A methodology for validating a 3D imaging modality for brain AVM delineation: Application to 3DRA.

Marie-Odile Berger, René Anxionnat, Erwan Kerrien, Luc Picard, Michael
Söderman

HAL Id: inria-00321688

https://inria.hal.science/inria-00321688

Submitted on 5 Nov 2008

A methodology for validating a 3D imaging modality for brain AVM delineation: Application to 3DRA.

Marie-Odile Berger, Ph.D.[a], René Anxionnat, M.D., Ph.D.[b,a], Erwan Kerrien, Ph.D.[a,*], Luc Picard, M.D.[b], Michael Söderman, M.D, Ph.D.[c]

[a] Projet Magrit – Bat. C
    LORIA & INRIA Nancy-Grand Est
    615, rue du Jardin Botanique – BP 101
    54602 Villers-lès-Nancy cedex, France
    Fax: +33 3 83 27 83 19 or +33 3 83 41 30 79
    Phone: (Berger) +33 3 54 95 85 01 – (Kerrien) +33 3 83 59 20 67
    E-mail: Marie-Odile.Berger@loria.fr – Erwan.Kerrien@loria.fr

[b] Service de Neuroradiologie
    CHU de Nancy – Hôpital Neurologique
    29, avenue du Maréchal de Lattre de Tassigny
    54000 Nancy, France
    Fax: +33 3 83 85 22 36
    Phone: (Anxionnat) +33 3 83 85 95 27 – (Picard) +33 3 83 85 14 56
    E-mail: r.anxionnat@chu-nancy.fr – l.picard@chu-nancy.fr

[c] Faculty of medicine
    Institution for clinical neurosciences – section of neuroradiology
    Karolinska Institute
    Stockholm, Sweden
    Fax: +46 8 5177 7072
    Phone: +46 8 5177 4037 (pager)
    E-mail: michael.soderman@karolinska.se

Shortened running title: Validating a 3D modality for AVM delineation

* Corresponding author:
    Erwan KERRIEN
    Projet Magrit, Bâtiment C
    LORIA & INRIA Nancy-Grand Est
    615, rue du Jardin Botanique, BP 101
    54602 Villers-lès-Nancy, Cedex, France
    e-mail: Erwan.Kerrien@loria.fr
    Phone: +33 3 83 59 20 67
    Fax: +33 3 83 27 83 19 or +33 3 83 41 30 79

Abstract
A general methodology is described to validate a 3D imaging modality with respect to 2D digital subtracted angiography (DSA) for brain AVMs (BAVM) delineation. It relies on the assessment of the statistical compatibility of the radiosurgical target delineated in 3D with its delineations in 2D. This methodology is demonstrated through a preliminary evaluation of rotational 3D angiography (3DRA). Generally speaking, BAVM delineation cannot be performed on 3DRA alone. However, in our study, 3DRA showed similar performances to DSA for rather easy cases, and even better for 3 patients. Conversely, 3 problematic cases are identified and discussed.

1 Introduction

Brain arteriovenous malformations (BAVMs) are complex pial or subpial arteriovenous shunts which can pose therapeutic problems. The major threat from a BAVM is brain hemorrhage. The natural course is still not very well known, but it seems that the risk of hemorrhage increases with BAVM volume, patient age, central location and prior hemorrhage. Most patients with BAVM are therefore subjected to treatment with surgery, embolization or radiosurgery (RS), unless the risk posed by the remedy itself is considered to be unacceptable.

RS is unique in that the outcome can be predicted from mathematical models [1]. The chances for BAVM obliteration after RS depend on the dose of radiation given to the periphery of the target [2,3]. Adverse radiation effects (ARE), in the worst case radionecrosis, constitute the main potential drawback. The risk for ARE depends on the location of the lesion and total amount of energy delivered, i.e. BAVM volume and radiation dose [2]. There is also a risk of hemorrhage during the latency period before obliteration, a risk that varies according to treatment parameters, AVM volume and patient age [4].

RS may be carried out with proton beams, the Gamma Knife (Elekta, Stockholm, Sweden) or a linear accelerator. In the latter case multileaf collimators are required to irradiate complex-shaped targets, such as BAVMs. The definition of the target volume is an essential stage of the treatment regardless of equipment. The totality of the nidus (which is the active part of the AVM) and as little as possible of the surrounding brain tissue should be included within the prescription isodose, in order to limit the risks of ARE. Indeed, the reasons for failed AVM obliteration are various but partial volume irradiation is put forward as a major cause by many authors [5-9].

Whereas in the case of tumours this volume is obtained from computed tomography (CT) or magnetic resonance (MRI) imaging, digital subtracted angiography (DSA) is still the gold standard for determining the target volume in the case of BAVMs [9-11]. The AVM nidus consists of a complex tangle of abnormal vessels with arteriovenous shunting. Only the detailed analysis of different times of opacification of an angiographic series allows us to precisely determine the contours of this nidus and to differentiate it from the related arteries and the draining veins. The spatial and especially temporal resolutions of DSA are better than those of CT and MRI and thus allow for a more precise target definition.

Using DSA, the target volume is reconstructed from the contours drawn independently on two or more perpendicular or oblique views (see figure 1), in stereotactic conditions. Due to the nature of DSA, where the image is projected onto two dimensions, the three-dimensional (3D) target volume cannot be reconstructed with complete accuracy [12]. This limitation of DSA has long been acknowledged, especially in the case of complexly shaped BAVMs [13]. Methods have been developed that aim at better estimating the nidus shape from two perpendicular angiographic projections [14,15]. These techniques basically assume the target presents an ellipse shape in every axial cut plane. Though they were shown to improve the accuracy of the target volume computation (in the sense of containment), their ability to provide an accurate target shape has not been demonstrated.

The aim of this paper is twofold:
  i.   In section 2.1, we propose a method to validate AVM targets determined from a 3D imaging modality.
  ii. In section 2.2, we describe a clinical pilot study investigating the use of ("3D") rotational angiography (3DRA) to delineate the AVM radiosurgical target.

A preliminary account of these findings has been presented in [16].

2 Material and methods

2.1 Statistical methods

2.1.1 Basic assumptions and outline of the method
The general method for validating a given 3D modality in the framework of RS was designed assuming the following experimental conditions. A set of patients was involved in the experiment. For each patient, a set of DSA views and one 3D image (e.g. MRI, CT or 3DRA) were available. The DSA views were corrected for geometric distorsions and the 3D image was spatially registered with the DSA views.

Two segmentation processes were to be compared:
   i.   A test process, applied on the 3D images. This process could be either manual, with a new delineation protocol, or automatic, thanks to a new image processing algorithm.
   ii. A reference process, applied on the 2D images. In our case, this process was the manual contouring of the AVMs in the DSA views by experts.

Each process generates 3D targets. Consequently, for each patient, a set of 3D targets was to be compared to a set of reference 3D targets by experts.

