Published in final edited form as:

J Neurooncol. 2014 January ; 116(1): 127–133. doi:10.1007/s11060-013-1266-4.

Measurements of the pons as a biomarker of progression for pediatric DIPG

Emilie A. Steffen-Smith,

Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Building 10/Room 1-5750, 9000 Rockville Pike, Bethesda, MD 20892, USA

Eva H. Baker,

Department of Radiology and Imaging Sciences, Clinical Center, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA

David Venzon,

Biostatistics and Data Management Section, National Cancer Institute, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA

Shaefali Shandilya,

Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Building 10/Room 1-5750, 9000 Rockville Pike, Bethesda, MD 20892, USA

Robyn S. Bent, and

Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Building 10/Room 1-5750, 9000 Rockville Pike, Bethesda, MD 20892, USA

Katherine E. Warren

Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Building 10/Room 1-5750, 9000 Rockville Pike, Bethesda, MD 20892, USA

Abstract

Treatment of pediatric diffuse intrinsic pontine glioma (DIPG) remains challenging, and reliable bio- markers of response are lacking. Radiographic response is a primary endpoint in many investigational studies of brain tumors, but there is no standard method of tumor measurement for DIPG, significant inter-observer variability exists given the invasive nature of these tumors, and tumor measurements are not predictive of outcome. Because DIPGs involve a significant portion of the pons, we evaluated the reliability and prognostic value of one-dimensional (1D) and two-dimensional (2D) pons measurements using anatomical landmarks rather than tumor boundaries. Patients with DIPG ($n = 75$) were evaluated longitudinally at our institution using MRI. Four readers independently performed 1D and 2D measurements of the pons using FLAIR images. Agreement and inter-reader variability were evaluated using differences among the six reader pairs and the coefficient of variation (CV). Prognostic value of pons measurements was calculated using Cox proportional hazards models, where relative hazard (RH) represents risk of death. Readers

steffene@mail.nih.gov, warrenk@mail.nih.gov.

Conflict of interest

The authors have no conflicts of interest to report.

evaluated 384 exams. Agreement of readers' 1D and 2D measurements was strong (median difference between reader pairs 3.1 and 5.4 %, respectively), with low inter-reader variability (median CV = 3.1 % and median CV = 4.8 %, respectively). Increases in 1D and 2D pons measurements over time indicated poorer prognosis (RH = 2.29, $p = 0.0025$ and RH = 1.13, $p = 0.0016$, respectively), with shorter overall survival. Pons measurements had low inter-reader variability compared to previously reported tumor measurement techniques and correlated with outcome in children with DIPG. Measurements of the pons (as opposed to direct measurements of tumor) are a viable in vivo bio- marker for DIPG.

Keywords

Pediatric brain tumors; Diffuse intrinsic pontine glioma; Magnetic resonance imaging; Measurements; Prognosis

Introduction

Diffuse intrinsic pontine gliomas (DIPGs) represent the majority (>75 %) of pediatric brainstem tumors and carry the worst prognosis of all childhood brain tumors [1, 2]. Current treatment strategies for DIPG remain unsuccessful with a median survival of less than 12 months and less than 10 % survival 2 years after diagnosis [3, 4]. With limited standard treatment options (i.e., radiation therapy) and a dismal prognosis, many patients enroll in investigational trials. Endpoints for assessing treatment response and tumor progression in most investigational studies of brain tumors are based on radiographic response, typically calculated by one-dimensional (1D) or two-dimensional (2D) tumor measurements using signal differences found on standard clinical MRI sequences. However, DIPGs are infiltrative tumors with indistinct borders, and tumor measurements are difficult to obtain reliably and consistently, even among experienced readers [5]. MRI findings such as contrast enhancement are variable among patients, non-specific and not predictive of outcome [6–8]. Advanced imaging techniques such as magnetic resonance spectroscopy (MRS) and perfusion imaging have been used to evaluate DIPG and have yielded potential biomarkers. Increased Cho:NAA from MRS and increased perfusion are both associated with poor patient outcome [9, 10]. However, these techniques require additional scanning time and specialized analysis techniques, which are not standardized across multiple institutions. The purpose of this study was to evaluate a potential biomarker for DIPG using commonly acquired clinical MRI data.

