Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2008 April 1; 70(5): 1607–1610. doi:10.1016/j.ijrobp.2007.12.022.

INCORPORATE IMAGING CHARACTERISTICS INTO AN ARTERIOVENOUS MALFORMATION RADIOSURGERY PLAN EVALUATION MODEL

Pengpeng Zhang, Ph.D.^{*,†,‡}, Leester Wu, M.D.^{*,†,‡}, Tian Liu, Ph.D.^{*}, Gerald J. Kutcher, Ph.D.^{*}, and Steven Isaacson, M.D.^{*}

^{*}Department of Radiation Oncology, Columbia University, New York, NY

[†]Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY

[‡]Department of Radiation Oncology, South Nassau Communities Hospital, Oceanside, NY

Abstract

Purpose—To integrate imaging performance characteristics, specifically sensitivity and specificity, of magnetic resonance angiography (MRA) and digital subtraction angiography (DSA) into arteriovenous malformation (AVM) radiosurgery planning and evaluation.

Methods and Materials—Images of 10 patients with AVMs located in critical brain areas were analyzed in this retrospective planning study. The image findings were first used to estimate the sensitivity and specificity of MRA and DSA. Instead of accepting the imaging observation as a binary (yes or no) mapping of AVM location, our alternative is to translate the image into an AVM probability distribution map by incorporating imagers' sensitivity and specificity, and to use this map as a basis for planning and evaluation. Three sets of radiosurgery plans, targeting the MRA and DSA positive overlap, MRA positive, and DSA positive were optimized for best conformality. The AVM obliteration rate (ORAVM) and brain complication rate served as endpoints for plan comparison.

Results—In our 10-patient study, the specificities and sensitivities of MRA and DSA were estimated to be (0.95, 0.74) and (0.71, 0.95), respectively. The positive overlap of MRA and DSA accounted for $67.8\% \pm 4.9\%$ of the estimated true AVM volume. Compared with plans targeting MRA and DSA-positive overlap, plans targeting MRA-positive or DSA-positive improved ORAVM by $4.1\% \pm 1.9\%$ and $15.7\% \pm 8.3\%$, while also increasing the complication rate by $1.0\% \pm 0.8\%$ and $4.4\% \pm 2.3\%$, respectively.

Conclusions—The impact of imagers' quality should be quantified and incorporated in AVM radiosurgery planning and evaluation to facilitate clinical decision making.

Keywords

Arteriovenous malformation; Radiosurgery; ROC; Treatment planning

© 2008 Elsevier Inc.

Reprint requests to: Pengpeng Zhang, Ph.D., Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, 1275 York Ave., New York, NY 10021. Tel: (212) 639-8768; Fax: (212) 717-3010; zhangp@mskcc.org.

Presented at the 49th Annual Meeting of the American Society for Therapeutic Radiology and Oncology (ASTRO), October 28–November 1, 2007, Los Angeles, CA.

Conflict of interest: none.

INTRODUCTION

Arteriovenous malformation (AVM) is clinically identified with both magnetic resonance angiography (MRA) and orthogonal digital subtraction angiography (DSA) for radiosurgery planning (1). However, there are often target discrepancies between the three-dimensional MRA images and the two-dimensional DSA images. First, the ambiguity comes from the performance uncertainties of the imaging modality, such as the amount of contrast agent injected, the timing of sampling the image, and the windowing/leveling of the image. Second, the ambiguity comes from the observers (i.e., neurosurgeons and radiation oncologists). Different training background and clinical experience may result in different delineations of the presence of blood vessels, bone, or embolization materials, which could amplify the imaging performance uncertainty. Indeed, substantial interobserver discrepancies can exist, as reported by Buis *et al.* (2,3). These existing uncertainties in the imaging and image interpretation process may cause underdosage of the AVM and thus, treatment failure. The purpose of this project was to develop a theoretic framework that integrates performance characteristics of the imaging/image interpretation process into AVM radiosurgery planning and evaluation, and thus improve the quality of the treatment.

The performance characteristics of medical image modalities or imagers can be quantitatively modeled by means of receiver operating characteristic curves (4), but this approach has not been incorporated into the AVM treatment-planning process. As Zhang *et al.* (5) reported in their article regarding ultrasound tissue typing-guided prostate dose escalation, instead of accepting the observation from imaging as a binary (yes or no) mapping of target location, an alternative is to translate the image into a target probability distribution map by incorporating the sensitivity and specificity of the imaging devices and use such a map as a basis for plan optimization and evaluation. In this article we extend this method by investigating how two independent imaging modalities, MRA and DSA, affect the design and evaluation of AVM radiosurgery plans.

