A novel, reproducible, and objective method for volumetric magnetic resonance imaging assessment of enhancing glioblastoma

Clinical article

Charles W. Kanaly, M.D., Ankit I. Mehta, M.D., Dale Ding, M.D., Jenny K. Hoang, M.B.B.S., Peter G. Kranz, M.D., James E. Herndon II, Ph.D., April Coan, M.P.H., Ian Crocker, M.D., Anthony F. Waller, Allan H. Friedman, M.D., David A. Reardon, M.D., and John H. Sampson, M.D., Ph.D.

Division of Neurosurgery, Department of Surgery, Department of Radiology, Department of Biostatistics and Bioinformatics, and Department of Pathology, Duke University Medical Center, Durham, North Carolina; Department of Neurosurgery, University of Virginia, Charlottesville, Virginia; Department of Radiation Oncology, Emory University, Atlanta, Georgia; and Dana-Farber Cancer Institute, Boston, Massachusetts

Object. Robust methodology that allows objective, automated, and observer-independent measurements of brain tumor volume, especially after resection, is lacking. Thus, determination of tumor response and progression in neurooncology is unreliable. The objective of this study was to determine if a semi-automated volumetric method for quantifying enhancing tissue would perform with high reproducibility and low interobserver variability.

Methods. Fifty-seven MR images from 13 patients with glioblastoma were assessed using our method, by 2 neuroradiologists, 1 neurosurgeon, 1 neurological resident, and 1 medical student. The 2 neuroradiologists also performed traditional 1-dimensional (1D) and 2-dimensional (2D) measurements. Intraclass correlation coefficients (ICCs) assessed interobserver variability between measurements. Radiological response was determined using Response Evaluation Criteria in Solid Tumors (RECIST) guidelines and MacDonald criteria. Kappa statistics described interobserver variability of volumetric radiological response determinations.

Results. There was strong agreement for 1D (RECIST) and 2D (Macdonald) measurements between neuroradiologists and neurosurgeons, but the agreement using the authors' novel automated approach was significantly stronger (ICC = 0.97). The volumetric method had the strongest agreement with regard to radiological response (κ = 0.96) when compared with 2D (κ = 0.54) or 1D (κ = 0.46) methods. Despite diverse levels of experience of the users of the volumetric method, measurements using the volumetric program remained remarkably consistent in all users (0.94).

Conclusions. Interobserver variability using this new semi-automated method is less than the variability with traditional methods of tumor measurement. This new method is objective, quick, and highly reproducible among operators with varying levels of expertise. This approach should be further evaluated as a potential standard for response assessment based on contrast enhancement in brain tumors.

(http://thejns.org/doi/abs/10.3171/2014.4.JNS121952)

Key Words • brain neoplasm • disease progression • glioma • magnetic resonance imaging • recurrence • oncology

Progression-free survival time based on contrast enhancement has become an accepted end point in glioma trials. However, current methods to assess radiological response in brain tumors are highly subjective and have high interobserver variability. To both accurately detect small changes in complex tumor configurations, especially after resection, and to strengthen the results from clinical trials that assess progression-free survival, there is a dire need for reproducible determination of radiological change in gliomas.

Abbreviations used in this paper: ICC = intraclass correlation coefficient; RANO = Response Assessment in Neuro-Oncology Working Group; RECIST = Response Evaluation Criteria in Solid Tumors.
Malignant gliomas are often very challenging to measure because they are commonly irregularly shaped, have cystic or hemorrhagic regions, have satellite lesions, are too small to be classified as measurable, or demonstrate a thin rim of enhancement around resection cavities. The Macdonald criteria are commonly used and apply 2-dimensional (2D) measurements. Many authors, however, have noted that the Macdonald criteria are still not comprehensive enough to describe subtle changes in complex tumor volumes, particularly when present in surgically created resection cavities, and a number of studies have questioned the validity of the criteria because they are subject to considerable interobserver variability. Precise and reproducible volumetric assessment is expected to be able to detect smaller changes in tumors and, therefore, could detect tumor response or progression sooner than 1D or 2D techniques. Prior studies have shown as much as a 26.1% discrepancy between response assessments made based on 2D or volumetric techniques. By taking the entire tumor volume into account, volumetric measurements should be more accurate and augment the power of clinical trials by reducing the variance in the measurements. Furthermore, results from previous studies demonstrate that radiological assessment by volumetric techniques exhibits decreased measurement variability when compared with 1D and 2D methods. Unfortunately, most previously published methods of volumetric assessment have required a considerable level of operator skill and experience because the operator must manually outline the specific region of tumor to calculate a tumor volume. As a result, volumetric assessments are prone to a significant amount of subjectivity and produce a high degree of interobserver variability. Additionally, these methods are hampered by the time required to perform the analysis and the difficulties specific to the CNS location, including frequently necrotic or cystic tumors, blood or other obscuring lesions intrinsically bright on T1-weighted MRI, trouble determining where tumor margins end, and differences between scans in the slice acquisition and the timing of contrast boluses. Because the relationship between MRI contrast enhancement and Gd concentration is not linear, the degree of brightness on standard MR images cannot be easily correlated with quantitative numbers.