For each patient, all the test 3D targets and all the expert 3D targets were reprojected onto all DSA views. A set of test 2D curves together with a set of expert 2D curves were thus provided for each patient and each DSA view.

The general strategy of our method consisted in first designing a statistical test for a given test 2D curve to be accepted as valid, given a reference curve distribution. An acceptance rate was then computed for the test curves and compared to an expected acceptance rate estimated from the reference data.

2.1.2 Building a statistical shape model
A statistical shape model was built as introduced by Cootes [17]. This mathematical tool makes it possible to learn what variations in shapes are plausible, based on a set of expert reference curves. One model was built per DSA view, by analyzing shape variations over the training set of M 2D expert curves $[C_i]_{1 \leq i \leq M}$

Cootes' method [17] requires that all curves are resampled on the same number N of points so that they are set into a one-to-one correspondence. This matching step must be done carefully for the shape model to make sense [18]. For our particular clinical application, dense point-to-point correspondences between curves were established using a semi-interactive interface (see figure 2). Segments were drawn by the user and linear interpolation was performed between these cuts. In practice, cuts were chosen so as to delimit areas where the expert curves were in good agreement from areas which presented high variability.

Each curve C can therefore be represented as a $2N$ dimensional vector, concatenating the x and y coordinates of the N sample points. The average curve is defined as the average vector:

$$\bar{C} = \frac{1}{M} \sum_i C_i \tag{1}$$

with the covariance of the curve data:

$$\Sigma = \frac{1}{M} \sum_i \left(C_i - \bar{C}\right)\left(C_i - \bar{C}\right)^t \tag{2}$$

Still following Cootes [17], a linear eigen-model was built by applying principal component analysis (PCA). The eigenvectors of Σ corresponding to the t largest eigenvalues represent the t principal modes of variation. Let P be the $2N \times t$ matrix gathering these eigenvectors. A particular curve C is then approximated by:

$$C' \approx \bar{C} + Pb \tag{3}$$

where $b = P^t(C - \bar{C})$. The squared error of this approximation is:

$$r^2(C) = \|C - C'\|^2 = (C - \bar{C})^t(C - \bar{C}) - b^t b \tag{4}$$

2.1.3 Acceptance test for one curve

Once the model had been built, an acceptance test had to be designed to decide how a test curve matched the model. A first measure of the quality of fit is classically deduced from the PCA:

$$F_m(C) = \sum_{k=1}^{t} \frac{b_k^2}{\lambda_k} \tag{5}$$

where λ_k is the k-th largest eigenvalue in the PCA and b_k the k-th component of b. This criterion compares, for each mode k, the observed variation b_k^2 on curve C to the expected variation λ_k. This value only measures the distance by which the shape deviates from the mean along the modes of the model. It does not consider the distance $r^2(C)$ of the test curve from its linear approximation (see equation (4)). Ideally, this distance should be small, meaning that the linear model given in equation (3) is able to closely model any curve bound to be an acceptable AVM contour.

In order to define what "small" means, an expected residual distance was evaluated for each linear model using leave-one-out experiments: for each expert curve C_i, a reduced PCA model was built from all the expert curves except C_i. The residual of this reduced PCA is denoted r_i^2. The expected residual was given by:

$$r_e^2 = \frac{1}{M} \sum_{i=1}^{M} r_i^2(C_i) \tag{6}$$

Besides, the residual observed on a new curve C in this experiment was:

$$r_o^2 = \frac{1}{M} \sum_{i=1}^{M} r_i^2(C) \tag{7}$$

The compatibility of the new curve C was given by:

$$F_r(C) = \frac{r_o^2}{r_e^2} = \frac{\sum_{i=1}^{M} r_i^2(C)}{\sum_{i=1}^{M} r_i^2(C_i)} \tag{8}$$

The entire measure of fit was then defined as the sum

$$F(C) = F_m(C) + F_r(C) \tag{9}$$

2.1.4 Study of the distribution of F

F theoretically obeys a chi-square law. The actual distribution could be estimated through leave-one-out experiments on the expert curves as follows. For each expert curve C_i, a reduced set of reference curves was considered, gathering all expert curves but C_i. A measure of fit F_i (see equation (9)) was then defined as above for the curve C_i, based on this reduced set. Testing of all expert curves in all DSA views provided an estimate of what the actual distribution of the fit measure was amongst experts.

2.1.5 Global acceptance rate for all test curves

Thanks to the above chi-square test, an accept/reject decision could be taken for each DSA view. Consideration of all the DSA views led to a percentage of acceptance of test curves. Unless this percentage was convincingly high, the question arose of the significance of the computed acceptance rate.

An expected acceptance rate could be estimated through leave-one-out experiments on the expert curves as above. An accept/reject decision was taken for each expert curve C_i

, based on F_i . Consideration of all expert curves provided an estimate of what the "natural" acceptance rate was amongst experts, setting a reference for the acceptance rate of new contours.

2.2 Application to 3DRA
The method formally described above was applied to a preliminary investigation: 3DRA was tested to delineate complex BAVM target shapes.

2.2.1 Clinical material
Eleven patients with BAVM were consecutively treated using embolization and RS at the University Hospital of Nancy (France) between June 2001 and January 2002. According to the complexity of the AVM shape, between 2 to 5 (2D) DSA views were considered for delineation (antero-posterior, lateral and oblique views). Overall, 37 DSA series were available for this study. In addition, one 3DRA volume was available for each patient. These images were acquired in stereotactic conditions, allowing for DSA/3DRA image registration.

2.2.2 Reference data
Three experts in interventional neuroradiology from the University Hospital of Nancy (France) and Karolinska Institute (Stockholm, Sweden) were involved in this study. They manually delineated twice at a one month interval the AVMs in the 37 DSA images. The delineation was done with the software used in clinical practice (Advantage Windows, GE Healthcare). The full DSA series was available so that hemodynamics could be leveraged to analyze the angioarchitecture (feeding arteries, nidus and draining veins) and better assess the actual nidus contours. Each time, the AVM nidus was reconstructed, based on the expert delineation in all the available views. This reference 3D shape was then reprojected on each available view. Thus, 6 contours of reference were available on each DSA image (see figure 3).

2.2.3 Test data
For each patient Maximum Intensity Projection (MIP) images of the 3DRA were generated with the same viewing parameters as used for the acquisition of DSA images (see figure 4). The same three experts were asked to delineate the AVMs in these MIP views, using the same delineation software as for DSA images. In the absence of hemodynamics information in 3DRA, the 3DRA volume was available to display using volume rendering software (Volume Viewer, GE Healthcare), in a view to analyze the angioarchitecture using full 3D information. Again, each AVM nidus was reconstructed using cones intersection and this new 3D shape reprojected onto the available views. These contours were then tested using the statistical model built from the six expert delineations.

