Comparison of segmentation software packages for in-hospital 3D print workflow

Niek Wijnen,[a,*] Lars Brouwers,[b] Erik Groot Jebbink,[c] Jan M. M. Heyligers,[b] and Mike Bemelman[b]

[a]University of Twente, Technical Medicine, Enschede, The Netherlands
[b]Elisabeth-Tweesteden Hospital, Department of Surgery, Tilburg, Noord-Brabant, The Netherlands
[c]University of Twente, Technical Medical Centre, Multi-Modality Medical Imaging Group, Enschede, The Netherlands

Abstract

Purpose: In-hospital three-dimensional (3D) printing of patient-specific pathologies is increasingly being used in daily care. However, the efficiency of the current conversion from image to print is often obstructed due to limitations associated with segmentation software. Therefore, there is a need for comparison of several clinically available tools. A comparative study has been conducted to compare segmentation performance of Philips IntelliSpace Portal® (PISP), Mimics Innovation Suite (MIS), and DICOM to PRINT® (D2P).

Approach: These tools were compared with respect to segmentation time and 3D mesh quality. The dataset consisted of three computed tomography (CT)-scans of acetabular fractures (ACs), three CT-scans of tibia plateau fractures (TPs), and three CTA-scans of abdominal aortic aneurysms (AAAs). Independent-samples *t*-tests were performed to compare the measured segmentation times. Furthermore, 3D mesh quality was assessed and compared according to representativeness and usability for the surgeon.

Results: Statistically significant differences in segmentation time were found between PISP and MIS with respect to the segmentation of ACs ($p =< 0.001$) and AAAs ($p = 0.031$). Furthermore, statistically significant differences in segmentation time were found between PISP and D2P for segmentations of AAAs ($p = 0.008$). There were no statistically significant differences in segmentation time for TPs. The accumulated mesh quality scores were highest for segmentations performed in MIS, followed by D2P.

Conclusion: Based on segmentation time and mesh quality, MIS and D2P are capable of enhancing the in-hospital 3D print workflow. However, they should be integrated with the picture archiving and communication system to truly improve the workflow. In addition, these software packages are not open source and additional costs must be incurred.

 [DOI: 10.1117/1.JMI.8.3.034004]

Keywords: 3D print; acetabular fracture; abdominal aortic aneurysm; segmentation; tibia plateau fracture.

Paper 20338R received Dec. 11, 2020; accepted for publication Jun. 4, 2021; published online Jun. 30, 2021.

1 Introduction

Medical application of three-dimensional (3D) printing technology has increased significantly in the past few years. The added value of pre-operative 3D printing is endorsed by many physicians.[1] With in-hospital 3D printing, physical models of patient-specific pathologies can be manufactured without the time and cost of involving a commercial company. These models

*Address all correspondence to Niek Wijnen, n.wijnen97@gmail.com

2329-4302/2021/$28.00 © 2021 SPIE

can be used to establish a comprehensive pre-operative workup. In comparison with conventional two-dimensional (2D) imaging, 3D printing provides additional spatial information about pathology, which could lead to an increased 3D perception of the surgeon.[2] Brouwers et al. demonstrated that interobserver consensus of the classification of acetabular fractures (ACs) is significantly higher when a preoperative 3D print is used instead of 2D medical imaging. In addition, Brouwers et al.[3] observed that 3D-printed models establish a better 3D understanding of complex fracture patterns in comparison with 3D reconstructions (3D-CT). 3D-CT is seen on a 2D screen, which does not provide an actual 3D view and limits rotation of the 3D reconstruction, resulting in less 3D understanding of fracture patterns. Furthermore, Shilo et al.[2] showed the added value of 3D printing in preadaptation of surgical instruments in craniomaxillofacial surgery. Therefore, the use of 3D printing could increase the effectiveness and efficiency of pre-operative surgical planning. In addition to pre-operative benefits of 3D printing, it is also proven to be valuable peri-operatively. In operations with complex bone fractures, a 3D print can function as surgical guidance for the surgeon.[4] Furthermore, the chance of inducing peri-operative complications, such as internal bleedings, could be potentially reduced when a 3D print is used.[5] Last, the use of 3D printing could result in a significantly decreased operation time and improved precision of surgical performance.[2] In-hospital pre-operative 3D printing of patient-specific pathologies associated with vascular and trauma procedures, such as complex bone fractures and abdominal aortic aneurysms (AAAs), is frequently performed. Threshold-based segmentation is commonly utilized to segment pathologies visualised by computed tomography (CT) scans. Threshold procedures are simple image segmentation tools. In this method, a specific Hounsfield unit (HU) value is set as threshold. Tissues represented by voxels with HU-values lower than the threshold are erased. As a result, soft tissue can be easily separated from bone tissue in case of bone fractures. However, threshold based tools have several limitations that obstruct segmentation efficiency and decrease quality of 3D prints. This can result in time consuming procedures and insufficient 3D prints. Currently, numerous segmentation tools are developed for medical image segmentation. Therefore, it was our purpose to conduct a comparative study and investigate the performance of different segmentation tools for in-hospital 3D printing.

2 Methods

A segmentation modality should provide several features to be applicable for the in-hospital workflow from 2D-CT to 3D print. First of all, segmentation software should have the possibility to connect to the picture archiving and communication system (PACS). Without this connection, every scan has to be manually uploaded in the segmentation software, making it time-consuming. Second, the software should be capable of exporting segmented volumes as Standard Tessellation Language (STL) files. An STL-file describes geometry of 3D objects with a polygonal 3D mesh, where the orientation of triangular faces defines the surface of 3D objects.[6] This conversion is essential since an STL-file is 3D print compatible and enables direct import of volumes in mesh modification software. The following segmentation software packages that suffice these requirements were available to include in our study.

1. Philips IntelliSpace Portal® (PISP) is developed by Philips Healthcare, a medical imaging device manufacturer. PISP is frequently used for in-hospital medical image segmentation, because it is incorporated in the purchase of Philips medical imaging devices. Currently, PISP is used as segmentation software for the in-hospital workflow from 2D-CT to 3D print in the Elisabeth-Twocities hospital Tilburg, The Netherlands.
2. Mimics Innovation Suite® (MIS) is an image processing tool for 3D design and engineering, developed by Materialise NV. MIS was included, because it provides the possibility to vary and investigate numerous settings for medical imaging segmentation. Unlike PISP, MIS is a software package that is separately purchased.
3. DICOM to PRINT® (D2P) is a software package that is specifically developed for the in-hospital 2D-CT to 3D print workflow, developed by 3D Systems. D2P focuses on creating a workflow consisting of as few steps as possible to provide a segmentation tool that is