In this study, we evaluate the performance of a new method of assessing CNS tumor volume that is highly reproducible and has low interobserver variability, even when used by examiners with vastly different levels of experience. The determination of radiological progression also had high reproducibility and low interobserver variability.

**Methods**

**Study Population**

A total of 57 MR images were analyzed from 13 patients with recurrent WHO Grade IV glioblastomas on 4 different experimental chemotherapy protocols. The 13 patients with recurrent glioblastoma used in our study all received standard of care therapy, including surgery (n = 11) or biopsy (n = 2), external beam radiation therapy, and temozolomide chemotherapy prior to recurrence. In addition, in all cases recurrence was documented by biopsy (n = 4) or unequivocal radiological progression according to standard protocols (n = 9). The patients were then followed-up using serial MRI, and placed on these chemotherapy protocols after radiological, clinical, or tissue diagnosis of progression had been made by the treating neurooncologist. These experimental protocols included topotecan with temozolomide, erlotinib with rapamycin, erlotinib with dasatinib, and an experimental chemotherapeutic agent BIBW 2992. A total of 13 baseline MRI scans and 44 follow-up scans were performed on these patients between August 2004 and December 2008. Baseline MRI scans were determined to be the first scan after resection. The median number of follow-up scans was 3. Patients ranged in age from 27 to 71 years old, with a median age of 48 years. The majority of the patients (85%) had a Karnofsky Performance Scale score ≥ 80 and some degree of resection (93%). A total of 38% of the patients were alive at the time of analysis, with a median follow-up duration of 14.6 months from study enrollment. All 57 MRI scans were assessed with our novel volumetric method by 2 neuroradiologists (J.K.H., P.G.K.), 1 neurosurgeon (J.H.S.), 1 neurosurgical resident (C.W.K.), 1 nurse practitioner, and 1 medical student; the 2 neuroradiologists assessed the scans with the 1D and 2D criteria. All 57 MRI scans were assessed with our novel volumetric method by 2 neuroradiologists (J.K.H., P.G.K.), 1 neurosurgeon (J.H.S.), 1 neurosurgical resident (C.W.K.), 1 nurse practitioner, and 1 medical student; the 2 neuroradiologists assessed the scans with the 1D and 2D criteria. The study was conducted in accordance with the Declaration of Helsinki and the US guidelines on good clinical practice. All data was collected at Duke University.

**Volumetric Method**

The volumetric method used in this study to calculate changes in enhancing tumor volume was a semi-automated, atlas-based segmentation program called VelocityAI (Velocity Medical Solutions). The detailed methodology involved in this new software has been previously reported (Fig. 1). In brief, each user was provided with a simple typed outline of the steps required to use the program and given a brief sample demonstration of the program that lasted less than 5 minutes. Afterward, the users were left to perform the volumetric analysis independently.