METHODS AND MATERIALS

Patient selection

Ten patients who underwent AVM radiosurgery at Columbia University Medical Center were chosen in this retrospective study. The volume of the AVM target ranged from 1.2 to 7.6 cm³ with a median of 3.6 cm³. Prescription doses ranged from 18 to 22 Gy with a median of 20 Gy. The AVMs were all located in or near critical brain areas, such as the occipital and temporal lobes, where limiting the volume of normal tissue radiation is desired. All patients underwent MRA and DSA before their radiosurgery. Treatment planning and analysis were performed with GammaPlan 5.34 software (Elekta, Stockholm, Sweden).

Theory of probability-based AVM mapping

An imager classifies a region of interest into two categories: one is the detected positive zone with a volume of v_t , and the other is the detected negative zone with a volume of v_n . Specificity and sensitivity are generally used to mark the characteristics of an imager. Specificity is the true negative fraction (TNF) and sensitivity is the true positive fraction (TPF) of the imager. In a population base, these two volumes are linked to the true positive volume V_t and the true negative volume V_n :

$$\begin{cases} v_t = V_t \times TPF + V_n \times (1 - TNF) \\ v_n = V_t \times (1 - TPF) + V_n \times TNF \end{cases} \quad (1)$$

Because v_t and v_n are revealed by an imager, but V_t and V_n are not unless a whole-mount study is performed, V_t and V_n can be derived by:

$$\begin{cases} V_t = \frac{TNF \times v_t - (1 - TNF) \times v_n}{TPF - (1 - TNF)} \\ V_n = v_t + v_n - V_t \end{cases} \quad (2)$$

Equation 2 is used to estimate the true AVM volume delineated by MRA and DSA. However, the two volumes obtained, V_{tMRA} and V_{tDSA} , may differ. If we assume that the two imaging modalities are equally validated against the gold standard (i.e., the sensitivity and specificity marked are equally accurate), the average of the two can be used to estimate the true AVM volume:

$$\overline{V_t} = (V_{tDSA} + V_{tMRA}) / 2 \quad (3)$$

The reasons that the estimated true AVM volumes according to different imagers differ are (1) individual patient variation, and (2) uncertainties produced when the imaging technique is evaluated against the gold standard. The sensitivity and specificity of an imager are obtained with a patient-population database. Although it is fitted for the entire patient population, for an individual patient, the sensitivity may vary at a given specificity. Furthermore, if this imager is applied to a different patient group other than the validation study group, the specificity marked may also vary. Equation 4 models this phenomenon in a broader sense:

$$\begin{cases} v_t = \overline{V_t} \times (TPF + \delta) + \overline{V_n} \times (1 - (TNF + \xi)) \\ v_n = \overline{V_t} \times (1 - (TPF + \delta)) + \overline{V_n} \times (TNF + \xi) \end{cases} \quad (4)$$

where δ accounts for the sensitivity variation for an individual patient, and ξ accounts for the specificity variation due to uncertainty produced when the imaging technique is evaluated against the gold standard. In this article we assumed the latter to be negligible (i.e., $\xi = 0$). Therefore, the individual patient variation can be calculated as follows:

$$\delta = \frac{v_t - (1 - TNF) \times (v_t + v_n - \overline{V_t}) - TPF \times \overline{V_t}}{\overline{V_t}} \quad (5)$$

The next step is to obtain an AVM probability distribution map according to the image findings and imagers' sensitivity and specificity. For each voxel inside the region of interest that is identified as AVM by an imager, the probability that it is truly AVM, P_{tt} , is:

$$P_{tt} = \overline{V_t} \times (TPF + \delta) / v_t \quad (6)$$

If a voxel is not identified as AVM, the probability that it is misidentified (should be true AVM), P_{nt} , is:

$$P_{nt} = \overline{V_t} \times (1 - TPF - \delta) / v_n \quad (7)$$

As shown in Figure 1, there are four regions defined by the MRA and DSA. Zone 1 is the area where the MRA- and DSA-positive overlap; Zone 2 represents MRA-positive but DSA-negative; Zone 3 is MRA-negative but DSA-positive; and Zone 4 is the MRA- and DSA-

negative overlap. If we assume MRA and DSA are independent when used to identify AVMs, the probability of an AVM being located in the overlap and mismatch regions should be adjusted according to the individual imaging findings. In the positive overlap region Zone 1, the true AVM probability is:

