Clinical validation of a software for quantitative follow-up of abdominal aortic aneurysm maximal diameter and growth by CT angiography

Claude Kauffmann[a,*], An Tang[b,1], Alexandre Dugas[a,2], Éric Therasse[c,3], Vincent Oliva[b,4], Gilles Soulez[b,5]

[a] Department of Medical Imaging, Hôpital Notre-Dame, Centre Hospitalier Universitaire de Montréal, 1560 Sherbrooke Est, Montréal, Québec, Canada H2L 4M1

[b] Department of Medical Imaging, University of Montreal, Hôpital Saint-Luc, Centre Hospitalier Universitaire de Montréal, 1058 rue Saint-Denis, Montréal, Québec, Canada H2X 3J4

[c] Department of Medical Imaging, University of Montreal, Hôpital Hôtel-Dieu, Centre Hospitalier Universitaire de Montréal, 3840 rue Saint-Urbain, Montréal, Québec, Canada H2W 1T8

Received 11 May 2009
Received in revised form 21 July 2009 Accepted 22 July 2009

Keywords:
Abdominal aortic aneurysm
Quantitative analysis
D-max follow-up
3D segmentation CTA

Corresponding author. Tel.: +1 514 890 8000x26131; fax: +1 514 412 7547.

E-mail addresses: claude.kauffmann@gmail.com (C. Kauffmann), duotango@gmail.com (A. Tang), alexandre.dugas@elf.mcgill.ca (A. Dugas), eric.therasse.chum@ssss.gouv.qc.ca (É. Therasse), vincent.oliva.chum@ssss.gouv.qc.ca (V. Oliva), gilles.soulez.chum@ssss.gouv.qc.ca (G. Soulez).

1 Tel.: +1 514 890 8000x28530/36400; fax: +1 514 412 7547.

2 Fax: +1 514 412 7547.

3 Tel.: +1 514 890 8150; fax: +1 514 412 7547.

4 Tel.: +1 514 890 8000x28530; fax: +1 514 412 7547.

5 Tel.: +1 514 890 8000x26522; fax: +1 514 412 7547.

Purpose: To compare the reproducibility and accuracy of abdominal aortic aneurysm (AAA) maximal diameter (D-max) measurements using segmentation software, with manual measurement on double- oblique MPR as a reference standard.

Materials and methods: The local Ethics Committee approved this study and waived informed consent. Forty patients (33 men, 7 women; mean age, 72 years, range, 49–86 years) had previously undergone two CT angiography (CTA) studies within 16 ± 8 months for follow-up of AAA ≥35 mm without previous treatment. The 80 studies were segmented twice using the software to calculate reproducibility of automatic D-max calculation on 3D models. Three radiologists reviewed the 80 studies and manually measured D-max on double-oblique MPR projections. Intra-observer and inter-observer reproducibility were calculated by intraclass correlation coefficient (ICC). Systematic errors were evaluated by linear regression and Bland–Altman analyses. Differences in D-max growth were analyzed with a paired Student's t-test. Results: The ICC for intra-observer reproducibility of D-max measurement was 0.992 (≥0.987) for the software and 0.985 (≥0.974) and 0.969 (≥0.948) for two radiologists. Inter-observer reproducibility was 0.979 (0.954–0.984) for the three radiologists. Mean absolute difference between semi-automated and manual D-max measurements was estimated at 1.1 ± 0.9 mm and never exceeded 5 mm.

Conclusion: Semi-automated software measurement of AAA D-max is reproducible, accurate, and requires minimal operator intervention.

1. Introduction

Abdominal aortic aneurysm (AAA) therapeutic management, either surgical or endovascular and follow-up, currently rely on measurement of the maximal diameter (D-max) [1,2]. While AAA volumetric analysis is an area of active research, D-max remains a widely accepted clinical tool [3,4]. Although ultrasound is aradiation free, widely available modality for D-max screening and follow-up, it is not as accurate as computed tomographic (CT) angiography for measurement of aneurysm diameter [5,6]. Because of its volumetric acquisition mode suitable for multiplanar and three-dimensional reconstructions of complex anatomy, as well as ability to distinguish between lumen and thrombus, there is an inherent benefit to use CT angiography for modeling of aneurysms. In addition to standard measurements of diameter, length and angulation [7] used for patient selection before EVAR, three-dimensional reconstructions allows the study of thrombus volume [8] and predictors of risk rupture such as expansion rate of aneurysmal volume [3].

