Eur J Vasc Endovasc Surg (2021) 62, 869—877

EJVES Open Access

Pre-surgical and Post-surgical Aortic Aneurysm Maximum Diameter Measurement: Full Automation by Artificial Intelligence

Chloé Adam [a], Dominique Fabre [b], Justine Mougin [b], Marc Zins [c], Arshid Azarine [c], Roberto Ardon [a], Gaspard d'Assignies [a], Stephan Haulon [b,*]

[a] Incepto Medical, Paris, France
[b] Aortic Centre, Hôpital Marie Lannelongue, Groupe Hospitalier Paris Saint Joseph, Université Paris Saclay, Paris, France
[c] Radiology Department, Groupe Hospitalier Paris Saint Joseph, Paris, France

WHAT THIS PAPER ADDS

A fully automated solution has been developed with an artificial intelligence (AI) start up company providing automatic cross sectional outer to outer wall aortic measurements along the entire aorta, from the ascending aorta to the iliac arteries. This AI solution has been proven to be accurate and will provide automatic monitoring of all segments of the aorta before and after treatment.

Objective: The aim of this study was to evaluate an automatic, deep learning based method (Augmented Radiology for Vascular Aneurysm [ARVA]), to detect and assess maximum aortic diameter, providing cross sectional outer to outer aortic wall measurements.

Methods: Accurate external aortic wall diameter measurement is performed along the entire aorta, from the ascending aorta to the iliac bifurcations, on both pre- and post-operative contrast enhanced computed tomography angiography (CTA) scans. A training database of 489 CTAs was used to train a pipeline of neural networks for automatic external aortic wall measurements. Another database of 62 CTAs, including controls, aneurysmal aortas, and aortic dissections scanned before and/or after endovascular or open repair, was used for validation. The measurements of maximum external aortic wall diameter made by ARVA were compared with those of seven clinicians on this validation dataset.

Results: The median absolute difference with respect to expert's measurements ranged from 1 mm to 2 mm among all annotators, while ARVA reported a median absolute difference of 1.2 mm.

Conclusion: The performance of the automatic maximum aortic diameter method falls within the interannotator variability, making it a potentially reliable solution for assisting clinical practice.

Keywords: Aortic aneurysm, Automatic measurements, Deep learning, Outer to outer wall diameters, Pre-operative/post-operative CT scans
Article history: Received 23 February 2021, Accepted 11 July 2021, Available online 10 September 2021

INTRODUCTION

The prevalence of abdominal aortic aneurysm (AAA) is 4 — 8% in screening studies,[1] affecting predominantly men > 60 years of age. AAA rupture occurs in 1% — 3% of men aged > 65 with an associated mortality rate of > 70% and up to 90% if rupture occurs outside of the hospital. AAA results in 15 000 deaths annually in the USA.[2]

According to current recommendations, the surgical decision to treat an aortic aneurysm to prevent rupture is based on its maximum external diameter and its evolution over time, as measured on ultrasound or computed tomography (CT) images. Beyond 55 mm in diameter and /or beyond a 10 mm increase over a 12 month period, aortic repair is recommended to prevent rupture.[3] Thus, accurate aortic diameter measurements are key. These repeated measures can be challenging and are time consuming because of tortuous aortic anatomy, especially in elongated aneurysmal aortas. The currently available three dimensional workstations provide several tools to facilitate aortic anatomy analysis, such as semi-automatic segmentations of the aorta and drawing of its central lumen line. They are accurate for contrast enhanced lumen diameter measurements but are unable to measure the outer to outer wall diameter because of the thrombus filling the aortic lumen; this inner lumen automatic diameter measurement is of limited use as only outer to outer wall diameters are relevant in clinical practice.

Therefore, there is a need to develop new tools for automatic and standardised maximum external diameter measurement all along the aorta. Fortunately, algorithmic methods and computing power have positively evolved,

---

* Corresponding author. Aortic Centre, Hôpital Marie Lannelongue, Groupe Hospitalier Paris Saint Joseph, Université Paris Saclay, 92350 Le Plessis-Robinson, France.

E-mail addresses: haulon@hotmail.com; haulon@ccml.fr (Stephan Haulon).

https://doi.org/10.1016/j.ejvs.2021.07.013

especially with the recent advances in artificial intelligence (AI), and in particular deep learning (DL). These tools have generated great improvements in algorithm performance, but clinically relevant results are still lacking.

Incepto Medical, a French AI start up company, in collaboration with the authors' vascular surgery department, developed a fully automated solution named Augmented Radiology for Vascular Aneurysm (ARVA), providing automatic cross sectional outer to outer wall aortic measurements. To the authors' knowledge, this is the first algorithm to provide automatic measurements of external aortic diameters along the entire aorta, from the ascending aorta to the iliac arteries. The aim of this study was to assess the ability of the algorithm to provide robust measurements of pre- and post-surgical aorta *vs.* human measurements.

MATERIALS AND METHODS

The study was approved by the Institutional Review Board of Groupe Hospitalier Paris Saint Joseph, Hôpital Marie Lannelongue. Informed consent was obtained from each participant (from both training and validation cohorts) before study participation (IRB 00012157).

Study design

A dataset was assembled retrospectively to develop and validate ARVA, an algorithm that automatically measures maximum external aortic diameters on thoracic and/or abdomen contrast enhanced (angio) CT scans of patients from the authors' institution. Patients $<$ 18 years old were excluded.