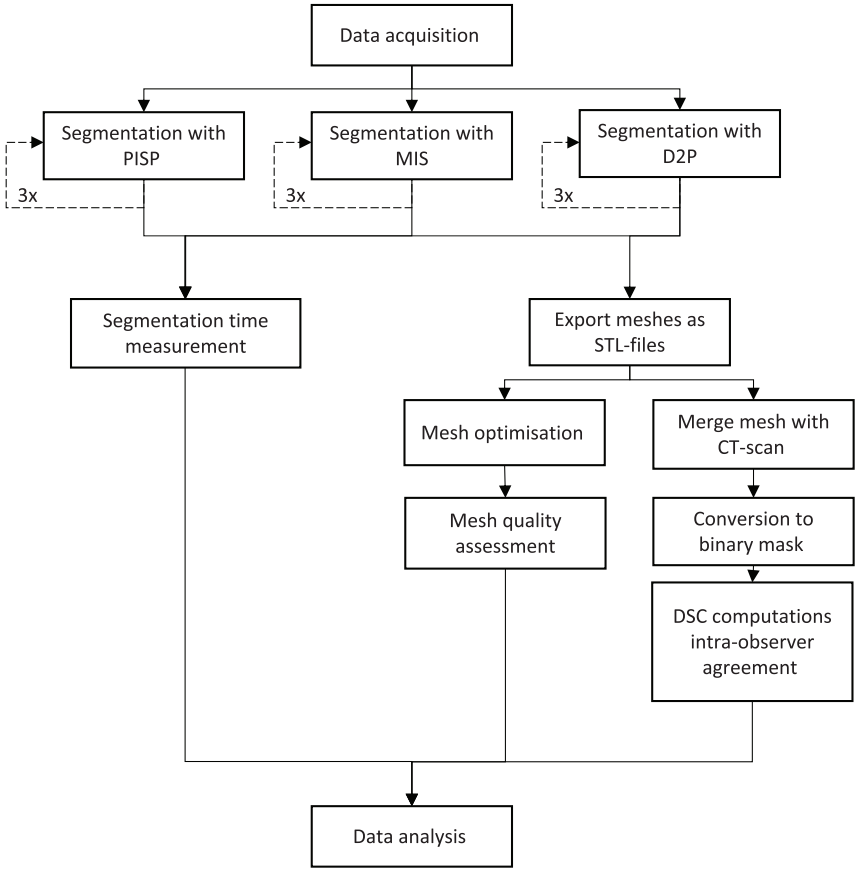

Fig. 1 Flowchart of method of segmentation study.

useable for clinicians with limited experience in image processing. Similar to MIS, D2P is acquired separately.

We compared the efficiency and effectiveness of segmentations performed in PISP, MIS, and D2P. Segmentation efficiency and effectiveness were expressed as segmentation time and quality of 3D meshes, respectively. The segmentation protocols of the different segmentation modalities for each pathology are elaborated in Appendix A. Furthermore, a flowchart of the method is shown in Fig. 1.

2.1 *Data Acquisition*

A dataset was collected to compare segmentation performance of the different software packages. The dataset consisted of three pathologies, each associated with a frequently occurring segmentation limitation.

1. Three axial bone CT-scans of ACs were collected. These were selected, to find out which segmentation modality was the most optimal for separating the femur head from the ipsilateral affected hemipelvis. This is essential for the trauma surgeon to facilitate assessment of ACs. All scans were made on a Philips CT scanner with the following parameters: 120 kVp; 109 to 201 mA; 0.80/1-mm slice thickness; 1024×1024 matrix size; $-650 - 1350/ - 425 - 2075$ window length; 2000/2500 window width.
2. Three axial bone 3D CT-scans of tibia plateau fractures (TPs) were included. These scans were obtained from patients with an age >70-years-old, to investigate which segmentation modality is the most optimal for osteoporotic bone segmentation in TPs.[7,8] Decreased bone mass results in a complicated segmentation and an insufficient 3D print when a simple threshold tool is used. All scans were made on a Philips CT scanner with the following

parameters: 100 kVp; 44–67 mA; 0.80/1-mm slice thickness; 512×512 matrix size; $-650 - 1350$ window length; 2000 window width.

3. Three iterative model reconstruction CTA-scans of AAAs were selected. AAA segmentation is often complicated by intraluminal thrombus. Voxels representing the intraluminal thrombus have a HU-value similar to that of surrounding soft tissue.[9,10] When the lower threshold limit is set too high to erase soft tissue surrounding the AAA, the intraluminal thrombus will also be expunged. This is undesirable since the assessment of an intraluminal thrombus is necessary to successfully utilise 3D prints for vascular surgery. Therefore, only CTA-scans of AAAs with an intraluminal thrombus covering the aortic wall were included. All scans were made on a Philips CT scanner with the following parameters: 100 kVp; 264–610 mA; 1-mm slice thickness; 512×512 matrix size; $-160 - 240$ window length; 400 window width.

To investigate which of the proposed segmentation software packages was the most applicable for the in-hospital 2D-CT to 3D print workflow, each CT-scan was segmented in the different software packages. The segmentations were performed by one observer who is experienced with all three software packages.

2.2 *Included Parameters*

2.2.1 *Intra-observer agreement*

The intra-observer agreement was determined to validate segmentations performed by the observer. The segmentation of each CT-scan in each software package was repeated three times. A minimum time interval of 24 h between each repeated segmentation was used to minimalize influence of a learning curve on segmentation performance. The similarity between the three segmented 3D meshes was determined by computing Dice similarity coefficients (DSC). DSC describes spatial overlap between segmented voxels of two volumes expressed in a number between 0 and 1, where 0 and 1 mean no similarity and a perfect similarity, respectively. Computations were performed in MATLAB R2019a® (MathWorks, Natick). To calculate the DSC, all segmented volumes were loaded in MIS alongside the corresponding CT-scan. Using MIS, the segmented volumes were superimposed on the corresponding CT-scans. Subsequently, these merged CT-scans were exported as DICOM-files, and imported in MATLAB R2019a. In this program, CT-scans were converted into binary masks, where voxels containing the segmented volume and voxels containing everything except the segmented volume obtained a value of 1 and 0, respectively. The binary masks were used for DSC calculation (Fig. 2).

First, DSCs of volumes of the first and second, first and third and second and third repeated segmentation were calculated. Next, the mean of the three DSCs was computed (Appendix C). The three segmentations of the same scan in the same software package should have a very high similarity. Therefore, the mean DSC should be ≥ 0.9 to validate the segmentation process. A DSC of 0.9 means an excellent similarity between different segmented volumes.[11]

2.2.2 *Segmentation time*

The duration of performed segmentations was quantified as segmentation time. This is the time period between the moment that the scan was imported in segmentation software, until the moment that the observer was satisfied with the segmented volume. Segmentation times were measured with a timer. This is an important parameter, because it determines segmentation efficiency. An efficient segmentation method is essential for a convenient in-hospital workflow from 2D-CT to 3D print. An independent-samples t-tests was used to statistically compare segmentation time measurements between the different software packages. A significance level of $\alpha = 0.05$ was utilised. Independent-samples t-tests were performed in IBM SPSS Statistics 27® (IBM, New York). Only the comparison between the different software packages is of interest. Therefore, segmentation times of the three repeated segmentations of the same scan used for intra-observer agreement computation were not separately statistically analysed.

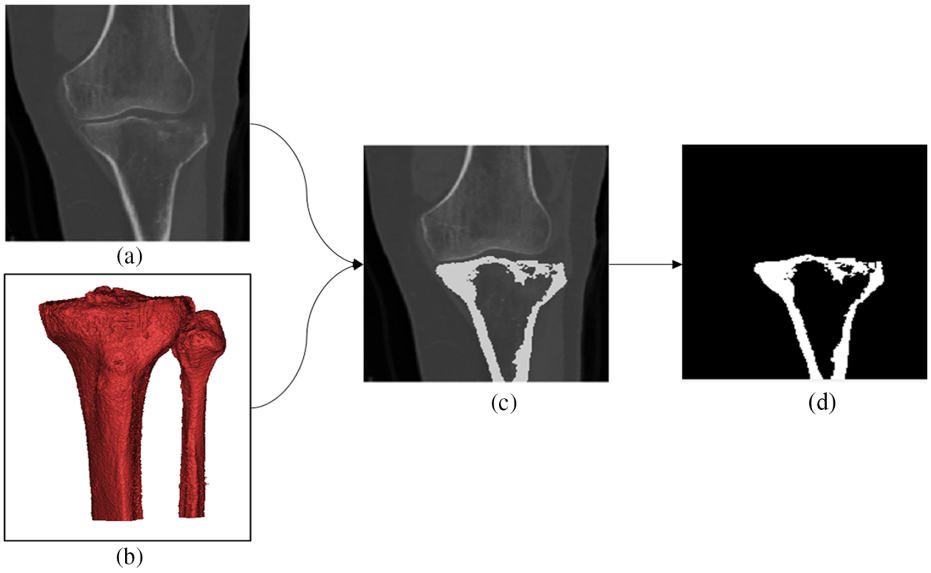

Fig. 2 (a) CT-scan of an TP is merged with the corresponding segmented volume (b) in MIS to obtain the combined CT-scan, where the uniform coloured connected voxels represent the segmented volume (c). This combined CT-scan is imported in MATLAB R2019a, where it is converted into a binary mask that represents the segmented volume (d). The similarity between this binary mask and another can be quantified using DSC.