2.2.4 AVM classification
Each expert was asked to label each AVM with A or B according to the difficulty encountered when delineating the nidus on DSA series: label A was given when no major difficulty was faced, whereas label B was affixed when the delineation required a very thorough analysis with difficult choices to make. Along with a B labelling, the experts were asked to list their reasons underlying their choice. Figure 4 shows an example of an A-labelled AVM, and a B-labelled AVM can be seen in figure 5.

3 Results

3.1 Test parameters

For each DSA view, a shape model was built based on the 6 reference contours. The number t of modes considered to compute F_m was chosen to represent 95% of the variation in the reference contours.

In order to study the actual distribution for F, the experiment described in section 2.1.4 was performed. The training set enabled us to test 222 expert curves (3×2 expert curves on 37 DSA series). Each measure of fit was associated to a number of modes for the underlying model: 163 models had 3 modes, 53 had 2 modes, and the last 6 models had 4 modes. The distribution of the F_i's based on 3 modes was therefore studied with the hypothesis that either 3 or 4 degrees of freedom should be considered for the actual chi-square distribution (number of modes vs number of summed terms in equation (9)). As shown in figure 6, the latter hypothesis is the one that was eventually retained.

Therefore, $t+1$ degrees of freedom and a p-value of 0.05 were considered in the experiment.

3.2 AVM classification

Each AVM was classified within type A or B, based on the consensus among the 3 experts: 5 AVMs were labelled as A (17 DSA series) and 6 as B (20 DSA series). Most noted difficulties were the lack of clarity of the nidus due to diffuse compartments, neoangiogenesis or previous embolization and the lack of a conspicuous nidus contour because vessels are superimposed.

3.3 Global acceptance rate

Results about the acceptance rates are summarized in table 1.

DSA acceptance rates call for two comments. First, in the case where the fit measure F would faithfully follow a chi-square distribution, a p-value of 0.05 would let us expect a reference acceptance rate of about 95%. It is not the case here, as is also apparent in figure 6: the tail of the observed distribution shows outliers. Second, the better acceptance rate for label A AVMs than for label B AVMs is consistent with the estimation by the experts of an easier, and therefore less variable, delineation of label A AVMs.

3DRA acceptance rates show that global AVM delineation cannot be performed on 3DRA alone (table 1, column for labels A&B) but should be considered as a complement to DSA for rather easy cases (table 1, column for A-labelled AVMs).

3.4 Per patient acceptance rate

A patient by patient analysis (table 2) reveals a more interesting situation. DSA acceptance rates demonstrate that our chi-square test enabled us to correctly label the delineation difficulty for all patients but A3 and B4: A-labelled AVMs correspond to a DSA acceptance rate above 85%.

The situation for 3DRA is somewhat different. 3DRA acceptance rates are similar to the reference DSA rate for 5/11 patients (A2, A3, A4, B2 and B3). The 3DRA rate is even higher than the DSA rate for 3 patients (A4, B2 and B3). In contrast, this rate is drastically low for 3 patients (B1, B5 and B6). These cases will be discussed in the next section.

4 Discussion

4.1 DSA as a reference

DSA still remains the benchmark imaging modality to delineate BAVM for RS. Besides having unequalled spatial resolution, DSA is the only modality that provides sufficient temporal resolution to study the hemodynamics. This point is key in analyzing BAVM: the nidus must be delineated at an early phase of the arterial time, when the nidus is completely opacified, just at the beginning of the venous drainage. Capturing this instant in time requires an image acquisition rate of at least 2 images per second. If taken too early, the DSA image might not display the whole nidus. If taken too late, superimposed venous drainage would degrade the nidus visibility.

However, it has long been acknowledged that DSA has severe limitations. Many views are required to delineate a 3D target but it is often difficult to find enough views where the nidus is clearly visible. Superimposed images of adjacent vessels, uncertain contours due to neoangiogenesis or previous embolization, or inhomogeneous opacification are common. As a consequence, the epipolar consistency of the contours between views is a condition that is seldom fulfilled and is not easy to control [19], in particular when oblique views are added (see figure 7).

Further, even when many views are available, the methods for reconstructing AVM from contours drawn in DSA are unable to deal with all types of concavities in the shape [9] (see figure 1). More generally speaking, an infinite number of 3D shapes fit a given set of any number of projections. For example, in figure 1 (right), both the bold line convex shape and the concave shape have the exact same reprojection contours in all 5 views.

4.2 Interest of 3D

3D imaging thus seems essential for the accurate determination of BAVM target shape. CT and/or MR alone do not seem to be reliable enough to differentiate the AVM nidus from the other adjacent vascular structures even when injected volume acquisitions are used [9,20]. Nevertheless these 3D imaging modalities are used by many teams to refine the initial approximate volume reconstructed from DSA [5,9,21].

Three-dimensional rotational angiography (3DRA) is a novel imaging technique [22] offering a selective vascular 3D analysis with a high spatial resolution (0.25 mm). 3DRA provides very useful 3D information that has been clinically validated for endovascular treatment of intracranial aneurysms [23]. It can also be used for a better understanding of the nidus shape as shown in a preliminary study [8] but this technique has not yet been validated for use on AVM.

All 3D modalities lack the temporal resolution that DSA provides. As a result, fast flow compartments or diffuse parts of the nidus might not be visible in the 3D image. On the other hand, vessels and nidus never become superimposed in 3D. In the framework of BAVM radiosurgery, validating a 3D modality comes down to validating the target shape that can be determined from the 3D image alone. The clinical outcome must definitely be investigated. However, such a study will take a long time since the incidence of BAVM is very low and the variability of cases is high [24,25]. In addition, only underestimation of target volume can be proven by lack of treatment effect. Overestimation of target volume may actually show up as an apparent increase in treatment efficacy, since the likelihood of obliteration will increase with the higher dose of radiation to the periphery of the BAVM. The concomitant increased risk for undue radiation effects will not be appreciated since the target definition was a priori correct.

A first validation step can be taken by comparing the target delineated in the 3D modality to the target reconstructed from reference standard DSA. Various criteria were proposed to that aim. The volume (expressed in cc) [14,15,20,26], the position of the center of mass [9,20,26-28], or the largest or mean dimension [8,20] of the target will

provide valuable information, but cannot ensure that it will be noticed when reconstructed targets have the wrong shape. More complex criteria were proposed to better take into account the target shape: ratio of the intersecting volume over the encompassing volume [26,28], overlapping contours [9], largest dimensions in several directions [27]. More recently, surface to surface distance computation, associated to local analysis along the target surface was proposed for lung cancer delineation [29].

Such studies failed to provide a definite answer about the actual use of the investigated 3D modality for AVM radiosurgery. Practical and clinical arguments were called upon in the end for recommending it as a complement to DSA. Whereas the above criteria tend to a better analysis of the target shape, the characterization of the correct shapes they provide is not local enough to discriminate wrong shapes. Furthermore, the above studies faced a fundamental difficulty in taking the target reconstructed in 3D from DSA as a reference: though great care was taken in using complex and state-of-the-art reconstruction methods, such a reference should still be acknowledged as possibly false and subject to high intra- and inter-observer variability [30].