Although the margins of DIPG itself are often indistinct, the anatomical boundaries of the pons generally remain visible and can be clearly demarcated. Based on this observation, we hypothesized that measurements of the pons using anatomical landmarks, rather than direct measurements of the tumor based on signal abnormalities within the pons and surrounding tissues, may be a more reliable in vivo biomarker for evaluating children with DIPG. We investigated the reproducibility and variability of 1D and 2D measurements of the pons among readers and the prognostic value of these pons measurements in children with DIPG.

Materials and methods

Patients

This retrospective study included children (21 years old or less at diagnosis) enrolled in an IRB-approved protocol at our institution between October 1997 and January 2012 for the evaluation or treatment of newly diagnosed, recurrent, or refractory DIPG. Diagnosis of all patients was confirmed at the time of evaluation based on clinical presentation and radiographic features. Patients having predominantly exophytic lesions were excluded, as the clinical course for these patients is not representative of classic DIPG. Additional inclusion criteria were as follows—1) Patients underwent at least one MRI at our institution that included pre- or post-contrast axial fluid attenuated inversion recovery (FLAIR) images, 2) MRI studies were archived and available for review on our institution's picture archiving and communication system (PACS) and 3) MRI studies were reviewed by a neuroradiologist at the time of imaging and found to be of acceptable quality for clinical use (e.g., no significant image distortion, no motion artifacts).

MRI

MRI exams were performed at our institution on a 1.5 or 3 T scanner (GE Medical Systems, Milwaukee WI and Philips, Best, Netherlands). Most patients were sedated during the procedure; propofol was used in the majority of these cases, as was supplemental oxygen. Clinical MRI sequences included pre- and post-contrast T1-weighted, pre-contrast T2-weighted and pre- or post- contrast FLAIR images. For patients with multiple MRI exams, clinical images were co-registered to the first available scan performed at our institution using an automated registration processing technique developed using MEDx (Medical Numerics, Inc., Germantown MD). Co-registration ensured that all relative changes in pons measurements between scans were real and not artifact due to variation in patient position relative to the slices. FLAIR images (maximum slice thickness = 5 mm, range = 2–5 mm, no interslice gap) were selected for pons measurements based on the superior inter-observer reliability of FLAIR measurements in patients with DIPG compared to other clinical imaging sequences [5].

Pons measurements

Four readers—one neuroradiologist (E.B.), one neurooncologist (K.W.), and two MRI analysis specialists with measurement training (E.S. and S.S.)—independently performed pons measurements using the standard measurement tools on our institution's PACS workstations. For each scan, each reader identified the slice on the axial FLAIR image with the maximum diameter of the pons in any direction (d_1) and the corresponding maximum perpendicular diameter to d_1 within the same axial slice (d_2), as shown in Fig. 1. d_1 represented the 1D measurement of the pons. 2D measurements of the pons were calculated using the product of d_1 and d_2 . Anatomical landmarks including the cerebellopontine angle and cistern and the sulcus formed by the junction of the pons and the middle cerebral peduncle were used to delineate the boundaries of the pons (Fig. 2). Cystic regions of the tumor were included in pons measurements when the entire cystic area was considered to be within the anatomical boundaries of the pons.

Statistical analysis

Inter-observer variability was estimated by first calculating the coefficient of variation (CV) over the four readers for each scan, where the CV is defined as the standard deviation divided by the mean, and then calculating the median and the mean of these CVs. Agreement of 1D and 2D measurements among readers was evaluated using a pair-wise Spearman rank correlation. Agreement among readers was further evaluated by calculating the median and the 95th percentile of the percent differences between each of six reader pairs for each scan, where the percent difference is defined as the absolute value of the difference divided by the mean of the two measurements of the reader pair. We also evaluated the percentage of studies in which the percent difference was greater than 25 %, as reported in a previous study of tumor measurement techniques in children with DIPG [5].