2.2.3 *3D mesh quality*

It is essential that the 3D mesh quality is optimal to manufacture sufficient 3D prints. Therefore, finalized 3D meshes of the performed segmentations were subjectively assessed (Fig. 3). These 3D meshes were modified in Meshmixer® (Autodesk Research, New York) and ready for 3D printing. 3D mesh quality was assessed according to the following parameters.

1. Amount of holes that the 3D mesh contains.
2. Characteristics associated with the pathology.
3. Overall representativity of the 3D mesh in comparison with the pathology visualized by the CT-scan.
4. Usability for the surgeon.

For both the ACs and TPs, it was important that holes in the 3D mesh were not interfering with the perceptibility of the fracture. Furthermore, parts of bone surrounding the fracture must

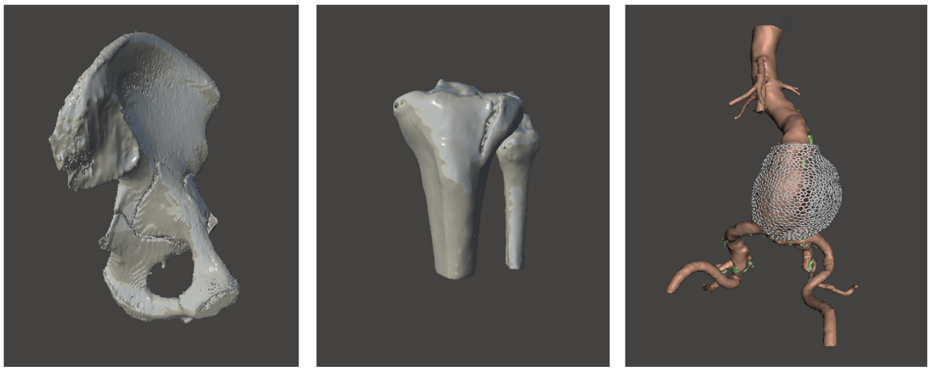

Fig. 3 (a) Hemipelvis with AC visualized in Meshmixer. (b) TP visualized in Meshmixer. (c) AAA visualized in Meshmixer with intraluminal thrombus as hexagonal mesh (white).

not fuse together. This is a commonly encountered problem. The 3D meshes of both ACs [Fig. 3(a)] and TPs [Fig. 3(b)] were assessed by a senior trauma surgeon according to the scoring system presented in Table 1. The trauma surgeon is acquainted with the utilisation of 3D prints.

The quality of 3D meshes of AAAs was mainly based on visualization of the intraluminal thrombus. It was essential that the intraluminal thrombus was well defined in the 3D mesh, to provide the vascular surgeon with sufficient morphological and spatial information about the aneurysm. The intraluminal thrombus was represented as a hexagonal mesh to provide an inside view of the aneurysm. The process of creating a hexagonal mesh is explained in Appendix B. Furthermore, celiac, renal, mesenteric and iliac arteries were included in the 3D mesh. These visceral branches suffice as anatomical landmarks for the vascular surgeon [Fig. 3(c)].

The 3D meshes were assessed by a senior vascular surgeon according to the scoring system presented in Table 2. The vascular surgeon is acquainted with the utilisation of 3D prints.

Table 1 Scoring system for AC and TP.

Score	Criteria
4	The 3D mesh does not contain any obvious holes and the bone surrounding the fracture is not fused together. The fracture is representative for the actual fracture visualized in the CT-scan. The 3D print would be useful for the trauma surgeon.
3	The 3D mesh does only contain several obvious holes, the bone surrounding the fracture is (almost) not fused together. The fracture is representative for the actual fracture visualized in the CT-scan. The 3D print would be useful for the trauma surgeon.
2	The 3D mesh may contain several holes, however, small parts of the bone surrounding the fracture are fused together; The fracture is representative for the actual fracture visualized in the CT-scan to a lesser extent. The 3D print would be useful for the trauma surgeon to a lesser extent.
1	The 3D mesh does contain many obvious holes, which make the fracture poorly recognizable and/or multiple parts of the bone surrounding the fracture are fused together. The fracture is barely representative for the actual fracture visualized in the CT-scan. The 3D print would barely be useful for the trauma surgeon.
0	The 3D mesh does contain many obvious holes, which make the fracture unrecognizable and/or the bone around the fracture is (almost) totally fused together. The fracture is not representative for the actual fracture visualized in the CT-scan. The 3D print would not be useful for the trauma surgeon.

Table 2 Scoring system for AAA.

Score	Criteria
4	The 3D mesh does not contain any obvious holes and the abdominal aorta and intraluminal thrombus are sufficiently visualized, including primary branches. The 3D mesh is representative for the actual AAA visualized in the CT-scan. The 3D print would be useful for the vascular surgeon.
3	The 3D mesh does only contain several obvious holes, the abdominal aorta and intraluminal thrombus are sufficiently visualized, including primary branches. The 3D mesh is representative for the actual AAA visualized in the CT-scan. The 3D print would be useful for the vascular surgeon.
2	The 3D mesh may contain several holes, however the intraluminal thrombus and/or primary branches are less sufficient visualized. The 3D mesh is representative for the actual AAA visualized in the CT-scan to a lesser extent. The 3D print would be useful for the vascular surgeon to a lesser extent.
1	The 3D mesh does contain many obvious holes, which make the AAA poorly recognizable and/or the intraluminal thrombus and/or primary branches are poorly visualized. The 3D mesh is barely representative for the actual AAA visualized in in the CT-scan. The 3D print would barely be useful for the vascular surgeon.
0	The 3D mesh does contain many obvious holes, which make the AAA unrecognizable and/or the intraluminal thrombus and/or primary branches are (almost) not visualized. The 3D mesh is not representative for the actual AAA visualized in the CT-scan. The 3D print would not be useful for the vascular surgeon.

To prevent observer bias, the trauma surgeon and vascular surgeon were blinded for which software package was applied for the segmentation of the shown 3D mesh, by randomly selecting the order of showing 3D meshes corresponding with different segmentation modalities. Assessment of 3D mesh quality of all segmented volumes was not possible due to time constraints. Therefore, only one of three repeated segmentations of each CT-scan was randomly selected and assessed by the surgeon.

3 Results

3.1 *Intra-Observer Agreement*

In Fig. 4, computed mean DSC's for each software package are presented per included CT-scan to validate the segmentation process. The red line represents a DSC of 0.9. In Table 3, the mean DSC for each pathology and overall mean DSC are displayed per segmentation modality.