Datasets

In total, 551 angioCT scans were collected from 345 patients, representing a large variety of cases encountered in a surgeon's or radiologist's real life practice (Table 1). Various aneurysm locations and morphologies, as well as healthy aortas, were included. Both pre- and post-operative angioCT scans were available for patients undergoing open or endovascular repairs. Patients with a large variety of grafts, endografts, coils, and other endovascular material (with possible strong metal artefacts) were enrolled. Most angioCT scans analysed were not electrocardiogram gated, resulting in a frequent blurring of the ascending aorta. Data diversity was ensured as the CT scans were performed by four different manufacturers (see Table 1). No specific acquisition protocol was required, except that angioCT scans with a slice thickness $>$ 2 mm were excluded. The dataset was split into two different subsets for training and validation.

Training dataset

The training dataset used to both train and finetune the different algorithms consisted of 489 CT scans from 236 males and 47 females. The majority of these images had an arterial phase, and only a few had a portal phase. The mean patient age in the training set was 71 years and the mean patient weight was 74 kg. Each of these CT scans was manually segmented by one of the three human experts (5 − 15 years of experience in vascular imaging) without overlap.

The aortic masks were created using an in house developed annotation tool (Incepto Medical, Paris, France). Based on the work of Mory *et al.*,[4] an online three dimensional segmentation tool was implemented. Expert annotators iteratively segmented the aorta by drawing and merging adaptive shapes. In homogenous regions, these shapes are spheres. A high contrast change in intensity induces a deformation of the sphere along the gradient. The user can generate additions to the current segmentation or delete regions. The iterative segmentation steps are illustrated in Fig. 1. A whole aorta (from ascending aorta to iliac arteries) can be interactively segmented in around 10 minutes. These CT scans and aortic masks were used as input and output data to train DL algorithms as explained in the "Automatic pipeline" subsection.

Validation dataset

The validation dataset used to test and evaluate the pipeline consisted of 62 CT scans from 43 males and 19 females. The mean patient age in the validation set was 69 years and the mean weight was 73 kg. The validation set was divided into three groups: healthy aortas; diseased aortas; and diseased aortas after treatment (Fig. 2). This last subgroup included follow up CT scans after endovascular aortic repairs. These three subsets were used to assess the algorithm performance in various clinical conditions: screening; pre-operative work up; and follow up. Each subset included

Table 1. Description of training and validation of dataset of aortic computed tomography (CT) angiography images for Augmented Radiology for Vascular Aneurysm evaluation

	Training (*n* = 489 CT scans from 283 patients)	**Validation (*n* = 62 CT scans from 62 patients)**
Patient demographics		
Male	236 (83)	43 (69)
Age − y	71 (23–90)	69 (32–89)
Weight − kg	74 (45–120)	73 (56–100)
CT parameters		
Scans − *n*	489	62
Manufacturers/no. of scans	Siemens/430, GE/59	Siemens/11, GE/48, Philips/2, Toshiba/1
Aortic characteristics		
Diseased aortas	464 (95)	42 (67)
Post-operative aortas with stents* − %	391 (80)	22 (35)

Data are presented as *n* (%) or mean (range) unless stated otherwise.
* Endovascular aneurysm repair (EVAR), branched EVAR, fenestrated EVAR, or thoracic EVAR.

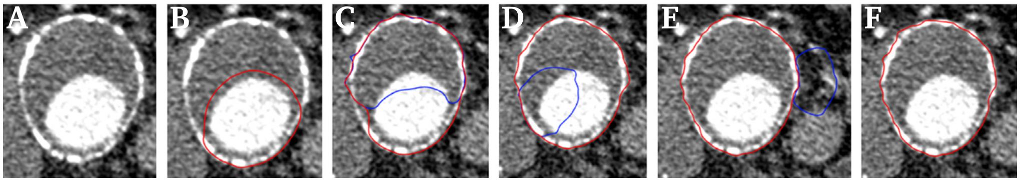

Figure 1. (A–F) Successive manual segmentation steps with the annotation tool. The red contour corresponds to the current segmentation and the blue contour to the gradient adaptive shape. The latter can either augment the current segmentation or reduce it by means of union or extrusion, respectively.

20, 20, and 22 angioCT scans, respectively, from different patients. These 62 patients were not used for training. Subset 1 contained “healthy aortas”, although some calcifications of the aortic wall were present. This subset was the only one in which CT acquisitions started at the level of the distal thoracic aorta, so only the descending thoracic and abdominal aorta were available for analysis (not the ascending or the arch). The validation process at the level of the ascending aorta and the arch was thus not possible. In the other two subsets, the complete aorta (from the ascending aorta to the iliac arteries) was processed. Subset 2 included patients with diseased aortas, including thoracic, thoraco-abdominal, and abdominal aortic aneurysms or dissections. Three of these patients had a prior history of open aortic repair. Subset 3 included 22 patients who underwent endovascular repair (endovascular aneurysm repair [EVAR], $n = 6$; thoracic EVAR, $n = 4$; branched EVAR, $n = 2$; fenestrated EVAR, $n = 10$); some angioCT scans had metal artefacts. The same human experts manually created aortic segmentation masks for all the angioCT scans from the validation set, without overlap.

Manual outer to outer wall diameter measurements

The maximum cross sectional diameter of the aorta was measured by seven clinicians: two experienced aortic surgeons (experts 1 and 2), three vascular surgery residents (residents 1, 2, and 3), and two general radiologists (radiologist 1 and 2). The two experts used the Aquarius Workstation (Terarecon, Foster City, CA, USA), whereas the other clinicians used the RadiAnt Dicom Viewer (Medixant, Poznan, Poland), which does not provide an automatic centreline, but allows multiplanar reconstructions (MPRs). The seven clinicians, blinded to any radiological report, independently measured the maximum external (outer wall to outer wall) aortic diameters on the 62 angioCT scans from the validation dataset.

Automatic pipeline

Pipeline steps. The proposed algorithms pipeline leverages neural network learning abilities and aims to reproduce diameter measurements performed by clinicians. It is composed of five successive steps.

