The International Journal of Cardiovascular Imaging (2019) 35:2067–2076
https://doi.org/10.1007/s10554-019-01646-1

ORIGINAL PAPER

Creating three dimensional models of the right ventricular outflow tract: influence of contrast, sequence, operator, and threshold

Barbara E. U. Burkhardt[1,2] · Nicholas K. Brown[1] · Jaclyn E. Carberry[3] · Marí Nieves Velasco Forte[3] · Nicholas Byrne[3] · Gerald Greil[1,3,4] · Tarique Hussain[1,3,4] · Animesh Tandon[1,4]

Received: 8 March 2019 / Accepted: 7 June 2019 / Published online: 15 June 2019

Abstract
The use of 3D printed models of the right ventricular outflow tract (RVOT) for surgical and interventional planning is growing and often requires image segmentation of cardiac magnetic resonance (CMR) images. Segmentation results may vary based on contrast, image sequence, signal threshold chosen by the operator, and manual post-processing. The purpose of this study was to determine potential biases and post-processing errors in image segmentation to enable informed decisions. Models of the RVOT and pulmonary arteries from twelve patients who had contrast enhanced CMR angiography with gadopentetate dimeglumine (GPD), gadofosveset trisodium (GFT), and a post-GFT inversion-recovery (IR) whole heart sequence were segmented, trimmed, and aligned by three operators. Geometric agreement and minimal RVOT diameters were compared between sequences and operators. To determine the contribution of threshold, interoperator variability was compared between models created by the same two operators using the same versus different thresholds. Geometric agreement by Dice between objects was high (intraoperator: 0.89–0.95; interoperator: 0.95–0.97), without differences between sequences. Minimal RVOT diameters differed on average by − 1.9 to − 1.3 mm (intraoperator) and by 0.4 to 1.4 mm (interoperator). The contribution of threshold to interoperator geometric agreement was not significant (same threshold: 0.96 ± 0.06, different threshold: 0.93 ± 0.05; $p = 0.181$), but minimal RVOT diameters were more variable with different versus constant thresholds (− 9.12% vs. 2.42%; $p < 0.05$). Thresholding does not significantly change interoperator variability for geometric agreement, but does for minimal RVOT diameter. Minimal RVOT diameters showed clinically relevant variation within and between operators.

Keywords Magnetic resonance angiography · Contrast agent · 3D model · Volume rendering · Image interpretation · Structural heart disease

Introduction

Technological developments in 3D printing and other advanced visualization technologies, such as virtual and augmented reality, have opened the door to novel ways of planning patient-specific surgical and catheter-based interventions for patients with congenital heart disease [1, 2], especially planning transcatheter interventions in the right ventricular outflow tract [3]. To create these 3D representations of anatomy, imaging datasets must undergo image segmentation, wherein the individual voxels of the data are assigned to the structure(s) of interest to create the 3D shapes of the structures.

Ideally, the process used to create these models should yield accurate and consistent models of patient anatomy. For the source images, this would entail a high signal-to-noise ratio, high edge sharpness of the blood pool, and low contamination of the structures of interest; while for image segmentation, there should also be low contamination of the structures of interest, low influence of the choice of segmentation threshold, and high interoperator reproducibility. In addition, because creation of 3D models can

✉ Barbara E. U. Burkhardt
Barbara.burkhardt@kispi.uzh.ch

[1] Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX, USA

[2] Pediatric Cardiology, Department of Surgery, Pediatric Heart Center, University Children's Hospital Zurich, Steinwiesstr. 75, 8032 Zurich, Switzerland

[3] Division of Imaging Sciences and Biomedical Engineering, King's College London, London, UK

[4] Department of Biomedical Engineering, University of Texas Southwestern Medical Center, Dallas, TX, USA

be time-consuming [4], it is important to identify areas for improvement to be incorporated into future automated segmentation methods. The ideal methods of imaging and segmentation for creating models of congenital heart disease have not been well-studied.