3.2 *Acetabular Fractures*

In Fig. 5, mean segmentation times and overall mean of the 3 CT-scans with ACs are displayed. In D2P, a graphical card error occurred when importing each of the three acetabular CT-scans. Therefore, the segmentation time in D2P could not be included. Overall mean segmentation time in PISP and MIS are 8.47 (± 0.54) min and 1.81 (± 1.14) min, respectively. The independent-samples t-test showed a statistically significant faster segmentation in MIS with respect to PISP ($p =< 0.001$) with an overall mean time difference of 6.66 minutes. In Fig. 6, accumulated 3D mesh quality scores are presented.

3.3 *Tibia Plateau Fractures*

In Fig. 7, mean segmentation times and overall mean of the three CT-scans with TPs for each segmentation modality are displayed. In Table 4, mean segmentation times and results of

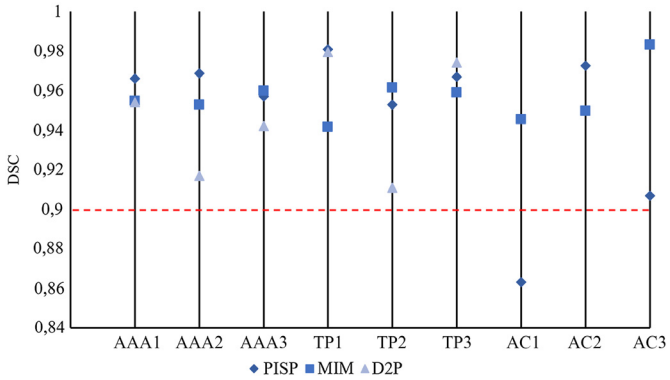

Fig. 4 Graph representing mean DSC of triple segmentation of the same scan in all of the segmentation modalities. The redline represents a DSC of 0.9. AAA = abdominal aortic aneurysm; TP = tibia plateau fracture; AC = acetabular fracture; DSC = Dice similarity coefficient.

Table 3 Intra-observer agreement expressed in DSC.

Software	Overall mean DSC of AAAs (std)	Overall mean DSC of TPs (std)	Overall mean DSC of ACs (std)	Overall mean DSC of entire dataset (std)
PISP	0.964 (0.006)	0.967 (0.014)	0.914 (0.055)	0.948 (0.038)
MIS	0.956 (0.004)	0.954 (0.011)	0.959 (0.021)	0.956 (0.012)
D2P	0.938 (0.019)	0.955 (0.038)	—	0.946 (0.029)

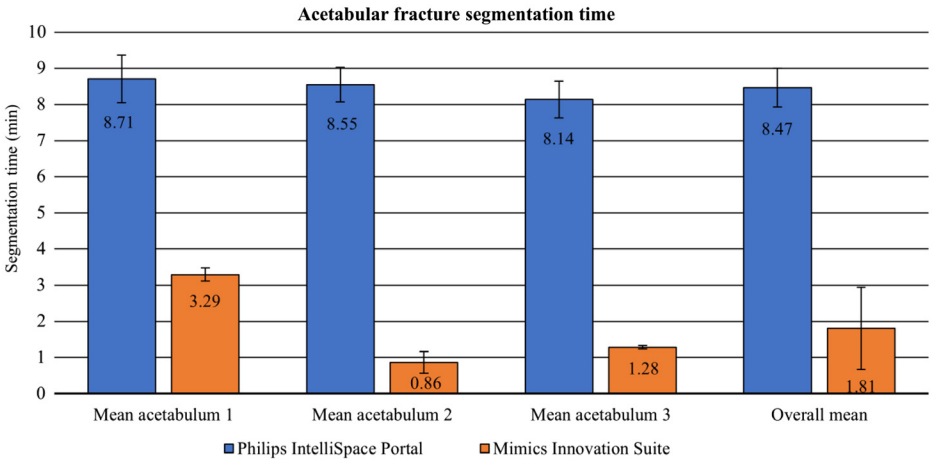

Fig. 5 Mean AC segmentation time comparison between segmentation software.

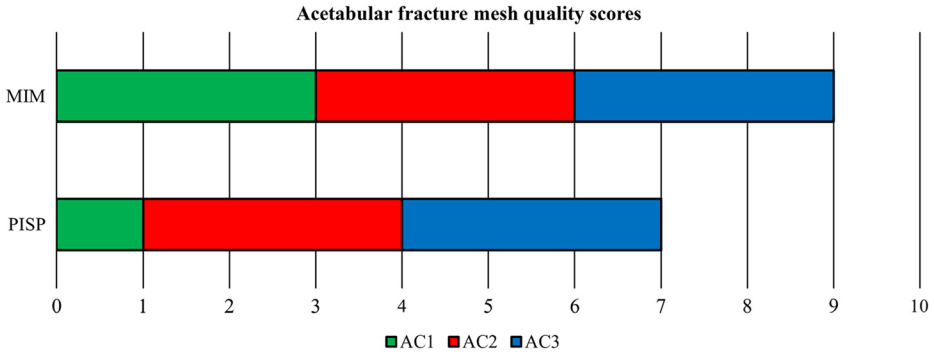

Fig. 6 Cumulative Acetabular 3D mesh quality assessment according to scoring system. AC = acetabular fracture.

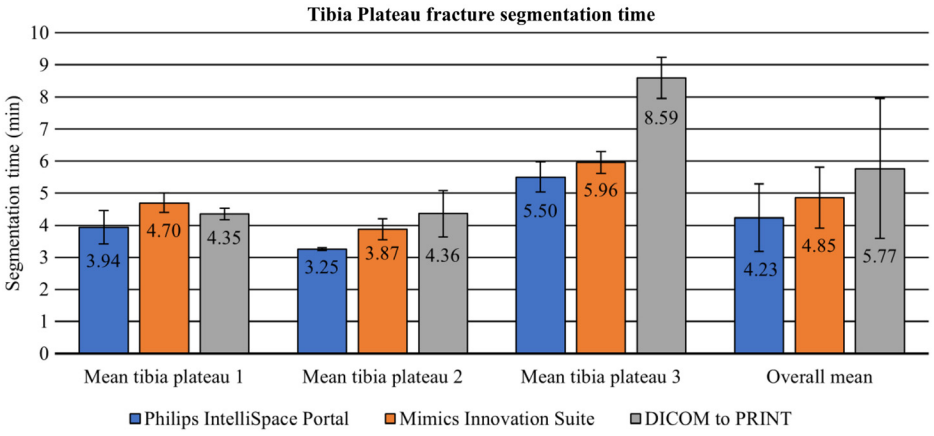

Fig. 7 Mean TP segmentation time comparison between segmentation software.

segmentation time comparison between the different software packages performed with independent-samples *t*-tests are represented. Furthermore, in Fig. 8, accumulated 3D mesh quality scores are presented.

3.4 *Abdominal Aortic Aneurysm*

In Fig. 9, mean segmentation times and overall mean of the three CTA-scans with AAAs are displayed for each segmentation modality. In Table 5, mean segmentation times and results of

Table 4 TP segmentation time comparison between different software.