Validation of a novel segmentation method requires comparison against established D-max measurement standards. A recent study [9] directly compared different methods of AAA D-max measurement and has shown that the manual double-oblique MPR method is theoretically closest to reality and provides the lowest inter- and intra-observer variability and is theoretically closer to reality than axial and orthogonal multiplanar reformations,

which both tend to overestimate the real diameter perpendicular to the centerline.

An ideal software method is commonly associated with auto- mated segmentation. However, fully automated segmentation of AAA generally fails because of low contrast between thrombus and surrounding structures such as psoas, bowel loops or unopacified inferior vena cava. With this approach, the solution is to correct segmentation errors on several hundred slices, a tedious and time- consuming task, incompatible with clinical workflow. For these reasons, a semi-automated or supervised method is more appropriate in clinical practice because segmentation errors can be avoided by introducing user input at specific steps during the segmentation process. Conceptual image understanding is used to initialize algorithmic tasks by the computer. The success of this approach lies in the complementarity between operator and machine tasks.

While published studies have addressed some issues, such as a proof-of-concept of automated D-max calculation [10] or aneurysm volume calculation [11–14], to our knowledge, no clinical validation of an integral solution has been reported for D-max measurement obtained from complete segmentation of AAA wall, lumen, thrombus and calcification.

Thus, the purpose of our study was to develop a semi-automated software for AAA segmentation on MDCT examinations and assess its reproducibility and accuracy to determine AAA maximal diameter and its progression in comparison with manual maximal diameter measurements on double-oblique MPR as the reference standard.

2. Materials and methods

2.1. Patients

Institutional review board approved this Health Insurance Portability and Accountability Act (HIPAA) compliant study. We performed a retrospective study of 40 patients with abdominal aortic aneurysm (AAA) followed by multi-detector computed tomography (MCDT). Patients were selected from the radiological information system. These patients were contacted by a research nurse and approval for the use of radiological imaging was obtained through patient's written consent. Between 2004 and 2006, 40 patients with AAA more than 35 mm, having at least two MDCT studies available on the local PACS with a minimal 6-month inter- val between exams were enrolled. If a patient had had more than two MDCT studies, the most remote and most recent exams were selected. A total of 80 exams were therefore analyzed in this study.

2.2. MDCT protocol

All 80 examinations were performed on 4 multi-detector CT scanners (Somatom Sensation 4, 16, 64, Siemens, Erlangen, Germany; Lightspeed 16, GE, Milwaukee, WI). The scanning parameters were the following: pitch 1–1.5, slice thickness 1–4 mm, collimation 0.75 and field of view 240–320. 79/80 exams were performed with a non-ionic contrast agent (iodine concentration 320–350 mg/ml) injected through an antecubital vein at 3–5 ml/s for total of 80–120 ml. The timing of acquisition was determined by an automatic bolus trigger positioned at the level of the thoraco- abdominal aorta.

2.3. Software measurement method

A software was written in *Interactive Data Language* (IDL) and C++ language to extract and quantify the volumetric component of AAA, distinguishing between lumen, thrombus, and calcifications. The task of the user is to interact with the software in order to segment the boundaries of the aneurysm on longitudinal reformations in a semi-automated process (Figs. 1 and 2). The pro- posed method has been previously reported in detail and consists of the following: (1) user identification of AAA lumen entry and exit points; (2) software calculation of 3D lumen; (3) creation of a curved-MPR following a luminal path with minimization of curvature; (4) automated aneurysm wall segmentation on 4–8 radial MPR reformations along the path axis initialized by the operator with an active contour based process; and (5) interactive contour editing on the same radial MPR reformations may be performed by the user, if needed. Once the segmentation was approved by the user, (6) a centerline based on the outer wall of the AAA and a 3D mathematical model of the AAA with distinct display for the thrombus and lumen were reconstructed and automatic calculation of D-max perpendicular to the new central line was processed. All CT examinations were anonymized and processed by an experimented CT technologist blinded to the radiology report. The time required to run this entire process (AAA segmentation and D-max calculation) on an IBM PC Pentium 4, CPU: 3.4 GHz, 2 Gb RAM, was recorded.

2.4. Manual measurement method

Axial images and multiplanar reformations (MPR) of the axial images were rendered and evaluated independently by two senior vascular and interventional radiologists and one junior staff, blinded to radiological reports. All diameters were measured from the aneurysm outside wall, using electronic calipers, with zooming and windowing liberally performed when judged pertinent on the same workstation (Impax, version 5.2; Agfa, Mortsel, Belgium).

