Multiple Aneurysms AnaTomy CHallenge 2018 (MATCH) - Phase I: Segmentation

List of authors:

Philipp Berg[1,4*], Samuel Voß[1,4], Sylvia Saalfeld[2,4], Gábor Janiga[1,4]
[1]Department of Fluid Dynamics and Technical Flows, University of Magdeburg, Magdeburg, Germany
[2]Department of Simulation and Graphics, University of Magdeburg, Magdeburg, Germany
[3]Institute of Neuroradiology, University Hospital Magdeburg, Magdeburg, Germany
[4]Forschungscampus *STIMULATE*, Magdeburg, Germany

*corresponding author: Email: berg@ovgu.de, Tel.: +49 391 67 181 95, ORCID: 0000-0003-4140-6771

Aslak W. Bergersen, Kristian Valen-Sendstad
Department of Computational Physiology, Simula Research Laboratory, Lysaker, Norway

Jan Bruening[1], Leonid Goubergrits[1], Andreas Spuler[2]
[1]Charité – Universitätsmedizin Berlin, Institute for Imaging Science and Computational Modelling in Cardiovascular Medicine, Berlin, Germany
[2]Helios Hospital Berlin Buch, Neurosurgery Department, Berlin, Germany

Nicole M. Cancelliere[1], David A. Steinman[2], Vitor M. Pereira[1,3]
[1]Division of Neuroradiology, Joint Department of Medical Imaging, Toronto Western Hospital, University Health Network, Toronto, ON, Canada
[2]Biomedical Simulation Laboratory, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON, Canada
[3]Division of Neurosurgery, Department of Surgery, Toronto Western Hospital, University Health Network, Toronto, ON, Canada

Tin Lok Chiu[1], Anderson Chun On Tsang[2]
[1]Department of Mechanical Engineering, The University of Hong Kong
[2]Division of Neurosurgery, Department of Surgery, The University of Hong Kong

Bong Jae Chung, Juan R. Cebral
Bioengineering Department, Volgenau School of Engineering, George Mason University, Fairfax, VA, USA

Salvatore Cito, Jordi Pallarès
Universitat Rovira I Virgili, Tarragona, Spain

Gabriele Copelli
Department of Industrial Engineering, University of Parma, Parma, Italy

Benjamin Csippa, György Paál
Department of Hydrodynamic Systems, Budapest University of Technology and Economics, Budapest, Hungary

Soichiro Fujimura[1,2], Hiroyuki Takao[1,2,3]
[1]Graduate School of Mechanical Engineering, Tokyo University of Science, Katsushika-ku, Tokyo, Japan
[2]Department of Innovation for Medical Information Technology, The Jikei University School of Medicine, Minato-ku, Tokyo, Japan
[3]Department of Neurosurgery, The Jikei University School of Medicine, Minato-ku, Tokyo, Japan

Simona Hodis
Department of Mathematics, Texas A&M University Kingsville, USA

Georg Hille
Department of Simulation and Graphics, University of Magdeburg, Germany

Christof Karmonik, Saba Elias
MRI core, Houston Methodist Research Institute, Houston TX, USA

Kerstin Kellermann
Dornheim Medical Images GmbH, Magdeburg, Germany

Muhammad Owais Khan, Alison L. Marsden
Stanford University, CA, USA

Hernán G. Morales[1,2]
[1]Philips Research Paris, France
[2]Centro de Investigación en Fisiología del Ejercicio, Facultad de Ciencias, Universidad Mayor, Chile

Senol Piskin[1,2], Ender A. Finol[1]
[1]Department of Mechanical Engineering, The University of Texas at San Antonio, One UTSA Circle, San Antonio, Texas, 78249, USA
[2]Department of Mechanical Engineering, Koc University, Rumelifeneri Kampusu, Istanbul, 34450, Turkey

Mariya Pravdivtseva
Department of Radiology and Neuroradiology, University Medical Center Schleswig-Holstein UKSH, Kiel, Germany

Hamidreza Rajabzadeh-Oghaz[1,2], Nikhil Paliwal[1,2], Hui Meng[1,2]
[1]Department of Mechanical and Aerospace Engineering, University at Buffalo, State University of New York, Buffalo, NY, USA
[2]Canon Stroke and Vascular Research Center, University at Buffalo, State University of New York, Buffalo, NY, USA

Santhosh Seshadhri[1], Matthew Howard[2]
Medtronic Engineering Innovation Centre, India
Synopsys Inc., CA, USA

Masaaki Shojima, MD
Saitama Medical University General Hospital
1981 Kamoda, Kawagoe-city, Saitama, 350-8655, Japan

Shin-ichiro Sugiyama, Kuniyasu Niizuma
Department of Neurosurgery, Tohoku University Graduate of Medicine, Sendai, Japan

Sergey Sindeev, Sergey Frolov
Department of Biomedical Engineering, Tambov State Technical University, Tambov, Russia

Thomas Wagner[1,2], Alexander Brawanski[1]
[1]University hospital Regensburg, Germany
[2]University of Applied Sciences Regensburg, Germany

Yi Qian
Macquarie University, Sydney, Australia

Yu-An Wu, Kent D. Carlson, Dan Dragomir Daescu
Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, USA

Oliver Beuing
Institute of Neuroradiology, University Hospital Magdeburg, Magdeburg, Germany

Acknowledgements

The authors acknowledge Thomas Hoffmann and Dr. Axel Boese (University of Magdeburg, Germany) for their assistance regarding the challenge design.

Abstract

Purpose: Advanced morphology analysis and image-based hemodynamic simulations are increasingly used to assess the rupture risk of intracranial aneurysms (IAs). However, the accuracy of those results strongly depends on the quality of the vessel wall segmentation.

Methods: To evaluate state-of-the-art segmentation approaches, the Multiple Aneurysms AnaTomy CHallenge (MATCH) was announced. Participants carried out segmentation in three anonymized 3D DSA datasets (left and right anterior, posterior circulation) of a patient harboring five IAs. Qualitative and quantitative inter-group comparisons were carried out with respect to aneurysm volumes and ostia. Further, over- and undersegmentation were evaluated based on highly resolved 2D images. Finally, clinically relevant morphological parameters were calculated.

Results: Based on the contributions of 26 participating groups, the findings reveal that no consensus regarding segmentation software or underlying algorithms exists. Qualitative similarity of the aneurysm representations was obtained. However, inter-group differences occurred regarding the luminal surface quality, number of vessel branches considered, aneurysm volumes (up to 20%) and ostium surface areas (up to 30%). Further, a systematic oversegmentation of the 3D surfaces was observed with a difference of approximately 10% to the highly resolved 2D reference image. Particularly, the neck of the ruptured aneurysm was overrepresented by all groups except for one. Finally, morphology parameters (e.g., undulation and non-sphericity) varied up to 25%.

Conclusions: MATCH provides an overview of segmentation methodologies for IAs and highlights the variability of surface reconstruction. Further, the study emphasizes the need for careful processing of initial segmentation results for a realistic assessment of clinically relevant morphological parameters.

1. Introduction

Image-based hemodynamic simulations are increasingly used to assess the rupture risk [1–4] or support the treatment planning [5–8] of intracranial aneurysms (IAs). To be able to carry out such virtual and risk-free predictions, several interdisciplinary working steps are required. These include mainly the reconstruction of clinical image data [9,10] and the subsequent segmentation [11,12] in order to generate three-dimensional surfaces that serve as a base for the computations. Further, realistic boundary conditions and appropriate simulation settings must be selected before reliable computational fluid dynamics (CFD) simulations can be carried out. Finally, in-vivo and in-vitro validation studies are essential before clinically relevant conclusions can be drawn [13–16].

With improved computational resources, the number of studies containing numerical blood flow simulations in IAs has strongly increased. On the one hand, this provided new insights and profound knowledge regarding cerebral disease [17,18]. However, controversial discussions regarding non-physiologic flow predictions as well as exaggerated parameter interpretation also evolved [19–22].

Hence, to enable a comparison and quantify the variability of image-based blood flow simulations, several international challenges with different foci were announced over the last ten years. The long history of competitions started with the Virtual Intracranial Stenting Challenge (VISC) in 2007. Radaelli et al. [23] demonstrated the ability of computational strategies in consistently quantifying the performance of commercial intracranial stents. Further VISCs followed in the context of the Interdisciplinary Cerebrovascular Symposium (Ankara 2008, Sendai 2009, Houston 2010, Shanghai 2011). Janiga et al. [24] and Cito et al. [25] documented the capability of virtually deploying flow-diverting devices to assist IA treatment planning.

To further address the aforementioned controversies and capture the “real world” variability of hemodynamic simulations, Steinman et al. [26] announced the “ASME 2012 Summer Bioengineering Conference CFD Challenge”. Its intent was the comparison of velocity and pressure results in a giant IA with a proximal stenosis at the parent artery. The study revealed that pressure predictions were consistent, independently of the applied numerical solver. However, for precise calculations of clinically relevant hemodynamic parameters, e.g., wall shear stresses, further investigations were required.

To create a snapshot of rupture risk prediction capabilities using CFD, Janiga et al. organized the “Computational Fluid Dynamics Rupture Challenge 2013”. In the first phase, participants were asked to predict the ruptured aneurysm among two cases and provide the corresponding rupture site based on their simulations [27]. 81% of the 26 participating groups selected the correct IA, but none was able to identify the exact site of rupture. A second phase followed that contained, similarly to Steinman’s challenge, a comparison of velocity and pressure results under given conditions as well as a comparison with in-vitro phantom measurements [28]. Again, good agreement among the groups was observed, confirming the usability of the underlying methodology.

Two years later, the “International Aneurysm CFD Challenge 2015“ was announced by Kono and Valen-Sendstad. Here, five middle cerebral artery aneurysm DICOM datasets were provided and segmentation, blood flow simulation as well as rupture prediction results were requested. The organizers compared rupture prediction between CFD teams and clinicians, and the preliminary results were indicative of CFD providing added value.

Last, the “Flow-Diverter CFD Challenge 2016”, again organized by Kono, focused on the outcome prediction of IA treatment using flow-diverter devices. For three treated internal carotid artery aneurysms, participants were asked to evaluate whether full, partial or no occlusion of the sac occurred. 13 teams submitted their solutions, but only one team correctly predicted the outcome in all cases.

In summary, it can be observed that a tradition of over 10 years of challenges evolved with slightly different central questions, but always including the aim of comparing computational capabilities with respect to their clinical applicability. It was clearly determined that the quality of such numerical results strongly depends on the prescribed boundary conditions, with a primary effect of the chosen vessel wall segmentation. However, this source of variability was not specifically addressed in the frame of such a broad comparison. Hence, to evaluate state-of-the-art segmentation approaches for neurovascular blood flow simulations, the Multiple Aneurysms AnaTomy CHallenge 2018 (MATCH) was announced. Participants were asked to carry out segmentation in three anonymized 3D DSA data sets (left and right anterior, posterior circulation) of a patient harboring five IAs. 26 research groups from 13 countries submitted their segmentation results, details regarding the applied software tools and the required processing time, respectively. In a second phase, image-based hemodynamic simulations were carried out to identify the rupture aneurysm. The results of that phase will be presented in a separate manuscript. Overall, this study provides an overview of recent segmentation approaches, and hence evaluates the influence of segmentation on corresponding morphological parameters that are used for the assessment of IA rupture probability.