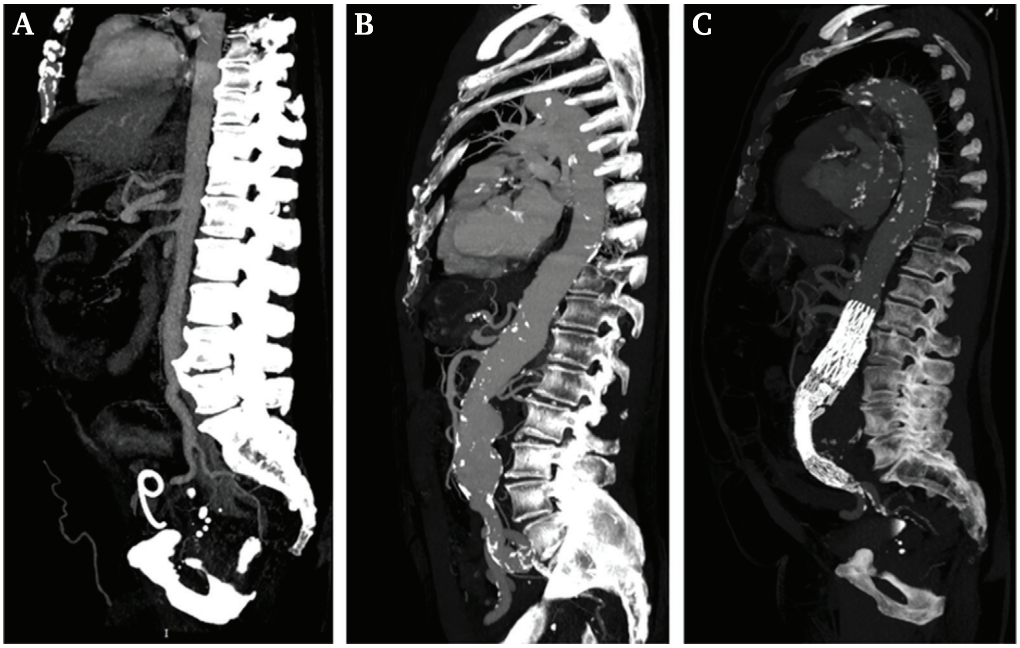

Figure 2. Representative images of (A) subset 1 of healthy aorta, (B) subset 2 of diseased aorta, and (C) subset 3 of diseased aorta after endovascular treatment.

Step 1: localisation. The aorta is, in general, a relatively small structure (in width) in the body, compared with other organs. To maintain enough resolution during segmentation without requiring excessive use of memory, first a fully convolutional network is used to regress the bounding box coordinates of the aorta. All volumes in the training set are resampled to a fixed size of (160, 160, 160). The architecture used is the following: four convolution blocks, each containing three convolutional layers, and ending with a global average pooling layer, resulting in six values corresponding to the coordinates of the upper left and the lower right corners of the aortic bounding box. The number of kernels used in the 12 resulting convolutional layers is (16, 16, 16); (32, 32, 32); (64, 64, 64); (128, 128, 6). All kernel sizes are fixed to (3, 3, 3). Each convolution layer is followed by a ReLU activation and a batch normalisation step. Maxpooling (factor 2) is applied after each convolution block, and final global average pooling is followed by a ReLU activation to output the final result. Optimisation is performed using an Adam optimiser, an L1 regression loss, and a learning rate starting at 1e-5 and being divided by two every 20 epochs. Optimisation was done over a total of 60 epochs. The training database was augmented using random translations and zooms limited to 10 mm and a factor variation of 0.1, respectively.

Step 2: segmentation. Following the localisation step, the predicted bounding box is used to perform a crop of the box containing the aorta at full resolution. This crop is further resampled to a fixed size of (160, 160, 160). The full segmentation of the aorta (lumen and thrombus) is performed using a V-Net architecture,[5] with a symmetrical structure. Both its descending and ascending branches are composed of three convolution blocks, each containing three convolutional layers. The number of kernels used in the nine resulting convolutional layers is (16, 16, 16); (32, 32, 32); (64, 64, 64). All kernel sizes are fixed to (3, 3, 3). Between convolutional layers and convolution blocks, the same layers as in the localisation network are used. The final activation function is a sigmoid to ensure output values in the range [0,1]. To train this network we use a combination of a Dice coefficient loss (one over the Dice coefficient) and a cross entropy loss. Optimisation is done using an Adam optimiser with a starting learning rate of 1e-4 and divided by two every 15 epochs. A total of 50 epochs was enough for convergence.

Step 3: segmentation refinement. This step uses the full resolution volume and the segmentation mask obtained in the previous step. It aims at correcting errors made by the V-Net generating by its inferior working resolution. A fixed number of iterations (10 in the current version) of the fuzzy region competition algorithm was performed.[6]

Step 4: centreline. Using a minimal path algorithm,[7] the centrelines of the main branches of the aorta (main trunk and iliac arteries) are computed based on both lumen and thrombus. This step is subtle, and crucial, as all subsequent measurements depend on it. Centrelines should be well positioned and smooth.

Step 5: diameter measurement. The current standard measurement in clinical practice is to measure the aortic diameter perpendicular to the centreline at the point where it is the largest. To replicate clinicians, ARVA measurements are made on the computed centreline points. For each point, the diameter of the intersection between the segmentation and the normal plane to the tangent of the centreline is measured. The maximum aortic diameter is given by the maximum of the measurements at each point. In the vicinity the iliac bifurcation, a circularity constraint is applied. If the intersection of the segmentation with the normal plane violates this constraint, another point is chosen for the measurement following the centreline in a downward direction.