The maximal diameter (D-max) was measured on a double- oblique reformation (DO) by determining a plane perpendicular to the aneurysm and the line of flow on the sagittal MPR, then on the coronal MPR, thus creating a “modified axial” plane. Maximal diameter of the aneurysm on this double-oblique plane was then measured (Fig. 3).

2.5. Repeat measurement and duration

In order to calculate intra-observer reproducibility of both methods, the same technologist and two of the three radiologists (one senior and one junior radiologist), all blinded to the results of the first reading session, independently repeated the D-max measurements on the 80 exams by the software and manual DO methods, respectively. Repeated readings were done with a 4-week minimal interval, using an identical protocol. The time required to measure the D-max manually was recorded during the second reading session.

2.6. Statistical analysis

2.6.1. Patient demographics

Descriptive statistics of patient baseline demographics, interval between D-max at baseline and follow-up, mean AAA D-max (all observers and two different methods) at baseline and follow-up were calculated.

2.6.2. Intra-observer and inter-observer reproducibility

The level of agreement for D-max between repeated measurements (reading sessions 1 and 2) by software method and by manual DO method was calculated by estimation of intra-observer intraclass correlation (ICC). The agreement between radiologists was also estimated by the inter-observer ICC for sessions 1 and 2. Values of up to 0.40 were considered to indicate positive but poor agreement; 0.41–0.60 good agreement; 0.61–0.80 very good agreement; and greater than 0.80 excellent agreement [15].

2.6.3 Intra-observer and inter-observer reproducibility

For the validity analysis, linear regression and Bland–Altman analysis were used to assess agreement between the two (software and manual DO) methods of measurement. Linear regression analysis was performed separately for measurement taken on baseline and follow-up examinations. Means of the two readings (sessions 1 and 2) were calculated for the software and for each of the two radiologists (1 and 3). The 95% CI for the slope and intercept are reported. If the slope of the line is close to unity and the intercept close to zero, this implies that the two methods of measuring D-max are in agreement. In conjunction with regression, Bland–Altman [16] range of agreement was also reported to support the conclusion of linear regression. The range of agreement was defined as the bias ± 2 SD, where SD is the corrected standard deviation of the differences between the two methods.

2.6.4 Responsiveness

To assess the responsiveness of the software method, that is, the ability to detect changes over time, a paired Student's *t*-test was used to compare D-max measurements taken at baseline with these taken at follow-up.

2.6.5 *Measurement time*

Descriptive statistics of measurement time by manual and semi-automated methods were calculated. The statistical analysis was performed by using a software package (SAS 9.1 for Windows).

3. Results

3.1. Patient demographics

Forty patients (33 men, 7 women; mean age, 72 years, range, 49–86 years) were included in our study. The interval between the two MDCTs at baseline and follow-up was 16 ± 8 months (mean ± SD), range 8–42 months. No patient was excluded from software segmentation for technical reasons.

The AAA diameters at baseline and follow-up are summarized in Table 1. Regardless of the measurement method and observer, the smallest aneurysm had a D-max of 35.8 mm and the largest 76.8 mm.

3.2. Reliability

Intra-observer reproducibility's for D-max measurement were excellent for the software and the two radiologists (1 and 3) with repeated measurements by manual double-oblique MPR method: 0.992 (≥0.987), 0.985 (≥0.974) and 0.969 (≥0.948), respectively (Table 2). It was significantly higher for the software when compared to observer 3 ($P < 0.05$). The D-max difference between the first and second reading sessions was not significant: −0.12 ± 0.86 mm ($P > 0.38$) for baseline examinations and 0.22 ± 1.14 mm ($P > 0.23$) for follow-up examinations. Inter-observer reproducibility (Table 3) for manual measurement of D-max by radiologists was excellent: 0.979 (0.954–0.984) at baseline and 0.975 (0.955–0.985) at follow-up.

3.3. *Assessing agreement between the two methods*

The slope and intercept for the regression model between soft- ware and manual measurements indicate that the estimates of the slope and intercept are very close to unity and zero. For the analysis of the mean of readings 1 and 2 between software and radiologist #1 at baseline (Fig. 4), the intercept was −0.490 (95% CI4.303–3.322) and slope equal to 1.022 (95% CI 0.947–1.097).

In a supportive manner, using a clinically meaningful limit of ≤5 mm, software user and radiologist #1 were within 4 mm in 40/40 instances, or interchangeable, 100% of the time for the mean of all reading on baseline (Fig. 5).

Similar results were obtained for all other analysis with the exception of the follow-up scan of radiologist #3. Although the scatter plot of the two measurements line up closely to the line of identity, only 92.5% (37/40) of the differences between the two methods were within the defined limit of agreement of 4 mm. How- ever, there was no difference of more than 5 mm.

