Data for MBI Workshop

Statistics of Time Warpings and Phase Variations

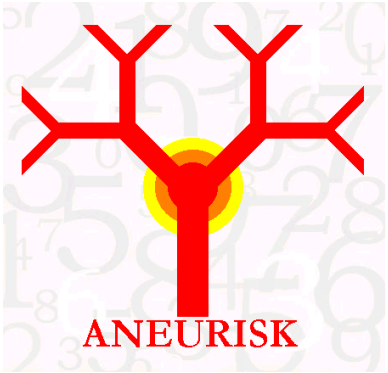
Mathematical Biosciences Institute, November 13-16, 2012

Three-dimensional vascular geometry dataset

May 30, 2012

1 Data background

These data have been collected within the AneuRisk project[1] is a scientific endeavour that aimed at investigating the role of vessel morphology, blood fluid dynamics and biomechanical properties of the vascular wall, on the pathogenesis of cerebral aneurysms. The project has gathered together researchers of different scientific fields, ranging from neurosurgery and neuroradiology to statistics, numerical analysis and bio-engineering.



[1]The project involved MOX Laboratory for Modeling and Scientific Computing (Dip. di Matematica, Politecnico di Milano), Laboratory of Biological Structure Mechanics (Dip. di Ingegneria Strutturale, Politecnico di Milano), Istituto Mario Negri (Ranica), Ospedale Niguarda Ca' Granda (Milano) and Ospedale Maggiore Policlinico (Milano), and has been supported by Fondazione Politecnico di Milano and Siemens Medical Solutions Italia.

More information about the project can be found at the AneuRisk webpage

http://mox.polimi.it/it/progetti/aneurisk/

AneuRisk data can be accessed from the AneuRisk Web Repository

http://ecm2.mathcs.emory.edu/aneurisk

managed by Emory University and Orobix.

Detailed descriptions of the project aims can be found e.g. in [6]. The issue of phase variation is described in [6]; how this issue interplays with data classification is described in [8] and [9]. In this report we give a short description of the problem and data.

Reference. When using these data, please reference the AneuRisk webpage and any among [6], [7] or [8].

1.1 Problem and data

Cerebral aneurysms are deformations of cerebral vessels characterized by a bulge of the vessel wall. This is a common pathology in the adult population, is usually asymptomatic and not disrupting. Epidemiological statistics (see, e.g., [5]) suggest that between 1% and 6% of adults develop a cerebral aneurysm during their lives. The rupture of a cerebral aneurysm, even if quite uncommon (about one event in every 10,000 adults per year), is usually a tragic event. Unfortunately, rupture-preventing therapies, both endovascular and surgical treatment, are not without risks; this adds to the fact that in clinical practice general indications about rupture risk are still missing. Even the origin of the aneurysmal pathology is still unclear. Possible explanations that have been discussed in the medical literature focus on interactions between the biomechanical properties of artery walls and hemodynamic factors, such as wall shear stress and pressure; the hemodynamics is in turn strictly dependent on vascular geometry. In particular, it has been conjectured that the pathogenesis of these deformations is influenced by the morphological shape of cerebral arteries, through the effect that the morphology has on the hemodynamics. For this reason, main goal of the AneuRisk project has been the study of relationships between vessel morphology and aneurysm presence and location.

These lesions may originate along the left or right Internal Carotid Artery (ICA, in short), two large arteries bringing blood to the brain, or at or after terminal bifurcation of the ICA, in the so-called Willis Circle. Each of the two ICAs sits for most of its length outside the skull, along the neck, surrounded by muscle tissues; just before its terminal bifurcation it enters inside the skull, passing through a dural ring (i.e., a hole in the skull bone). See Figures 1 and 2. Arteries downstream of ICA terminal bifurcation float