Statistics

Dice score to evaluate the segmentation. A common metric used to evaluate the segmentation performance between ground truth and predicted masks is the Dice Similarity Coefficient. It is equal to twice the number of pixels in common between the ground truth mask (green) and the predicted mask (red) divided by the sum of the number of pixels in each mask (Fig. 3). Therefore, a Dice score can range between 0 (no overlap between ground truth and predicted masks) and 1 for a perfect overlap. To assess the overall segmentation performance over the entire validation set, the average Dice is reported.

While the Dice score is widely used in the data science community, it does not fully represent clinical reality as medical decisions are based on the maximum aneurysm diameter. A high Dice score does not ensure an accurate measurement. Thus, the segmentation might be precise everywhere in the aorta except in the aneurysm sac where measurements are taken. This example leads to a high Dice score but to incorrect diameters. Therefore, maximum diameter measurement error is reported.

Measurement evaluation. The absolute error between the measurements made by clinicians and the measurements made automatically by ARVA were evaluated. As the subsets were relatively small, the 25th, 50th, and 75th percentiles of the absolute error distributions were recorded.

Training settings. The pipeline was implemented in Python using Keras DL library with the Tensorflow backend. Neural networks were trained on graphics processing units (NVIDIA Tesla V100) with 16 GM memory.

RESULTS

The results presented in this section were obtained from the validation dataset composed of 62 scans after training the algorithmic pipeline with the 489 scans of the training dataset.

Aorta segmentation

The aortic segmentation model achieved mean Dice scores of 0.84, 0.95, and 0.93, respectively, on healthy aortas, diseased

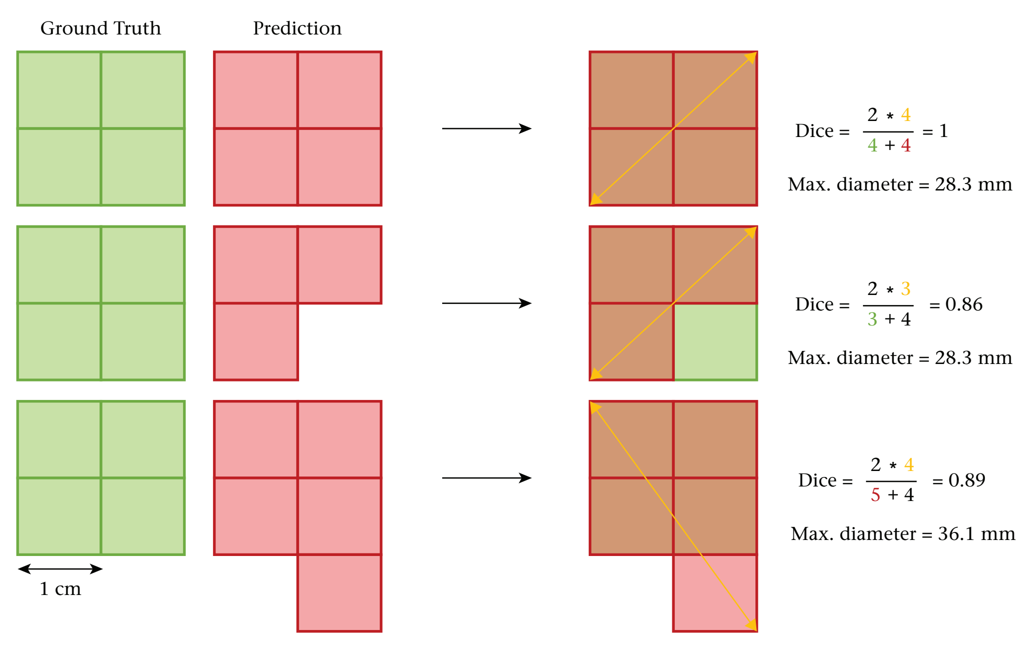

Figure 3. The Dice score is equal to twice the number of pixels in common between the ground truth mask (green) and the predicted mask (red) divided by the sum of the number of pixels in each mask. As illustrated, a high Dice score does not ensure an accurate diameter measurement.

aortas, and diseased aortas after endovascular treatment. (On healthy patients, the validation database contained mostly scans in the portal phase, highly affecting segmentation of the iliac arteries. Mean Dice scores were thus lower without affecting maximum cross sectional diameter measurement, which is not in the iliac region in this subset.)

Diameter measurements

Maximum cross sectional diameter measurements of two experienced aortic surgeons (experts 1 and 2), three vascular surgery residents (residents 1, 2, and 3), two general radiologists (radiologists 1 and 2), and ARVA on the validation dataset were compared (Fig. 4).

The median absolute difference with respect to the first expert is reported in Table 2. These ranged from 1 mm to 2 mm, while ARVA, on the validation set, reported 1.2 mm. When focusing on the different subsets, (1) on healthy aortas the median absolute difference with respect to the first expert ranged from 0.9 mm to 1 mm between the annotators, while ARVA reported a median absolute error of 0.8 mm; (2) on diseased aortas, the median absolute difference with respect to the first expert ranged from 1 mm to 2 mm between the annotators, while ARVA reported a median absolute error of 1.9 mm; and (3) on stented aortas, the median absolute difference with respect to the first expert ranged from 1 mm to 3 mm between the annotators, while ARVA reported a median absolute error of 2.2 mm.

Bland—Altman plots representing the difference and the absolute between expert 1 and ARVA (Fig. 4) showed that, in most cases, ARVA was within the variability of the annotators thus providing suitable measurements for clinical practice (see Fig. 5, examples A and B). In a few cases, however, ARVA measurements were outside acceptable ranges (outliers).