3.4. *Responsiveness*

Statistically significant D-max growth between baseline and follow-up examinations of 4.2 ± 3.2 mm (first reading session) and 4.5 ± 3.5 mm (second reading session) ($P < 0.0001$), respectively, were observed.

3.5. *Measurement time*

Average measurement time was 104.7 ± 24.9 s for the manual method limited to double-oblique D-max and 175.2 ± 100.9 s for the semi-automated method, including the entire AAA segmentation, of which D-max was calculated for clinical validation.

4. Discussion

The purpose of this study was to demonstrate that reproducible quantitative follow-up of AAA D-max by MDCT can be efficiently addressed by a software with minimal human intervention, limited to correction of the AAA segmentation generated by the software, only when needed. In contrast to prior methods for AAA segmentation at CT angiography, our method is the first one that clinically validates automated D-max calculation following complete wall, lumen and thrombus segmentation. The high reproducibility of the semi-automated method combined to its accuracy (1 mm mean error difference with manual measurements), makes this method valid for clinical use. The absence of measurement error higher than 5 mm and the ability to detect a 5 mm growth between studies with 100% confidence are important criteria to assume that clinical decision made by the semi-automated method are sound. Since it can be run by a CT technologist, it could assist a radiologist or a vascular surgeon in determining D-max and its progression over time in a reproducible way. Manual comparison of D-max between multiple studies can be tedious. Significant variability of D-max measurement has been reported previously with variation of more than 5 mm in 17% of patients [5]. Without standardization of the measuring process, Cayne et al. reported a mean variation between observers of 4 ± 5.1 mm [17]. In the same study, after standardization, a mean variability of 2.8 ± 4.4 mm was still observed [5,17]. Furthermore, multiple approaches have been proposed when measuring D-max manually. Some investigators recommend measurement of antero-posterior and transverse diameter on axial images [18], others have reported lower variability when measuring shorter axis on axial images [19]. Finally, lower variability and better inter-observer correlation were found with a standardized approach and a measurement perpendicular to the central line [9,17]. The semi-automated approach proposed in our study provides a highly reproducible D-max [9] measurement on a double-oblique plane perpendicular to the central line and pre- vent variation due to the methodology used by the radiologist or its interpretation.

To our knowledge, only one study [10] automatically calculated a D-max from AAA segmentation. In that study automated D-max computation and manual D-max measurements were performed in the same image plane (selected by automatic curved-MPR or manual double-oblique reformation) were compared. The D-max difference between the two methods was 0.342 ± 0.245 cm. However, since the study was limited to 4 patients, no valid statistical conclusion could be made.

The main limitation of AAA segmentation is the segmentation of the peripheral thrombus whose density is close to surrounding structures. Several segmentation methods based on exclusive or hybrid combinations of level-sets, geometric deformable mod- els, active contours or active shape models [11,12,14] have been described for AAA segmentation. The performance of these algorithms depends on the parameter settings obtained empirically. These algorithms either perform 2D (slice-by-slice) or 3D AAA seg- mentation. The output is AAA global volume which is compared to AAA volume obtained by manual slice-by-slice aneurysm delineation. Since D-max measurement is the recognized gold standard for AAA diagnosis, treatment and follow-up, the validation of our software was based on D-max measurement. Additional evaluation and validation of volume measurement will be performed in subsequent studies.

Use of minimal curvature path-based image reformation, and in particular for selection of radial planes along the path axis, was a straightforward and highly reproducible method. Its intra-observer reproducibility was better than manual measurements. The correction process by interactive contour editing is also original. This feature could explain the robustness and accuracy of this method compared to manual measurements. No published work details solutions to correct the segmentation results if the process failed, a potentially disruptive outcome in clinical workflow.

Double-oblique measurement was chosen as a reference because we have previously found it was more reproducible than measurements taken on axial slices [9]. Furthermore, this mea-surement is closer to reality than axial measurements that can be influenced by the obliquity of the aorta [9]. In our study, a small but noticeable reproducibility improvement between the semi-automated method and the double-oblique MPR manual measurements was observed.

Development of this AAA segmentation software opens the door to additional research implications. Conceptually, automated D-max measurement requires 3D modeling of AAA. The result- ing model can subsequently be analyzed for volume analysis and follow-up, topology, shape asymmetry and tissue anisotropy. By extension, these additional tools may be used for wall stress and rupture risk assessment not afforded by manual D-max.