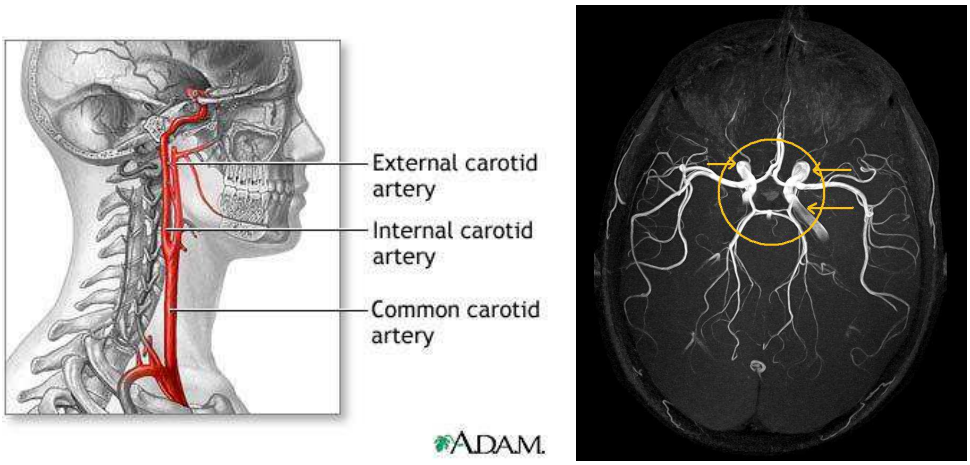


Figure 1: Left: draw of an Internal Carotid Artery (ICA); the ICA sits for most of its length outside the skull, surrounded by the neck muscle tissues; just before its terminal bifurcation it enters inside the skull, passing through a dural ring (i.e., a hole in the skull bone). Right: the Willis circle, located ad the base of the brain, inside the skull, is a net of small arteries and capillaries connecting the main arteries bringing blood to the brain; the terminal parts of the left and right ICAs, clearly visible in the image, are indicated by arrows.

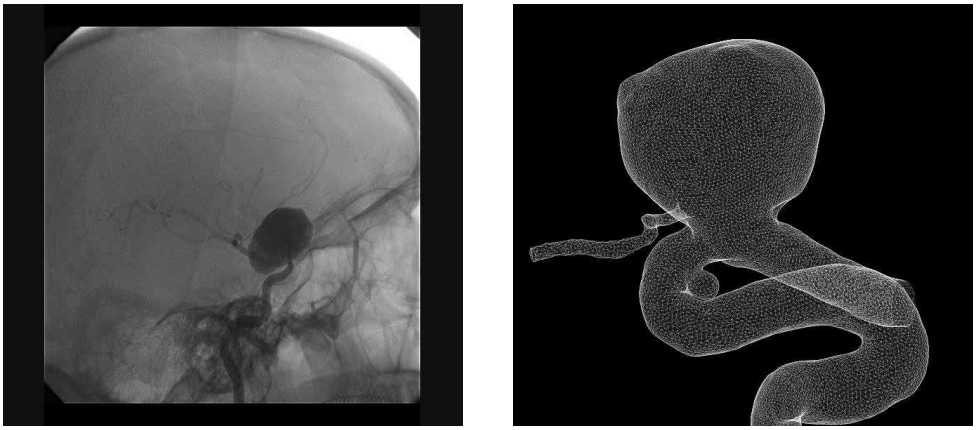


Figure 2: Left: X-rays image of an aneurysm along an ICA; the artery, with its siphon, is clearly visible in the image. Right: image reconstruction of an ICA with aneurysm (different subject with respect to left panel).

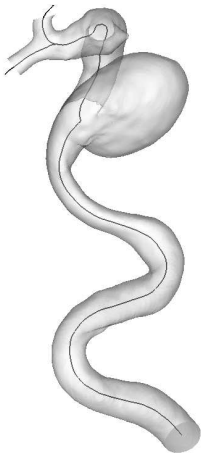


Figure 3: Three-dimensional image of an Internal Carotid Artery with an aneurysm [subject 1]; the black line inside the vessel is its centerline.

in the brain humor, inside the skull. For this reason, aneurysms located at or after ICA terminal bifurcation are more life-threatening; the possible rupture of one such aneurysms is fatal in most cases.