Software comparison	Overall mean time (std)	*p*-value	Segmentation time difference (min)
PISP and MIS	4.23 (1.06) and 4.85 (0.95)	0.212	0.62
PISP and D2P	4.23 (1.06) and 5.77 (2.17)	0.081	1.54
MIS and D2P	4.85 (0.95) and 5.77 (2.17)	0.268	0.92

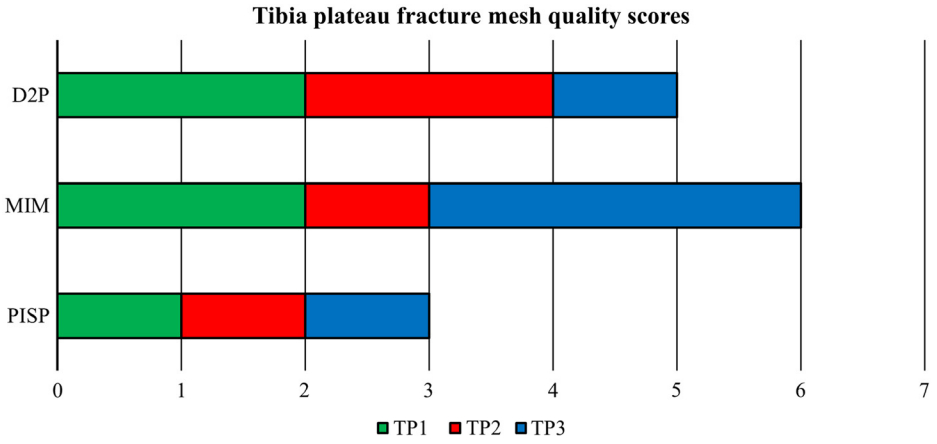

Fig. 8 TP 3D mesh quality assessment according to score system. TP = tibia plateau fracture.

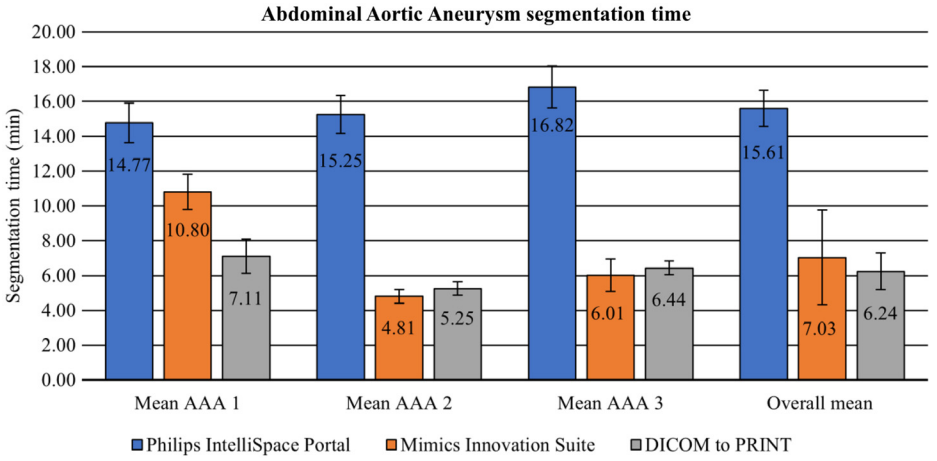

Fig. 9 Mean AAA segmentation time comparison between segmentation software.

Table 5 AAA segmentation time comparison between different software.

Software comparison	Overall mean time (std)	*p*-value	Segmentation time difference (min)
PISP and MIS	15.61 (1.04) and 7.03 (2.72)	0.031	8.58
PISP and D2P	15.61 (1.04) and 6.24 (1.05)	0.008	9.37
MIS and D2P	7.03 (2.72) and 6.24 (1.05)	0.321	0.79

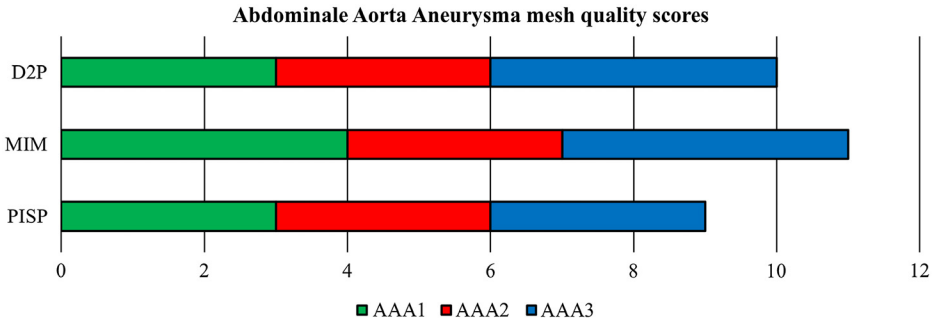

Fig. 10 AAA 3D mesh quality assessment according to score system. AAA = abdominal aortic aneurysm.

segmentation time comparison between the different software packages performed with independent-samples *t*-tests are represented. Furthermore, in Fig. 10, accumulated 3D mesh quality scores are presented.

4 Discussion

In this study, we compared the efficiency and effectivity of 3D mesh segmentation using PISP, MIS, and D2P. Results showed a statistically significant lower 3D mesh segmentation time of ACs and AAAs using MIS when compared with PISP. Similarly, the segmentation of AAAs was significantly faster using D2P when compared to PISP. Furthermore, the accumulated quality scores of the 3D meshes is for all pathologies the highest in MIS, followed by D2P.

4.1 *Literature Review*

Other studies that compared MIS with other software packages for segmentations of 3D mesh volumes of pathologies in terms of efficiency and effectivity had similar results to our study.[12,13] To the authors best knowledge, no similar segmentation software package comparisons were performed for PISP and D2P. Only one other study was found that compared MIS with other commercial software packages.[14] Furthermore, multiple studies were found that use PISP for their 3D print workflow.[15,16]

Particularly, the segmentation step in the workflow from 2D-CT to 3D print was investigated in our study. Brouwers et al.[15] showed in a 3D print validation study that the segmentation step is an important determinant for 3D model accuracy. Hence, the segmentation step should be optimal.

4.2 *Intra-Observer Agreement*

Every included scan was segmented three times using each software modality to assess intra-observer agreement. All individual DSC's were $\geq$0.9, except for the first AC segmented using PISP (DSC: 0.863) due to scatter caused by a metallic implant. In PISP the scatter needed to be manually erased, which resulted in three less identical volumes. However, the mean DSC of the three CT-scans with ACs segmented in PISP was 0.914. Furthermore, the overall mean DSC's of the entire dataset for PISP, MIS, and D2P are 0.948, 0.956, and 0.946, respectively. This indicates an excellent segmented volume similarity and a good intra-observer agreement.

4.3 *Acetabular Fractures*

Statistically significant differences in segmentation time were found between the segmentation using PISP and MIS, indicating that the segmentation of ACs in MIS was more efficient. This significant difference in 3D mesh segmentation time is the result of the split mask function that is available in MIS. This function automatically separates the femur head from the acetabulum by

roughly indicating both anatomical structures on just one axial/coronal/sagittal slice with a brush tool. In PISP, it is often not possible to automatically separate the femur from the ipsilateral hemipelvis. To achieve this, a manual separation has to be done in a 2D environment, slice for slice. This can be a very time-consuming procedure. Furthermore, our study showed a higher accumulated 3D mesh quality score for the segmentation of ACs using MIS when compared to PISP. This is, again, due to the split mask function in MIS, which reduced influence of scatter caused by a metallic implant on the 3D mesh. Therefore, using the split mask function in MIS in case of metallic implants could result in a higher quality 3D mesh.

The segmentation of ACs was not possible in D2P due to a local graphical card error. The volume of a pelvis is substantially larger with respect to tibial or aortic volumes. Therefore, a pelvic mesh consists of significant more faces, which requires a large computation power. A reason for the crash could be that D2P used the laptop Intel graphic processor instead of the NVIDIA GeForce RTX 2070. Therefore, AC segmentation performance in D2P could not be evaluated in our study.