In children, the normal pons continues to grow from infancy until adulthood [11]. Given the broad age inclusion criteria of this study (e.g., 21 years) and length of follow-up in some of our cases (>500 weeks), changes in measurements of the pons could be the result of both tumor growth and normal, age-related growth. Age-matched normal controls were not obtained for comparison in our study. Rather, we examined the relationship between pons measurements and patient age using a Spearman rank correlation.

Averaged 1D and 2D measurements from all four readers at each exam were used to assess the prognostic value of 1D and 2D pons measurements. Overall survival time was calculated as the time from the first scan to the date of death or the last follow-up for those patients still under evaluation at the time of analysis. Median 1D and 2D pons measurements were used to compare survival among patients using the Kaplan–Meier method and log-rank test. The likelihood ratio test of univariate and multivariate Cox proportional hazards models was used to assess the prognostic value of measurements from each patient's first exam at our institution. We then applied the same models with time-dependent covariates to evaluate the prognostic value of changes in 1D and 2D measurements over time during subsequent exams. Changes over time were evaluated as incremental changes in measurements (10 mm for 1D, 100 mm² for 2D) and percent increases (C10 % for 1D, C25 % for 2D). We chose to include the relationship between outcome and percent changes in pons measurements as this type of evaluation is more commonly used for measurement techniques than absolute changes. Patient prognosis associated with changes in 1D and 2D measurements was expressed as a relative hazard (RH), where $RH > 1$ represents increased risk of death (i.e., poorer prognosis). Each estimated RH is adjusted for the effects of the other risk factors in the model. $p < 0.05$ was considered statistically significant for all analyses. Data were analyzed using SAS version 9.1.3 (SAS Institute Inc., Cary, NC).

Results

Patients and imaging

Seventy-five patients had MRI exams that met inclusion criteria for this study. Median age at diagnosis was 6 years (range = 2–17 years). Median overall survival from the first exam was 38 weeks (range = 4–520 + weeks). 384 exams with FLAIR imaging (251 pre-contrast, 133 post-contrast) were evaluated, with a median of five exams per patient (range = 1–32 exams

per patient). Time between MRI exams was variable among patients (median = 5 weeks, range = 0.1–108.8 weeks).

Measurement

Each reader performed a total of 384 1D and 2D measurements. For one patient, the 2D measurements on each of his exams ($n = 3$) were substantially different for one reader compared to the other three readers. These three cases were included in our evaluations. 1D and 2D measurements averaged across all four readers are presented in Table 1.

Variability and agreement of measurements among readers

Median CV for 1D measurements across all exams was 3.1 % (mean = 3.7 %, interquartile range = 1.8–4.8 %). Median CV for 2D measurements across all exams was % (mean = 5.8 %, interquartile range = 3.3–7.1 %). We found a strong correlation of both 1D and 2D measurements among the four readers (range of Spearman rank correlation coefficient, $\rho = 0.84$ – 0.94 for 1D measurements and $\rho = 0.93$ – 0.97 for 2D measurements). Median percent difference among reader pairs was 3.1 % for 1D and 5.4 % for 2D (Table 2). Percentage of studies in which readers' 1D and 2D measurements differed by more than 25 % was low (0.7 and 2.6 %, respectively). These results show that variation among readers was relatively low for both 1D and 2D measurements compared to a previous analysis of 1D and 2D tumor measurements in DIPG which reported significant variability among readers using FLAIR images [5]. Also, as in previous studies, we found lower Variability in 1D measurements compared to 2D measurements.

Patient age and pons measurements

Overall, we observed no significant trend between pons measurements and patient age (Spearman rank correlation coefficients 0.18 for 1D measurements and 0.15 for 2D, $p > 0.10$ for each). 1D and 2D measurements from the patients' first exams were slightly larger in patients over 6 years of age (medians 48.8 mm and 1,741 mm², respectively, vs 45.0 mm and 1,532 mm² for patients under 6 years), but the differences were not statistically significant ($p = 0.10$ for each) and within each age group, correlation coefficients were close to zero (range = -0.06 – 0.04 , $p > 0.7$).