The AneuRisk data set is based on a set of three-dimensional angiographic images taken from 65 subjects, hospitalized at Niguarda Ca' Granda Hospital (Milan), who were suspected of being affected by cerebral aneurysms. Out of these 65 subjects, 33 subjects have an aneurysm at or after the terminal bifurcation of the ICA ("Upper" group), 25 subjects have an aneurysm along the ICA ("Lower" group), and 7 subjects were not found any visible aneurysm during the angiography ("No-aneurysm" group). As commented above, Upper group subjects are those with the most dangerous aneurysms; for this and other clinical reasons, for some statistical analyses it might make sense to join the Lower and No-aneurysm groups in a unique group, to be contrasted to the Upper group. Percentages of females and males and of right and left ICAs do not differ significantly from 50% (the p-values of the test for equal proportions are 0.14 and 0.78, respectively). Age, apart from a superior outlier, appears normally distributed (the p-value of the Shapiro-Wilk test is 0.29), with a sample mean equal to 55.85 years and a sample standard deviation equal to 13.45 years. Gender, ICA side, and age are not included in the statistical analysis because they are supposed to be related to the aneurysmal pathology only through their effect on geometry. Notice that due to the high radiation quantity implied by the scan, this exam is only performed in case of acute symptoms and if the doctor strongly suspect an aneurysm may be the cause; also, no follow-up is performed.

The analyses conducted within the AneuRisk project have focussed on the ICA, which is clearly recognizable in each of the 65 angiographies. Starting from the three-dimensional array of grey-scaled voxels that is gener-

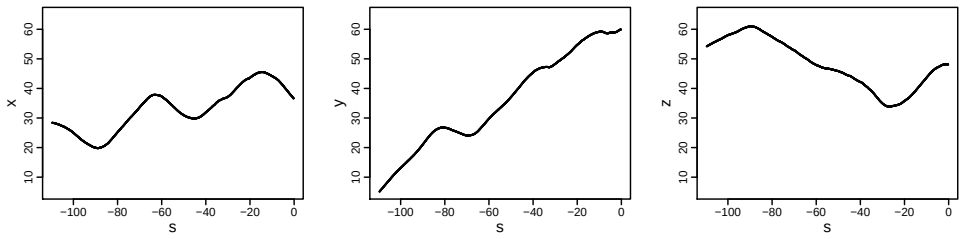


Figure 4: Reconstructed space coordinates of ICA centerline for subject 1, (x_{1j}, y_{1j}, z_{1j}), versus the abscissa parameter s_{1j}, for $j = 1, \ldots, n_1$ ($n_1 = 1350$).

ated by the angiography (with lighter voxels showing the presence of flowing blood), the artery lumen (i.e., the volume that is occupied by flowing blood) is identified by a reconstruction algorithm that is coded in the "Vascular Modeling ToolKit" (VMTK). See [3] and [4]. Figure 3 shows the draw of the reconstruction of the ICA of the first subject in the dataset, and also displays the reconstructed centerline of the vessel. The centerline is computed as the set of centers of maximal spheres inscribed in the artery lumen. In particular, for every subject i in our dataset ($i = 1, \ldots, 65$), VMTK reconstruction of ICA centerline is a set of points in $\mathbb{R}^3$, $\{(x_{ij}, y_{ij}, z_{ij}) : j = 1, 2, \ldots, n_i\}$, where x, y, and z denote respectively the left/right, up/down and front/back coordinates of each point. It should though be noticed that the x, y, and z coordinates are not absolute, but are relative to the cubic volume analyzed during angiography, that in turn depends on where the angiographic image has been centered; this means that these coordinates are not directly comparable across subjects, since they vary with the location of the scanned volume. Points along the centerline are ordered moving downward along the ICA, from the point closest to its terminal bifurcation (detected by VMTK) towards the proximal districts, i.e., aorta and heart. The reason for this choice is that the terminal bifurcation of the ICA is present in each angiography, even if the portion of ICA captured by the angiography varies from subject to subject (depending on where the angiographic image has been centered). For each subject i, we can associate the set of space coordinates with an index set $\{s_{ij} : j = 1, 2, \ldots, n_i\}$, which measures an approximate distance along the ICA centerline, thus providing an approximate curvilinear abscissa. More precisely, $-s_{i1}$ is the distance of the point (x_{i1}, y_{i1}, z_{i1}) from the terminal bifurcation of the ICA (as determined by VMTK), and, for $j = 2, \ldots, n_i$,

$$s_{ij} - s_{ij-1} = -\sqrt{(x_{ij} - x_{ij-1})^2 + (y_{ij} - y_{ij-1})^2 + (z_{ij} - z_{ij-1})^2}\,.$$