4.3 Paper contributions

4.3.1 A new validation methodology

Evaluating a segmentation with respect to expert delineations has long been recognized as a difficult problem [31,32]. Various methods have been proposed for comparing given segmentation with those of a group of experts. The simplest approaches, in 2D, use metrics on curve distances [33]. Other techniques compute the agreement of the curve under consideration with the joint agreement of the experts, provided that a convenient metric on curves is available [31]. More sophisticated methods try to define a statistical model of the shape [17] or a probabilistic framework allowing the definition of the true segmentation and a measure of performance to be associated to each segmentation [32].

In our application, manual delineations suffer from observer bias and large inter- and intra-observer variability (figure 5, left): areas where the consensus is good alternate with areas where the expert delineations are highly variable, preventing us from defining a reliable metric between curves and from using the methods described in [31]. The STAPLE methodology [32] aims at estimating the true segmentation from the expert delineations. This methodology is well suited to region segmentation but needs to incorporate a model describing spatially correlated structures. Such assumptions are clearly not fulfilled in our case: due to large image inhomogeneities in the AVM appearance, AVM detection cannot be formulated as a region segmentation problem. Moreover, due to the large variability of the AVM shape, it is impossible to build a model that provides prior information on the expected shape.

The first contribution of this paper is a methodology for validating a 3D imaging modality for BAVM delineation. The strong points of this methodology are the following.

The argument that the comparison should be made in 2D whenever DSA is considered as a reference. Indeed, 3D targets reconstructed from delineations in DSA are known not to be trustworthy everywhere: only the projections of these 3D shapes back onto the DSA views (i.e. 2D curves) can actually be considered as reference data. As a consequence, making the comparison in 3D is not valid.

The contour curves to serve as reference should be consistent in 3D. It occurred that some medical experts chose to draw a larger contour for the nidus in areas where the image was not clear, thus making use of the fact that a different DSA view could bring a correction to the current view. As a consequence, the original DSA contours were not to be considered as reference data but rather the reprojection of the reconstructed 3D target. Figure 3 demonstrates how this reduced the variability in the delineation.

The design of a statistical acceptance test for a given test curve, based on a PCA of the expert curves. This statistical test relies on the computation of a measure of fit F as the sum of two terms (see equation (9)). The first term F_m controls variations along the modes of the model, that is, along the parts of the nidus contour where the experts disagree, binding the distance from the mean curve to remain within the observed limits. The second term F_r penalizes any variation in ways not seen in the training set, that is, where the experts agree [17]. As a result, this test makes it possible to locally adapt the decision of acceptance/rejection.Theoretically, F should follow a chi-square law with t degrees of freedom. In practice, the global chi-square shape is usually a valid hypothesis but the number of degrees of freedom must be adapted to the training set [17]. Leave-one-out experiments showed that $t+1$ degrees of freedom should be considered for the chi-square distribution modeling F in the case of BAVM contours.

In this paper, the clinical experiment focused on 3DRA but this methodology can be applied to any 3D modality.

4.3.2 Preliminary investigation of 3DRA
The second contribution is clinical. A database of AVM cases has been constructed and the delineation process has been carefully studied. Our experiments showed that the variability between experts can be very large. Physicians should be more aware of the importance of this variability and protocols for delineation should be standardized. Moreover, computed acceptance rates showed that BAVM delineation could not be performed on 3DRA alone. However, this modality is promising and should be considered as a complement to DSA.

The per-patient analysis (see table 2) gave example cases of this complementarity. Cases B2 and B3 were considered as difficult ones, but present with a better acceptance rate using 3DRA than DSA. These were AVMs with a rather large nidus, strongly remodeled by previous embolization but with limited venous drainage. Conversely, 3DRA was unable to faithfully depict the nidus of 3 B-labelled patients (B1, B5 and B6). The corresponding AVMs had in common to present with an important and fast venous drainage that superimposed onto the nidus in DSA views, and was reconstructed in 3DRA volumes. In other words, 3DRA is a better aid to understanding complex 3D shapes, such as remodeled nidii, but lacks the temporal resolution of DSA, which is necessary to distinguish the draining veins from the nidus.

However, this last point highlights one limitation of this study: the views chosen to delineate the nidus in 3DRA were the same as in DSA since our aim was to demonstrate the potential of our methodology. In order to fully demonstrate the capacity of 3DRA to display the nidus with all the required conspicuity, and in particular to distinguish the nidus from the draining veins, the experts should be allowed to take better advantage of 3DRA by choosing carefully adapted views, even, and especially, views that are not allowed by DSA, such as axial (or top) views. We believe that the third spatial dimension of 3DRA can compensate for its lack of temporal resolution. Moreover, the contrast medium injection protocol for 3DRA should be improved for cases with a large draining vein.

4.4 Future works

Several improvements and extensions can be made to our approach. The database must be enriched, especially with label A malformations. It would also be convenient if the construction of the statistical model of the shape could be facilitated. Short-term improvements consist in allowing curves instead of segments for cuts in our software interface used to match contours (see dark segments on figure 2). More robust and automated matching could be used but it either requires to have similar curves [34] or to

define expected properties of the dense curve matching [18]. None of these conditions are met in our case.

Multimodality methods are under investigation in order to improve the current delineation process. One research direction is to use 3DRA to build virtual 2D views. Hence, the physician could potentially use an arbitrary number of views to delineate the AVM without further dose exposure to the patient. The accuracy of the target shape will be improved as the precision of the shape recovered with cone intersections is tightly related to the number of views that can be considered.

Outlining a 3D shape is classically performed by manually drawing contours in 2D cut-planes through the pathology. With AVMs, this is a very difficult task (see figure 8). In this paper, the 3DRA target volume was built by cone intersection from contours drawn in MIP views. Concavities in the target cannot be fully recovered with this method. Among other advantages, 3D modalities, and in particular 3DRA, should make it easier to segment the AVM nidus than when DSA is used. A further step to improve the target delineation is to use quasi-automated algorithms to reduce variability. We are currently working on evaluating active surfaces to that aim.

References

[1] Söderman M, Volume determination and predictive models in the management of cerebral arteriovenous malformations, Medical thesis, Karolinska Insitute; 2000.

[2] Karlsson B, Lax I, Söderman M. Can the probability for obliteration after radiosurgery for arteriovenous malformations be accurately predicted? Int J Radiat Oncol Biol Phys. 1999; 43(2):313 9.

[3] Bollet M, Anxionnat R, Buchheit I, et al. Efficacy and morbidity of arc-therapy radiosurgery for cerebral arteriovenous malformations: a comparison with the natural history. Int J Radiat Oncol Biol Phys. 2004; 58:1353 63.