$$P_{it}O = P_{it}MRA + P_{it}DSA - P_{it}MRA \times P_{it}DSA \quad (8)$$

Similarly, in the negative overlap region Zone 4, the true AVM probability is:

$$P_{nt}O = P_{nt}MRA \times P_{nt}DSA \quad (9)$$

Given the AVM probability obtained for the overlap regions, we can update the true AVM volume in the regions in which the two imaging modalities do not agree with each other (i.e., Zone 2 and Zone 3). In this adjustment, the AVM volume in Zone 2 is the difference between the true AVM identified in the entire MRA-positive region and the AVM in the overlap region. Therefore, the probability of being true AVM in Zone 2 is calculated by:

$$P_tMRA = (\bar{V}_t \times (TPF_{MRA} + \delta_{MRA}) - V_o \times P_{it}O) / (V_tMRA - V_o) \quad (10)$$

where V_o is the imaging positive overlap volume. Similarly, for Zone 3, the probability of being true AVM is:

$$P_tDSA = (\bar{V}_t \times (TPF_{DSA} + \delta_{DSA}) - V_o \times P_{it}O) / (V_tDSA - V_o) \quad (11)$$

If the two imaging modalities improve toward perfection, $P_{it}DSA$ and $P_{it}MRA$ approach 1, and $P_{nt}DSA$, $P_{nt}MRA$, P_tDSA , and P_tMRA all approach 0. However, in reality, because of imaging uncertainty the AVM treatment planning and evaluation must be carried out according to the AVM distribution probability in the four regions.

Obtain imagers' sensitivity and specificity

Gauvrit *et al.* (6) reported that sensitivity reached 0.81 whereas specificity was 1 in their study using MRA to identify AVM. For the case of DSA, our experience suggests that whereas the sensitivity is very high, the specificity could be low owing to its two-dimensional nature. In our study we used two specificity and sensitivity pairs, (1, 0.81) and (0.7, 0.95), as the initial estimations of the MRA and DSA's specificity and sensitivity. Then the best-fitting (TNF TPF_{MRA} and (TNF TPF_{DSA}) for a group of N_p (10) patients were obtained by minimizing the least-square error of estimating the true AVM volume ($\bar{V}_t$) from both MRA (V_{tMRA}) and DSA (V_{tDSA}):

$$EI = \sum_{i=1}^{N_p} \left((\bar{V}_t i - V_{tDSA} i)^2 + (\bar{V}_t i - V_{tMRA} i)^2 \right) \quad (12)$$

AVM radiosurgery planning and evaluation

Three sets of radiosurgery plans, targeting the MRA- and DSA-positive overlap (TP₁), MRA-positive only (TP₂), and DSA-positive only (TP₃), were optimized for best conformity to the target with the GammaPlan system. The AVM obliteration rate (ORAVM) was calculated with the dose-volume histogram (DVH) and AVM probability distribution. Flickinger *et al.* (7)

estimated the AVM in-field obliteration rate by linking the prescription dose to the outcome assessment. To utilize the DVH information, we extended Flickinger's method as follows to estimate in-field AVM obliteration rate, OR_{AVM}^{in} :

$$OR_{AVM}^{in} = 1 - \sum_{i=1}^M (V_i P_i / V_{in} P_{in}) \exp(\ln K - \alpha V_i P_i D_i - \beta (V_i P_i D_i)^2) \quad (13)$$

where V_i , P_i , and D_i are the volume, probability of being true AVM, and dose of the i^{th} DVH bin, respectively, V_{in} is the target volume, P_{in} is the probability of being true AVM inside the target, and α , β , and $\ln K$ are parameters associated with the radiation biologic effect model.

As indicated by Flickinger *et al.* (7), some AVMs are often missed in the target. The out-of-field AVM obliteration rate, OR_{AVM}^{out} is similarly modeled as:

$$OR_{AVM}^{out} = 1 - \sum_{i=1}^M (V_i P_i / V_{out} P_{out}) \exp(\ln K - \alpha V_i P_i D_i - \beta (V_i P_i D_i)^2) \quad (14)$$

where the DVH used is generated for volume outside the primary AVM target. Therefore, the overall AVM obliteration rate is obtained as:

$$OR_{AVM} = \frac{V_{out} P_{out} OR_{AVM}^{out} + V_{in} P_{in} OR_{AVM}^{in}}{V_{out} P_{out} + V_{in} P_{in}} \quad (15)$$

The volume of brain tissue receiving more than 12 Gy (V_{12}) can be used to predict late treatment toxicity. Flickinger *et al.* (8) estimated the AVM complication rate as a function of V_{12} and plotted curves for different brain regions, such as the temporal and occipital lobes. In this article we derived potential complication rates from these curves on the basis of the obtained V_{12} for plan comparison.