Table 2. **Inter-annotator variability of clinicians and Augmented Radiology for Vascular Aneurysm (ARVA) performance for the measurement of the maximum aortic diameter in millimetres against expert 1 presented as 25th, 50th, and 75th percentiles of the distribution of the measurements absolute error**

	Healthy aortas (n = 20)			Diseased aortas (n = 20)			Aortas with stent (n = 22)			All (n = 62)		
	25th	50th	75th	25th	50th	75th	25th	50th	75th	25th	50th	75th
Expert 1	–	–	–	–	–	–	–	–	–	–	–	–
Expert 2	0.0	1.0	1.0	0.6	1.1	1.8	0.0	1.0	3.5	0.0	1.0	1.9
Resident 1	0.0	1.0	1.2	1.0	1.0	3.2	1.0	2.0	3.0	0.0	1.0	2.0
Resident 2	0.8	1.0	2.0	1.0	2.0	4.5	1.8	3.0	4.0	1.0	2.0	3.0
Resident 3	0.5	0.9	1.5	0.8	2.0	2.7	1.0	1.8	2.5	0.7	1.4	2.2
Radiologist 1	0.8	1.0	1.0	1.0	1.5	3.5	1.0	1.0	2.8	1.0	1.0	2.0
Radiologist 2	1.0	1.0	2.0	1.0	2.0	3.2	1.0	2.0	4.5	1.0	2.0	3.0
ARVA	0.3	0.8	1.3	0.8	1.9	3.5	0.6	2.2	4.8	0.5	1.2	3.2

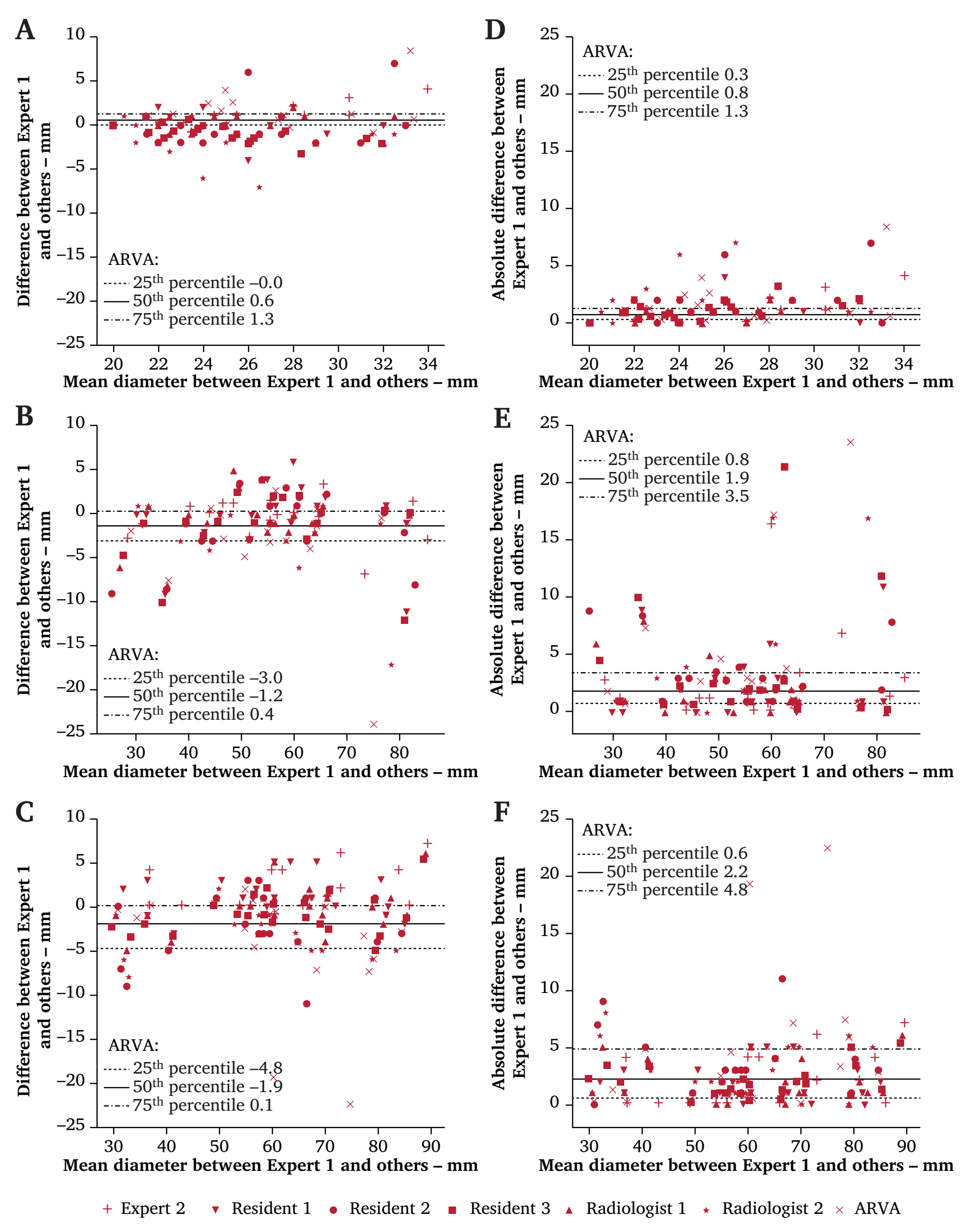

Figure 4. Bland–Altman plots representing the difference between expert 1 and others and Augmented Radiology for Vascular Aneurysm (ARVA) for (A) healthy aortas, (B) diseased aortas, and (C) aortas with stent; and the absolute difference between expert 1 and others and ARVA for (D) healthy aortas, (E) diseased aortas, and (F) aortas with stent. Dashed, solid and dash–dot lines represent the 25th, 50th, and 75th percentiles for ARVA, respectively.

These outliers were easily detected by clinicians on the provided automatic report, which includes detailed visualisation of the aortic segmentation (see Fig. 5, example C).