[4] Karlsson B, Lax I, Söderman M. Risk for hemorrhage during the 2-year latency period following gamma knife radiosurgery for arteriovenous malformations. Int J Radiat Oncol Biol Phys. 2001; 49(4):1045 51.

[5] Flickinger JC, Kondziolka D, Lunsford LD, et al. A multi-institutional analysis of complication outcomes after arteriovenous malformation radiosurgery. Int J Radiat Oncol Biol Phys. 1999; 44:67 74.

[6] Gallina P, Merienne L, Meder J, et al. Failure in radiosurgery treatment of cerebral arteriovenous malformations. Neurosurgery. 1998; 42:996 1004.

[7] Pollock BE, Flickinger JC, Lunsford LD, et al. Factors associated with successful arteriovenous malformations radiosurgery. Neurosurgery 1998; 42:1239 47.

[8] Colombo F, Cavedon C, Francescon P, et al. Three-dimensional angiography for radiosurgical treatment planning for arteriovenous malformations. J Neurosurg. 2003; 98:536 43.

[9] Zhang XQ, Shirato H, Aoyama H, et al. Clinical significance of 3D reconstruction of arteriovenous malformation using digital subtraction angiography and its modification with CT information in stereotactic radiosurgery. Int J Radiat Oncol Biol Phys. 2003; 57(5):1392 9.

[10] Guo W. Application of MR in stereotactic radiosurgery. JMRI 1998; 8:415 20.

[11] 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 91.

[12] Laurentini A. How far 3D shapes can be understood from 2D silhouettes. IEEE Trans Pattern Anal Machine Intell. 1995; 17(2):188 95.

[13] Bova FJ, Friedman WA. Stereotactic angiography: an inadequate database for radiosurgery? Int J Radiat Oncol Biol Phys. 1991; 20:891 5.

[14] Söderman M, Karlsson B, Launay L, et al. Volume measurement of cerebral arteriovenous malformations from angiography. Neuroradiology. 2000; 42:697 702.

[15] Coste E, Gibon D, Leclercq X, et al. 3D reconstruction of the encapsulating contour of arteriovenous malformations for radiosurgery using digital subtraction angiography. Int J Radiat Oncol Biol Phys. 2001; 50(1):247 55.

[16] Berger MO, Anxionnat R, Kerrien E. A methodology for validating a new imaging modality with respect to a gold standard imagery: Example of the use of 3DRA and MRI for AVM delineation. In: Barillot C, Haynor DR, Hellier P editors. Proc. of MICCAI'04 (Saint Malo, France). Lecture Notes in Computer Science No. 3216: Springer Verlag; 2004. p. 516 24.

[17] Cootes T, Page C, Jackson C, et al. Statistical grey-level models for object location and identification. Image and Vision Computing. 1996; 14:533 40.

[18] Davies RH, Twining CJ, Cootes TF, et al. A minimum description length approach to statistical shape modeling. IEEE Trans Med Imag. 2002; 21(5):525 37.

[19] Anxionnat R, Berger MO, Kerrien E, et al., Intra- and inter-observer variability in the angiographic delineation of brain arterio-venous malformations (AVMs), in: Proc. of

the 17th International Congress and Exhibition on Computer Assisted Radiology and Surgery - CARS'2003. (London, UK), Elsevier, 2003, pp. 1297 8.

[20] Aoyama H, Shirato H, Katoh N, et al. Comparison of imaging modalities for the accurate delineation of arteriovenous malformation, with reference to stereotactic radiosurgery. Int J Radiat Oncol Biol Phys. 2005; 62(4):1232 8.

[21] Bednarz G, Downes B, Werner-Wasik M, et al. Combining stereotactic angiography and 3D time-of-flight magnetic resonance angiography in treatment planning for arteriovenous malformation radiosurgery. Int J Radiat Oncol Biol Phys. 2000; 46:1149 54.

[22] Wiesent K, Barth K, Navab N, et al. Enhanced 3-D reconstruction algorithm for C-arm systems suitable for interventional procedures. IEEE Trans Med Imag. 2000; 19(5):391 403.

[23] Anxionnat R, Bracard S, Ducrocq X, et al. Clinical value of three-dimensional digital subtraction angiography (3D DSA) in the therapeutic decision and endovascular treatment of intracranial aneurysms. Radiology. 2001; 218:799 808.

[24] Ogilvy C, Stieg P, Awad I, et al., Recommendations for the management of arteriovenous malformations: a statement for healthcare professionals from a special writing group of the stroke council, Stroke 32 (2001) 1458 71.

[25] Al-Shahi R, Warlow C, A systematic review of the frequency and prognosis of arteriovenous malformations of the brain in adults, Brain 124 (2001) 1900 26.

[26] 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; 61(1):246 52.

[27] Tanaka H, Numaguchi Y, Konno S, et al. Initial experience with helical CT and 3D reconstruction in therapeutic planning of cerebral AVMs: Comparison with 3D time-of-flight MRA and digital subtraction angiography. JCAT. 1997; 21(5):811 7.

[28] Yamamoto M, Nagata Y, Okajima K, et al. Differences in target outline delineation from CT scans of brain tumours using different methods and different observers. Radiother Oncol. 1999; 50:151 6.

[29] Steenbakkers RJHM, Duppen JC, Fitton I, et al. Reduction of observer variation using matched CT-PET for lung cancer delineation: A three-dimensional analysis. Int J Radiat Oncol Biol Phys. 2006; 64(2):435 48.

[30] Al-Shahi R, Pal N, Lewis SC, et al. Observer agreement in the angiographic assessment of arteriovenous malformations of the brain. Stroke. 2002; 33:1501 8.

[31] Chalana V, Kim Y. A methodology for evaluation of boundary detection algorithms on medical images. IEEE Trans Med Imag. 1997; 16(5):642 52.

[32] Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. IEEE Trans Med Imag. 2004; 23(7):903 21.

[33] Williams G. Comparing the joint agreement of several raters with another rater. Biometric. 1976; 32:619 27.