Correlation between measurements and survival

Kaplan–Meier two-group analysis found patients with 2D pons measurements above the median (1,642 mm²) had significantly shorter survival ($p = 0.0022$) compared to patients with measurements below the median (Fig. 3). Differences in survival for patients with 1D pons measurements above and below the median were not significant ($p = 0.087$). Table 3 summarizes the results of the correlation of pons measurements with patient prognosis as assessed using Cox proportional hazards models. Patients had shorter overall survival times and a poorer prognosis in proportion to higher 2D FLAIR measurements at their first exam ($p = 0.0004$). Both 1D and 2D measurements taken from each patient's first exam were associated with overall survival ($p = 0.006$ and $p = 0.0004$, respectively). However, 2D measurements had a much stronger association with patient outcome compared to 1D measurements. Change in pons measurement over time (i.e. during therapy) was also

significant. Patients with increasing 1D and 2D pons measurements had a much poorer prognosis, estimated as a >1.1 increased risk of death (i.e., relative hazard ratio) for each 100 mm² increase and a >2 -fold increased risk of death for each 10 mm increase in 1D measurement, compared to patients with stable or decreasing pons measurements (Table 3). Larger increases in 2D measurements were associated with a greater risk of death. For example, a 600 mm² increase was associated with a >2 -fold risk of death and an 850 mm² increase was associated with a >3 -fold risk of death. As shown in Table 4, a 10 % increase in 1D measurements or a 25 % increase in 2D measurements gave a patient a much poorer prognosis and shorter expected overall survival. In some cases, pons measurements decreased compared to the previous exam. A 5 % decrease in 1D measurements or a 10 % decrease in 2D measurements tended toward a better prognosis for patients, but these findings did not reach statistical significance (Table 4).

Discussion

Conventional MRI is typically used to monitor changes and response in children with DIPG. However, response criteria (e.g., contrast enhancement and tumor measurements) are not standardized and are difficult to obtain given the widespread infiltration, irregular borders and heterogeneous appearance of DIPG on MRI. 1D (e.g., response evaluation criteria in solid tumors, RECIST criteria) and 2D (e.g., MacDonald criteria) tumor measurement techniques are commonly used for determining radiographic response, but as previously reported, without distinct tumor boundaries, measurements of DIPG are difficult to obtain and highly-variable among readers [5]. For patients with DIPG, the majority of the tumor burden is, by definition of the diagnosis, centralized in the pons. Despite invasion and expansion of the pons, many of the anatomical landmarks remain intact and visible on MRI. We hypothesized that changes in the size of the pons reflect changes in tumor burden and that measurements of the pons may be a more reliable alternative to direct tumor measurements for monitoring treatment-related changes and predicting outcome. Unlike previously reported tumor measurement techniques in which tumor boundaries can be difficult to define, pons measurements as defined by the anatomical boundaries of the pons, can be obtained consistently and reliably among readers and serve as clinically relevant markers of outcome. In this study, we found lower inter-reader variability using pons measurements compared to previous analyses of tumor measurements in children with DIPG [5].

Treatment of children with DIPG remains challenging. Recent histological studies of these tumors at autopsy have revealed several potential therapeutic targets [12–18]. However, the eloquent location of these tumors severely limits tissue sampling in these patients, and in vivo biomarkers are needed for clinically relevant endpoints in clinical trials to determine drug efficacy and treatment response. Accurate, reproducible assessment of tumor burden is important in gauging response to therapy. In this study, increasing pons measurements were strongly associated with patient prognosis. Results from our analysis indicate that patients with a large increase in 2D measurements will have a shorter survival time compared to patients with smaller increases, suggesting that changes in size of the pons during therapy may be indicative of response or lack of response, and 1D and 2D pons measurements are potential candidates for surrogate endpoints in clinical trials. Agreement among readers in

our study was strong, with low inter-reader variability and few cases with more than 25 % disagreement, despite the varying levels of experience in the analysis and interpretation of brain MRIs among the four readers. These findings suggest that pons measurements should be reproducible at multiple institutions.