2. Methods

2.1 Case details and image acquisition

All five intracranial aneurysms that were the subject of MATCH were found in a 53-year-old woman who had acute subarachnoid hemorrhage. She suffered from severe headaches but had no focal neurological deficits. Two aneurysms were located at the right M1-segment, one on the left M1-segment, another on the left MCA-bifurcation, and the fifth on the left posterior inferior cerebellar artery (PICA). Four of the five aneurysms were of similar size (between 4.4mm and 5.6mm) and all were complexly shaped. Both M1-aneurysms on the right side were clipped, the remaining 3 aneurysms were treated by coiling. Figure 1 illustrates the investigated IAs and visualizes the corresponding locations.

The 3D-RAs were acquired on an Artis Q (Siemens Healthineers, Forchheim, Germany) with a spatial resolution of 0.28 x 0.28 x 0.28 mm. After a native run, a fill run was performed during manual injection of 15-20ml iodinated contrast material (Imeron 300, Bracco Imaging, Konstanz, Germany). Subsequently, the native run was subtracted from the fill run and the resulting data was reconstructed on a syngo Workplace (Siemens Healthineers, Forchheim, Germany) with the “HU auto” kernel [9].

2.2 Participating groups

MATCH was initially announced on November 03 in 2017, and interested research groups were able to receive detailed information from the associated website (https://www.ics2018.de). Participants were asked to submit their segmentation results until January 08 in 2018, while the following points were requested:

- 3D segmentations of the aneurysm datasets using the original coordinate system
- Submission of the segmentation results of the aneurysms and adjacent vasculature in STL format
- Consideration of at least 15 nominal vessel diameters proximal and 10 diameters distal to each aneurysm is required
- Submission of an informal abstract (max. 1 page) containing author names, affiliations and segmentation details ((1) segmentation method, (2) software that was used, and (3) processing time from data import to final segmentation)

Please see the complete announcement in the supplementary files for further details. In total, 26 groups from 13 different countries followed the call and submitted three segmented datasets each. The groups had the following origins: Europe (Germany: 6; France, Hungary, Italy, Norway, Russia, Spain: 1); North America (USA: 7; CAN: 1); Asia (Japan: 3; India, Hong Kong: 1); Australia: 1.

2.3 Segmentation approaches

In the following, segmentation software solutions, underlying algorithms and the corresponding processing times are summarized based on the submitted abstracts.

2.3.1 Segmentation software

The 26 participating groups reported that they used 20 different software packages in total to segment and manipulate the provided datasets. This large variety highlights that no common agreement with respect to segmentation software exists (as opposed to image-based CFD simulations, which are mainly carried out by a handful of well-known solvers). Most workflows were highly individualized, and were based on long experiences regarding clinical image data or on evolved workarounds in order to obtain plausible surface segmentations.

Table 1 summarizes all software packages, their corresponding type and the number of uses among the participants, respectively. It must be noted that groups usually applied a set of different software for different processing steps. Interestingly, more than half (53%) of the tools were open-source solutions, while the Vascular Modeling ToolKit (VMTK [29]) was used seven times (19%), followed by Meshmixer (11%) and MATLAB (8%) as commercial options. Further, only three (14%) in-house solutions were applied, demonstrating that most groups rely on freely or commercially available software packages.

2.3.2 Segmentation algorithm

A clear variability exists with respect to the applied segmentation algorithms. These can be categorized as follows: Most groups (11 groups; 42%) used a *threshold*-based approach for their segmentation. The second most popular selection was the *level set* method, which was selected by 10 groups (39%). The remaining 5 groups either applied a *region growing* (4 groups; 15%) or a *watershed* algorithm (1 group; 4%).

Noteworthy is that most groups only used the algorithms mentioned above to receive initial segmentation results. Further manipulations such as smoothing, cropping or treating undesired artifacts are reported by almost all groups. The extent of those manipulations might be associated with the subsequent section on segmentation times.

2.3.3 Segmentation times

In order to measure the duration required from the initial reading of the corresponding DICOM data to the final segmented surface areas, participants were requested to provide this information. Table 2 contains the processing times (in hours) for both anterior and the posterior circulation. Since groups were aware of the subsequent application of the data (assessing rupture risk probability based on morphology and image-based hemodynamic simulation results), careful post-processing of initial segmentation results was imperative. Hence, although group 13 provided overall the fastest approach, they still required 43 minutes. On the other hand, group 11 invested approximately 26 hours per case, resulting in a total processing time of 78 hours. Overall, a median total processing time of 6 hours was required, underlining the fact that manual correction steps were needed to obtain satisfactory results.

2.3.4 Detailed description of an individual segmentation process

Since group 11 spent by far the longest time on each of the three segmentations and ended up with high quality segmentation results, a detailed description of their working process is presented in the following:

In Mimics, they created successive masks with various CT grayscale threshold values. In each mask, they performed region growing to select and retain only the connected vasculature to be included in the model, and they went slice-by-slice and compared the boundaries of the mask to the underlying CT image, by turning the mask on and off. They also extracted 3D objects from the masks to judge the correctness of the anatomy. The neurosurgeon on their team informed judgements about whether or not to include questionable regions from CT slices in the mask. Once the optimal mask based on the threshold values was selected, they performed slice-by-slide manual editing, which consisted of erasing and adding pixels to the mask to ensure that only the relevant anatomy was included. Again, 3D objects were extracted from the masks to judge the 3D quality of the model at each step of manual editing. The 3D object extracted from the manually edited optimal mask was then smoothed and 3D wrapped to ensure that the model had smooth walls. After these operations, they again compared the results of the boundaries after smoothing and wrapping to the original CT gray scale in each slice, and made further manual adjustments to keep the boundaries confined to the anatomy that was deemed relevant. At every stage of segmentation, special attention was paid to the aneurysm necks, as they considered these regions to be very important.

The resulting 3D models were then remeshed in 3-matic to create a triangular surface mesh. This initial mesh was operated on using smoothing and quality preserving operations to ensure that the surface triangles (which were later used in the CFD software to seed the volume mesh) were as uniform and close to equilateral as possible. The final mesh from 3-matic was brought back into Mimics as a surface (.stl file), and the contours from this surface were again compared slice-by-slice to the original CT grayscale to ensure that the final surface was still representative of the anatomy observed from CT. If necessary, all these operations were repeated until the team was satisfied with the final .stl surface, again paying particular attention to the aneurysm neck region.

2.4 Analysis

The comparison of 78 datasets (26 groups with three segmentations each) was divided into qualitative as well as quantitative analysis, and required the existence of an identical coordinate system. However, some software packages did not necessarily preserve the original coordinate system of the image data. Hence, co-alignment was carried out using the iterative closest point (ICP) algorithm, introduced by Chen and Medioni [30] as well as Besl and McKay [31]. The algorithm minimizes the difference between two point clouds and is often used to reconstruct 2D or 3D surfaces from independent image scans.

2.4.1 Qualitative comparison

In order to qualitatively compare the segmentation results, surfaces were evaluated with respect to their number of considered outlet cross-sections, the individual aneurysm representations as well as the shapes and sizes of representative ostium planes.

2.4.2 Quantitative comparison

To quantify the differences occurring among the submitted 3D segmentation results, measurements of characteristic features were carried out. First, the individual aneurysm volumes were calculated. Here, identical planes were used to truncate the aneurysm from its parent artery.

Second, the variability of ostium surface areas was assessed. It must be noted that these cut-planes do not necessarily represent the real aneurysm ostium, since the selection of an appropriate ostium strongly depends on the segmentation results. However, to ensure an objective comparison, the identical plane was considered for each aneurysm.

Third, beside the inter-group variability, comparisons to a reference solution were conducted. Here, two 2D DSA (Siemens Healthcare GmbH, Forchheim, Germany) datasets (containing IAs C and E) with a spatial resolution of 0.077 mm x 0.077 mm served as a base for a manual segmentation. Hence, the 2D images were 3.5-fold higher resolved compared to the provided 3D DICOM images. The coordinate system of the 3D segmentations was manually registered to the 2D DSA using EnSight 10.1.6 (CEI Inc., Apex, NC, USA). As illustrated in Figure 2, the projected view of each group's 3D segmentation was then exported and compared to the 2D DSA. Afterwards, the over- and undersegmentation were rated pixel by pixel using MATLAB 2016a (The MathWorks, Inc., Natick, Massachusetts, USA). Finally, the 3D segmentation with minimum distance to the 2D reference was determined. Beside the quantitative error, locations prone to segmentation errors are of interest. Therefore, the

over- and undersegmented pixels of all groups are superimposed in order to obtain an error map. Both manual steps, 2D segmentation and registration to 3D, were performed independently by three domain experts. Their corresponding maximum variability between the reference segmentation and registration was 2.1% and 2.0% for aneurysm C and E, respectively.

Fourth, the in- and outflow cross-sections of the global segmentation results were analyzed, taking into account only those areas that were available for all groups.

Last, morphological parameters were computed based on Saalfeld et al. [32] and Niemann et al. [33]. These comprise the maximum height of the aneurysm H_{max}, the maximum width W_{max}, the height of the aneurysm approximated as the length of the ray perpendicular to the ostium plane H_{ortho}, the maximum width parallel to the projected ostium plane W_{ortho} and the maximum diameter of the aneurysm D_{max}. Further, the aspect ratio AR, the volume of the convex hull V_{CH}, the ellipticity index EI, the non-sphericity index NSI and the undulation index UI were analyzed [34,35].

It should be noted that group 3 (left anterior and posterior) was partly rejected from quantitative analysis since the segmented domain size was smaller than initially requested in the MATCH announcement, and thus the co-alignment was partly unsuccessful due to distortion issues.

3. Results

3.1 Qualitative comparison of the segmentation results

The submitted segmentations, which are all based on the identical DICOM datasets, revealed clear visual differences. While some exhibited a very continuous and smooth surface, others contained various spots of increased gradients or even discontinuities. Interestingly, rather strong variations occurred with respect to the number of considered outlets. For the anterior right dataset (containing aneurysms A and B), between three and seven distal branches were considered (3 outlets: 3 groups; 12%; 4 outlets: 7 groups; 27%; 5 outlets: 6 groups; 23%; 6 outlets: 5 groups; 19%; 7 outlets: 5 groups; 19%). Hence, the number of segmented side branches is almost equally distributed. For the case “anterior left” (containing aneurysms C and D) half of the groups considered six outlets (13 groups; 50%). However, variations occurred as well with the following distribution: 4 outlets: 2 groups; 8%; 5 outlets: 5 groups; 19%; 7 outlets: 5 groups; 19%; 8 outlets: 1 group; 4%.

In the posterior circulation, this effect was not as prominent as in the two previous cases. Most groups (23/88%) segmented three distal branches, while only three groups (12%) chose two. Figure 3 illustrates representative groups considering a low and a high number of outlets, respectively.

Beside these variations within the overall segmentation results, clear focus lies on the assessment of differences regarding the aneurysms. Figure 4 contains all surfaces of aneurysm E, which was the one that ruptured in this study. The segmentation results for aneurysms A to D can be observed in the supplementary Figures A1 and A2, respectively.