[34] Tagare HD. Shape-based nonrigid correspondence with application to heart motion analysis. IEEE Trans Med Imag. 1999; 18(7):570 9.
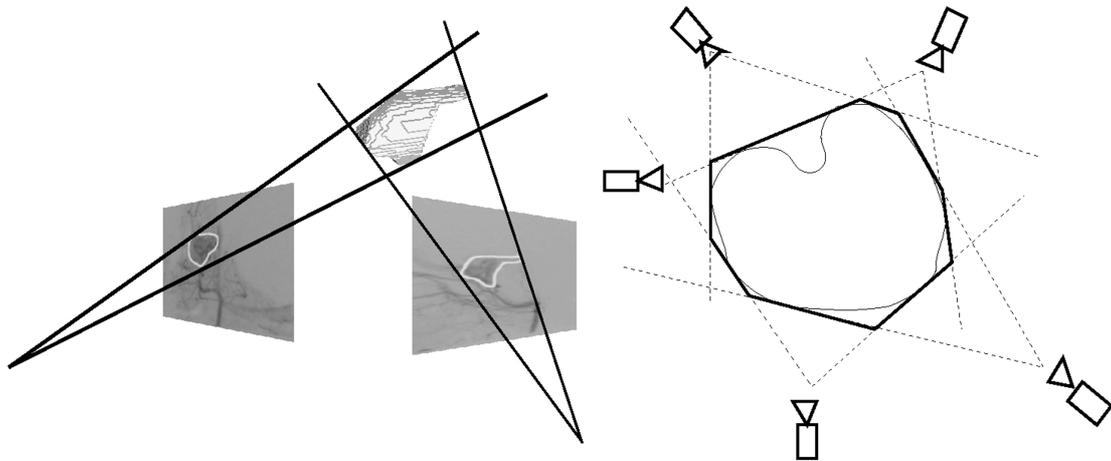
Figures and tables

Fig. 1: AVM reconstruction using conventional angiography. Left: setup with 2 orthogonal views. Right: concavities cannot be recovered even with many views.
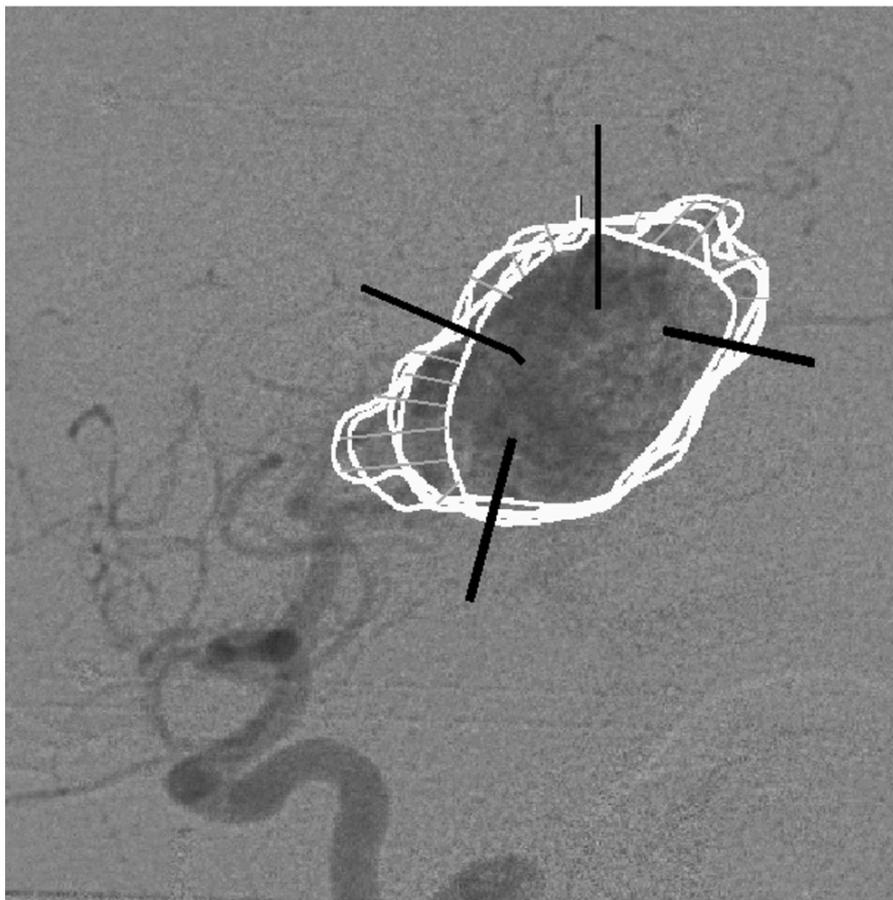