We also demonstrate that larger pons size at the patient's initial evaluation and changes in 1D and 2D pons size over time are prognostic. Overall, 1D measurements generally had weaker associations with survival than 2D measurements. Few prognostic factors have been identified for DIPG, and most are based on advanced imaging techniques such as MRS [9, 19–21] or dynamic susceptibility contrast perfusion [10], which require additional time for image acquisition and analysis and (in the case of perfusion imaging) additional exposure to MR contrast agents. The 1D and 2D pons measurements used in our study can be made on images that are part of the standard clinical MRI exam performed at most institutions, without the need for additional imaging sequences, specialized equipment or specialized analysis software.

Limitations

This study was performed retrospectively and evaluated a heterogeneous group of patients with DIPG seen at our institution during a 15-year period. As such, there was some variation in scan sequences parameters, such as slice thickness due to changes in our imaging protocols over time. Post-contrast FLAIR was used for measurements in studies where pre-contrast FLAIR was not obtained. Use of both pre- and post-contrast FLAIR images may present some additional variation in pons measurements, but the contribution is probably minor as anatomical boundaries, rather than signal characteristics, were used to guide the measurements. As patients were referred to our institution at varying stages of their disease course, timing of scans used for pons measurements was variable relative to time from diagnosis and clinical course. Timing of follow-up scans was also variable, depending on the imaging requirements of the investigational study in which the patient was enrolled and the availability of families to return for follow-up imaging at our institution rather than their home institution. Such variability in treatment course and timing of imaging exams precluded an evaluation of growth rates in our study population. We examined the relationship between pons measurements and patient age using a Spearman rank correlation rather than evaluate age- matched controls. We expect the contribution of normal growth to be minimal as the growth of the normal pons is small compared to tumor growth in these patients and the majority of these patients survived less than a year.

Conclusions

We demonstrate that measurement of the entire pons is feasible, reproducible, and prognostic for children with DIPG. In this study, 1D and 2D pons measurements were obtained with a strong correlation and much lower variability among multiple readers compared to previous reports of tumor measurements in DIPG. Both actual and percentage changes in 1D and 2D measurements over time had a significant association with patient survival. Based on the findings of this study, pons measurements are a reliable, in vivo biomarker for DIPG.

Acknowledgments

This work was presented in part at the 2012 International Society of Pediatric Neuro-Oncology meeting in Toronto, Ontario Canada. This research was supported in part by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. The views expressed do not necessarily represent the views of the National Institutes of Health or the United States Government.