The conventional negative sign highlights that we are moving upstream, i.e., in opposite direction with respect to blood flow. Figure 4, for instance,

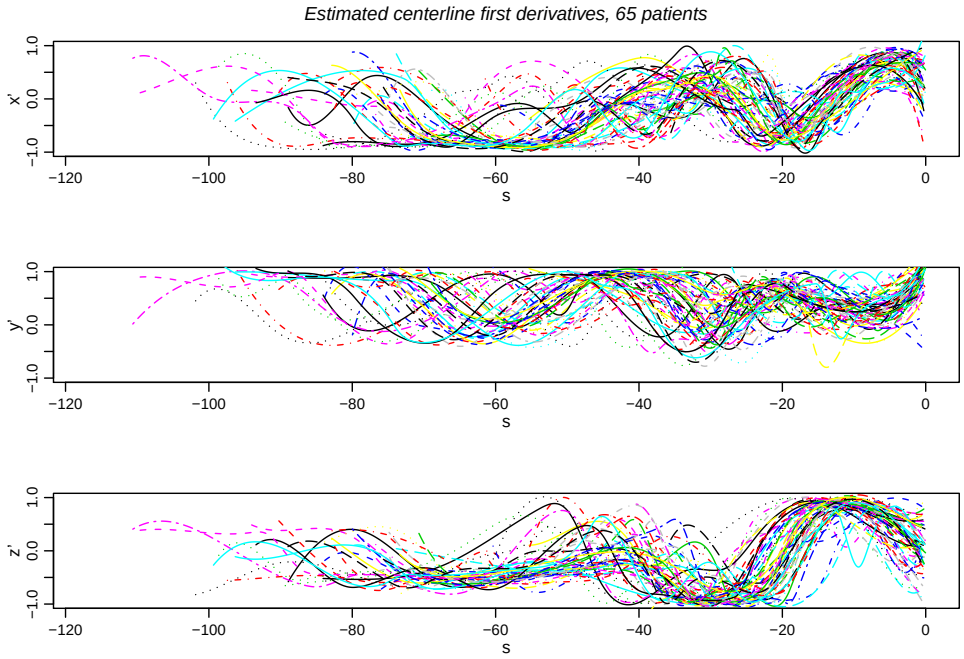


Figure 5: First derivatives $\{x'(s), y'(s), z'(s)\}$ of estimated ICA centerlines for the 65 subjects.

displays the reconstructed space coordinates of ICA centerline for subject 1, (x_{1j}, y_{1j}, z_{1j}), versus the approximate curvilinear abscissa parameter s_{1j}, for $j = 1, \ldots, n_1$ ($n_1 = 1350$). The number n_i of data points available for each subject ranges from 350 to 1380, and is almost perfectly correlated to the approximate length $|s_{i\,n_i} - s_{i\,1}|$ of the reconstructed centerline (correlation coefficient=0.999), which in turn varies from 27.219mm to 110.136mm. In other words, the grid density of the 65 reconstructions is the same, even if the 65 grids are different. The grids are not equispaced; their average step is 0.079mm. The reconstruction algorithm also provides, for each of the grid points, the radius R_{ij} of the vessel lumen section, computed as the radius of the local maximal inscribed sphere.