This study had several limitations. The software segmentation time appears longer than the manual double-oblique MPR method. However, the additional processing time is fully justified since the software allowed complete 3D modeling of the AAA, of which D-max calculation was performed primarily for clinical validation. Since the software can be operated by the CT technologist, it will not involve physician time. Furthermore, the primary focus of our study was to show the feasibility and accuracy of a semi- automated segmentation and no effort was made to optimize the source code or exploit state-of-the-art workstation. We believe it can be shortened within the time of a manual measurement after optimization.

We did not include any patient after EVAR and included only one unenhanced studies. We need to validate in the future the accuracy of the software in the follow-up of patient after

EVAR and also investigate on a larger population if it will give the same reliability in unenhanced

5. Conclusion

Validation of this software to measure AAA maximal diameter showed higher intra-observer reproducibility than manual measurements. While this improvement was statistically significant, it might not be clinically significant. The main benefit of this automated method lies in the possibility of delegating the segmentation process, thus saving precious time for the radiologist.

Compared to manual measurements a very good accuracy was obtained for baseline and follow-up measurements and estimation of D-max growth. Since no error of more than 5 mm between both manual and software methods was observed, this algorithm is sufficiently robust to be used in various AAA morphology. This alleviates the uncertainty related to identification of the maximal diameter with conventional manual methods by providing a standardized method and would be of tremendous use to readers with less experience.
Since D-max measurement is the recognized gold standard for AAA diagnosis, treatment and follow-up, the validation of our software was based on D-max measurement. Additional evaluation and validation of volume measurement will be performed in sub- sequent studies.

Acknowledgements

This work was supported by a clinical research scholarship (to G.S.) from Fonds de la recherche en santé du Québec (FRSQ).

References

[1] Nevitt MP, Ballard DJ, Hallett JW. Prognosis of abdominal aortic aneurysms. A population-based study. New Engl J Med 1989;321(15):1009–14.

[2] Brewster DC, Cronenwett JL, Hallett Jr JW, et al. Guidelines for the treatment of abdominal aortic aneurysms. Report of a subcommittee of the Joint Council of the American Association for Vascular Surgery and Society for Vascular Surgery. J Vasc Surg 2003;37(5):1106–17.

[3] Hatakeyama T, Shigematsu H, Muto T. Risk factors for rupture of abdom- inal aortic aneurysm based on three-dimensional study. J Vasc Surg 2001;33(3):453–61.

[4] Lutz AM, Willmann JK, Pfammatter T, et al. Evaluation of aortoiliac aneurysm before endovascular repair: comparison of contrast-enhanced magnetic resonance angiography with multidetector row computed tomo- graphic angiography with an automated analysis software tool. J Vasc Surg 2003;37(3):619–27.

[5] Lederle FA, Wilson SE, Johnson GR, et al. Variability in measurement of abdominal aortic aneurysms. Abdominal Aortic Aneurysm Detection and Management Veterans Administration Cooperative Study Group. J Vasc Surg 1995;21(6):945–52.

[6] Wilmink AB, Forshaw M, Quick CR, et al. Accuracy of serial screening for abdom- inal aortic aneurysms by ultrasound. J Med Screen 2002;9(3):25–7.

[7] de Gracia MM, Rodriguez-Vigil B, Garzon-Moll G, et al. Correlation between the measurement of transverse diameter in the proximal neck on computed tomography and on aortography before endovascular treatment of infrarenal aortic aneurysm. Ann Vasc Surg 2006;20(4):8–95.

[8] Golledge J, Wolanski P, Parr A, et al. Measurement and determinants of infrarenal aortic thrombus volume. Eur Radiol 15 2008.

[9] Dugas A, Therasse E, Oliva V, et al. Abdominal aortic aneurysm follow-up with MDCT: reproducibility of maximal diameter measurement. In: Radiological Society of North America Scientific Assembly and Annual Meeting Program. Oak Brook, IL: Radiological Society of North America; 2007. pp. 632–633.

[10] Bodur O, Grady L, Stillman A, et al. Semi-automatic aortic aneurysm analysis. In: Manduca A, Hu X, editors. Medical imaging 2007: physiology, function, and structure from medical images. 2007.

[11] Das B, Mallya Y, Srikanth S, et al. Aortic thrombus segmentation using nar- row band active contour model. Conf Proc IEEE Eng Med Biol Soc 2006;1: 408–11.