Thus, we focused on two questions essential to improving image segmentation: (1) is one of the common cardiac magnetic resonance (CMR) imaging techniques superior for 3D model creation, and (2) what are the sources of variation in the models that are created? To evaluate model variation, we used Dice coefficients (DC) to assess overall geometric agreement [5] and the minimum circumference or, derived from this, the minimal idealized diameter of the right ventricular outflow tract (RVOT), since this would be important for transcatheter intervention planning [6].

CMR is a recommended imaging modality for evaluation of patients with congenital heart disease [7, 8]. A respiratory-triggered, post-contrast inversion-recovery 3D balanced steady-state free precession whole heart acquisition (3D IR bSSFP) obtained in the end-systolic or mid-diastolic cardiac rest phase was previously found to have improved vessel wall sharpness, contrast-to-noise ratios, and image quality compared to non-electrocardiogram (ECG) gated, contrast-enhanced, time-resolved MR angiography (TRMRA) [9]. We therefore hypothesized that 3D IR bSSFP would yield the most reproducible models, when compared to TRMRA with extravascular or intravascular contrast media. We expected the IR sequence to show less variation of resulting geometric agreement scores and minimal circumferences, because of its electrocardiographic triggering in a cardiac rest period.

In terms of intra- and interoperator reproducibility, there is little previous literature in cardiology about the reproducibility of image segmentation. We showed previously that manual segmentation results in higher interoperator disagreement than automatic segmentation, especially in contaminated datasets [10]. We hypothesized that two independent segmenters would have similar endpoints for model creation regardless of the image brightness threshold used to start, suggesting that thresholding was not the most important part of the process.

It is vital to understand the potential pitfalls and sources of variation and error in image segmentation, because it is becoming more commonly used today for patient-specific procedural planning, and there is no standard method of performing it.

Materials and methods

Study design

A schematic of the overall study design is shown in Fig. 1.