1.2 Data preprocessing

Reconstructed ICA centerlines are of course affected by measurement and reconstruction errors. The multidimensional free-knot spline technique detailed in [7] might for instance be used to obtain accurate estimates of these three-dimensional curves and their derivatives. This technique is used to obtain the pre-processed data provided for the workshop, including, for each subject $i = 1, \ldots, 65$, the estimated ICA centerlines $\{x_i(s), y_i(s), z_i(s)\}$, their

first two derivatives $\{x_i'(s), y_i'(s), z_i'(s)\}$ and $\{x_i''(s), y_i''(s), z_i''(s)\}$, and the corresponding curvature of the centerline $Curv_i(s)$. The first derivatives of the estimated ICA centerlines of the 65 subjects are displayed in Figure 5; in the picture, left carotids have been left-right reflected (so that the orientation in the three-dimensional space of all reconstructed ICAs is the same). The pre-processed data include also left-right reflected coordinates of the estimates of centerlines and their derivatives.

1.3 Phase variation

Figure 5 shows that the three-dimensional centerlines display a considerable misalignment. This misalignment is the expression of a strong phase variability present among the data, largely due to the different dimensions of the ICA of the various subjects; if not taken properly into account, this misalignment acts as a confounding factor in the data analyses. To enable meaningful comparisons across subjects, it is thus necessary to efficiently decouple the phase and the amplitude variability, the former being mainly due to the differences in the dimensions of subjects carotids and the latter instead to the differences in their morphological shapes. [6], [8] and [9] describes the answers to this problem given within the AneuRisk project, considering also the issue of classification of these three-dimensional curves.

It should be stressed again that the portion of ICA captured by the angiography varies from subject to subject, depending on where the angiographic image has been centered (moreover, even if the terminal bifurcation of the ICA is present in each angiography, the bifurcation point identified by VMTK depends on bifurcation angle and other geometrical quantities). For this reason, it seems unappropriate to use directly on these data a registration method that forces the starting and ending abscissas of the various curves to be the same.

1.4 Some goals of the analyses

Overall goal of the analysis should be the study of relationships between ICA morphology and aneurysm presence and location.

Two geometrical quantities that strongly influence the hemodynamics, and hence may in turn play a role on the aneurysm pathogenesis, are the artery radius and the artery curvature (the latter might be identified by the curvature of the artery centerline). Hence these are good quantities to have a look at. In general, one would of course like to consider the three-dimensional shape of the carotid; in particular, the shape of the siphon characterizing the distal part of the ICA is certainly fundamental in determining the hemodynamics.

Because of the complexity of the data, various levels of analysis are possible. For instance, analytic groups interested only in one-dimensional regi-

stration might focus their attention on one-dimensional curves such as the ICA curvature functions (we never carried out direct registration of AneuRisk data based directly on the curvature profiles). It would be then of interest to see how the registered curvature (or radius) functions relate to the presence and location of the aneurysms, and to the grouping of subjects in the Upper, Lower and No-aneurysm group.

Certainly of big interest is the classification (unsupervised clustering) of ICAs depending on their morphological shape ([8] and Section 5 of [9] report some of the results in this direction, given within the AneuRisk project). Again, depending on the interest of the analytic group, this might perhaps be done by looking only at one-dimensional curves, or instead by working on the full three-dimensional geometries. Studying how the obtained ICA classification relates to the subject groups, Upper, Lower and No-aneurysm, can shed some light on the pathology. When dealing with the problem of classification of ICAs morphologies, it might also be of interest to know that one of the classifications mostly used in the medical literature, proposed by [1], discriminates between Γ-shaped, Ω-shaped, and S-shaped ICAs, according to the form of siphon in their distal part, which may resemble the letters Γ, Ω or S (in presence of zero, one, or two large bends in the siphon, respectively).

It should be mentioned that, besides the data provided for the workshop analysis, the AneuRisk data-warehouse also include the full 3D reconstructions of the ICA walls, and of the connecting arteries, as well as data concerning hemodynamical quantities, such as wall shear stress and pressure, simulated via computational fluid dynamics in the real subject-specific ICA geometries (see [2]).