References

1. Donaldson SS, Laningham F, Fisher PG (2006) Advances toward an understanding of brainstem gliomas. *J Clin Oncol* 24:1266–1272 [PubMed: 16525181]
2. Warren KE (2012) Diffuse intrinsic pontine glioma: poised for progress. *Front Oncol* 2:205 [PubMed: 23293772]
3. Jansen MH, van Vuurden DG, Vandertop WP, Kaspers GJ (2012) Diffuse intrinsic pontine gliomas: a systematic update on clinical trials and biology. *Cancer Treat Rev* 38:27–35 [PubMed: 21764221]
4. Hargrave D, Bartels U, Bouffet E (2006) Diffuse brainstem glioma in children: critical review of clinical trials. *Lancet Oncol* 7:241–248 [PubMed: 16510333]
5. Hayward RM, Patronas N, Baker EH, Vezina G, Albert PS, Warren KE (2008) Inter-observer variability in the measurement of diffuse intrinsic pontine gliomas. *J Neurooncol* 90:57–61 [PubMed: 18587536]
6. Hargrave D, Chuang N, Bouffet E (2008) Conventional MRI cannot predict survival in childhood diffuse intrinsic pontine glioma. *J Neurooncol* 86:313–319 [PubMed: 17909941]
7. Liu AK, Brandon J, Foreman NK, Fenton LZ (2009) Conventional MRI at presentation does not predict clinical response to radiation therapy in children with diffuse pontine glioma. *Pediatr Radiol* 39:1317–1320 [PubMed: 19657635]
8. Kornreich L, Schwarz M, Karmazyn B, Cohen IJ, Shuper A, Michovitz S, Yaniv I, Fenig E, Horev G (2005) Role of MRI in the management of children with diffuse pontine tumors: a study of 15 patients and review of the literature. *Pediatr Radiol* 35:872–879 [PubMed: 15918050]
9. Steffen-Smith EA, Shih JH, Hipp SJ, Bent R, Warren KE (2011) Proton magnetic resonance spectroscopy predicts survival in children with diffuse intrinsic pontine glioma. *J Neurooncol* 105:365–373 [PubMed: 21567301]
10. Hipp SJ, Steffen-Smith E, Hammoud D, Shih JH, Bent R, Warren KE (2011) Predicting outcome of children with diffuse intrinsic pontine gliomas using multiparametric imaging. *Neuro Oncol* 13:904–909 [PubMed: 21757444]
11. Raininko R, Autti T, Vanhanen SL, Ylikoski A, Erkinjuntti T, Santavuori P (1994) The normal brain stem from infancy to old age. A morphometric MRI study. *Neuroradiology* 36:364–368 [PubMed: 7936176]
12. Warren KE, Killian K, Suuriniemi M, Wang Y, Quezado M, Meltzer PS (2012) Genomic aberrations in pediatric diffuse intrinsic pontine gliomas. *Neuro Oncol* 14:326–332 [PubMed: 22064882]
13. Oka H, Utsuki S, Tanizaki Y, Hagiwara H, Miyajima Y, Sato K, Kusumi M, Kijima C, Fujii K (2012) Clinicopathological features of human brainstem gliomas. *Brain Tumor Pathol* 30(1):1–7 [PubMed: 22484454]
14. Grill J, Puget S, Andreiulo F, Philippe C, MacConaill L, Kieran MW (2012) Critical oncogenic mutations in newly diagnosed pediatric diffuse intrinsic pontine glioma. *Pediatr Blood Cancer* 58:489–491 [PubMed: 22190243]
15. Puget S, Philippe C, Bax DA, Job B, Varlet P, Junier MP, Andreiulo F, Carvalho D, Reis R, Guerrini-Rousseau L, Roujeau T, Dessen P, Richon C, Lazar V, Le Teuff G, Sainte-Rose C, Geoerger B, Vassal G, Jones C, Grill J (2012) Mesenchymal transition and PDGFRA amplification/mutation are key distinct oncogenic events in pediatric diffuse intrinsic pontine gliomas. *PLoS One* 7:e30313 [PubMed: 22389665]
16. Barrow J, Adamowicz-Brice M, Cartmill M, MacArthur D, Lowe J, Robson K, Brundler MA, Walker DA, Coyle B, Grundy R (2011) Homozygous loss of ADAM3A revealed by genome-wide

analysis of pediatric high-grade glioma and diffuse intrinsic pontine gliomas. *Neuro Oncol* 13:212–222 [PubMed: 21138945]