DISCUSSION

Outer to outer wall aortic diameter measurement is a complex task that requires a high degree of precision to make appropriate clinical decisions. A large variety of methods addressing automatic aortic segmentation have been proposed in the past,[8] but only a limited number address the difficult task of automatic external wall diameter measurements.[9,10] The ARVA algorithm achieved a median diameter measurement error of 1.2 mm with expert1 on the validation set, while the two experts had a median error of 1 mm.

Aortic segmentation algorithms

Comparing previous work on automatic aortic segmentation is challenging as no public database or metrics establishing

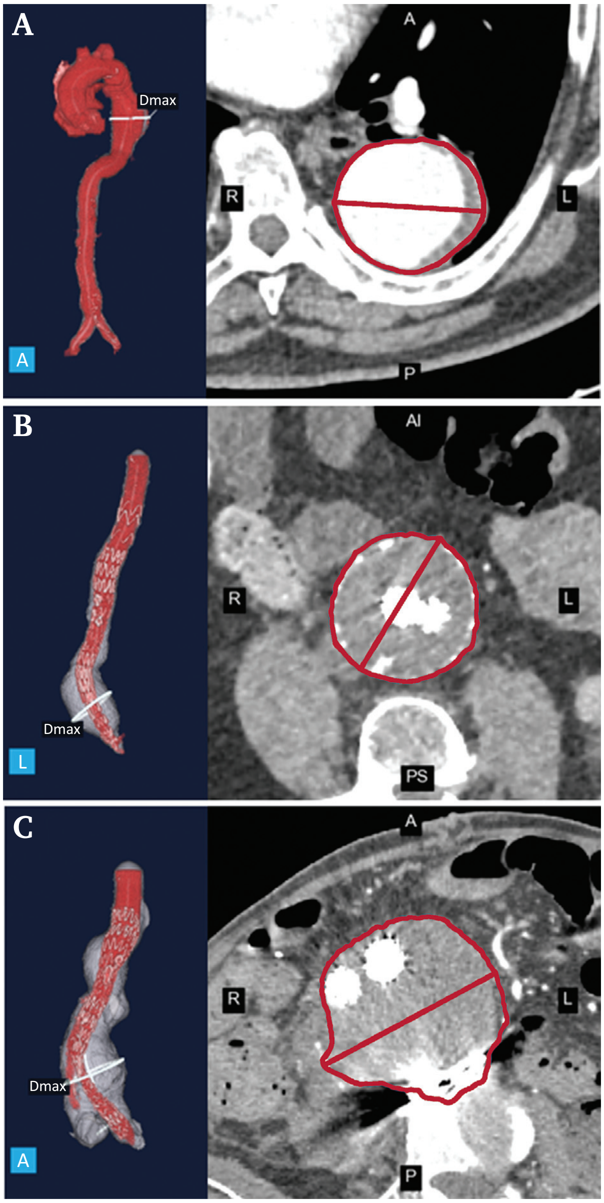

Figure 5. Examples of accurate measurements made by ARVA on cross sections of (A) thoracic aneurysm and (B) endovascular aneurysm repair (EVAR), and (C) of an outlier of EVAR, where incorrect segmentation contour is easily detected. Dmax = maximum cross sectional diameter localisation on automatically segmented aorta for corresponding cross section.

an evaluation protocol are available. Working datasets differ in various characteristics, namely database size, presence or absence of endografts, aneurysms, and contrast media. More importantly, most studies focus on the abdominal aorta alone.

In this study, a complete aortic analysis from the ascending aorta to the iliac arteries was performed, thus including various angulations and high anatomical variability making segmentation and diameter measurements more challenging.

Table 3 summarises previously published papers and ARVA performances.

Many of the published studies involve solutions based on classic computer vision techniques relying on intensity and/or shape constraints. These methods often require user intervention and are very sensitive to their initial parameterisation, making them less generalisable.

Lareyre *et al.*[11] proposed to segment the infrarenal aorta by sequentially segmenting two dimensional axial images. This method is based on rather simple computer vision models[12] on which ad hoc thresholding and or morphological steps are added. Limitations can be anticipated with suboptimal contrast injection or in presence of any type of artefacts, inducing high intensity patterns. The authors reported a Dice score of 0.88 ± 0.12 on a limited number of slices of the infrarenal aorta (34 pre- and six post-operative scans), without describing their selection criteria. Standardisation of evaluation is mandatory to overcome these limitations.

Also using variational optimisation methods, Lalys *et al.*[13] tackled the aortic segmentation challenges using a three dimensional shape based approach. Overall, they managed to treat a broader spectrum of CT scans (120 pre- and 25 post-operative) and anatomies (abdominal aorta and iliac arteries), with state of the art results: Dice scores of 0.86 ± 0.06 and 0.81 ± 0.06, respectively, for pre-operative abdominal aorta and iliac analysis, and a Dice score of 0.87 ± 0.03 for post-operative scans of the abdominal aorta. Manually tuned features and shape based deformable models can, nevertheless, lack precision in the presence of unusual anatomies or artefacts. The most important drawback of this study is the user interaction required.