2 Data files

The file "Patients.txt"contains:

```
patient        patient number, from 1 to 65

code           patient code

type           patient type:
               "U" (Upper group) if at least one aneurysm at or after
                   ICA bifurcation
               "L" (Lower group) if aneurysm before ICA bifurcation
               (and no visible aneurism at or after ICA bifurcation)
               "N" (No-aneurysm group) if no visible aneurysm

AN_Abscissa    location of the aneurysm along the ICA centerline or at
```

	ICA centerline bifurcation (not available if type = N or if aneurysm is after ICA bifurcation)
left_right	left or right carotid: "LC" Left Carotid "RC" Right Caotid

For each subject, from 1 to 65, the file "Rawdata-FKS-*patientnumber*.txt"contains both raw and preprocessed data:

Curv_Abscissa	curvilinear abscissa of ICA centerline
MISR	Maximum Inscribed Sphere Radius of the ICA, i.e., radius of the vessel lumen section (raw data)
X0_obs, Y0_obs, Z0_obs	observed values of the three space coordinates of ICA centerline (raw data)
X0_FKS, Y0_FKS, Z0_FKS	three space coordinates of free-knot-spline estimate of centerline (preprocessed data)
X0_FKS_ref	left-right reflected first coordinate of centerline estimate; for right carotids equals X0_FKS for left carotids equals -X0_FKS (preprocessed data)
X1_FKS, Y1_FKS, Z1_FKS	first derivative of centerline estimate (preprocessed data)
X1_FKS_ref	left-right reflected first coordinate of first derivative of centerline estimate; for right carotids equals X1_FKS for left carotids equals -X1_FKS (preprocessed data)
X2_FKS, Y2_FKS, Z2_FKS	second derivative of centerline estimate (preprocessed data)
X2_FKS_ref	left-right reflected first coordinate of second derivative of centerline estimate;

	for right carotids equals X2_FKS for left carotids equals -X2_FKS (preprocessed data)
Curvature_FKS	curvature of centerline estimate (preprocessed data).

Assistance when analyzing the data

Please, write to laura.sangalli@polimi.it.

References

[1] Krayenbuehl, H., Huber, P., and Yasargil, M. G. (1982), "Krayenbuhl/Yasargil Cerebral Angiography," *Thieme Medical Publishers,* 2nd ed.

[2] T. Passerini, L.M. Sangalli, S. Vantini, Marina Piccinelli, S. Bacigaluppi, L. Antiga, E. Boccardi, P. Secchi, A. Veneziani (2012), "An Integrated CFD-Statistical Investigation of Parent Vasculature of Cerebral Aneurysms," *Cardiovascular Engineering and Technology,* DOI: 10.1007/s13239-011-0079-x.

[3] M. Piccinelli, S. Bacigaluppi, E. Boccardi, B. Ene-Iordache, A. Remuzzi, A. Veneziani, and L. Antiga (2011), "Geometry of the ICA and recurrent patterns in location, orientation and rupture status of lateral aneurysms: an image-based computational study," *Neurosurgery,* 68, 5, 1270–1285.

[4] M. Piccinelli, A. Veneziani, D.A. Steinman, A. Remuzzi, and L. Antiga (2009), "A framework for geometric analysis of 852 vascular structures: applications to cerebral aneurysms," *IEEE Trans. Med. Imaging,* 28, 8, 1141–1155.

[5] Rinkel, G. J., Djibuti, M., Algra, A., and Van Gijn, J. (1998), "Prevalence and Risk of Rupture of Intracranial Aneurysms: A Systematic Review," *Stroke,* 29, 251–256.

[6] Sangalli, L. M., Secchi, P., Vantini, S., and Veneziani, A. (2009a), "A Case Study in Exploratory Functional Data Analysis: Geometrical Features of the Internal Carotid Artery," *J. Amer. Statist. Assoc.*, 104, 37–48.

[7] — (2009b), "Efficient estimation of three-dimensional curves and their derivatives by free-knot regression splines, applied to the analysis of inner carotid artery centrelines," *Journal of the Royal Statistical Society Ser. C, Applied Statistics,* 58, 3, 285–306.

[8] L.M. Sangalli, P. Secchi, S. Vantini and V. Vitelli (2010a), "K-means alignment for curve clustering," *Computational Statistics and Data Analysis*, 54, 1219–1233.

[9] — (2010b), "Classification of Functional Data: Unsupervised Curve Clustering When Curves are Misaligned," *2010 JSM Proceedings,* pp. 4034–4047.