4.4 *Tibia Plateau Fractures*

Mean segmentation times of TPs in each segmentation modality did not differ statistically significantly from each other. 3D meshes segmented in MIS, obtained the highest accumulated mesh quality score. Furthermore, segmentations in both MIS and D2P resulted in a more qualitative 3D mesh of TPs in contrast with PISP. Due to the threshold tool in PISP, the segmented volumes contained many holes as a result of the significantly reduced bone mass in osteoporosis. Higher 3D mesh quality in MIS is established using the smart fill function to fill gaps that are caused by decreased bone density while simultaneously applying 3D interpolation. This makes it a semi-automatic segmentation operation. Similarly, D2P provides a brush function that can be utilized to fill gaps in the 3D mesh. These are powerful tools to overcome segmentation problems associated with osteoporosis. However, utilisation of these tools increases segmentation time. Due to this, no statistically significant differences were found for the mean segmentation times.

4.5 *Abdominal Aortic Aneurysms*

Statistically significant differences in segmentation time of AAAs were found between the segmentation using PISP and MIS and PISP and D2P, indicating that the segmentation of AAAs is more efficient in both MIS and D2P when compared with PISP. However, no statistically significant difference in segmentation time was found between segmentations in MIS and D2P. The segmentation of AAAs in MIS and D2P was more efficient due to faster segmentation of the intraluminal thrombus with the 3D interpolation function. Every 5 to 10 axial slices, the intraluminal thrombus had to be segmented once. Subsequently, the function interpolated the segmented slices to obtain the intraluminal thrombus semi-automatically. Furthermore, the dynamic region growing function in MIS, which is a combination of a threshold procedure and a region growing function, required just one click on the contrast in the aorta to segment it automatically. D2P has an automatic vascular segmentation tool, which similarly required just one click on the contrast containing aorta. These tools in MIS and D2P make the segmentation of AAAs more efficient in comparison with PISP. In addition, when the lower threshold limit was decreased in PISP to segment the intraluminal thrombus, surrounding soft tissue was also included. This needed to be manually erased, which is aggravating and time-consuming. Mesh quality scores of the segmented AAAs are relatively similar for all software packages. Remarkably, the accumulated quality score is again the highest for the segmentations performed in MIS. This indicated that segmentations performed in MIS results in the highest quality 3D prints.

4.6 *Segmentation Time versus 3D Mesh Quality*

The comparison between segmentation time and 3D mesh quality is important to consider. The statistical significant overall mean segmentation time differences ranged from 6.66 to 9.37 min.

When 3D mesh quality is sufficient, this effect size can be inconsiderable. In our opinion, 3D mesh quality is the main determinant for a convenient 3D print workflow when the segmentation times differ up to ~ 10 min. Both MIS and D2P provide better 3D mesh quality.

4.7 *Limitations*

First of all, the number of included cases is limited, we included three CT-scans from different patients for three types of pathologies. In addition, the study focused on three types of pathology that are common to our daily practice, without the intention to cover all areas and technical issues connected to in-hospital 3D printing. However, the technical issues addressed in this paper are transferrable to other types of pathology, such as the fracture separation issue, which also applies to craniomaxillofacial trauma.[2] Furthermore, in the first CT-scan of an AC, a metallic implant placed in the pelvis was seen. This implant caused scatter in the scan, which influenced segmentation time and 3D mesh quality. However, the presence of a metallic implant is a relevant feature for a segmentation study. Therefore, it would be better to include more scans in the dataset where a metallic implant is present. Similarly, segmentation of bone fractures where a plaster cast is applied is complicated using PISP. This is because the plaster cast is represented by high gray value intensities similar to that of bone. Therefore, thresholding in PISP results in segmenting the plaster cast surrounding the bone fracture. Consequently, the plaster cast must be manually erased to obtain a sufficient 3D model of the bone fracture. In MIS and D2P the plaster cast can be automatically erased with the split mask function, saving time and work.

None of the segmentations performed in PISP are statistically significantly faster with respect to MIS and D2P. This is a result of the segmentation protocol used in PISP, which consists of a threshold procedure followed by manually erasing irrelevant structures. This segmentation protocol is obsolete, because the amount of developed algorithms for advanced medical imaging segmentation is increasing and the segmentation is becoming more efficient. For example, the split mask and smart fill algorithms in MIS increased efficiency and effectivity for bone fracture segmentation significantly. In addition, deep learning is increasingly being used for automatic medical image segmentation. Especially recent advances in convolutional neural networks (CNN) improve the performance of automated 3D semantic segmentation. CNNs are outperforming similar state-of-the-art segmentation methods for medical imaging.[17,18] These segmentation algorithms are much further in development in contrast with the used segmentation protocol in PISP. Although the segmentation method in PISP is outdated, it should be mentioned that the use of PISP is convenient for clinicians that have limited time to make 3D prints.

Finally, only one observer segmented all scans in all three software packages. This might have led to observer bias. However, we expect that the influence of this observer bias is very limited since the segmentation time was used as the outcome measure.

4.8 *Recommendations*

In the current in-hospital 2D-CT to 3D print workflow, the segmented 3D mesh is imported in Meshmixer, where it is optimized for 3D printing. Subsequently, the resulting 3D mesh is exported from Meshmixer and imported in the 3D print software to add print supports. Both MIS and D2P provide direct mesh modification options similar to those available in Meshmixer. Therefore, these software modalities have the possibility to surpass the step where the STL-file is imported and exported in Meshmixer to even further optimise the process. However, this should be investigated. In our study, we did not include this measure to focus on segmentation performance and be able to make a fair comparison between software packages.

Another critical aspect of this study is that none of the utilised software modalities is open source software. It would be interesting to similarly investigate capabilities of open source software for the in-hospital 3D print workflow, such as 3DSlicer and ITK-SNAP. However, the commercial software package PISP is often incorporated in the purchase of a Philips CT-scanner. In addition, PISP is then integrated with the PACS. Integration with PACS is an essential aspect since scans can be directly imported and segmented in PISP. In most open source software, PACS integration is not possible, which requires manual transfer of CT-scans from PACS to the

software package. This is a time-consuming process. Similar to PISP, both MIS and D2P are capable of overcoming this problem, because these software packages offer the possibility to connect to the PACS server. However, MIS and D2P must be purchased separately and additional costs must be incurred. The license prices for these software packages depend on license configuration and the amount of licenses bought. Therefore, the additional costs are variable. Although MIS and D2P enhanced segmentation time and 3D mesh quality, it should be considered whether this improvement is worth the investment. Therefore, we highly recommend experimenting with these programs using the free trial license prior to purchasing the software packages.

5 Conclusion

Three types of 3D segmentation software were investigated and compared in this study. Segmentation efficiency and 3D mesh quality were measured. Statistically significant differences in segmentation time were found, indicating that segmentation using MIS and D2P is more efficient in contrast with PISP as a result of (semi-)automatic segmentation tools. In addition, segmentation in MIS and D2P improved 3D mesh quality. This indicates that MIS and D2P could improve the in-hospital 3D print workflow. However, the software packages should be integrated with the PACS server to truly enhance it. In addition, both MIS and D2P must be separately purchased resulting in additional costs.

6 Appendix A: Segmentation Protocols

6.1 *Acetabular fracture*

6.1.1 *Philips intellispace portal*

First, the CT-scan was converted to 3D-CT using the volume rendering function. Next, thresholding was performed until (almost) all soft tissue disappeared. Subsequently, the proximal femur and contralateral hemipelvis were erased by cutting through the pubic symphysis and sacroiliac joint. When there were too many holes in the presented volume, the minimum threshold limit was decreased and soft tissue structures surrounding the affected hemipelvis were manually erased. Next, the 2D cut plane was initiated. To erase the femur head, the correct angle in the 2D cut plane needed to be found to prevent deleting parts of the acetabulum. Finally, the femur head was manually erased slice by slice until it was totally deleted.

