme.** *Radiology* 1979;132:469–471
3. Wood JR, Green SB, Shapiro WR. **The prognostic importance of tumor size in malignant gliomas: a computed tomographic scan study by the Brain Tumor Cooperative Group.** *J Clin Oncol* 1988;6:338–343
4. Miller AB, Hoogstraten B, Staquet M, Winkler A. **Reporting results of cancer treatment.** *Cancer* 1981;47:207–214
5. Macdonald DR, Cascino TL, Schold SC Jr, Cairncross JG. **Response criteria for phase II studies of supratentorial malignant glioma.** *J Clin Oncol* 1990;8:1277–1280
6. Therasse P, Arbuck SG, Eisenhauer EA, et al. **New guidelines to evaluate the response to treatment in solid tumors: European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada.** *J Natl Cancer Inst* 2000;92:205–216
7. Gregor A, Lind M, Newman H, et al. **Phase II studies of RMP-7 and carboplatin in the treatment of recurrent high grade glioma: RMP-7 European Study Group.** *J Neurooncol* 1999;44:137–145
8. Cox DR. **Regression models and life tables.** *J R Statist Soc B* 1972;34:187–220
9. Wong ET, Hess KR, Gleason MJ, et al. **Outcomes and prognostic factors in recurrent glioma patients enrolled onto phase II clinical trials.** *J Clin Oncol* 1999;17:2572
10. Bland M. *An introduction to medical statistics.* Oxford: Oxford University Press;2000:272
11. Warren KE, Patronas N, Aikin AA, et al. **Comparison of one-, two-, and three-dimensional measurements of childhood brain tumors.** *J Natl Cancer Inst* 2001;93:1401–1405
12. Kelly PJ, Daumas-Duport C, Scheithauer BW, et al. **Stereotactic**

histologic correlations of computed tomography- and magnetic resonance imaging-defined abnormalities in patients with glial neoplasms. *Mayo Clin Proc* 1987;62:450–459

13. Gurland J, Johnson RO. **How reliable are tumor measurements?** *JAMA* 1965;194:973–978

14. Chisholm RA, Stenning S, Hawkins TD. **The accuracy of volumetric measurement of high-grade gliomas.** *Clin Radiol* 1989;40:17–21

15. Warr D, McKinney S, Tannock I. **Influence of measurement error on response rates.** *Cancer Treat Rep* 1985;69:1127–1132

16. Clarke LP, Velthuizen RP, Clark M, et al. **MRI measurement of brain tumor response: comparison of visual metric and automatic segmentation.** *Magn Reson Imaging* 1998;16:271–279

17. Fornage BD. **Measuring masses on cross-sectional images.** *Radiology* 1993;187:289

18. Medical Research Council Brain Tumour Working Party. **Prognostic factors for high-grade malignant glioma: development of a prognostic index: a report of the Medical Research Council Brain Tumour Working Party.** *J Neurooncol* 1990;9:47–55