Briefly, each user loaded the DICOM files of both pre- and postcontrast T1-weighted axial MRI sequences of the brain into the computer program. The program automatically adjusted for motion and aligned the images between the two sequences. After alignment, the program automatically subtracted the precontrast images from the postcontrast images, so that only contrast-enhancing objects were used in later analyses and not intrinsically bright T1 objects, such as hemorrhage in a resection cavity. An atlas was loaded that automatically located and conforms to the nasal mucosa to normalize for Gd signal between scans. Next, the user manually used a drawing tool to grossly outline a region of interest of any size that contained the tumor and a separate small region of normal brain parenchyma on the same axial slices as
the previously outlined tumor. This rough outline of the area of interest consistently took only a few minutes per scan by every user to complete, regardless of the user’s prior level of expertise. The computer program then automatically compared the normal region of the precontrast scans with that of the postcontrast scans to determine a correction factor that was applied to all subsequent analyses to normalize between the two scans. Subsequently, the normalized, subtracted values in the enhancing nasal mucosa structure were used by the program to determine a quantitative cutoff for enhancement by excluding the top 5% of enhancement values and then calculating the 25% value of this remaining maximum value. Finally, the volume of pixels in the outlined tumor structure that met or exceeded this normalized enhancement threshold was automatically calculated by the program to be the enhancing tumor volume.

The 2 neuroradiologists also performed traditional 1D (Response Evaluation Criteria In Solid Tumors [RECIST])9 and 2D (Macdonald)18 measurements (Fig. 2). The radiological response assessment criteria used percentage changes from baseline in 1D and 2D measurements. The RECIST guidelines state that a decrease of at least 30% from baseline value is considered regression and an increase of 20% of the baseline is progression. The Macdonald criterion for regression was defined as a decrease of at least 50% from baseline values, and progression was defined as an increase of at least 25%. Our volumetric method defined regression as a decrease of at least 65% of the volume and progression as an increase in volume of at least 40%, based on previously published studies.7

Statistical Analysis

To determine the interobserver variability of MR image measurements (actual, percentage change from baseline, and percentage change from smallest) for each measurement method (1D, 2D, and novel volumetric), intraclass correlation coefficients (ICCs) were estimated from variance components of a hierarchical linear model analysis. The interobserver variability of radiological judgments (complete response/regression, stable disease, and progression) was examined by computing Cohen’s κ coefficient for the 1D and 2D methods (2 examiners) and Fleiss’ κ coefficient for the novel volumetric method (more than 2 examiners). Analyses were conducted under the assumption that each individual patient MRI scan was independent. Statistical analyses were conducted using SAS (version 9.2, SAS Institute Inc.) and an SAS macro “MKAPPA” developed by Westat Inc.

Results

Interobserver Variability

Two neuroradiologists who were blinded to the study objectives used our volumetric software or standard 1D (RECIST) and 2D (Macdonald) methods to calculate the tumor size for each patient. The traditional RECIST guidelines and Macdonald criteria determine progression based on the percentage change from baseline tumor size. Therefore, comparisons were made between the methods for percentage change from baseline tumor size. Calculated ICC is shown as an assessment of the interobserver variability of radiological measurements for each measurement method (Table 1). For the 3 methods used for assessing tumor size—1D, 2D, and our novel method—the actual volume measurement showed the percentage change from baseline to be 0.42, 0.61, and 0.97, respectively. Our novel method demonstrated the highest ICC.
between the 2 neuroradiologists in comparison with the other 1D and 2D methods.

Examiners of Varying Experience Levels

We had 6 examiners with various levels of experience use our volumetric software to assess the change in tumor volume using the volumetric, 1D, and 2D criteria. In Table 2, the actual volume measurements among the 6 examiners had a high ICC at 0.94. The percentage change from baseline of tumor volume was also consistent with a high ICC at 0.93. Even when used by examiners without clinical or radiological experience, our novel volumetric software method showed a strong ICC.

Radiological Response Assessment

Radiological response assessment was determined by neuroradiologists to be complete response, stable disease, or progression based on the 1D method, 2D method, or our volumetric method. The Cohen’s and Fleiss’ κ coefficients are presented in Table 3 as an assessment of the interobserver variability of determination of radiological progression or response for each of the measurement methods. The novel volumetric method had the strongest agreement between the 2 neuroradiologists (Cohen’s κ = 0.96) and also had strong agreement (Fleiss’ κ = 0.79) among all examiners. Cohen’s κ statistic for agreement between the 2 neuroradiologists’ findings for 2D and 1D was 0.54 and 0.46, respectively (Table 3).