RESULTS

In our 10-patient study, the positive overlap of MRA and DSA accounted for $90.5\% \pm 8.1\%$ of the MRA-positive-only volume and $61.9\% \pm 7.8\%$ of the DSA-positive-only volume. The specificities and sensitivities of MRA and DSA were estimated to be (0.95, 0.74) and (0.71, 0.95), respectively. The standard deviation of interpatient imager's sensitivity variation (δ) is 0.044 and 0.032 for MRA and DSA, respectively. The ratio of the MRA- and DSA-positive overlap volume over the estimated true AVM volume is 0.678 ± 0.049 . The probability of being true AVM was estimated to be 0.994 ± 0.005 in the MRA- and DSA-positive overlap Zone 1, 0.560 ± 0.261 in the MRA-positive and DSA-negative Zone 2, 0.642 ± 0.125 in the MRA-negative and DSA-positive Zone 3, and 0.083 ± 0.071 in the MRA- and DSA-negative overlap Zone 4.

The in-field, out-of-field, and overall AVM obliteration rates for plans targeting the MRA- and DSA-positive overlap region (TP_1) are $93.6\% \pm 5.9\%$, $65.2\% \pm 16.9\%$, and $80.7\% \pm 10.1\%$, respectively. Meanwhile V_{12} is $12.1 \pm 5 \text{ cm}^3$, and the corresponding complication rate is $8.7\% \pm 2.6\%$. Compared with TP_1 , plans targeting the MRA-positive (TP_2) and the DSA-positive (TP_3) improved overall obliteration rate OR_{AVM} by $4.1\% \pm 1.9\%$ and $15.7\% \pm 8.3\%$, respectively, while also increasing V_{12} by $1.88 \pm 1.36 \text{ cm}^3$ and $6.91 \pm 2.14 \text{ cm}^3$, and

complication rate by $1.0\% \pm 0.8\%$ and $4.4\% \pm 2.3\%$, respectively. Statistical analysis indicates that all differences are significant ($p < 0.05$).

DISCUSSION

In this article we demonstrated the feasibility of incorporating the imaging/imager's performance characteristics into AVM treatment planning and evaluation. We investigated the impact of using both MRA and DSA as guidance for AVM radiosurgery. As indicated by the result, the probability of being true AVM in the DSA- and MRA-positive overlap region is very close to 1; therefore, the positive overlap is the primary target. However, the positive overlap only accounts for 67.8% of the entire AVM volume. Because of the uncertainty of the imaging/image interpolation process, a significant amount of AVM is missed in the primary target. As a result, the out-of-field obliteration rate is only approximately 65%. Therefore, we emphasize that before the radiosurgery team does treatment planning, it is crucial to carry out studies to investigate how accurate the target delineation is, especially when the AVM volume is large (i.e., $>3 \text{ cm}^3$), and the obliteration rate could become compromised owing to the relatively larger true AVM volume missed. As the AVM volume gets even larger ($>10 \text{ cm}^3$), a staged treatment could be more practical (9), not only because late effects are better minimized with a fractionated regimen but also because more series of images are taken, and the probability of correctly identifying the true AVM increases.

Buis *et al.* (2) suggested that MRA alone could be used as the sole basis for AVM radiosurgery planning when the AVM volume is $<3 \text{ cm}^3$. Our dosimetry study indicates that aiming at the MRA-positive target is superior than aiming at the MRA and DSA overlap because overall AVM obliteration rate increased 4% at an expense of only 1% increased complication rate. The obliteration rate of treatment plans based on the DSA-positive region can reach an average of 96% but is often not a clinical option because too much brain tissue is damaged. Therefore, in our radiosurgery center we used MRA-positive as the primary target initially. However we insist that DSA should not be excluded even though it has its own risk and often makes the radiosurgery procedure longer. First, the interpatient sensitivity variation is smaller for DSA than for MRA, which suggests that DSA is more consistent and robust for target identification. Despite DSA's poorer specificity, DSA is still necessary for the purpose of confining the AVM target because of its high sensitivity. Second, DSA provides guidance for dose planning. Spread of therapeutic dose outside the radiosurgery target is unavoidable owing to the nature of planning (i.e., using ellipsoidal "shots" to cover a target with irregular shape). In AVM radiosurgery planning, the therapeutic dose spread outside the MRA-defined target should fall into the surrounding DSA-positive but MRA-negative region to improve the out-of-field AVM obliteration rate, while avoiding the surrounding normal/critical structures. Third, the estimation of the AVM distribution probability, and thereafter the in-field and out-of-field obliteration rate, will be more accurate and credible with two imaging modalities present because they can cross-check each other. This is especially applicable if the AVM is located in a critical brain area and the radiosurgery team needs as much complication/treatment benefit information as possible to facilitate clinical decision making.