6.1.2 *Mimics innovation suite*

First, a threshold procedure was performed with the pre-set bone threshold. Also, "keep largest part" was selected, to automatically erase irrelevant noisy parts. Next, the region of interest (ROI) of the created mask was defined. The ROI contained the affected hemipelvis and was bounded to the middle of the pubic symphysis. Subsequently, the split mask function was performed to automatically separate the femur head and sacrum from the affected hemipelvis. The split mask function works as follows:

1. Search for a slice (axial, coronal, or sagittal) where femur and hemipelvis are visible.
2. Select region A and draw roughly over the hemipelvis with the brush.
3. Select region B and draw roughly over the femur with the brush.
4. Then search for a slice where the sacrum is visible. Select region B and draw roughly over the sacrum.
5. Then let the program automatically segment the regions and delete region B. As a result, the hemipelvis is automatically separated from the femur and sacrum.

If necessary, holes were filled with smart fill. The smart fill tool semi-automatically filled selected holes while 3D interpolating.

6.2 *Tibia Plateau Fracture*

6.2.1 *Philips intellispace portal*

First, the CT-scan was converted to 3D-CT using the volume rendering function. Next, thresholding was performed until (almost) all soft tissue was disappeared. Subsequently, the distal femur was erased to obtain a more sufficient top view of the tibia plateau. When there were too many holes in the presented volume, the minimum threshold limit was decreased and soft tissue structures surrounding the tibia and fibula were manually erased.

6.2.2 *Mimics innovation suite*

First, a threshold procedure was performed with the pre-set bone threshold. Also, keep largest part was selected, to automatically erase irrelevant noisy parts. Subsequently, the split mask function was performed to semi-automatically separate the distal femur from the tibia plateau similarly to the segmentation of ACs. Since all of the CT-scans of TPs were derived from patients with an age ranging from 70- to 80-years-old, the bone density was significantly decreased. Therefore, the mesh contained a lot of holes. These holes were filled with smart fill tool. The smart fill tool semi-automatically filled selected holes while 3D interpolating. Due to this function, gaps occurring due to osteoporosis could be filled to obtain a more sufficient 3D mesh.

6.2.3 *DICOM to PRINT*

First, a threshold procedure was performed with the pre-set bone threshold. Subsequently, with the select option the tibia and fibula were selected. Then the selection was inverted and deleted to only contain the tibia and fibula. Since all of the CT-scans of TPs were from patients with an age ranging from 70- to 80-years-old, the bone density was significantly decreased. Therefore, the mesh contained a lot of holes. These holes were manually filled on the axial slices, using the brush function. Due to this function, gaps occurring due to osteoporosis could be filled to obtain a more sufficient 3D mesh.

6.3 *Abdominal Aortic Aneurysm*

6.3.1 *Philips intellispace portal*

The segmentation of AAAs in PISP was divided in two parts. First, the aorta was segmented. As a result of the administered contrast in arterial phase, the abdominal aorta obtained a high HU value. Consequently, thresholding could be used to erase soft tissue without deleting the aorta. The upper boundary and lower boundary of the segmented volume were set just above the coeliac trunk and past the aortic bifurcation, respectively. After exporting the volume as STL-file, the volume rendering was performed again to segment the intraluminal thrombus. Using the 2D cut plane, the intraluminal thrombus was segmented. The right angle of the cut plane needed to be found. Since the contrast was not situated in the intraluminal thrombus, the HU-values were similar to that of soft tissue. Therefore, thresholding could not be applied.

6.3.2 *Mimics innovation suite*

First, the contrast containing aorta was segmented using the dynamic region growing function. This tool is a threshold procedure combined with a region growing function. The threshold value was set at the bone pre-set. Subsequently, a seed was placed somewhere in the contrast containing aorta to automatically segment it. Peripheral ends of the aortic branches were erased using the cut option. Next, the intraluminal thrombus was semi-automatically segmented using the 3D interpolation option. The intraluminal thrombus needed to be segmented once every 5 to 10 axial slices. Finally, the 3D interpolation function interpolated the segmented slices to obtain the intraluminal thrombus.

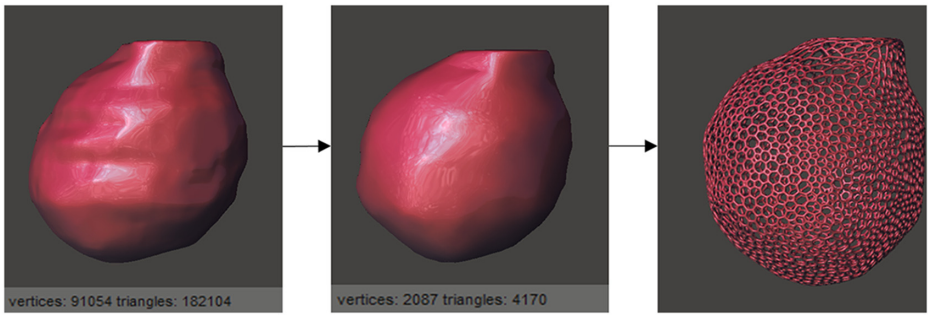

Fig. 11 (a) The intraluminal thrombus mesh imported in Meshmixer after segmentation consists of 91,054 and 182,104 vertices and triangles, respectively. (b) The 3D mesh is down sampled in Meshmixer with the reduce option. The vertices and triangles are reduced by 97.7%. (c) Subsequently, the obtained mesh is converted into a hexagonal mesh with the 'make pattern' option in Meshmixer.

6.3.3 *DICOM to PRINT*

First, the contrast containing aorta was automatically segmented using the automatic vascular segmentation option. Before using this function, the ROI was set just above the coeliac trunk and past the aortic bifurcation. Next, the automatic vascular segmentation tool was used to automatically segment the contrast containing aorta. The intraluminal thrombus was segmented similarly to the segmentations in MIS, using the 3D interpolation function.

7 Appendix B: Process of Converting Intraluminal Thrombus Volume to Hexagonal Mesh

To obtain the hexagonal mesh, the intraluminal thrombus was separately segmented. Subsequently, the amount of triangular faces of the intraluminal thrombus mesh was significantly decreased by down sampling it in Meshmixer. Consequently, a hexagonal mesh with a sufficient grid resolution could be obtained (Fig. 11).