Fig. 2: Semi-interactive interface to match curves: the segments in black were indicated by the user to define homologous sections of curves; the segments in gray show the point-to-point correspondence established between two curves within the set.
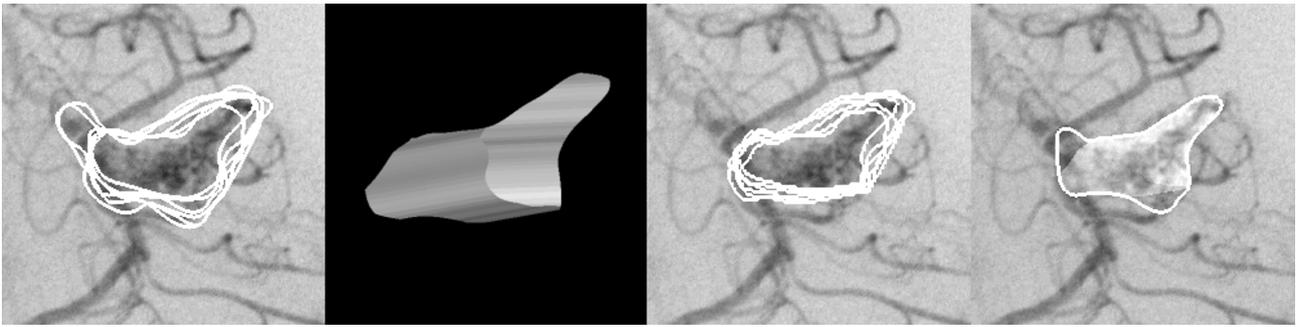
Fig. 3: (From left to right) 1) The set of 6 original expert delineations. 2) reconstructed surface in the same orientation as the DSA view 3) The new database using the reprojection of the 3D shape recovered from expert delineations for the same view. 4) Reprojection of the recovered 3D shape on a reference image in semi-transparent light-gray along with the expert delineation. Data for patient A2 in table 2.
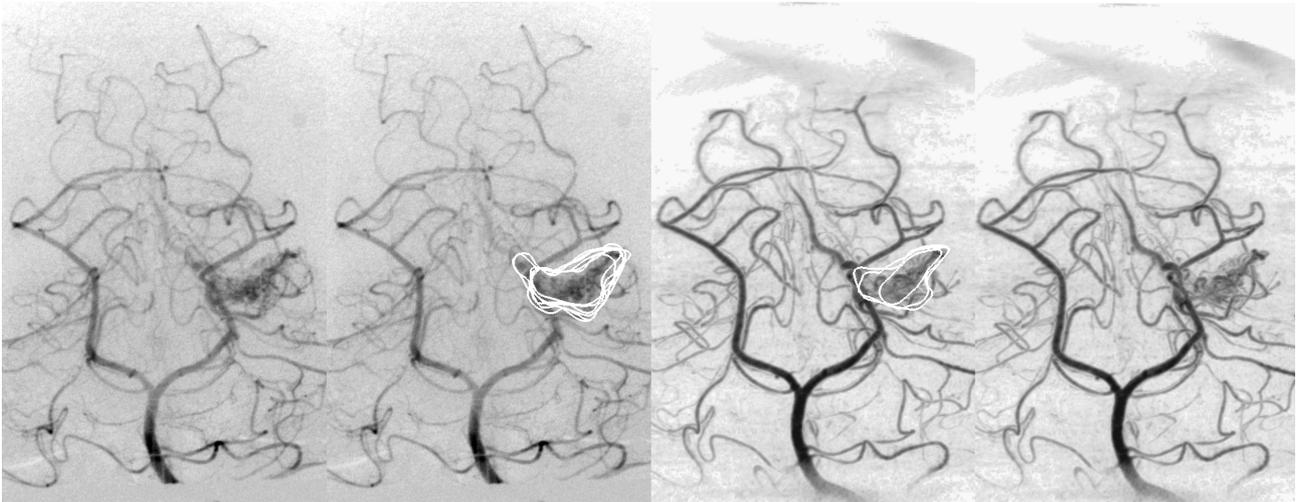
Fig. 4: Example of A-labelled AVM. From left to right in AP orientation: DSA image; DSA image with all original expert delineations; 3DRA MIP image with all original expert delineations; 3DRA MIP. Little vessel superimposition and a limited venous drainage for a nidus with moderate flow speed. A high quality 3DRA reconstruction with no artefact displays an apparently even more conspicuous nidus. Data for patient A2 in table 2.
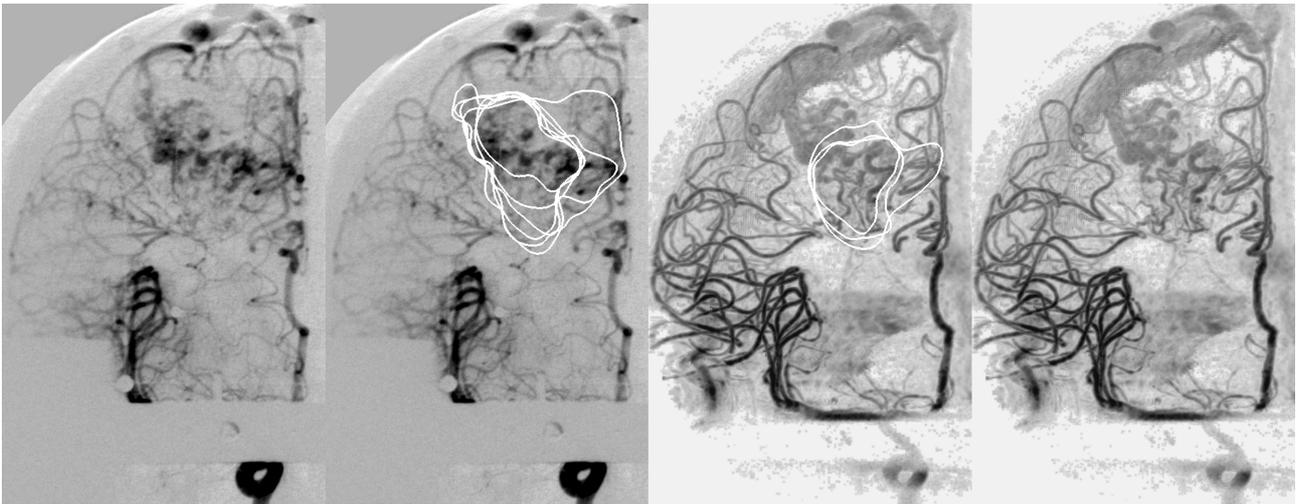
Fig. 5: Example of B-labelled AVM. From left to right in AP orientation: DSA image; DSA image with all original expert delineations; 3DRA MIP image with all original expert delineations; 3DRA MIP. This large nidus appears unclear on most of its contour due to artery superimposition and large previously embolized areas (right and bottom). The venous drainage hampers the analysis of the upper part of the nidus in 3DRA. Experts excluded this part in 3DRA whereas it was always included in DSA. Moreover, the 3D reconstruction of the nidus presents with a poor quality especially in the previously embolized areas. Data for patient B3 in table 2.
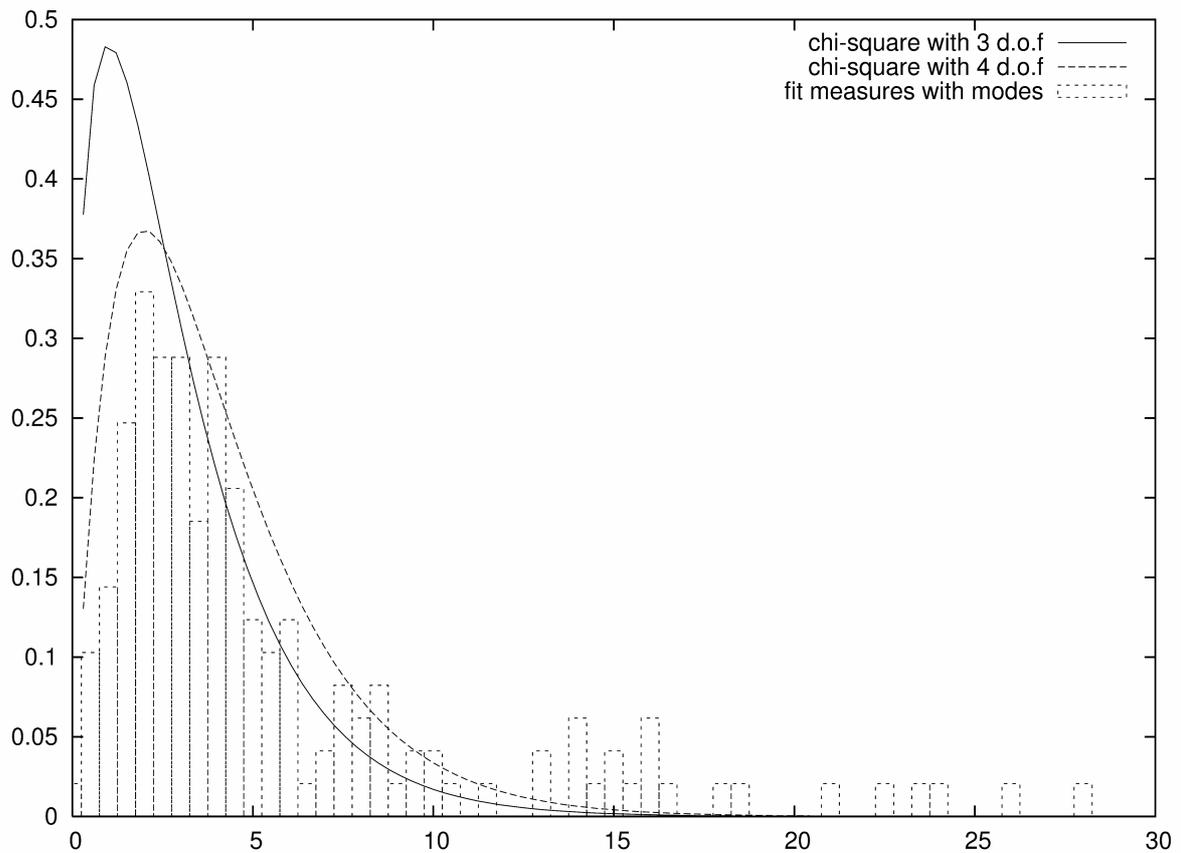