Classic approaches with manual features and handmade models rely only on the operator's good understanding of the segmentation task without leveraging data. On the contrary, DL techniques have demonstrated their potential for this task. To the authors' knowledge, none of the classical approaches have so far performed segmentation of the complete aorta; published series have only focused on the abdominal segment. Lu *et al.*[14] proposed a three dimensional V-Net combined with ellipse fitting, to perform aortic segmentation and aneurysm diameter measurements. They trained a network on a dataset of 321 scans, with both contrast enhanced and non-injected CT scans, and achieved an average Dice score of 0.89 ± 0.05 with contrast enhanced CT scans. However, this work only evaluated pre-operative CT scans. Using the same type of network combined with data modelling, Caradu *et al.*[9] recently reported a Dice score of 0.95 ± 0.01 for the segmentation of pre-operative infrarenal AAAs. Their method provides a median absolute measurement error of the maximum infrarenal aortic diameter of 1.8 mm. ARVA achieves a median absolute error of 1.2 mm for the entire aorta on both pre-operative and post-operative scans. Finally, López-Linares *et al.*[10] proposed tackling both pre- and post-operative CT scans, and trained a pipeline of neural networks on 50 CT scans (22 pre-operative, 28 post-operative). With their method, they reported a Dice score of 0.87 ± 0.062 on 28 CT scans (12 pre-operative, 16 post-operative) with an absolute axial diameter measurement error of

Table 3. Performance comparison of methods for aortic segmentation and maximum aortic diameter measurements in earlier studies and by Augmented Radiology for Vascular Aneurysm (ARVA)

Reference	Aortic section	Segmentation performance ± SD (Dice)		Measurement performance, absolute error – mm
		Pre-operative	Post-operative	
Lareyre *et al.*[11]	Abdominal aorta	0.88 ± 0.12 (40 CT scans)		–
Lalys *et al.*[13]	Abdominal Aorta	0.86 ± 0.06	0.87 ± 0.03 (25 CT scans)	–
	Iliac sections	0.81 ± 0.06 (120 CT scans)	–	
López-Linares *et al.*[10]	Abdominal aorta	0.84 ± 0.068 (12 CT scans)	0.89 ± 0.044 (16 CT scans)	Mean absolute error 3.309 ± 6.03 mm. $n = 28$ manual measurements. Axial diameters
Lu *et al.*[14]	Abdominal aorta	0.89 ± 0.05 (57 CT scans)	–	No absolute error reported
Caradu *et al.*[9]	Abdominal aorta	0.95 ± 0.01 (100 CT scans)	–	Median absolute error 1.8 mm. $n = 100$ manual measurements made by two clinicians
ARVA	Thoracic aorta Abdominal aorta Iliac sections	Diseased aortas: 0.95 Healthy aortas (portal phase): 0.84 (40 CT scans)	0.93 (22 CT scans)	Median absolute error 1.2 mm. $n = 62$ manual measurements made by 7 clinicians

SD = standard deviation; CT = computed tomography.

3.309 mm and a standard deviation of 6.03 mm. However, axial measurements do not answer clinical need, as it is necessary to measure the diameter in the plane perpendicular to the centreline, in the way ARVA does.

Potential clinical impact of ARVA

In patients with aortic disease, imaging of the aorta plays a key role in determining timing of surgical intervention and during follow up after open or endovascular repair. As each aortic segment evolves and remodels independently before and after the repair, ARVA will provide automatic monitoring of these segments independently and detect focal reductions (occluded sac) and/or focal enlargements of the aortic diameter.

Many studies have reported interobserver variability in aortic measurements in clinical practice. Tortuous aortic anatomy is the main reason for such variability because overestimation of aortic aneurysm size is frequent when performing axial measurements (which often pass obliquely through the aorta). Standardisation of aortic measurements with MPRs are thus required (oriented perpendicular to the axis of each aortic segment) for accurate measurements.[15] Of note, the maximum diameter is a valid clinical parameter only for aneurysms, before or after treatment. In non-aneurysmal aortas, the maximum diameter is located in a non-diseased, often straight, segment and thus more reproducible to measure. Moreover, the diameter of a normal aorta is more constant over a small segment, so the maximum diameter can be measured at more locations around the true maximum diameter than that in case of an aneurysm, where there will be a very specific site with the absolute maximum diameter. Being off just a little but along the longitudinal axis or the MPR plane in the aneurysmal aorta creates more error. This is the reason why there is more variability in measuring pathological aortas.

MPR standardised measurement protocols can decrease measurement variability, in order to avoid reporting size changes that may represent measurement errors rather than true aneurysm evolution or regression. This is critical, as indication for treatment will be connected directly to maximum diameter size or growth rate (for the index repair or secondary interventions during follow up). Such standardised protocols reduce interobserver measurement differences,[16] but do not eliminate them. In addition, accurate measurements of the various aortic segments and their comparison with previous CT examinations are time consuming. Although mandatory, these multiple measurements cannot be performed during a busy consultation. ARVA automatically forwards a report approximately 10 minutes after acquisition, allowing the clinician to access it while opening the exam for interpretation. Fully automatic measurements reduces interobserver variability to a minimum with a standard measurement protocol, and instantly provides accurate data on diameter evolution, in each segment of the aorta. Future developments will include automatic aortic volume assessment. Volume data are not currently analysed in everyday practice,[17] because of the technical challenge associated with its evaluation and mostly because of the time required to evaluate it accurately. Future clinical evaluations will determine if it could have a critical role in evaluating rupture risk and enhance follow up.

In the European Society for Vascular Surgery guidelines,[18] ultrasound is considered the gold standard for measurement of AAA diameter to determine the threshold for elective intervention. The use of CT scan measurement will not influence the decision to operate in most centres but will be used for surgical planning once the threshold has been met on ultrasound. Furthermore, post-EVAR surveillance is commonly done by duplex with selective CT scans for patients with endoleaks or sac expansion. In the future, if three dimensional ultrasound becomes routine practice, it will probably benefit from ARVA technology. Nowadays, thoracic (ascending, arch, and descending) and thoraco-abdominal aneurysm surveillance before and or after open or endovascular repair requires CT or magnetic resonance imaging. The clinical benefit of ARVA for this later cohort of patients needs be determined.