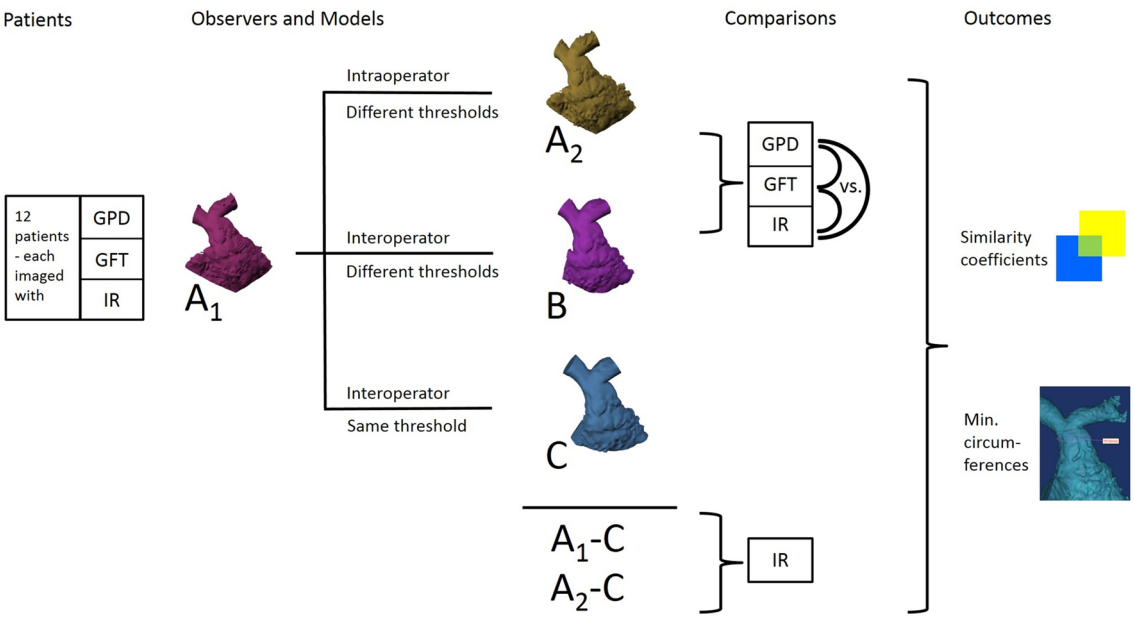

Fig. 1 Study design. *GPD* gadopentetate dimeglumine (Magnevist®), *GFT* gadofosveset trisodium (Vasovist®), *IR* post-GFT inversion recovery three-dimensional balanced steady state free precession whole heart sequence

Patients and image acquisition

All patients provided informed consent to participate in the protocol approved by the local ethics board. CMR datasets from twelve patients with congenital heart disease who had undergone CMR imaging on two subsequent days at a single center for a contrast and sequence comparison study were anonymized for analysis. Inclusion criteria were repaired or unrepaired congenital heart defects, exclusion criteria were stents or other metal implants in the RVOT or the pulmonary arteries. All patients had 3D acquisitions by a contrast-enhanced TRMRA using gadopentetate dimeglumine (Magnevist®)(GPD), a TRMRA with gadofosveset trisodium (Vasovist®)(GFT), and a post-GFT IR whole heart MR angiography sequence as described previously. Images were obtained with a 1.5 T magnetic resonance scanner (Achieva, Philips Healthcare, Best, Netherlands) using a 32-channel cardiac coil.

Imaging parameters were the following: For contrast-enhanced TRMRA, repetition time (TR) was 4.2 ms, echo time (TE) 1.3 ms, flip angle (FA) 40 degrees, bandwidth 434 Hz/pixel, field of view (FOV) 340 mm × 269 mm, voxel size $1.8 \times 1.8 \times 1.8$ mm^3, acceleration factor for sensitivity encoding acquisition 4, acquisition time 12 s. IR-SSFP was obtained with TR 4.5 ms, TE 2.1 ms, TI 260–280 ms, FA 90 degrees, bandwidth 542 Hz/pixel, FOV 340 mm × 340 mm, voxel size $1.4 \times 1.4 \times 1.4$ mm^3, acceleration factor 4, acquisition time 230 – 248 s, navigator efficiency $53 \pm 9\%$, with diastolic triggering in 10 cases, systolic triggering in 1 case, and unknown in 1 case.

Image segmentation process

The 3D datasets were oriented in Mimics, version 19.0 (Materialise Medical NV, Belgium). The overall image segmentation process has been described in depth previously [10]. Briefly, first, lower and upper image brightness thresholds were chosen and visually inspected to allow as much of the structures of interest to remain and noise to be eliminated as possible. Then, seeds were placed in the cardiac chambers, including the right atrium, right ventricle, and pulmonary artery, and initial segmentation masks were generated by the CT Heart tool in Mimics. The segmenters then manually and independently modified structures where automatic detection was considered suboptimal compared to the underlying source data, until the objects gave the best representation of the anatomy in their individual judgment. The masks were combined to form a single right heart blood pool mask. From the resulting mask, a 3D model was then calculated with a contour interpolation method preferring continuity, XY resolution of 0.8854 mm, Z resolution of 0.75 mm, and advanced edge triangle reduction (tolerance 0.1107 mm, edge angle 10 degrees), and exported into 3-matic, version 11.0 (Materialise, Belgium).

Image sequence comparisons

Intraoperator comparisons

For intraoperator comparisons, the same segmenter (NB) performed image segmentation semi-automatically for each dataset twice, one month apart, yielding objects A1 and A2 (Fig. 1).

Interoperator comparisons

For interoperator comparisons, a second segmenter (BB) performed image segmentation on each dataset with independent thresholds, yielding objects B (Fig. 1).