Fig. 6: Distribution of the measures of fit for the expert curves in DSA views corresponding to models with 3 modes: the data histogram fits the chi-square distribution with 4 degrees of freedom.
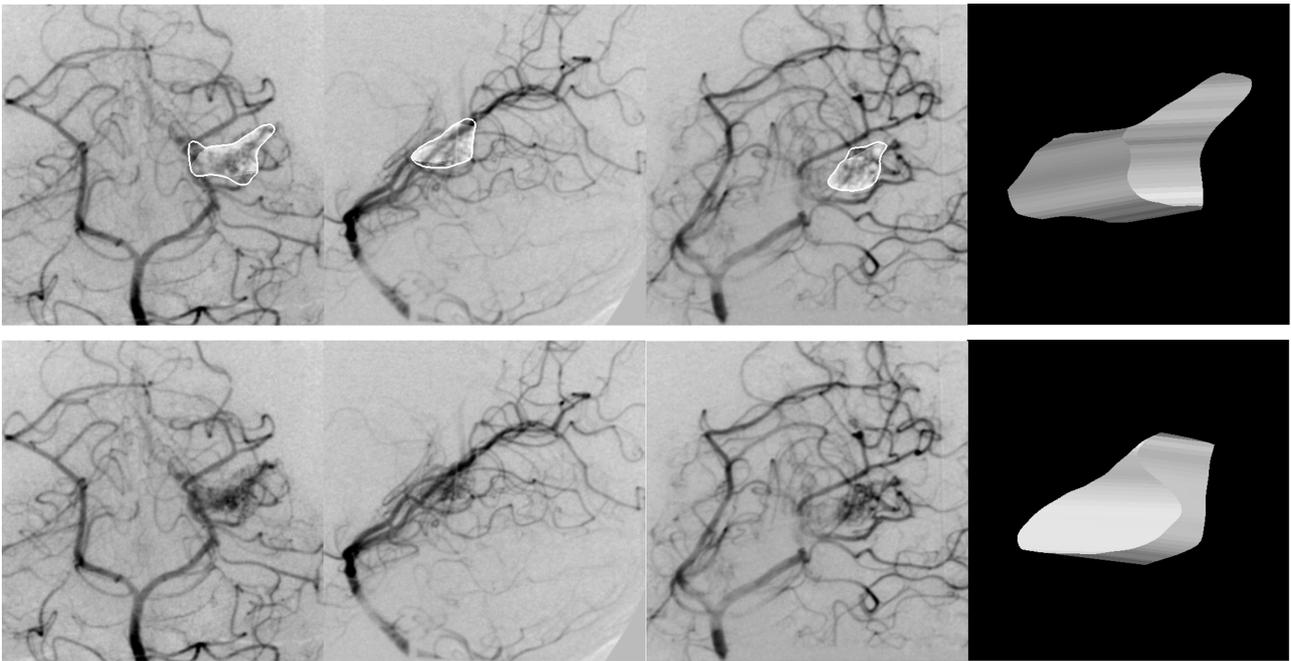
Fig. 7: Contour consistency. Top row: DSA views with one original expert manual delineation (curve in white) and the reconstructed target (semi-transparent light grey) in antero-posterior, lateral and oblique orientations (from left to right). The reconstructed target in AP orientation is shown on the right. Second row: the DSA views with no graphics superimposed for reference. The reconstructed target in lateral orientation is shown on the right. Even though every contour seems correct independently on each DSA view, the large difference in the AP view between the initial contour and the reprojected target demonstrates their inconsistency in 3D. Data for patient A2 in table 2.
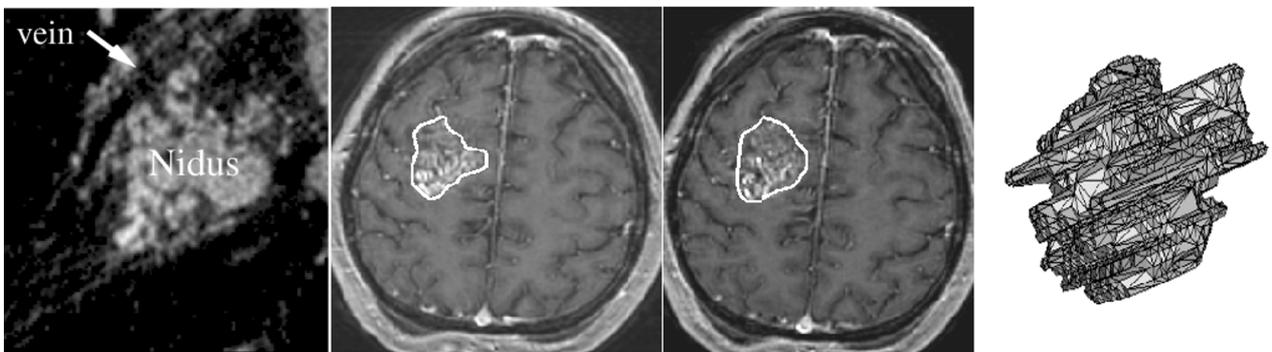

Fig. 8: Difficulties encountered when reconstructing an AVM 3D shape from manual outlines in a stack of axial cut planes. (Left) A 3DRA axial plane through an AVM: the different vascular structures cannot be easily distinguished. (Middle) Two consecutive axial planes in T1 Gd-enhanced MRI together with the manually drawn AVM (Right) Reconstruction of an AVM by stacking up these manual outlines: the smoothness of the 3D shape is not ensured in the longitudinal direction.

Table 1

	global acceptance rate (labels A & B)	acceptance for label A AVMs	acceptance for label B AVMs
3DRA	69%	82%	58%
DSA	84%	87%	81%

Table 1: Acceptance rate of 3DRA contours (second line) compared to the acceptance rate of DSA (third line) computed as explained in section 2.1.5. DSA rates were established on 222 contours and 3DRA rates on 111 contours.

Table 2

Patient	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5	B6
Nb of DSA series	3	3	4	3	4	4	3	4	3	3	3
DSA acceptance rate	89%	94%	79%	89%	88%	83%	72%	79%	89%	83%	83%
3DRA acceptance rate	78%	89%	75%	100%	75%	17%	89%	100%	67%	33%	44%

Table 2: Per patient acceptance rates: DSA rates are computed on 18 or 24 curves (6 curves per DSA series) and 3DRA rates on 9 or 12 curves (3 curves per 3DRA MIP view). The DSA data consisted of a total of 37 series; Patients A1 to A5 are A-labelled and B1 to B6 are B-labelled. Note the AP data for patient A2 are presented in figure 4 and AP data for patient B1 are presented in figure 5.