When comparing the individual panels, one can notice that visual deviations regarding the aneurysm size occur. Particularly, the aneurysms segmented by groups 13 and 17 are clearly larger than those by groups 11 and 22, which directly influences the representation of the aneurysm neck. While some groups suggest a rather broad neck, others found smaller and narrower ostia. Further, there is clear variation regarding the adjacent side

branches. For instance, while almost half of the groups solely segmented aneurysm A (12; 46%), another half considered the existing small proximal vessel (12; 46%). Two groups (19 and 24) even included two small side branches in their segmentation result (recall Figure A1). Also, the thickness of those branches varies considerably, which partly results in the creation of pseudo-stenoses (e.g., Figure 4, group 22).

Another phenomenon that occurred was the inappropriate manipulation of initial segmentation results. This became prominent in particular for aneurysm C (recall Figure A2, group 16), where the aneurysm and the parent vessel melted, drastically increasing the ostium surface area. And even when lower threshold values were chosen initially, direct contact between the aneurysm dome and the main artery can still occur (recall Figure A2, group 03), which needs to be corrected prior to subsequent analysis steps.

In addition to differences in size, differences with respect to the morphology can also be observed. Due to the application of smoothing filters or too coarse resolutions of the surface meshes, contours might not represent existing features such as blebs or sharp edges. However, these are particularly important for the assessment of complex shapes (see Section 3.3 for further details).

A more precise qualitative comparison is possible by comparing 2D instead of 3D segmentation results. Hence, Figure 5 contains the contours of the ostium surface of aneurysm E. Overall, a good agreement with respect to the almost circular shape exists. However, as assumed from the previous observations, variations regarding the size are present. Further, some groups possess an overlapping with the adjacent side branch, indicating an overestimation of the vasculature. The extent of potential over- or underestimation will be addressed in the following section.

3.2 Quantitative comparison

The observed differences with respect to sizes and shapes are quantified in the following based on mean values and standard deviations. First, aneurysm volumes were compared. As introduced in Section 2.1, aneurysm A is the largest, with an average volume provided by the groups of 90.02 ± 15.26 mm^3. Aneurysms C, D and E possess a comparable size with values of 27.15 ± 4.5 mm^3, 18.48 ± 3.46 mm^3 and 28.53 ± 5.32 mm^3, respectively. Aneurysm B is clearly the smallest, with a mean volume of 3.43 ± 2.36 mm^3. When the relative differences (standard deviation over average value) for each aneurysm volume are assessed, it is interesting to see that they are in a similar range. While the values differ by 17% for aneurysms A and C, the overall deviation for aneurysms D and E is 19%. Only aneurysm B shows increased relative differences of 69%, which relates to its actual size and hence the difficulty to objectively find an appropriate cut-plane for comparison. A reason for these differences might also be the application of strong global smoothing algorithms, which clearly affect the shape of the very small aneurysm.

Second, the individual surface area of each ostium was calculated to estimate the corresponding variability. The box-plots in Figure 6 enable an assessment of reported minimum and maximum values (e.g., from 4.98 mm^2 (group 22) to 19.34 mm^2 (group 14) for ostium A). On average, the ostium of aneurysm B exhibits the smallest surface area (4.22 ± 1.91 mm^2). Due to its small neck, aneurysm C possesses the second-smallest ostium (6.2 ± 1.79 mm^2), while aneurysms A, D and E show an almost identical size (9.3 ± 3.21 mm^2; 10.72 ± 1.3 mm^2; 11.23 ± 2.25 mm^2). However, the box-plots indicate how strong the actual variation appears, again with the largest

differences for aneurysms A and E, respectively. The assessment of the relative differences reveals values ranging from 12% to 45%: 12% (aneurysm D), 20% (aneurysm E), 29% (aneurysm C), 35% (aneurysm A), 45% (aneurysm B). Again, the last value should be treated with care due to the actual size and location of the ostium.

The presented inter-group variability further requires a comparison to a reference solution to evaluate the individual accuracies. Here, highly resolved 2D datasets were used (see Section 2.4.2) and the corresponding results are illustrated in Figure 7 for aneurysms C and E.

Overall, it can be observed, that each group rather overestimates the lumen compared to the reference image. On average, the overestimation (underestimation) was 8.79% (1.92%) for aneurysm C and 14.9% (0.72%) for aneurysm E in the analyzed region of interest. In this regard, it is important to point out that the overestimation mainly occurred at the neck of each aneurysm leading to a clear overrepresentation of almost every ostium (recall Figure 4 for qualitative comparison). However, only group 11 edited the neck of aneurysm E manually and hence reconstructed it appropriately. In addition to this finding, oversegmentation was present at the dome of aneurysm C, where strong variations in the shape occurred.

In contrast, 19 of the 26 groups (73%) decided not to segment the small vessel next to aneurysm C, which leads to a clear underestimation of the lumen in this region. Further, the 3D results were smaller compared to the 2D image at sites of aneurysm E where strong changes in shape (e.g., due to blebs) occur. Hence, those important morphological features are underrepresented.

In addition to the quantitative analysis of aneurysm volumes and ostium surface areas, the variability of in- and outflow cross-sections was also addressed. Table A3 (see supplemental spreadsheet) contains the corresponding values for all submitted data and enables an overview of which vessel side branches were considered and which were rejected, either on purpose or due to insufficient segmentation techniques.

The inlet cross-sections represent the most proximal locations in the datasets considered, and thereby possess the largest surface areas orthogonal to the vessel centerlines. The calculation revealed average values of 20.44 ± 2.34 mm^2 (11.5%) for the anterior right case, 18.11 ± 1.89 mm^2 (10.4%) for the anterior left case and 11.52 ± 2.35 mm^2 (20.4%) for the posterior case, respectively.

A more distal analysis within the vasculature is enabled by comparing the outlet cross-sections that were considered by each individual group. Here, the effect of segmentation becomes more prominent, since the cumulated cross-sectional areas sum up as follows: anterior right 6.78 ± 1.59 mm^2 (23.4%); anterior left 10.44 ± 2.52 mm^2 (24.1%); posterior 3.79 ± 1.29 mm^2 (34%).

3.3 Morphological parameters

After evaluating integral parameters of the segmentation results, clinically relevant variables are now compared in addition. Here, five length measures based on the 3D surfaces, one volumetric quantity and four dimensionless ratios were selected to cover a range of relevant parameters. Table 3 contains the corresponding values for each group for the ruptured aneurysm E, and also presents the minimum, mean and maximum values as well as the standard deviations. It can be seen that the maximum lengths (diameter, height and width) vary between 4.6% and 7.1% among all results. This variation increases with respect to the orthogonal height and width, respectively

(8.07% and 8.12%). With increasing complexity of the morphological measure, stronger discrepancies among the groups can be observed. For instance, the aspect ratio, a clinically used quantity, was calculated as 0.483 for group 13 (minimum), while group 11 had a value almost three times larger of 1.422 (maximum).

Overall, this tendency resulted in a relative standard deviation of 19.74% for all groups. The same trend appears for advanced dimensionless ratios that aim at describing the complexity of shape, e.g. NSI and UI. Here, clear deviations among the groups are noted with a factor of around three between the lowest and highest values. Interestingly, ellipticity (EI) of the aneurysms, was not as variable as the previously mentioned parameters with a relative standard deviation of 4.23%.

4. Discussion

With advanced engineering techniques and increasing computational resources, physicians are supported prior to or during their interventions. One example is the application of image processing techniques to obtain three-dimensional representations of neurovascular diseases such as stenoses or IAs. Specifically, for IAs, a reliable rupture risk assessment is desired to be able to differentiate between stable and unstable aneurysms. However, in order to precisely evaluate both the existing morphology as well as the occurring hemodynamic features, accurate segmentation of the luminal surface is mandatory. This is typically based on different imaging modalities, which already differ with respect to spatial resolution (e.g., subtracted 3D rotational angiography, computed tomography angiography or magnetic resonance angiography) [36,37].

In order to evaluate the influence of segmentation approaches exclusively, the international MATCH 2018 was announced. Here, 26 groups from 13 countries participated and carried out segmentations of three 3D DSA datasets acquired in one female patient harboring five IAs. Within this study, comparisons regarding the segmentation methodologies focusing on applied algorithms and the segmentation durations were performed. Further, the inter-group variation with respect to aneurysm volumes as well as ostium size was assessed, and over- and underestimation were evaluated based on comparisons to highly resolved 2D images. Finally, clinically relevant morphology parameters were computed, which allows for a critical reflection of potential rupture risk predictions.

4.1 Segmentation methodology

It was noted that groups created highly individualized workflows to obtain segmentation results, which are usable for morphological analysis or image-based simulations. Those include multi-software applications and hence, no gold-standard with respect to segmentation strategy was observed. Overall, no consensus regarding an established software solution exists. Interestingly, mostly open-source software was used with a preference for VMTK (7 groups; 27%), which is based on a level set approach. However, accurate, non-outlying segmentation results were also obtained based on threshold methods and might not depend on the selection of the software or the underlying algorithm. Rather, they are related to subsequent modification steps after initialization, which involve experience as well as thoroughness in the finalization of the luminal surfaces.

Regarding segmentation times, it was found that no (semi-) automatic procedure exists, which creates suitable 3D representations of the intracranial vasculature within seconds (and hence, would be clinically applicable). Careful post-processing of initial segmentations still requires experienced manipulation, which is iterative and time-consuming. Although no correlation between the total time invested for the three segmentations and the resulting accuracies was found, it is noteworthy that group 11 performed best with respect to the aneurysm representation (recall Figure 4). This team used the commercial software Mimics (Materialise, Leuven, Belgium) for the initial segmentation, but invested by far the longest duration (26 hours per case) to finalize the segmentations. The following statement was provided in the corresponding abstract: "Once only important vasculature and the aneurysm(s) remained, the aneurysms were edited manually (by adding or removing individual pixels) to ensure good agreement between the CT images and the corresponding 3D model. Similarly, small cavities and spikes on the vasculature were removed via manual editing.".

Interestingly, group 11 is associated with one of the most experienced research centers regarding intracranial aneurysms, and was capable of reconstructing the neck of the ruptured aneurysm E appropriately, while almost all other groups overestimated its size. Hence, not only image quality of the 3D DSA but also the experience is crucial for obtaining plausible results.

These comparisons emphasize that realistic segmentations indeed required an investment of time. In contrast to the subsequent hemodynamic simulations, computational power is not a limiting factor. Here, case- and imaging-specific conditions dictate the amount of time required to achieve satisfactory results. Nevertheless, it is desired in the future to obtain accurate and reliable 3D vessel surfaces using a fully-automated process in a reasonable and clinically applicable time frame.

4.2 Segmentation variability

As an initial observation, a strongly varying number of outflow branches considered was utilized among the participants (e.g., ranging from four to eight). In particular, small side branches close to the aneurysms were either missing or underrepresented. As a consequence, this circumstance would certainly lead to differences with respect to potential intra-saccular blood flow predictions. As one can imagine, flow splitting techniques for outflow distributions highly depend on the underlying segmentation [38–40]. Therefore, it is suggested to keep as many side branches as possible, when blood flow investigations are desired (e.g., to assess the effect of side branch occlusion [41–43]) and the contrasted images allow for a reliable segmentation.