[12] Olabarriaga SD, Rouet JM, Fradkin M, et al. Segmentation of thrombus inabdominal aortic aneurysms from CTA with nonparametric statistical grey level appearance modeling. IEEE Trans Med Imaging 2005;24(4):477–85

[13] Subasic M, Loncaric S, Sorantin E. Model-based quantitative AAA image analysis

using a priori knowledge. Comput Methods Programs Biomed 2005;80(2):103–14.
[14] Zhuge F, Rubin GD, Sun S, et al. An abdominal aortic aneurysm segmen- tation method: level set with region and statistical information. Med Phys 2006;33(5):1440–53.
[15] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–74.
[16] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307–10.

[17] Cayne NS, Veith FJ, Lipsitz EC, et al. Variability of maximal aortic aneurysm diameter measurements on CT scan: significance and methods to minimize. J Vasc Surg 2004;39(4):811–5.
[18] Abada HT, Sapoval MR, Paul JF, et al. Aneurysmal sizing after endovascular repair in patients with abdominal aortic aneurysm: interobserver variabil- ity of various measurement protocols and its clinical relevance. Eur Radiol 2003;13(12):2699–704.
[19] Dillavou ED, Buck DG, Muluk SC, et al. Two-dimensional versus three-dimensional CT scan for aortic measurement. J Endovasc Ther 2003;10(3): 531–8.

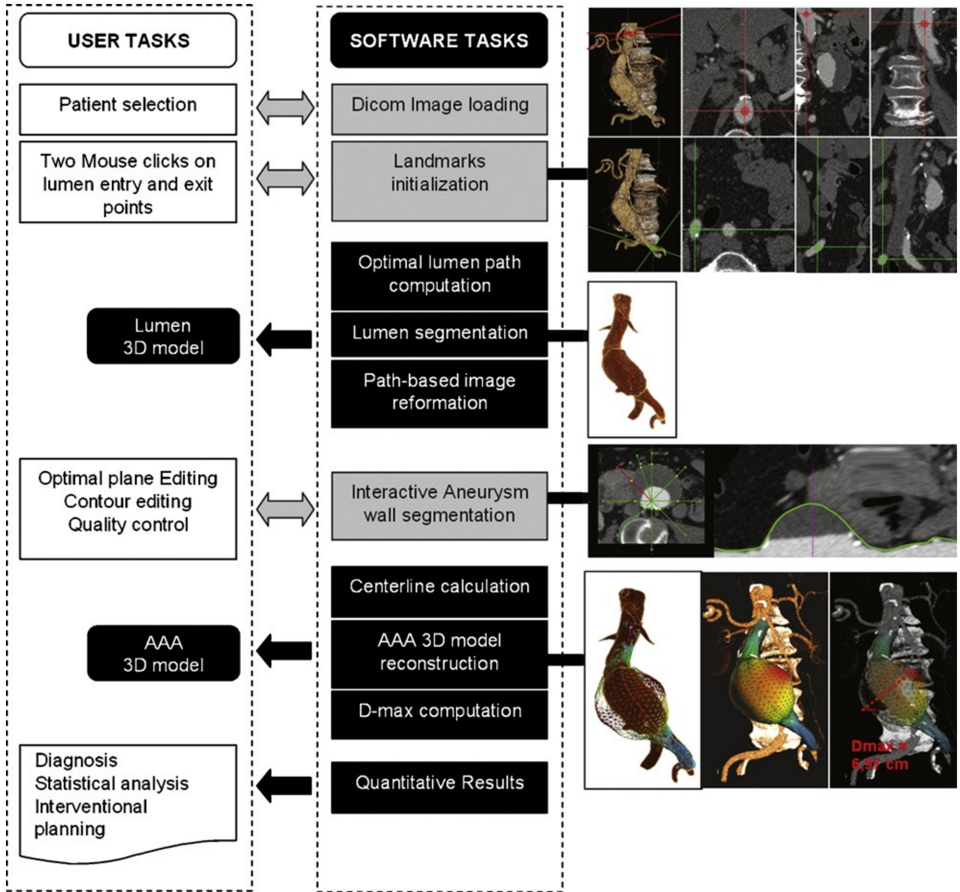

Fig. 1. Overview of software interaction. User tasks, software tasks and graphic display are illustrated in the left, middle and right column, respectively.