17. Zarghooni M, Bartels U, Lee E, Buczkowicz P, Morrison A, Huang A, Bouffet E, Hawkins C (2010) Whole-genome profiling of pediatric diffuse intrinsic pontine gliomas highlights platelet-derived growth factor receptor alpha and poly (ADP-ribose) polymerase as potential therapeutic targets. *J Clin Oncol* 28:1337–1344 [PubMed: 20142589]
18. Paugh BS, Qu C, Jones C, Liu Z, Adamowicz-Brice M, Zhang J, Bax DA, Coyle B, Barrow J, Hargrave D, Lowe J, Gajjar A, Zhao W, Broniscer A, Ellison DW, Grundy RG, Baker SJ (2010) Integrated molecular genetic profiling of pediatric high-grade gliomas reveals key differences with the adult disease. *J Clin Oncol* 28:3061–3068 [PubMed: 20479398]
19. Yamasaki F, Kurisu K, Kajiwaru Y, Watanabe Y, Takayasu T, Akiyama Y, Saito T, Hanaya R, Sugiyama K (2011) Magnetic resonance spectroscopic detection of lactate is predictive of a poor prognosis in patients with diffuse intrinsic pontine glioma. *Neuro Oncol* 13:791–801 [PubMed: 21653595]
20. Panigrahy A, Nelson MD, Jr, Finlay JL, Spoto R, Krieger MD, Gilles FH, Bluml S (2008) Metabolism of diffuse intrinsic brainstem gliomas in children. *Neuro Oncol* 10:32–44 [PubMed: 18003889]
21. Laprie A, Pirzkall A, Haas-Kogan DA, Cha S, Banerjee A, Le TP, Lu Y, Nelson S, McKnight TR (2005) Longitudinal multivoxel MR spectroscopy study of pediatric diffuse brainstem gliomas treated with radiotherapy. *Int J Radiat Oncol Biol Phys* 62:20–31 [PubMed: 15850898]

Fig. 1.
Pons measurement included the maximum diameter in the pons (d_1) and its perpendicular (d_2) using the slice of the axial FLAIR image where the pons was the largest

Fig. 2. Anatomical landmarks used to guide pons measurements— **a** basilar artery, **b** fourth ventricle, **c** cerebellopontine angle cistern, and **d** sulcus formed by the junction of the pons and the middle cerebral peduncle

Fig. 3. Kaplan–Meier curve showing differences in survival for patients with pons measurements above (lower line, black) or below (upper line, grey) the median 2D pons measurement. Patients with 2D pons measurements above the median had significantly shorter survival ($p = 0.0022$)

Table 1

Pons measurements

Pons Measurement ^a	Median	Range
1D (mm)		
First scan	48.1	37.1–61.9
All scans	46.4 ^b	32.5–67.5
2D (mm ²)		
First scan	1642.4	883.1–2916.6
All scans	1585.3 ^b	814.7–3060.5

^a Averaged across all four readers

^b Decrease in median is likely reflective of 1) relatively fewer follow-up scans for patients with larger tumors versus a larger number of follow-up scans for patients with smaller, more stable tumors or 2) treatment-related effects

Table 2

Inter-reader variation

Measurement	Median % difference between reader pairs	95th Percentile % difference between reader pairs (%)	% of reader pairs that disagreed by more than 25 %
1D (mm)	3.1	13.0	0.7
2D (mm ²)	5.4	19.3	2.6

Table 3

Relationship of pons measurements and overall survival

Pons measurement ^a	Relative hazard ^b	<i>p</i> value ^c	Combined <i>p</i> value
1D (mm)			
First exam	2.24	0.0060	
Changes at follow-up ^d (+10 mm)	2.29	0.0025	0.0011
2D (mm ²)			
First exam	1.14	0.0004	
Changes at follow-up ^d (+100 mm ²)	1.13	0.0016	<0.0001

^a Averaged across all four readers^b Adjusted for the effects of the other risk factors in the model^c Likelihood ratio test of the proportional hazards model^d Relative to the first exam

Table 4

Relationship of percent change in pons measurements and overall survival

Pons measurement ^a	Relative hazard ^b	<i>p</i> value ^c	Combined <i>p</i> value
1D (10 mm)			
First exam on study	2.29	0.0045	
C10 % increase at follow-up ^d	2.05	0.030	0.0050
C5 % decrease at follow-up ^d	0.51	0.063	
2D (100 mm ²)			
First exam on study	1.14	0.0004	
25 % increase at follow-up ^d	2.09	0.044	0.0007
10–25 % increase at follow-up ^d	1.37	0.38	
10 % decrease at follow-up ^d	0.56	0.13	

^a Averaged across all four readers^b Adjusted for the effects of the other risk factors in the model^c Likelihood ratio test of the proportional hazards model^d Relative to the first exam