Furthermore, the qualitative and quantitative surface and volume comparisons between the submitted segmentation results revealed unexpected differences. Overall, the variability among the groups were around 20% for the volume calculations and around 30% for the ostium surface areas (ignoring the strongest deviations for the smallest aneurysm B). Additionally, it was found that with decreasing vessel diameter, the variability among the groups increases (inlets versus outlets). Roughly, the existing trend can be expressed as follows: Cross-sectional areas of approximately 20 mm^2 (≈ 5 mm average vessel diameter) exhibit variations of about 10%, while the difference for areas of 10 mm^2 (≈ 3.5 mm) increases to 20%. With further decrease in vessel diameter, deviations of even 30% were calculated.

Regarding the assessment of over- and underestimation of the 3D solutions compared to 2D images, interesting tendencies were revealed. First, groups rather oversegmented aneurysms C and E (the aneurysms considered in Fig. 7), with a difference of 9% and 15%, respectively. Especially, the aneurysm neck area was noticeably overrepresented, leading to potential misinterpretations. Second, small adjacent side branches were ignored, which affects the reconstruction of the true aneurysm ostium and in consequence the calculation of established hemodynamic parameters [18,44]. Third, undersegmentation occurred on the dome area of the aneurysms, which might be an effect of smoothing during the post-processing of initial segmentation results. Hence, important morphological features, such as existing blebs or irregularities in general, were diminished or even eliminated. These over- and underestimations may result from insufficient image quality or inappropriate time point of the acquisition (e.g., when the contrast agent is delayed), and certainly have an impact on subsequent CFD simulation results.

Overall, the smallest deviations occurred for groups using the open-sources software package VMTK, which is based on a level sets approach (to obtain the initial segmentation). This tool was specifically developed for vascular applications and hence appears to be suited for the considered sort of image data. However, due to its specialization, it might also be used by rather experienced researchers, who are aware of mandatory post-processing steps. Actually, manual modifications of initial segmentation results are often required, particularly when aneurysms possess narrow but short neck regions, or when an insufficient propagation of the contrast agent (e.g., in small side branches) exists.

4.3 Morphological parameters

As mentioned earlier, research groups have associated morphological properties with the rupture risk of IAs [45–47]. Further, the choice of treatment strongly depends on the individual phenotype, e.g., the existence of wide-neck bifurcation aneurysms [48]. Those size and morphology values are either obtained from two-dimensional images, which are exposed to vessel superpositions and perspective errors, or from three-dimensional segmentation of the intracranial vasculature. However, their sensitivity has not been objectively assessed in a broader comparison.

The analysis of ten clinically relevant morphology parameters for aneurysm E demonstrates that differences of simple measures such as maximum diameters or heights were below 10% (recall Table 3). However, with increasing complexity, deviations increased to 25% among the groups. Further, for quantities such as aspect ratio, non-sphericity and undulation, a factor of three lay between the corresponding minimum and maximum values. Taking this into account, previous studies comparing morphological differences between ruptured and unruptured aneurysms might be questioned, if inappropriate segmentation was performed.

4.4 Limitations

Besides the findings derived in this work based on qualitative as well as quantitative comparisons, several limitations of this study have to be mentioned. First, the co-alignment of the submitted segmentation results was inevitable in order to carry out the desired comparisons. This was achieved using an iterative closest point algorithm with a special focus on the existing aneurysms in each domain. Here, very good agreements between

the 3D surfaces were obtained, but minor misalignments in the most distal regions (at the borders of each dataset) might be possible. However, for the present analyses, each dataset was cropped and only the regions close to the aneurysms were considered.

Second, for the quantitative comparison of the aneurysm volumes as well as the ostium surfaces constant plane sections were selected. These had to be positioned with a slight distance to the parent vessel to capture all segmentation extents. Hence, those planes do not necessary represent the real ostium, and may lead to smaller volume values compared to reality. However, for the morphological analysis in Section 3.3, a semi-automatic ostium reconstruction based on Saalfeld et al. [32] was used, and therefore contains clinically realistic results.

Third, for the evaluation of the segmentation accuracy, a reference solution based on highly resolved 2D images was generated for aneurysms C and E, which can be seen as the clinical gold standard with respect to spatial resolution. However, due to incompatible calibration matrices, differing coordinate systems as well as unequal perspectives, the alignment between 2D and 3D was not trivial. Hence, three domain experts manually segmented the 2D images and co-registered the data with an inter-user-variability below 2%. Nevertheless, an automatic registration to match the existing data is desired in the future.

Finally, quantitative analysis included the data of all groups (except for partly group 3) and hence potential outliers might increase the value of the reported differences. However, no extreme differences occurred among the groups with respect to the analyzed quantities, which would have required an exclusion from the comparison.

5. Conclusions

To demonstrate and compare state-of-the-art segmentation techniques for IAs, which are required to assess clinically relevant morphology parameters, an international challenge (MATCH) was announced. Overall, 26 research groups around the world segmented five different IAs and the surrounding vasculature in the anterior and posterior circulation. Based on inter-group and reference comparisons, three main findings were derived:

1) Qualitative similarity of the segmentations was reached, but considerable variations occurred regarding the number of segmented vessels, and the predicted volumes as well as the sizes and shapes of the aneurysm ostia;

2) A comparison to highly resolved 2D reference images for two aneurysms revealed an overestimation of approximately 10% by the 3D surfaces. Specifically, the neck region of the ruptured aneurysm was overrepresented by almost all groups except for one; 3) The assessment of ten clinically relevant morphology parameters for the ruptured aneurysm revealed differences of up to 25% between the participants, which highlights the importance of careful and precise imaging, segmentation and subsequent measurement. Additionally, an accurate geometry is important for obtaining consistent CFD results, if quantification of the individual hemodynamics is desired.

Future work should include the development of an improved segmentation strategy that is tailored to the application of neurovascular diseases. Specifically, those methods should address effects that occur in the context of intracranial aneurysms, e.g., the overestimation of the aneurysm neck or the existence of melting artifacts. Here, the consideration of highly resolved 2D DSA images can be helpful. Further, the effect of the

presented segmentation variability on 1) the occurring hemodynamics and 2) on the aneurysm rupture risk assessment will be addressed in separate studies.

Compliance with Ethical Standards:

Funding: This study was funded by the Federal Ministry of Education and Research in Germany within the Forschungscampus *STIMULATE* (grant number 13GW0095A) and the German Research Foundation (grant number 399581926).

Conflict of Interest: Authors Philipp Berg, Samuel Voß, Sylvia Saalfeld, Gábor Janiga, Aslak W. Bergersen, Kristian Valen-Sendstad, Jan Bruening, Leonid Goubergrits, Andreas Spuler, Nicole M. Cancelliere, David A. Steinman, Vitor M. Pereira, Tin Lok Chiu, Anderson Chun On Tsang, Bong Jae Chung, Juan R. Cebral, Salvatore Cito, Jordi Pallarès, Gabriele Copelli, Benjamin Csippa, György Paál, Soichiro Fujimura, Hiroyuki Takao, Simona Hodis, Georg Hille, Christof Karmonik, Saba Elias, Kerstin Kellermann, Muhammad Owais Khan, Alison L. Marsden, Hernán G. Morales, Senol Piskin, Ender A. Finol, Mariya Pravdivtseva, Hamidreza Rajabzadeh-Oghaz, Nikhil Paliwal, Hui Meng, Santhosh Seshadhri, Matthew Howard, Masaaki Shojima, MD, Shin-ichiro Sugiyama, Kuniyasu Niizuma, Sergey Sindeev, Sergey Frolov, Thomas Wagner, Alexander Brawanski, Yi Qian, Yu-An Wu, Kent Carlson, Dan Dragomir Daescu, and Oliver Beuing declare that they have no conflicts of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors. Institutional review board approval was obtained from University Hospital Magdeburg for sharing of the anonymized images.

References

[1] Cebral JR, Mut F, Weir J, Putman C. Quantitative characterization of the hemodynamic environment in ruptured and unruptured brain aneurysms. AJNR Am J Neuroradiol 2011; 32: 145–151.

[2] Xiang J, Natarajan SK, Tremmel M, et al. Hemodynamic-morphologic discriminants for intracranial aneurysm rupture. Stroke 2011; 42: 144–152.

[3] Valen-Sendstad K, Steinman DA. Mind the gap: impact of computational fluid dynamics solution strategy on prediction of intracranial aneurysm hemodynamics and rupture status indicators. AJNR Am J Neuroradiol 2014; 35: 536–543.

[4] Cebral JR, Vazquez M, Sforza DM, et al. Analysis of hemodynamics and wall mechanics at sites of cerebral aneurysm rupture. J Neurointerv Surg 2015; 7: 530–536.

[5] Cebral JR, Mut F, Raschi M, et al. Aneurysm rupture following treatment with flow-diverting stents: computational hemodynamics analysis of treatment. AJNR Am J Neuroradiol 2011; 32: 27–33.

[6] Ma D, Xiang J, Choi H, et al. Enhanced aneurysmal flow diversion using a dynamic push-pull technique: an experimental and modeling study. AJNR Am J Neuroradiol 2014; 35: 1779–1785.

[7] Janiga G, Daróczy L, Berg P, Thévenin D, Skalej M, Beuing O. An automatic CFD-based flow diverter optimization principle for patient-specific intracranial aneurysms. J Biomech 2015; 48: 3846–3852.

[8] Berg P, Iosif C, Ponsonnard S, Yardin C, Janiga G, Mounayer C. Endothelialization of over- and undersized flow-diverter stents at covered vessel side branches: An in vivo and in silico study. J Biomech 2016; 49: 4–12.

[9] Berg P, Saalfeld S, Voß S, et al. Does the DSA reconstruction kernel affect hemodynamic predictions in intracranial aneurysms? An analysis of geometry and blood flow variations. J Neurointerv Surg 2018; 10: 290–296.

[10] Lauric A, Hippelheuser JE, Malek AM. Critical role of angiographic acquisition modality and reconstruction on morphometric and haemodynamic analysis of intracranial aneurysms. J Neurointerv Surg 2018.

[11] Firouzian A, Manniesing R, Flach ZH, et al. Intracranial aneurysm segmentation in 3D CT angiography: method and quantitative validation with and without prior noise filtering. Eur J Radiol 2011; 79: 299–304.

[12] Wang Y, Zhang Y, Navarro L, et al. Multilevel segmentation of intracranial aneurysms in CT angiography images. Med Phys 2016; 43: 1777.

[13] Berg P, Stucht D, Janiga G, Beuing O, Speck O, Thévenin D. Cerebral blood flow in a healthy Circle of Willis and two intracranial aneurysms: computational fluid dynamics versus four-dimensional phase-contrast magnetic resonance imaging. J Biomech Eng 2014; 136.