Determination of Progression

In the 13 study patients, glioblastomas were determined to have progressed during the study period in 8 patients per an assessment by the treating neurooncologist. A comparison was performed between the times at which the initial determination (or call) of progression was made by each neuroradiologist to the times determined to be progression by the treating neurooncologist, using the 3 methods. If simultaneous determinations of progression occurred with different methods, it was counted as the initial call of progression for all methods that made that simultaneous call. Comparing across the 3 methods, the neurooncologist detected progression first in 1 of 8 patients, while 1 neuroradiologist using the novel volumetric method detected progression first in 6 of the 8 patients. The same neuroradiologist detected progression first in 3 of 8 patients using the 1D method and detected progression first in 4 of 8 patients with the 2D method. A similar comparison with the calls of progression made by the second neuroradiologist to the neurooncologist’s calls yielded similar results: 2 of 8 patients for the neurooncologist first, 5 of 8 patients using the novel volumetric method, 2 of 8 patients using the 1D method, and 4 of 8 patients using the 2D method.

Discussion

Our semi-automated volumetric method for quantifying enhancing tissue appears to represent a significant improvement over previous measurement techniques for numerous reasons. Because our method measures tumor volume, it incorporates more detail than 1D and 2D methods, whereby it is also able to calculate changes in tumors even in complicated, postresection shapes or in the presence of residual blood or other intrinsically bright T1 entities. This method is able to automatically correct for motion between the 2 sequences and automatically subtracts the precontrast images from the postcontrast images, which eliminates intrinsically bright T1 objects from the calculation of enhancement. This subtraction enables our method to detect enhancing tumor in the midst of any intrinsically bright T1 signals. This is a particularly important advantage for postoperative scans in which residual tumor may be hidden within blood in the resection cavity. Other studies have shown that similar subtraction techniques resulted in improved contrast detection. Additionally, because the relationship between MRI contrast enhancement and Gd concentration is not linear, the degree of brightness on MRI scans cannot be easily correlated with quantitative values. Our program overcomes this difficulty by normalizing the scans with comparisons of normal brain parenchyma between precontrast and postcontrast sequences, and it determines an individual enhancement threshold for each scan based on the automated nasal mucosa atlas. The top 5% of enhancement values in the nasal mucosa were excluded to

TABLE 1: Correlation of standard and volumetric measurements between 2 neuroradiologists

<table>
<thead>
<tr>
<th>Method</th>
<th>ICC*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>0.42</td>
</tr>
<tr>
<td>2D</td>
<td>0.61</td>
</tr>
<tr>
<td>volumetric</td>
<td>0.97</td>
</tr>
</tbody>
</table>

* The calculated ICCs are shown as an assessment of the interobserver variability of radiological measurements for each measurement method. There were 57 total scans (44 comparisons made to 13 baseline scans).
## TABLE 2: Correlation of volumetric measurements*

<table>
<thead>
<tr>
<th>Measurement†</th>
<th>Examiner No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>actual (n = 57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
</tr>
<tr>
<td>median (cm³)</td>
<td></td>
<td>6.75</td>
<td>13.35</td>
<td>11.17</td>
<td>8.50</td>
<td>11.61</td>
<td>10.35</td>
<td></td>
</tr>
<tr>
<td>range (cm³)</td>
<td></td>
<td>0.4–53.22</td>
<td>1.6–73.56</td>
<td>0.16–65.59</td>
<td>1.3–59.37</td>
<td>1.29–65.28</td>
<td>1.11–89.85</td>
<td></td>
</tr>
<tr>
<td>% change from baseline (n = 44)</td>
<td></td>
<td>-12.91</td>
<td>5.67</td>
<td>2.11</td>
<td>-15.01</td>
<td>5.16</td>
<td>-4.06</td>
<td>0.93</td>
</tr>
<tr>
<td>median</td>
<td></td>
<td>-85.06 to 1189.66</td>
<td>-79.53 to 1691.55</td>
<td>-93.43 to 1006.39</td>
<td>-86.22 to 1749.38</td>
<td>-91.94 to 1495.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>range</td>
<td></td>
<td>97.25 to 1212.98</td>
<td>1197.25 to 1212.98</td>
<td>1197.25 to 1691.55</td>
<td>1197.25 to 1749.38</td>
<td>1197.25 to 1495.39</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The volume measurements between the 6 examiners had a high ICC at 0.94; the percentage change from baseline of tumor volume was also consistent with a high ICC at 0.93. Even when used by examiners without clinical or radiological experience, there was a strong ICC using our novel volumetric software method.
† Interobserver agreement on MR image measurements of both expert and nonexpert examiners.