8 Appendix C: MATLAB Script for DSC Computations

```
%% Volumetric comparison segmentation study - intra-observer
agreement
  close all
  clear all

  %% Volume 1
  % step 1: create 3D matrix
  seg_dir1=uigetdir([],...
  'Please select folder containing DICOM files of volume 1'); % import
    DICOM files of first segmented volume
  seg1=dicomreadVolume(seg_dir1);    % create 4D matrix from
    dicomfiles
  Seg1=squeeze(seg1);    % squeeze dimensions

  %% step 2: make binary
  Seg1_bin=logical(Seg1==2967);    % convert in binary volume

  %% Volume 2
  % step 1: create 3D matrix
  seg_dir2=uigetdir([],...
```

```
'Please select folder containing DICOM files of volume 2');
   % import DICOM files of first segmented volume
seg2=dicomreadVolume(seg_dir2);    % create 4D matrix from
   dicomfiles
Seg2=squeeze(seg2);    % squeeze dimensions

%% step 2: make binary
Seg2_bin=logical(Seg2==3278);    % convert in binary volume

%% Volume 3
% step 1: create 3D matrix
seg_dir3=uigetdir([],...
'Please select folder containing DICOM files of volume 3');
   % import DICOM files of first segmented volume
seg3=dicomreadVolume(seg_dir3);    % create 4D matrix from
   dicomfiles
Seg3=squeeze(seg3);    % squeeze dimensions

%% step 2: make binary
Seg3_bin=logical(Seg3==3585);    % convert in binary volume

%% Calculate DICE-coefficient
% DICE 1-2
VoxelsNumber1=sum(Seg1_bin(:));    % total sum of voxels in seg 1
VoxelsNumber2=sum(Seg2_bin(:));    % total sum of voxels in seg 2
CommonArea12=sum(Seg1_bin(:) & Seg2_bin(:));    % find common
   area between seg 1 & 2
DICE12=(2*CommonArea12)/(VoxelsNumber1+VoxelsNumber2);    % cal-
   culate dice seg 1 & 2

% DICE 1-3
VoxelsNumber1=sum(Seg1_bin(:));    % total sum of voxels in seg 1
VoxelsNumber3=sum(Seg3_bin(:));    % total sum of voxels in seg 3
CommonArea13=sum(Seg1_bin(:) & Seg3_bin(:)); % find common area
   between seg 1 & 3
DICE13=(2*CommonArea13)/(VoxelsNumber1+VoxelsNumber3);  %  cal-
   culate dice seg 1 & 3

% DICE 2-3
VoxelsNumber2=sum(Seg2_bin(:)); % total sum of voxels in seg 2
VoxelsNumber3=sum(Seg3_bin(:)); % total sum of voxels in seg 3
CommonArea23=sum(Seg2_bin(:) & Seg3_bin(:));     % find common
   area between seg 2 & 3
DICE23=(2*CommonArea23)/(VoxelsNumber2+VoxelsNumber3);    % cal-
   culate dice seg 2 & 3
% mean DICE of DICE 1, 2 and 3
meanDICE=(DICE12+DICE13+DICE23)/3      % mean DICE of DICE 1, 2
   and 3
```

Disclosures

The authors declare that they have no conflict of interest.

Acknowledgments

The authors received no financial support for the research and/or publication of this article.

References

1. A. Aimar, A. Palermo, and B. Innocenti, "The role of 3D printing in medical applications: a state of the art," *J. Healthcare Eng.* **2019**, 1–10 (2019).
2. D. Shilo et al., "Printing the future—updates in 3D printing for surgical applications," *Rambam Maimonides Med. J.* **9**(3), e0020 (2018).
3. L. Brouwers, "The value of 3D printed models in understanding acetabular fractures," *3D Print. Addit. Manuf.* **5**(1), 37–46 (2018).
4. V. Bagaria et al., "Use of rapid prototyping and three-dimensional reconstruction modeling in the management of complex fractures," *Eur. J. Radiol.* **80**(3), 814–820 (2011).
5. W. Zheng et al., "The feasibility of 3D printing technology on the treatment of pilon fracture and its effect on doctor-patient communication," *Biomed. Res. Int.* **2018**, 8054698 (2018).
6. Y. H. Chen, C. T. Ng, and Y. Z. Wang, "Generation of an STL file from 3D measurement data with user-controlled data reduction," *Adv. Manuf. Technol.* **15**(2), 127–131 (1999).
7. N. B. Watts, "Postmenopausal osteoporosis: a clinical review," *J. Women's Heal.* **27**(9), 1093–1096 (2018).
8. T. M. Link, "Radiology of osteoporosis," *Can. Assoc. Radiol. J.* **67**(1), 28–40 (2016).
9. R. Canellas et al., "Characterization of portal vein thrombosis (neoplastic versus bland) on CT images using software-based texture analysis and thrombus density (Hounsfield units)," *Am. J. Roentgenol.* **207**(5), W81–W87 (2016).
10. A. Piechota-Polanczyk et al., "The abdominal aortic aneurysm and intraluminal thrombus: current concepts of development and treatment," *Front. Cardiovasc. Med.* **2**, 1–14 (2015).
11. F. Milletari, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *Fourth Int. Conf. 3D Vision*, pp. 567–573 (2016).
12. N. Byrne et al., "A systematic review of image segmentation methodology, used in the additive manufacture of patient-specific 3D printed models of the cardiovascular system," *JRSM Cardiovasc. Dis.* **5**, 1–9 (2016).
13. M. Jermyn et al., "Fast segmentation and high-quality three-dimensional volume mesh creation from medical images for diffuse optical tomography," *J. Biomed. Opt.* **18**(8), 086007 (2013).
14. G. L. Presti et al., "Assessment of DICOM viewers capable of loading patient-specific 3D models obtained by different segmentation platforms in the operating room," *J. Digit. Imaging* **28**(5), 518–527 (2015).
15. L. Brouwers et al., "Validation study of 3D-printed anatomical models using 2 PLA printers for preoperative planning in trauma surgery, a human cadaver study," *Eur. J. Trauma Emerg. Surg.* **45**(6), 1013–1020 (2018).
16. E. Arts et al., "The value of 3D reconstructions in determining post-operative reduction in acetabular fractures: a pilot study," *Eur. J. Trauma Emerg. Surg.* (2019).
17. H. R. Roth et al., "Deep learning and its application to medical image segmentation," *Med. Imaging Technol.* **36**(2), 63–71 (2018).
18. H. R. Roth et al., "Anatomy-specific classification of medical images using deep convolutional nets," in *Proc. Int. Symp. Biomed. Imaging*, pp. 101–104 (2015).

Niek Wijnen received his BS degree in technical medicine from the University of Twente, The Netherlands, in 2019. During his internship in 2020, he researched the in-hospital 3D print workflow efficiency and effectivity at the Department of Surgery at the Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands. Currently, he is pursuing his MSc degree in both medicine and clinical research at the University of Utrecht, The Netherlands.

Lars Brouwers is a surgeon-in-training and a PhD candidate in the Department of Trauma Surgery at the Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands. He is co-founder of the medical 3D lab in the Elisabeth-Tweesteden Hospital, co-supervisor of the Radboud University Medical Center PhD project: 3D printing in Sierra Leone, founder of the 3D lab Masanga Sierra Leone and co-founder of the Dutch surgical podcast. His current research interest includes medical 3D printing and virtual/augmented reality in health care.

Erik Groot Jebbink received his BS, MS, and PhD degrees from the University of Twente, Enschede, The Netherlands, 2009, 2013, and 2017, respectively. He is currently an assistant professor at the M3i group, University of Twente and technical physician at the Department of Vascular Surgery at the Rijnstate Hospital, Arnhem. His research interests focus on the interaction between blood flow and stents, based on *in vitro* and *in vivo* measurements. He received a NWO-ZonMW-VENI personal grant in 2020.

Jan M.M. Heyligers is a vascular surgeon at the Department of Surgery at the Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands. He received his PhD in durability of artificial blood vessels at the Department of Vascular Surgery, University Medical Center Utrecht, The Netherlands, in 2006. In 2008, he became fellow of the European Board of Vascular Surgeons. He also is cofounder of the medical 3D lab in the Elisabeth-Tweesteden Hospital.

Mike Bemelman is a trauma and military surgeon at the Department of Surgery at the Elisabeth-Tweesteden Hospital, Tilburg, The Netherlands since 2013. He also is co-founder of the medical 3D lab in the Elisabeth-Tweesteden Hospital and a pioneer in the application of 3D printing in trauma surgery and shared decision making. In addition, he is chairman of the disaster & military section for the European Society of Trauma and Emergency Surgery.