[14] Bouillot P, Brina O, Ouared R, Lovblad KO, Pereira VM, Farhat M. Multi-time-lag PIV analysis of steady and pulsatile flows in a sidewall aneurysm. Exp Fluids 2014; 55: 145.
[15] Bouillot P, Brina O, Ouared R, Lovblad K, Farhat M, Pereira VM. Hemodynamic transition driven by stent porosity in sidewall aneurysms. J Biomech 2015; 48: 1300–1309.
[16] Liang F, Liu X, Yamaguchi R, Liu H. Sensitivity of flow patterns in aneurysms on the anterior communicating artery to anatomic variations of the cerebral arterial network. J Biomech 2016; 49: 3731–3740.
[17] Meng H, Tutino VM, Xiang J, Siddiqui A. High WSS or low WSS? Complex interactions of hemodynamics with intracranial aneurysm initiation, growth, and rupture: toward a unifying hypothesis. AJNR Am J Neuroradiol 2014; 35: 1254–1262.
[18] Xiang J, Tutino VM, Snyder KV, Meng H. CFD: computational fluid dynamics or confounding factor dissemination? The role of hemodynamics in intracranial aneurysm rupture risk assessment. AJNR Am J Neuroradiol 2014; 35: 1849–1857.
[19] Fiorella D, Sadasivan C, Woo HH, Lieber B. Regarding "Aneurysm rupture following treatment with flow-diverting stents: computational hemodynamics analysis of treatment". AJNR Am J Neuroradiol 2011; 32: E95-7; author reply E98-100.
[20] Kallmes DF. Point: CFD--computational fluid dynamics or confounding factor dissemination. AJNR Am J Neuroradiol 2012; 33: 395–396.
[21] Cebral JR, Meng H. Counterpoint: realizing the clinical utility of computational fluid dynamics--closing the gap. AJNR Am J Neuroradiol 2012; 33: 396–398.
[22] Steinman DA. Computational modeling and flow diverters: a teaching moment. AJNR Am J Neuroradiol 2011; 32: 981–983.
[23] Radaelli AG, Augsburger L, Cebral JR, et al. Reproducibility of haemodynamical simulations in a subject-specific stented aneurysm model--a report on the Virtual Intracranial Stenting Challenge 2007. J Biomech 2008; 41: 2069–2081.
[24] Janiga G, Rössl C, Skalej M, Thévenin D. Realistic virtual intracranial stenting and computational fluid dynamics for treatment analysis. J Biomech 2013; 46: 7–12.
[25] Cito S, Geers AJ, Arroyo MP, et al. Accuracy and reproducibility of patient-specific hemodynamic models of stented intracranial aneurysms: report on the Virtual Intracranial Stenting Challenge 2011. Ann Biomed Eng 2015; 43: 154–167.
[26] Steinman DA, Hoi Y, Fahy P, et al. Variability of computational fluid dynamics solutions for pressure and flow in a giant aneurysm: the ASME 2012 Summer Bioengineering Conference CFD Challenge. J Biomech Eng 2013; 135: 21016.
[27] Janiga G, Berg P, Sugiyama S, Kono K, Steinman DA. The Computational Fluid Dynamics Rupture Challenge 2013—Phase I: prediction of rupture status in intracranial aneurysms. AJNR Am J Neuroradiol 2015; 36: 530–536.
[28] Berg P, Roloff C, Beuing O, et al. The Computational Fluid Dynamics Rupture Challenge 2013--Phase II: Variability of Hemodynamic Simulations in Two Intracranial Aneurysms. J Biomech Eng 2015; 137: 121008.
[29] Antiga L, Piccinelli M, Botti L, Ene-Iordache B, Remuzzi A, Steinman DA. An image-based modeling framework for patient-specific computational hemodynamics. Med Biol Eng Comput 2008; 46: 1097–1112.
[30] Chen Y, Medioni G. Object modelling by registration of multiple range images. Image and Vision Computing 1992; 10: 145–155.
[31] Besl PJ, McKay ND. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. 1992; 14: 239–256.
[32] Saalfeld S, Berg P, Niemann A, Luz M, Preim B, Beuing O. Semi-automatic neck curve reconstruction for intracranial aneurysm rupture risk assessment based on morphological parameters. Proc. of Computer assisted radiology (CARS), Berlin, Germany 2018.
[33] Niemann U, Berg P, Niemann A, et al. Rupture Status Classification of Intracranial Aneurysms Using Morphological Parameters. 31st IEEE CBMS International Symposium on Computer-Based Medical Systems, Karlstad, Sweden 2018.
[34] Raghavan ML, Ma B, Harbaugh RE. Quantified aneurysm shape and rupture risk. J Neurosurg 2005; 102: 355–362.
[35] Dhar S, Tremmel M, Mocco J, et al. Morphology parameters for intracranial aneurysm rupture risk assessment. Neurosurgery 2008; 63: 185-96; discussion 196-7.
[36] Geers AJ, Larrabide I, Radaelli AG, et al. Reproducibility of image-based computational hemodynamics in intracranial aneurysms: Comparison of CTA AND 3DRA. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI). p. 610–613 .

[37] Ren Y, Chen G, Liu Z, Cai Y, Lu G, Li Z. Reproducibility of image-based computational models of intracranial aneurysm: a comparison between 3D rotational angiography, CT angiography and MR angiography. Biomed Eng Online 2016; 15: 50.
[38] Murray CD. The Physiological Principle of Minimum Work: I. The Vascular System and the Cost of Blood Volume. Proc Natl Acad Sci U S A 1926; 12: 207–214.
[39] Chnafa C, Valen-Sendstad K, Brina O, Pereira VM, Steinman DA. Improved reduced-order modelling of cerebrovascular flow distribution by accounting for arterial bifurcation pressure drops. J Biomech 2017; 51: 83–88.
[40] Chnafa C, Brina O, Pereira VM, Steinman DA. Better Than Nothing: A Rational Approach for Minimizing the Impact of Outflow Strategy on Cerebrovascular Simulations. AJNR Am J Neuroradiol 2018; 39: 337–343.
[41] Cebral JR, Raschi M, Mut F, et al. Analysis of flow changes in side branches jailed by flow diverters in rabbit models. Int J Numer Method Biomed Eng 2014; 30: 988–999.
[42] Iosif C, Berg P, Ponsonnard S, et al. Role of terminal and anastomotic circulation in the patency of arteries jailed by flow-diverting stents: animal flow model evaluation and preliminary results. J Neurosurg 2016; 125: 898–908.
[43] Iosif C, Berg P, Ponsonnard S, et al. Role of terminal and anastomotic circulation in the patency of arteries jailed by flow-diverting stents: from hemodynamic changes to ostia surface modifications. J Neurosurg 2017; 126: 1702–1713.
[44] Chung B, Cebral JR. CFD for evaluation and treatment planning of aneurysms: review of proposed clinical uses and their challenges. Ann Biomed Eng 2015; 43: 122–138.
[45] Cai W, Hu C, Gong J, Lan Q. Anterior Communicating Artery Aneurysm Morphology and the Risk of Rupture. World Neurosurg 2018; 109: 119–126.
[46] Skodvin TØ, Evju Ø, Sorteberg A, Isaksen JG. Prerupture Intracranial Aneurysm Morphology in Predicting Risk of Rupture: A Matched Case-Control Study. Neurosurgery 2018.
[47] Mocco J, Brown RD, Torner JC, et al. Aneurysm Morphology and Prediction of Rupture: An International Study of Unruptured Intracranial Aneurysms Analysis. Neurosurgery 2018; 82: 491–496.
[48] Mihalea C, Caroff J, Rouchaud A, Pescariu S, Moret J, Spelle L. Treatment of Wide-Neck Bifurcation Aneurysm Using "WEB Device Waffle Cone Technique". World Neurosurg 2018; 113: 73–77.

Tables

Table 1 Summary of all software packages used for segmentation and post-processing (e.g., mesh correction) of the provided clinical datasets. The table contains the corresponding software name, its type with respect to licensing and the number of usages in total.

Software	**type**	**# of uses**
VMTK	open-source	7
MeshMixer	commercial	4
MATLAB	commercial	3
3D Slicer	open-source	2
MevisLab	open-source	2
ScanIP	commercial	2
Amira	commercial	2
ParaView	open-source	1
MeshLab	open-source	1
Blender	open-source	1
OpenFOAM	open-source	1
Solid Edge	commercial	1
AneuFuse	open-source	1
Mimics	commercial	1
Dornheim Segmenter	commercial	1
SC/Tetra	commercial	1
ZMD	in-house	1
Clinical research prototype	in-house	1
No-name	in-house	1

Table 2 Processing times (as reported by each group) from opening the clinical image data until finishing the three-dimensional surface representation of each lumen. The longest duration was required for the case "anterior right" due to the small distance between aneurysm A and the adjacent side branch. Notice the range of processing times: the fastest total segmentation time by group 13 was 43 minutes, and the longest duration by group 11 was 78 hours.

Group No.	**Processing times in hours**			
	anterior left	anterior right	posterior	**TOTAL**
1	2	2	2	**6**
2	2	2	2	**6**
3	3	3	3	**9**
4	0.67	0.67	0.67	**2**
5	0.5	0.5	0.5	**1.5**
6	0.75	0.75	0.75	**2.25**
7	2	2	2	**6**
8	3.8	2.7	2.2	**8.7**
9	3	4	2	**9**
10	1	1	1	**3**
11	26	26	26	**78**
12	3	3	3	**9**
13	0.18	0.4	0.13	**0.71**
14	3	3	3	**9**
15	2	2	2	**6**
16	0.67	2	0.67	**3.34**
17	6	9	7	**22**
18	1.5	7.33	1.33	**10.16**
19	0.25	0.42	0.25	**0.92**
20	0.33	0.22	0.15	**0.7**
21	2	2.5	3	**7.5**
22	0.3	0.3	0.3	**0.9**
23	1	1	1	**3**
24	6	6	6	**18**
25	0.5	0.67	0.33	**1.5**
26	0.25	0.25	0.25	**0.75**

Table 3 Morphological parameters for the ruptured aneurysm E including the following quantities: maximum height of the aneurysm H_{max}, maximum width W_{max}, height of the aneurysm approximated as length of the ray perpendicular to the ostium plane H_{ortho}, maximum width parallel to the projected ostium plane W_{ortho}, maximum diameter of the aneurysm D_{max}, aspect ratio AR, volume of the convex hull V_{CH}, ellipticity index EI, non-sphericity index NSI, undulation index UI. For comparison, minimum, mean and maximum values are presented as well as the absolute and relative standard deviations (SD). To reduce the effect of potential outliers, median values as well as the first (Q1) and third (Q3) quartiles are provided.