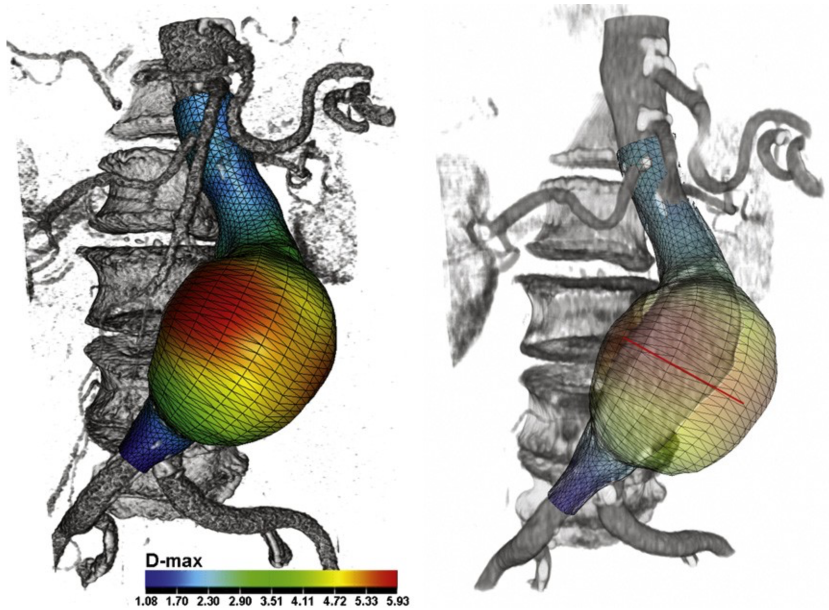

Fig. 2. Abdominal aortic aneurysm in a 79-year-old female. Left picture shows 3D volume rendering displays with 3D AAA model overlay. Different diameter values are color-coded, the smallest diameters are represented in blue and the largest in red. The automatically calculated D-max is displayed in the right picture by the red line.

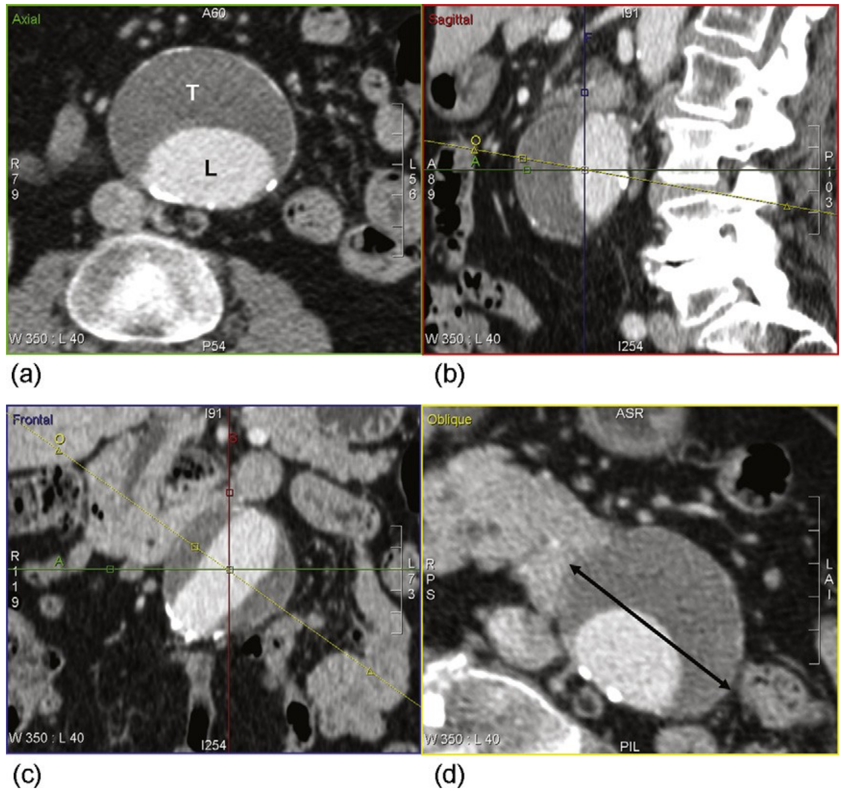

Fig. 3. Sequential approach to double-oblique (DO) reformation method illustrated in a 79-year-old female. (a) Axial contrast-enhanced CT in arterial phase shows abdominal aortic aneurysm. Lumen (L) is opacified by IV contrast. Large mural thrombus (T) fills part of the aneurysm. (b) and (c) MPR views of the aneurysm. Red line represents sagittal plane (b); blue line, coronal plane, green line, axial plane and yellow line, the user-defined DO plane (d), which is perpendicular to aneurysm wall in sagittal and coronal planes. D-max is measured manually on DO (line with double-arrows).