## TABLE 3: Assessment of interobserver variability*

| Neuroradiologist 2 (no. of patients) | 1D | 2D | Volumetric | Neuroradiologist 1 | Comp | Stable | Prog | Total | Comp | Stable | Prog | Total | Comp | Stable | Prog | Total |
|---------------------------------------|----|----|------------|---------------------|------|--------|------|-------|------|--------|------|-------|------|-------|
| comp                                  |    |    |            | Comp                | 4    | 3      | 4    | 11    | 5    | 2      | 4    | 11    | 5    | 0     | 0    | 5     |
| stable                                |    |    |            | Stable              | 2    | 17     | 1    | 20    | 1    | 15     | 2    | 18    | 0    | 18    | 0    | 18    |
| prog                                  |    |    |            | Prog                | 2    | 3      | 8    | 13    | 2    | 2      | 11   | 15    | 0    | 1     | 20   | 21    |
| concordance                           | 65.9% (29/44) | 70.5% (31/44) | 97.7% (43/44) | Cohen's κ†          | 0.46 | 0.54   | 0.96 |       |       |       |       |       |       |

* Cohen’s and Fleiss’ κ coefficients are presented as an assessment of the interobserver variability of determining radiological progression or response for each of the methods. The novel volumetric method had the strongest agreement between the 2 neuroradiologists (Cohen’s κ = 0.96) and also had strong agreement (Fleiss’ κ = 0.79) among all examiners. Cohen’s κ statistic for agreement between the 2 neuroradiologists for 2D and 1D were 0.54 and 0.46, respectively. Comp = complete response/regression; Prog = progression.
† Interobserver agreement on radiological evaluations of examiners for the 1D, 2D, and volumetric methods.
avoid error because of the nonlinear relationship between enhancement and Gd concentration. The enhancement threshold was selected to be 25%, based on the expert opinion of the radiologist as to which threshold value most closely corresponded with his or her subjective judgment of tissue enhancement. Our prior publication on this method presented data showing that as long as the threshold value is consistently maintained, the results are strongly correlated. These data suggest that response assessment conclusions are largely insensitive to a specific threshold level. This threshold method compensates for differences between scans in contrast boluses and timing, and is similar to a method that is commonly used and has been validated with PET scans.

One possible concern about our method is that although the volumes generated by this method are highly precise, this does not necessarily mean that the measurements are accurate. This is unfortunately a limitation of all radiological measurements, as there is no definite “gold standard” volumetric measurement technique. The only way to absolutely know the volume of tumor is to resect it entirely and physically measure the volume. Comparisons between our method and any other volumetric method would simply be measuring agreement between two different estimations, neither of which has been proven to be accurate. However, it does appear that increasing automation of radiological measurements improves accuracy and reproducibility when compared with manual techniques. Computer-assisted methods of volume calculation have been shown to have less variability than manual calculations. Furthermore, our method does not require the time or expertise that manual volumetric methods require. The increased automation of our approach significantly improves the reproducibility of tumor volumes and determination of progression obtained across different users. Our results suggest that reliable tumor volumes can be determined even by users with minimal expertise in the field of neuroradiology. Therefore, our method does not rely on the subjective interpretation of scans that is inherent in many other volumetric methods. The program calculates the amount of enhancing tissue within the grossly outlined tumor region. The processing steps are highly automated, which minimizes both the time required for analysis and the expertise needed by the user. When comparing neuroradiological response assessments in gliomas, Shah et al. demonstrated ICCs for 1D, 2D, and 3D criteria of 0.874, 0.822, and 0.889, respectively. Our volumetric method demonstrates a superior interobserver ICC at 0.97, whereas 1D and 2D were similar (0.42 and 0.61, respectively). In a study of the radiological assessment of response to chemotherapy in glioma, Vos et al. showed a 2D ICC of 0.64 (0.61 in our study) and a response classification \( \kappa \) of 0.51 (0.54 in our study).