Group	Morphological parameters									
	D_{max} [mm]	H_{max} [mm]	W_{max} [mm]	H_{ortho} [mm]	W_{ortho} [mm]	AR [-]	V_{CH} [mm³]	EI [-]	NSI [-]	UI [-]
1	5.343	3.685	4.451	3.118	5.335	0.919	36.978	0.258	0.142	0.066
2	5.576	3.592	4.684	2.994	5.339	0.747	42.596	0.264	0.125	0.088
3	4.305	2.992	3.889	2.760	3.972	0.840	22.078	0.258	0.094	0.088
4	5.625	3.597	4.516	2.899	5.556	0.690	41.101	0.270	0.113	0.096
5	5.333	3.637	4.404	3.049	5.319	0.891	35.480	0.260	0.138	0.064
6	5.596	3.480	4.569	2.887	5.529	0.685	41.218	0.269	0.114	0.103
7	5.323	3.904	4.661	3.273	5.291	0.950	40.637	0.252	0.144	0.073
8	5.057	3.723	4.193	3.105	4.816	0.917	33.539	0.246	0.116	0.053
9	5.369	3.362	4.476	2.859	5.210	0.712	37.706	0.262	0.079	0.053
10	4.772	3.672	4.202	3.095	4.575	1.076	26.846	0.253	0.151	0.056
11	5.146	4.170	4.440	3.806	4.939	1.422	36.699	0.250	0.180	0.043
12	5.216	3.471	4.369	2.968	4.949	0.833	33.228	0.257	0.124	0.086
13	5.786	3.238	4.554	2.302	5.630	0.483	33.775	0.304	0.054	0.128
14	5.738	3.557	4.093	2.864	5.699	0.705	39.151	0.274	0.103	0.067
15	4.988	3.480	4.358	3.005	4.866	0.926	29.387	0.258	0.146	0.068
16	5.343	3.591	4.350	2.973	5.237	0.847	35.219	0.261	0.129	0.067
17	5.476	3.444	4.838	2.923	5.410	0.690	41.058	0.261	0.086	0.065
18	5.914	3.830	4.399	3.035	5.896	0.804	42.595	0.271	0.153	0.100
19	4.774	3.488	4.312	3.101	4.659	1.026	28.160	0.252	0.143	0.076
20	5.074	3.552	4.335	2.997	5.004	0.911	31.428	0.259	0.133	0.055
21	5.368	3.845	4.618	3.243	5.344	0.949	39.166	0.253	0.139	0.062
22	4.711	3.589	4.209	3.066	4.653	1.046	27.100	0.255	0.150	0.061
23	4.729	3.250	4.108	2.816	4.649	0.894	25.609	0.260	0.129	0.066
24	5.122	3.427	4.253	2.971	5.093	0.858	32.770	0.260	0.131	0.078
25	5.306	3.718	4.451	3.185	5.291	0.916	37.878	0.254	0.129	0.057
26	4.985	3.665	4.336	3.192	4.815	1.037	32.663	0.249	0.154	0.086
min	4.305	2.992	3.889	2.302	3.972	0.483	22.078	0.246	0.054	0.043
max	5.914	4.170	4.838	3.806	5.896	1.422	42.596	0.304	0.180	0.128
mean	5.230	3.575	4.387	3.019	5.118	0.876	34.772	0.260	0.127	0.073
Q1	5.005	3.473	4.268	2.905	4.828	0.761	31.737	0.253	0.115	0.062
median	5.315	3.590	4.384	3.001	5.224	0.893	35.350	0.259	0.130	0.067
Q3	5.449	3.682	4.506	3.104	5.343	0.943	39.162	0.262	0.144	0.086
SD	0.371	0.228	0.202	0.244	0.416	0.173	5.575	0.011	0.027	0.019
%	**7.10**	**6.39**	**4.61**	**8.07**	**8.12**	**19.74**	**16.03**	**4.23**	**20.90**	**25.62**

Figures

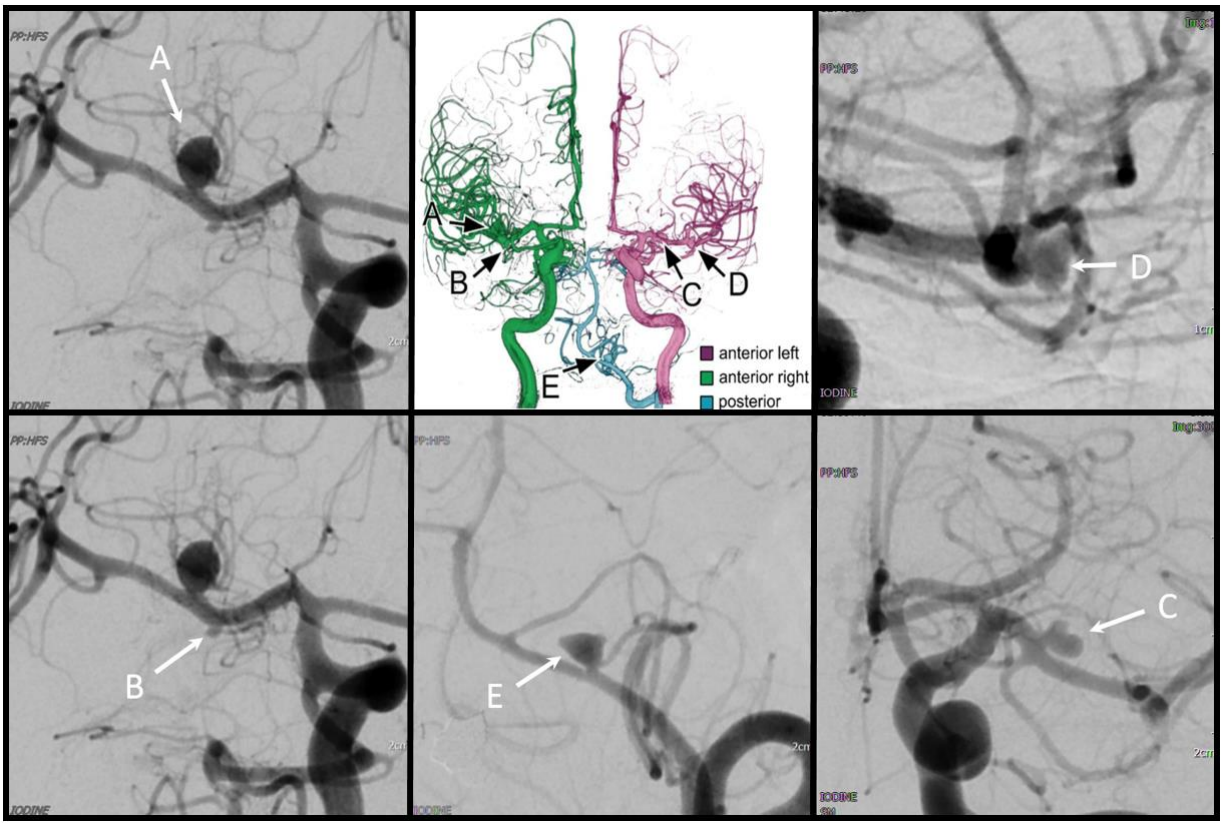

Fig. 1 Illustration of the provided MATCH case containing five intracranial aneurysms. Arrows indicate the corresponding locations, denominated A (right middle cerebral artery M1), B (right middle cerebral artery M1), C (left middle cerebral artery M1), D (left middle cerebral artery bifurcation) and E (left posterior inferior cerebellar artery).

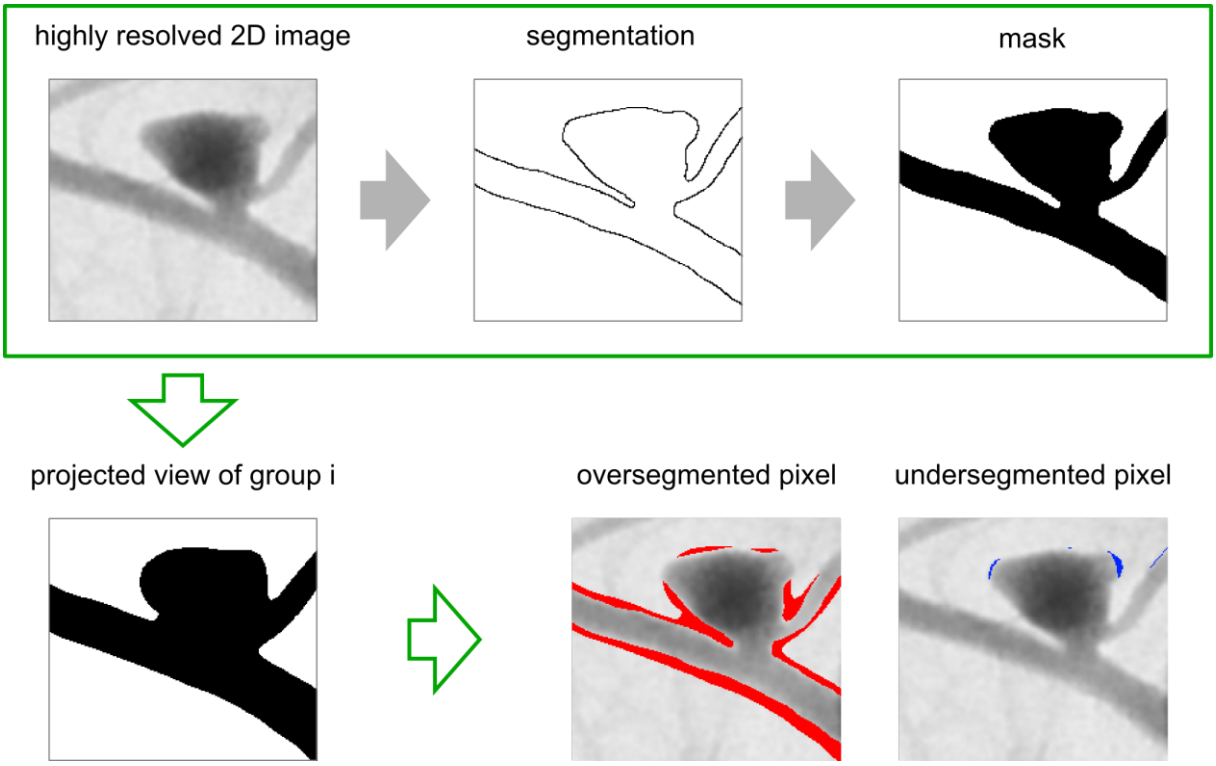

Fig. 2 Illustration of the comparison technique developed to quantify the difference between the highly resolved 2D images and the individual segmentation result from each participating group. After segmenting the 2D image, a mask was created, which was compared pixel-wise with the projected view of the individual group. Hence, over- and underestimation were evaluated as exemplified on the bottom right.

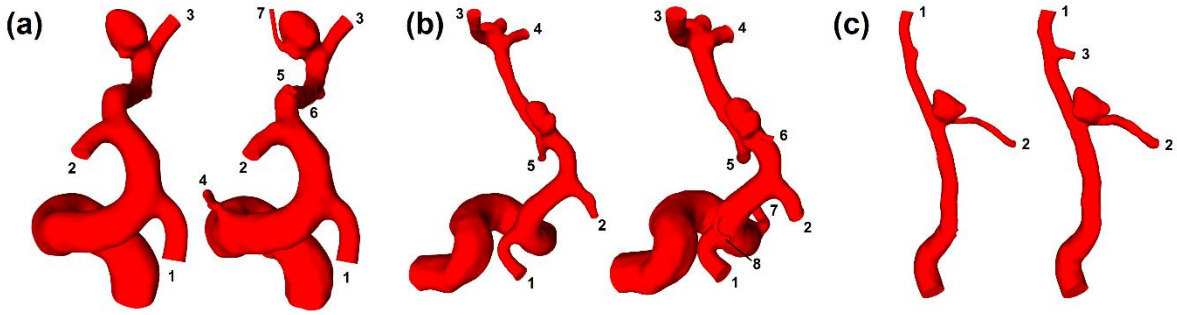

Fig. 3 Illustration of representative groups providing segmentation results with the lowest and the highest number of considered outlet cross-sections: (a) right anterior circulation (group 21 (3 outlets) versus group 23 (7 outlets)); (b) left anterior circulation (group 18 (5) versus group 7 (8)); (c) posterior circulation (group 15 (2) versus group 7 (3)).