Contribution of the image thresholds

To determine the contribution of the image segmentation threshold, we took the best dataset from the image sequence comparisons, and performed segmentation again, with the same thresholds as objects A1, yielding objects C, made by another, independent segmenter (JC). Comparisons were then made for the models A1 vs. B (different threshold) and A1 vs. C (same threshold). This allowed the determination of the contribution of the threshold in an interoperator comparison.

Calculation of geometric agreement (relative spatial overlap)

In 3-matic, two objects made from the same 3D image of the RVOT were imported together. Heart structures were resected to leave only the right ventricle, main, right, and left pulmonary artery. Then, the objects were globally aligned with a goal of 80% overlap and 80 iterations, which was repeated until two instances of 80 iterations each resulted in the same average distance error (no further improvement in alignment). The distal pulmonary arteries, right ventricular inflow, and apex were then trimmed in both objects simultaneously, and loose parts were removed. The alignment was repeated as above, followed by another step of simultaneous object trimming in all areas of previous trimming. Finally, any coincident surfaces of both objects were expanded by a negligible amount (0.001 ml) so as to avoid interference with Boolean operations.

The objects were then compared by Boolean subtractions (A – B and B – A), and the resulting objects of disagreement were united by Boolean union to yield an object of aggregate error. Boolean union was used to create a summation object

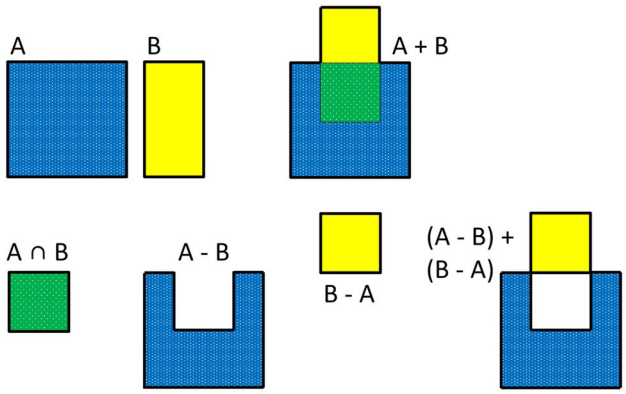

Fig. 2 Two-dimensional representation of shape overlap and error principles. Shapes A and B overlap partially. A + B depicts the sum of both shapes. A ∩ B shows the overlap between objects A and B. A − B and B − A are subtraction objects of the originals. (A − B) + (B − A) is the aggregate error

of both segmentations combined (A + B). Overlap between two objects was calculated as their summation volume minus their aggregate error. Figure 2 shows the principles of operations and the resultant objects on two-dimensional examples.

DC also take into account the spatial alignment between two objects. They can range from 0 (no similarity) to 1 (identical). DC were calculated as the overlap between two objects divided by the arithmetic mean of the volumes of the two objects [11].

$$DC = |A \cap B| / 0.5 \times (|A| + |B|)$$

Calculation of minimal diameter of RVOT

A centerline was fit in Mimics inside the right ventricular outflow tract (RVOT) and main pulmonary artery (MPA) with a resolution of 0.5 mm, and the minimum circumference of each object was measured. The idealized minimal diameter d (assuming a round cross-section as would result from implantation of a prosthetic valve) was calculated as $d = c/\pi$, with c = minimal vessel circumference.

Statistical methods

Statistical analysis was carried out with IBM SPSS Statistics 24 (IBM Corporation, Armonk, NY, United States). Continuous variables were tested for normal distribution by the one-sample Kolmogorov–Smirnov test and are expressed as mean ± standard deviation (SD), or as median and interquartile range, as appropriate.

Bland–Altman graphs were generated for assessment of intra- and interoperator bias. Thresholds, minimal RVOT circumferences and geometric agreement coefficients were compared by paired sample t-tests for groups with normal distribution and by Wilcoxon signed ranks tests for non-normally distributed variables.