There are four AVM distribution probability zones according to two imaging modalities. If a separate radiosurgery team is involved, interobserver variance will further divide the region of interest into 16 zones; or in the future if another imaging modality, such as three-dimensional ultrasound angiography, is applied for AVM detection as suggested by Mathiesen *et al.* (10), the AVM distribution probability will be even more complicated. A voxel and probability-based computer optimization algorithm (11,12) would be a better approach than manual forward planning to achieve the maximum AVM obliteration rate while minimizing the complication rate. To apply such an algorithm to probability distribution-based AVM radiosurgery planning is our future subject of study.

CONCLUSIONS

Imager's performance has important implications in the radiosurgery of AVM. The impact of imager quality should be quantified by a systematic approach and incorporated in AVM radiosurgery planning and evaluation to facilitate clinical decision making.

References

1. Niranjana A, Maitz A, Kondziolka D, et al. Radiosurgery: Current techniques. *Tech Neurosurg* 2003;9:119–127.
2. Buis DR, Lagerwaard FJ, Clemens MF, et al. Delineation of brain AVMs on MR-angiography for the purpose of stereotactic radiosurgery. *Int J Radiat Oncol Biol Phys* 2007;67:308–316. [PubMed: 17189079]
3. Buis DR, Lagerwaard FJ, Barkhof F, et al. Stereotactic radiosurgery for brain AVMs: Role of interobserver variation in target definition on digital subtraction angiography. *Int J Radiat Oncol Biol Phys* 2005;62:246–252. [PubMed: 15850928]
4. Zhou, X.; Obuchowski, N.; McClish, D. Statistical methods in diagnostic medicine. New York: Wiley; 2002.
5. Zhang P, Osterman KS, Liu T, et al. How does performance of ultrasound tissue typing affect design of prostate IMRT dosepainting protocols? *Int J Radiat Oncol Biol Phys* 2007;67:362–368. [PubMed: 17236962]
6. Gauvrit JY, Oppenheim C, Nataf F, et al. Three-dimensional dynamic magnetic resonance angiography for the evaluation of radiosurgically treated cerebral arteriovenous malformations. *Eur Radiol* 2006;16:583–591. [PubMed: 16220210]
7. Flickinger JC, Kondziolka D, Maitz AH, et al. An analysis of the dose-response for arteriovenous malformation radiosurgery and other factors affecting obliteration. *Radiother Oncol* 2006;63:347–354. [PubMed: 12142099]
8. Flickinger JC, Kondziolka D, Lunsford LD, et al. Development of a model to predict permanent symptomatic postradiosurgery injury for arteriovenous malformation patients. Arteriovenous Malformation Radiosurgery Study Group. *Int J Radiat Oncol Biol Phys* 2000;46:1143–1148. [PubMed: 10725624]
9. Sirin S, Kondziolka D, Niranjana A, et al. Prospective staged volume radiosurgery for large arteriovenous malformations: Indications and outcomes in otherwise untreatable patients. *Neurosurgery* 2006;58:17–27. [PubMed: 16385325]
10. Mathiesen T, Peredo I, Edner G, et al. Neuronavigation for arteriovenous malformation surgery by intraoperative three-dimensional ultrasound angiography. *Neurosurgery* 2007;60:345–350. [PubMed: 17415173]
11. Zhang P, Wu J, Dean D, et al. Plug pattern optimization for gamma knife radiosurgery treatment planning. *Int J Radiat Oncol Biol Phys* 2003;55:420–427. [PubMed: 12527055]
12. Zhang P, Dean D, Metzger A, et al. Optimization of Gamma knife treatment planning via guided evolutionary simulated annealing. *Med Phys* 2001;28:1746–1752. [PubMed: 11548945]

Fig. 1.
Illustration of defining arteriovenous malformation target with the guidance of magnetic resonance angiography and digital subtraction angiography.