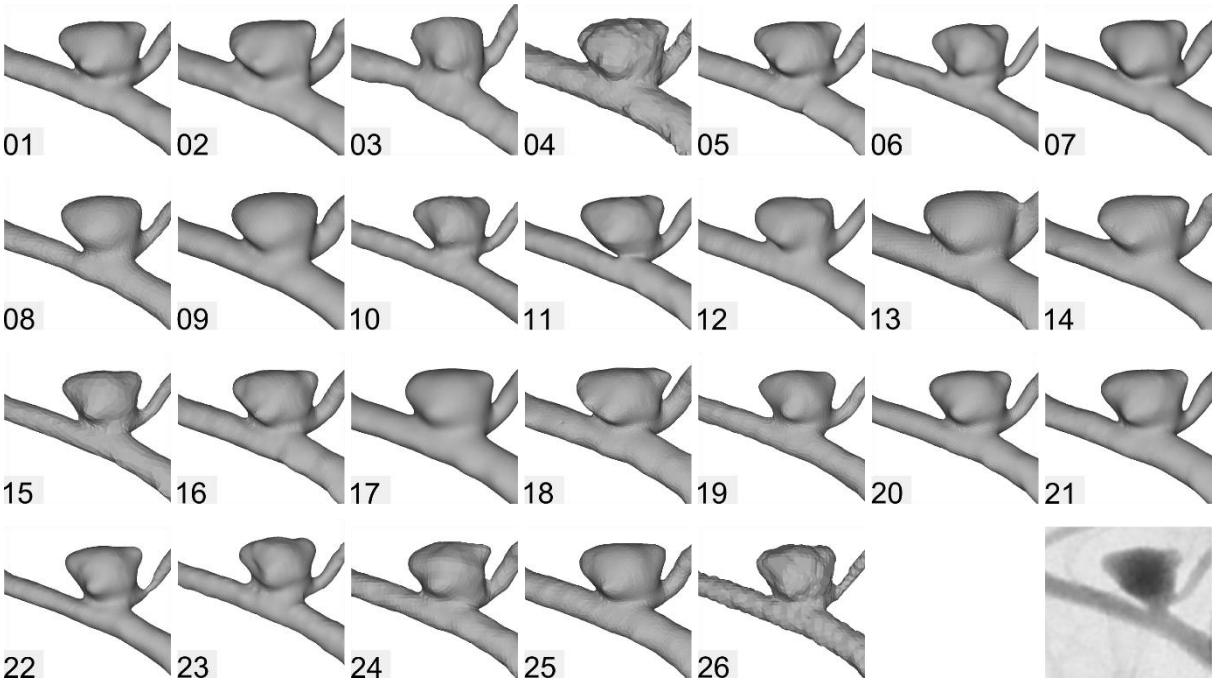

Fig. 4 Segmentation results of each group (1-26) for the ruptured aneurysm E located at the left PICA. Notice the inconsistencies with respect to surface smoothness as well as the aneurysm shape and neck representation, respectively. The panel in the right bottom shows the highly resolved 2D DSA image, which serves as a reference for the evaluation of over- and underestimation of the segmentations (see Section 3.2).

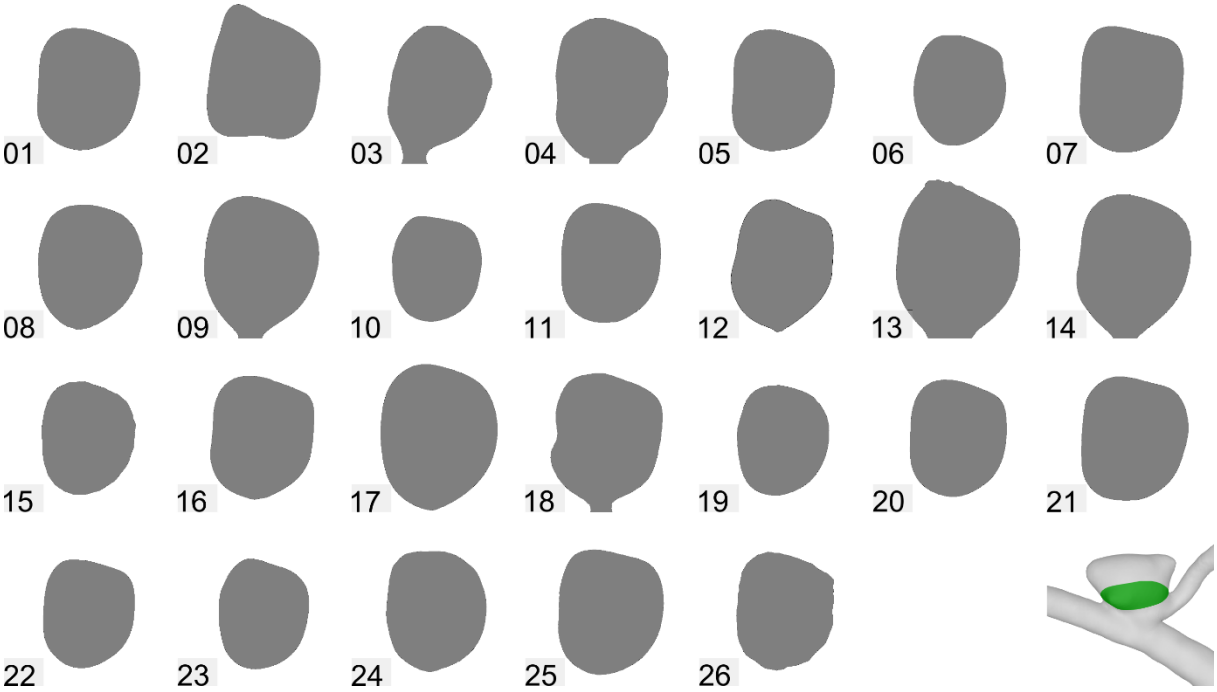

Fig. 5 Illustration of the ostium cut-plane for the ruptured aneurysm E located at the left PICA. Notice the inconsistencies with respect to size and shape.

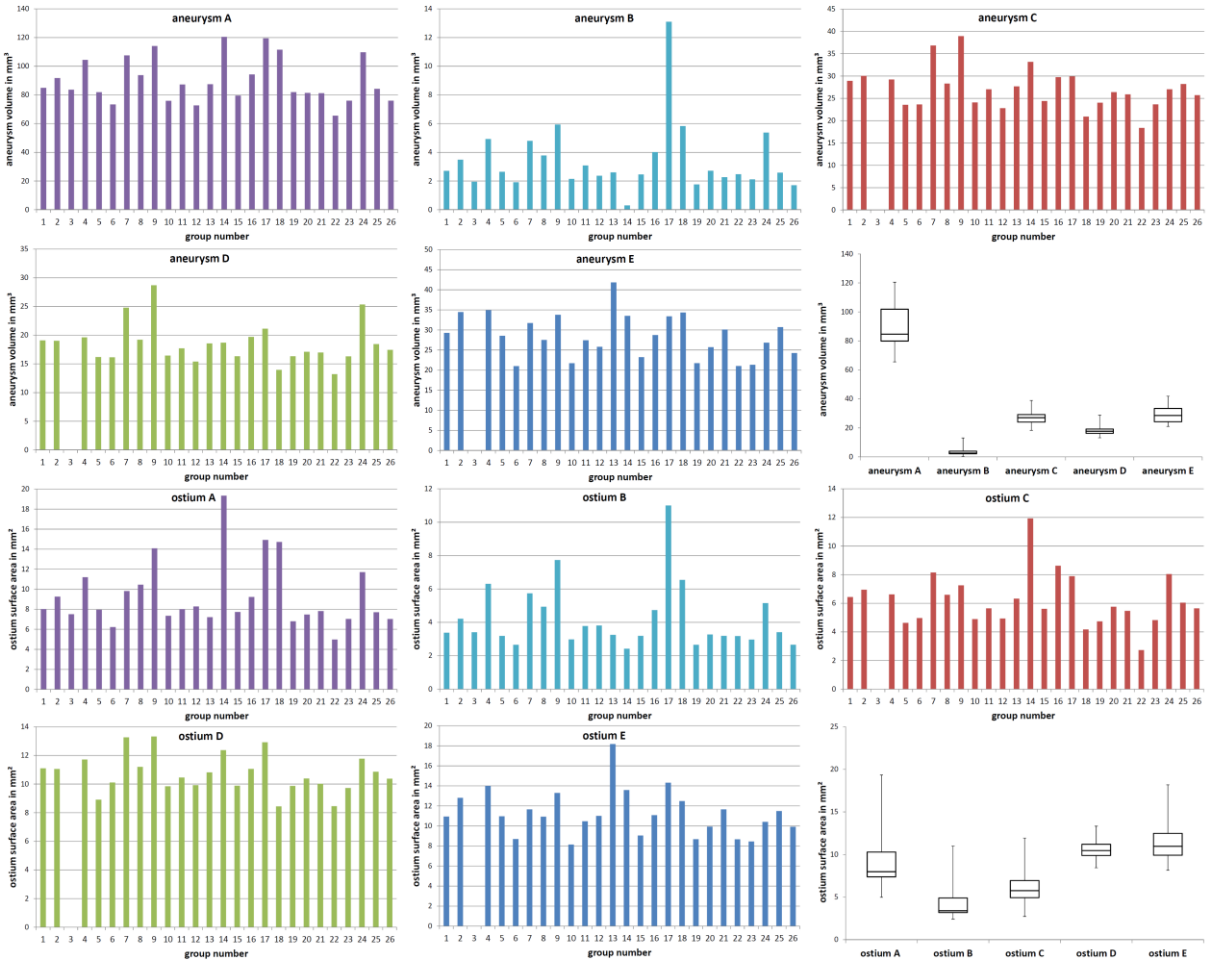

Fig. 6 Bar graphs and box-plots of the individual aneurysm volumes (top) and the representative aneurysm cut-planes in the vicinity of each aneurysm ostium (bottom).