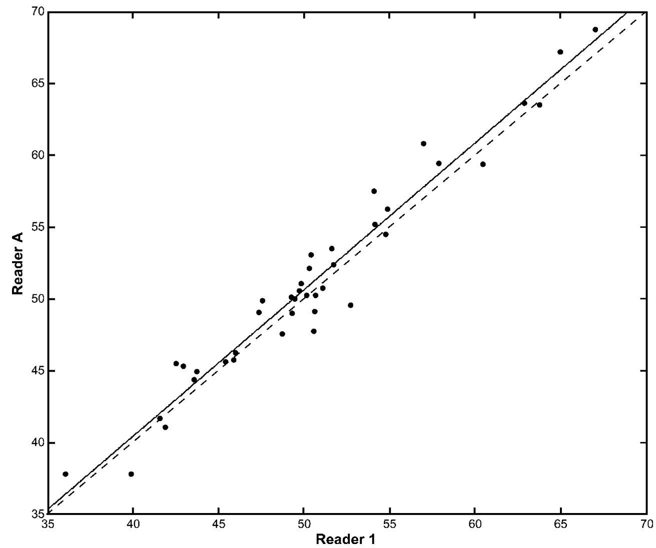

Fig. 4. Regression model between software and manual: mean of readings 1 and 2 between software (reader A) and radiologist #1 at baseline.

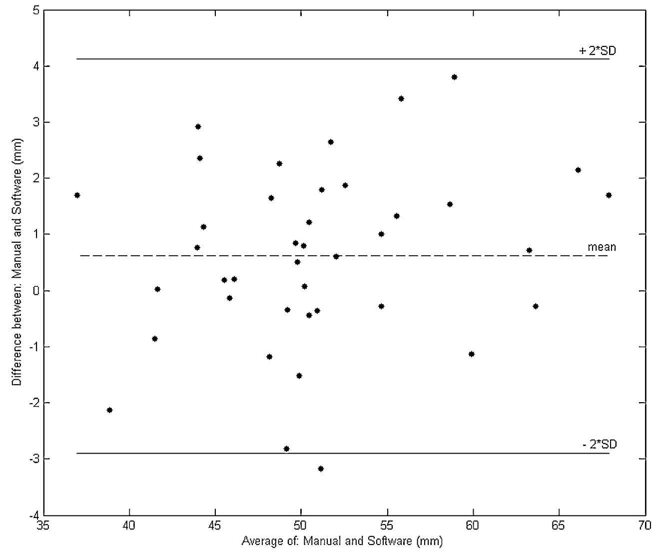

Fig. 5. Bland–Altman plot of the mean of readings 1 and 2 for software and radi- ologist #1 at baseline. The average of each pair of measurements is plotted against their difference. The range of agreement (solid lines) was defined as the bias ± 2 SD, where SD is the corrected standard deviation of the differences between the two methods.

Table 1
Maximal AAA diameter at baseline and follow-up MDCT.

	Baseline		Follow-up	
	Mean ± SD (mm)	Range (mm) (mm)	Mean ± SD (mm)	Range
First reading				
Radiologist 1	51.2 ± 7.4	37.7–69.9	55.2 ± 8.4	42.2–74.9
Radiologist 2	50.1 ± 7.4	37.0–69.0	54.3 ± 8.7	39.0–75.0
Radiologist 3	51.2 ± 7.5	37.0–71.2	54.9 ± 8.7	38.1–75.7
Software	50.6 ± 6.9	36.4–67.2	54.8 ± 7.9	42.3–76.8
Second reading				
Radiologist 1	51.2 ± 7.1	37.1–67.6	55.4 ± 8.7	39.0–75.1
Radiologist 2	–	–	–	–
Radiologist 3	50.9 ± 7.3	36.3–69.1	55.2 ± 8.5	38.9–75.2
Software	50.5 ± 6.9	35.8–66.9	55.0 ± 8.3	41.5–76.4

Table 2

Intra-observer reproducibility of software and manual D-max measurements.

Observers	Baseline		Follow-up	
	ICC (95%)	CI (one-sided)	ICC (95%)	CI (one-sided)
Radiologist 1	0.985	0.974	0.984	0.973
Radiologist 3	0.969	0.948	0.979	0.965
Software	0.992	0.987	0.990	0.983

Table 3
Inter-observer reproducibility of manual D-max measurements.

Reading	Baseline		Follow-up	
	ICC (95%)	CI (two-sided)	ICC (95%)	CI (two-sided)
First	0.979	0.954–0.984	0.975	0.955–0.985
Second[a]	0.981	0.964–0.990	0.987	0.9765–0.993

[a] Radiologist 2 had not repeated the measurements twice.