Although our method represents a considerable advance in radiological assessment for high-grade gliomas, there are a number of remaining difficulties that it does not yet solve. Most significantly, this method was developed to measure the volume of enhancing tumor and, therefore, will not detect nonenhancing tumor. This drawback will be of increasing importance as more patients begin antiangiogenic therapy with agents such as bevacizumab that normalize vasculature and dramatically decrease enhancement. For patients in this situation, the measurement of enhancing tumor is no longer an adequate assessment of tumor burden, although even in these studies tumor response burden measurements of enhancing tumor predict survival. The measurement of nonenhancing tumor is also extremely important for the assessment of low-grade tumors. The Response Assessment in Neuro-Oncology Working Group (RANO) developed new standardized response criteria for high-grade gliomas in clinical trials that account for nonenhancing tumor. The RANO criteria use T2/FLAIR imaging to incorporate an increase in nonenhancing tumor burden to determine progression. Adaptation of our method for use with T2/FLAIR imaging sequences is currently being attempted to assess nonenhancing tumor volumes. Although technology will need to be developed to accurately measure nonenhancing tumor volume, no other currently accepted method is able to accurately address this issue. This method is able to quantify enhancing volume, but this calculated volume is not necessarily tumor and the measured enhancement could be secondary to radiation necrosis or pseudoprogression. This is an inherent limitation of any method that measures enhancement, as a measurement change does not necessarily imply a specific diagnosis of tumor recurrence. Clinical judgment or even possibly tissue biopsy may be required to determine the true meaning of these changes in enhancement. It is possible that knowledge of the time course of these enhancement changes or serial imaging to determine trends in these changes may be able to distinguish between treatment effect and recurrence. Future investigations are planned to determine if recurrence can be distinguished from other enhancement changes based on the serial measurements of enhancement over time.

**Conclusions**

In this study we demonstrate that our semi-automated method of quantifying enhancing intracranial tumor volume has very low variability among users with widely divergent levels of expertise. Even when comparing the measurements made by 2 neuroradiologists, this new method had greater agreement than traditional 1D and 2D methods. This technique has the potential to gain wide acceptance in clinical practice because users with all levels of expertise appear to obtain highly similar results, it only requires a few minutes to perform the analysis, and there is no need for special MRI sequences or computers. Such an approach may also then be applicable to the analysis of benign tumors such as meningioma, and might also be useful to follow radiosurgical responses over time as well. Future studies with larger numbers of patients will need to be performed to validate this technique.

**Disclosure**

Anthony F. Waller has a consulting relationship with Velocity Medical Solutions. Dr. Ian Crocketer is a co-founder of Velocity Medical Solutions and is entitled to royalties on sales based on the intellectual property agreement between Velocity Medical Solutions and
Volumetric MRI assessment of glioblastoma

Emory University, but had no role in data collection and analysis. Dr. Crocker did provide technical support and design modifications to the software to adapt it to this novel use. This work was supported in part by NIH R25 grants to Drs. Mehta and Sampson, NIH grant no. 5R01-CA135272-03 to Dr. Sampson, the Goldhirsh Foundation, Pediatric Brain Tumor Foundation of the United States, NIH grant no. 5R21-NS067975-02, and The Ben and Catherine Ivy Foundation.

Author contributions to the study and manuscript preparation include the following. Conception and design: Sampson, Kanaly. Acquisition of data: Kanaly, Ding, Hoang, Kranz. Analysis and interpretation of data: Kanaly, Mehta, Ding, Hoang, Kranz. Drafting the article: Sampson, Kanaly, Mehta. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Sampson. Statistical analysis: Herndon, Coan. Administrative/technical/material support: Sampson. Study supervision: Sampson.

References


Manuscript submitted February 5, 2013. Accepted April 15, 2014.

Please include this information when citing this paper: published online July 18, 2014; DOI: 10.3171/2014.4.JNS121952.

Address correspondence to: John H. Sampson, M.D., Ph.D., M.H.Sc., Division of Neurosurgery, P. O. Box 3050, Rm. 220, Sands Building, Research Dr., Duke University Medical Center, Durham, NC 27710. email: john.sampson@duke.edu.