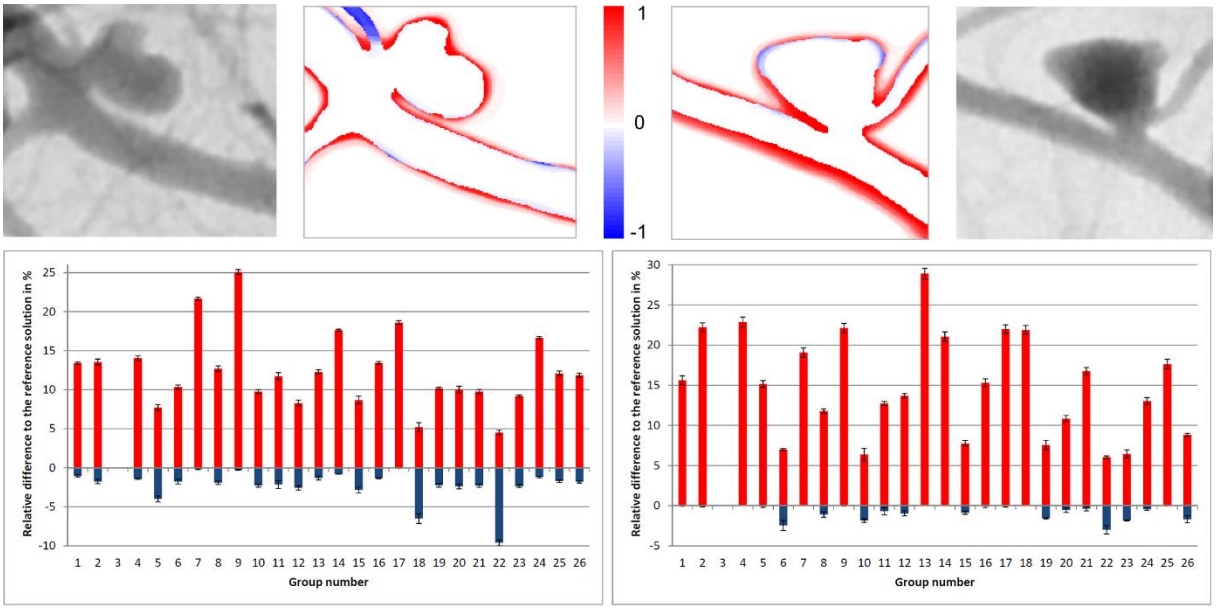

Fig. 7 Quantification of the segmentation differences between each group and the highly resolved 2D reference images for representative aneurysms C (left) and E (right). *Top row*: Cumulative over- (red) and underestimation (blue) ranging from 1 (maximum overestimation) to -1 (maximum underestimation). Notice the differences occurring close to the aneurysm neck as well as the underestimation in areas of steep gradients along the aneurysm wall. Further, the ruptured aneurysm E (right) exhibits biggest diameter errors, presumably because it had a smaller diameter; *Bottom row*: Bar graphs of the over- and underestimation compared to the highly resolved 2D DSA reference solution for aneurysm C (left MCA) and aneurysm E (left PICA). Error bars indicate the domain expert variability in the registration and segmentation process.

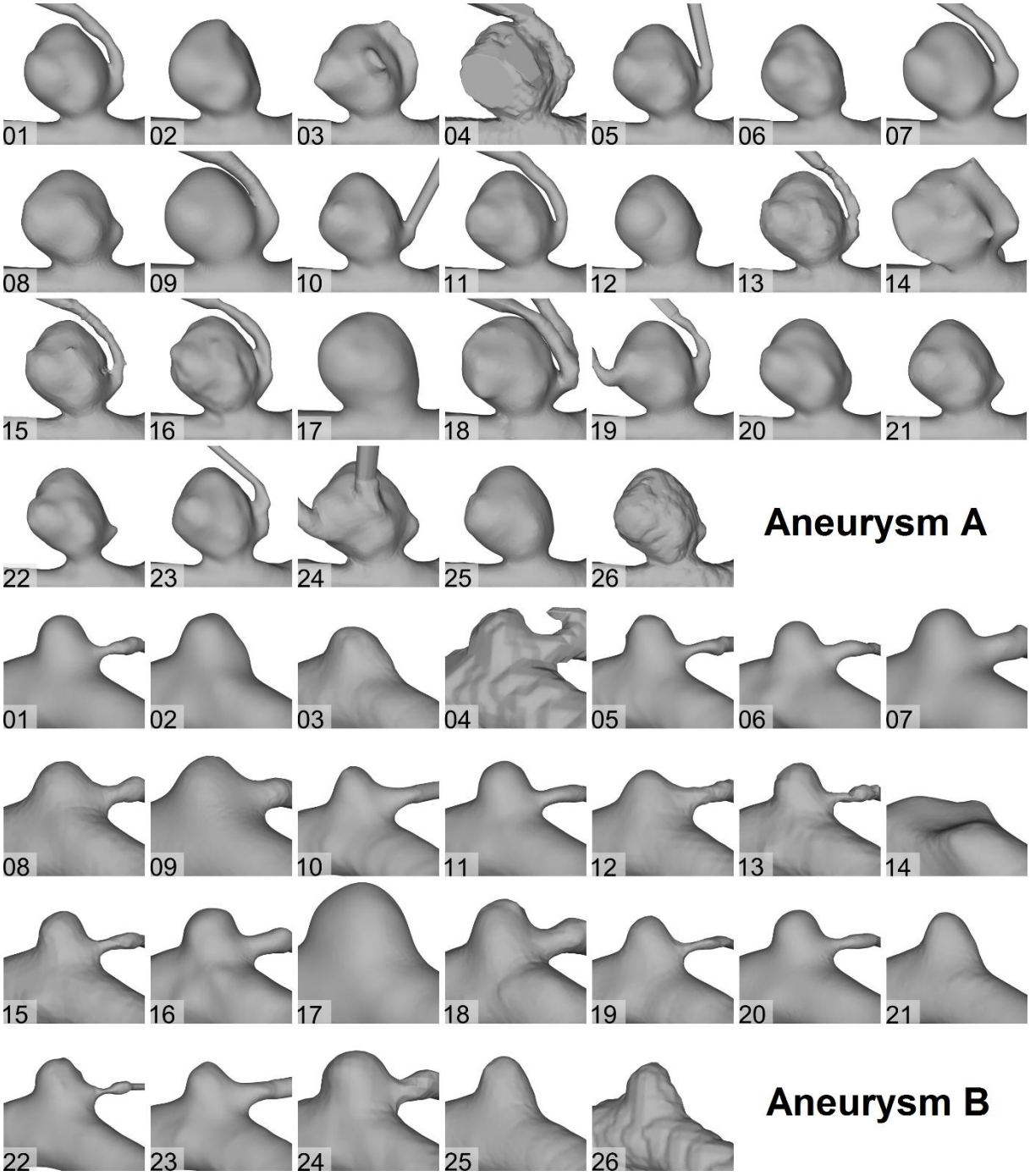

Fig. A1 Segmentation results of each group for aneurysms A and B (right MCA) from one perspective. Notice the inconsistencies with respect to surface smoothness, aneurysm neck representation and number of side branches considered.

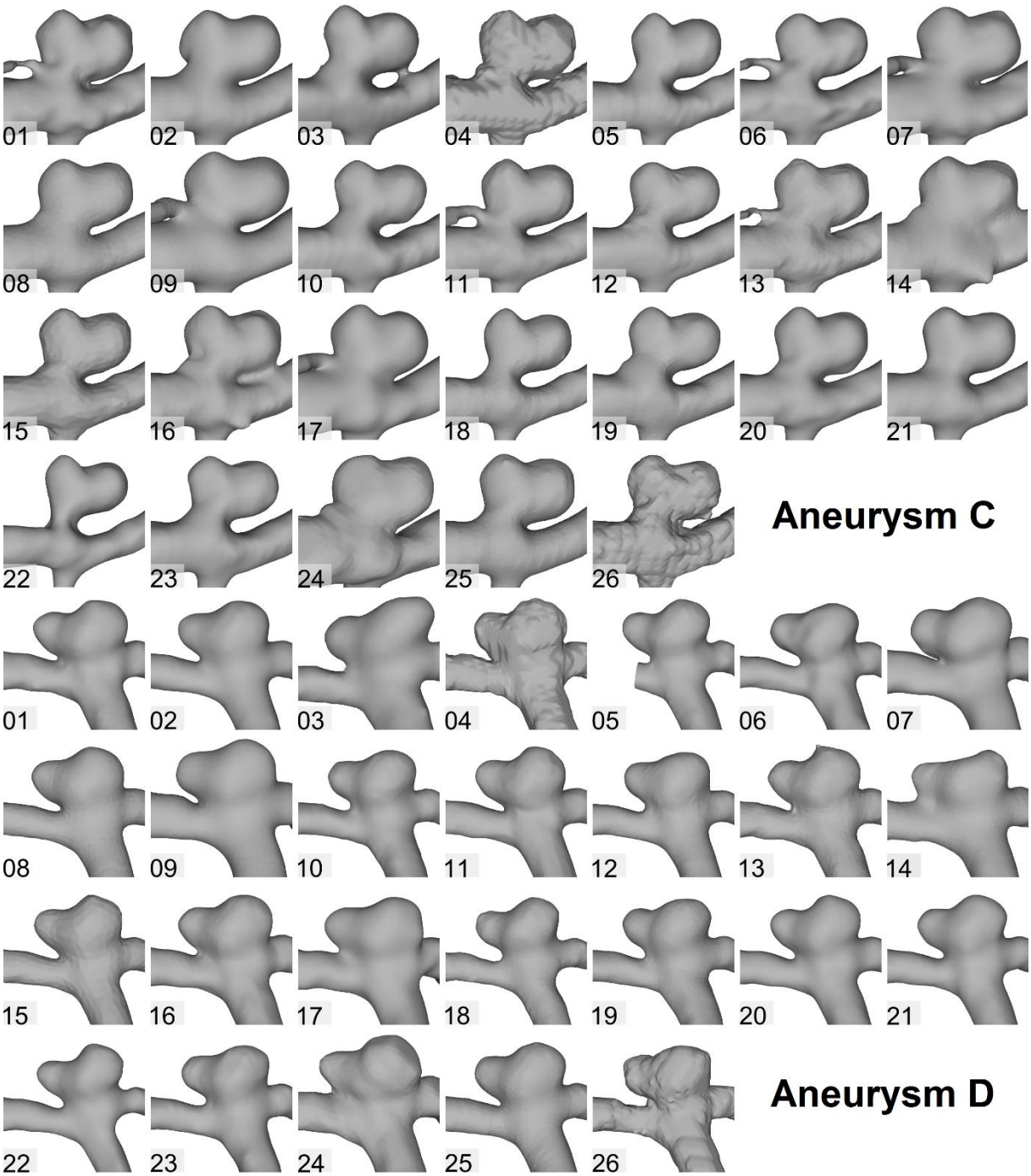

Fig. A2 Segmentation results of each group for aneurysm C and D (left MCA) from one perspective. Notice the inconsistencies with respect to surface smoothness, aneurysm neck representation and number of side branches considered.

Table captions

Table 1 Summary of all software packages used for segmentation and post-processing (e.g., mesh correction) of the provided clinical datasets. The table contains the corresponding software name, its type with respect to licensing and the number of usages in total.

Table 2 Processing times (as reported by each group) from opening the clinical image data until finishing the three-dimensional surface representation of each lumen. The longest duration was required for the case "anterior right" due to the small distance between aneurysm A and the adjacent side branch. Notice the range of processing times: the fastest total segmentation time by group 13 was 43 minutes, and the longest duration by group 11 was 78 hours.

Table 3 Morphological parameters for the ruptured aneurysm E including the following quantities: maximum height of the aneurysm H_{max}, maximum width W_{max}, height of the aneurysm approximated as length of the ray perpendicular to the ostium plane H_{ortho}, maximum width parallel to the projected ostium plane W_{ortho}, maximum diameter of the aneurysm D_{max}, aspect ratio AR, volume of the convex hull V_{CH}, ellipticity index EI, non-sphericity index NSI, undulation index UI. For comparison, minimum, mean and maximum values are presented as well as the absolute and relative standard deviations (SD). To reduce the effect of potential outliers, median values as well as the first (Q1) and third (Q3) quartiles are provided.

Table S3 Cross-sectional areas of the in- and outflow vessels in mm² for all three datasets. Empty fields indicate the absence of the corresponding vessel.