Limitations

ARVA has some limitations. Firstly, although this evaluation is one of the largest reported datasets in the literature, it is composed of only 62 scans and therefore it will need to be further increased in size while accounting for the variety of suppliers (Table 1). Secondly, performances were diminished on portal phase scanners. Portal phase scans were used for controls who underwent CT scans for non-vascular purposes, and unusually an arterial phase was performed in these patients free of aortic pathology. As this situation is also a source of variability for human annotators, it is anticipated that a trained AI classifier to alert the user of an increased risk of measurement errors will be used. However, this issue explains the poor Dice score concerning controls. Finally, the algorithm can output outlier measurements that are out of the variability range of annotators. Fortunately, these gross errors are very rare. Moreover, they are easily identified by the clinician by looking at the segmentation mask provided with the report. As ARVA is intended to help the clinician and not to replace him or her, these outliers will be easily eliminated by human proofreading.

Conclusion

ARVA, a fully automatic pipeline for maximum aortic diameter, and cross sectional outer to outer aortic wall measurements, achieves a performance level suitable for clinical practice on a validation database reflecting the daily caseload of a vascular surgeon/radiologist. Further accuracy improvements are expected by leveraging a large amount of remaining unsegmented images in the database. Future improvements will include automatic measurements of the maximum external diameter and volume for each aortic segment.

CONFLICTS OF INTEREST

C. Adam, R. Ardon, and G. d'Assignies are employees of Incepto Medical. Incepto Medical and Hôpital Marie Lannelongue and Groupe Hospitalier Paris Saint Joseph are bound by a co-creation contract.

FUNDING

None.

REFERENCES

1 Norman PE. Population based randomised controlled trial on impact of screening on mortality from abdominal aortic aneurysm. *BMJ* 2004;**329**:1259.

2 Sakalihasan N, Limet R, Defawe OD. Abdominal aortic aneurysm. *Lancet* 2005;**365**:1577–89.

3 Moll FL, Powell JT, Fraedrich G, Verzini F, Haulon S, Waltham M, et al. Management of Abdominal Aortic Aneurysms Clinical Practice Guidelines of the European Society for Vascular Surgery. *Eur J Vasc Endovasc Surg* 2011;**41**:S1–58.

4 Mory B, Ardon R, Yezzi AJ, Thiran J-P. Non-Euclidean image-adaptive Radial Basis Functions for 3D interactive segmentation. In: *2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan*; 2009. p. 787–94.

5 Milletari F, Navab N, Ahmadi S-A. V-Net: fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA*; 2016. p. 565–71.

6 Mory B, Ardon R. Fuzzy region competition: a convex two-phase segmentation framework. In: Sgallari F, Murli A, Paragios N, editors. *Scale Space and Variational Methods in Computer Vision. SSVM 2007. Lecture Notes in Computer Science*, Vol. 4485. Berlin, Heidelberg: Springer; 2007. p. 214–26.

7 Chen D, Zhang J, Cohen LD. Minimal paths for tubular structure segmentation with coherence penalty and adaptive anisotropy. *IEEE Trans Image Process* 2019;**28**:1271–84.

8 Raffort J, Adam C, Carrier M, Ballaith A, Coscas R, Jean-Baptiste E, et al. Artificial intelligence in abdominal aortic aneurysm. *J Vasc Surg* 2020;**72**:321–33.

9 Caradu C, Spampinato B, Vrancianu AM, Bérard X, Ducasse E. Fully automatic volume segmentation of infra-renal abdominal aortic aneurysm CT images with deep learning approaches versus physician controlled manual segmentation. *J Vasc Surg* 2021;**74**:246–56.

10 López-Linares K, García I, García-Familiar A, Macía I, Ballester MAG. 3D convolutional neural network for abdominal aortic aneurysm segmentation. Available at: https://arxiv.org/abs/1903.00879 [Accessed 23 July 2021].

11 Lareyre F, Adam C, Carrier M, Dommerc C, Mialhe C, Raffort J. A fully automated pipeline for mining abdominal aortic aneurysm using image segmentation. *Sci Rep* 2019;**9**:13750.

12 Chan TF, Vese LA. Active contours without edges. *IEEE Trans Image Process* 2001;**10**:266–77.

13 Lalys F, Yan V, Kaladji A, Lucas A, Esneault S. Generic thrombus segmentation from pre- and post-operative CTA. *Int J Comput Assist Radiol Surg* 2017;**12**:1501–10.

14 Lu J-T, Brooks R, Hahn S, Chen J, Buch V, Kotecha G, et al. DeepAAA: clinically applicable and generalizable detection of abdominal aortic aneurysm using deep learning. Available at: https://arxiv.org/abs/1907.02567 [Accessed 23 July 2021].

15 Mendoza DD, Kochar M, Devereux RB, Basson CT, Min JK, Holmes K, et al. Impact of image analysis methodology on diagnostic and surgical classification of patients with thoracic aortic aneurysms. *Ann Thorac Surg* 2011;**92**:904–12.

16 Asch FM, Yuriditsky E, Prakash SK, Roman MJ, Weinsaft JW, Weissman G, et al. The need for standardized methods for measuring the aorta. *JACC Cardiovasc Imaging* 2016;**9**:219–26.

17 Paraskevas KI, Torella F, Swaelens C, England A, Chan TY, Shaikh U, et al. Temporal changes in intraluminal thrombus volume within abdominal aortic aneurysms: implications for planning endovascular aneurysm sealing. *J Endovasc Ther* 2018;**25**:47–51.

18 Wanhainen A, Verzini F, Herzeele IV, Allaire E, Bown M, Cohnert T, et al. Editor's Choice – European Society for Vascular Surgery (ESVS) 2019 Clinical Practice Guidelines on the Management of Abdominal Aorto-iliac Artery Aneurysms. *Eur J Vasc Endovasc Surg* 2019;**57**:8–93.