Interoperator variability of diameters was assessed by coefficients of variance (SD of the differences of measurements, divided by the mean of all measurements) and intraclass correlation coefficients using a two-way mixed effects model.

Comparisons between imaging sequences were carried out with the maximum number of images available, if one or more datasets were missing. Statistical significance was defined as $p < 0.05$.

Results

Patients and images

Datasets from twelve patients with congenital heart defects but without RVOT or pulmonary artery stents or other metal implants were available for segmentation in at least one modality (IR, GPD, GFT). Nine patients had all three datasets available, two patients had IR and GPD (Magnevist®) images, but the GFT (Vasovist®) dataset was corrupted, likely from mistriggering, in one and missing in the other, and one patient had a GPD (Magnevist®) dataset only. Diagnoses included left-sided obstructive lesions in five patients (aortic coarctation with or without bicuspid aortic valve), right-sided obstructions without metal stents or prostheses in five patients (tetralogy of Fallot, pulmonary atresia), d-transposition of the great arteries status post atrial switch operation in 1 (for this study, the subpulmonary outflow tract and the pulmonary arteries were used), and systemic vein anomaly status post atrial redirection in one patient. Patients' mean weight (± SD) was 76.8 (± 12.7) kg, heart rate 66 (± 9) beats per minute.

Image sequence comparisons

Intraoperator comparison

No intraoperator differences in DC were found between sequences regarding the attainable geometric agreement of objects (Table 1).

Intraoperator variability in minimal diameter Minimal diameters of the resulting 3D objects correlated quite well in intraoperator comparisons, especially those from IR, as shown in Table 2, albeit with large coefficients of variation and limits of agreement. The intraoperator limits of agreement for minimal RVOT diameters between segmentations were spread across zero with a span of 7.38 mm using the IR sequence; 12.81 mm using the GPD sequence; and 12.56 mm using the GFT sequence.

Table 1 High geometric agreement of objects and comparable agreement scores between sequences

	Intraoperator comparison (A1 vs. A2) Dice coefficient	Interoperator comparison (A1 vs. B) Dice coefficient
IR[a]	0.89 ± 0.10	0.97 (0.96–0.98)
GPD[b]	0.95 (0.9–0.95)	0.95 (0.93–0.97)
GFT[c]	0.89 ± 0.07	0.95 ± 0.03
IR vs. GPD	p = 0.594	p = 0.131
IR vs. GFT	p = 0.99	p = 0.086
GPD vs. GFT	p = 0.401	p = 0.214

Dice coefficients are compared between imaging sequences (Wilcoxon signed-rank or paired *t* tests). Normally distributed values are expressed as mean ± standard deviation, values with other distribution as median (interquartile range)

[a]IR—post-GFT inversion recovery sequence

[b]GPD—angiography with gadopentetate dimeglumine (Magnevist®)

[c]GFT—angiography with gadofosveset trisodium (Vasovist®)

Figure 3 shows Bland–Altman plots of intraoperator variability in minimal RVOT circumferences for the different sequences.

Interoperator comparison

No interoperator differences in DC were found between sequences regarding the attainable geometric agreement of objects (Table 1).

Interoperator variability in minimal diameter Minimal diameters of the resulting 3D objects correlated very well in interoperator comparisons, especially with IR and GFT, as shown in Table 2. However, coefficients of variation and limits of agreement were not negligible. The limits of agreement between observers for minimal RVOT diameters were spread across zero with a span of 3.1 mm using the IR sequence; 10.18 mm using the GPD sequence; and 3.94 mm using the GFT sequence.

Figure 4 shows Bland–Altman plots of interoperator variability in minimal RVOT circumferences for the different sequences. The smallest mean deviation between operators is achieved by the IR sequence, which has an only mildly higher standard deviation than the GFT sequence.

Blood pool thresholds

The lower thresholds (lowest desirable signal) and threshold ranges (from lowest up to highest desirable signal) chosen by the different segmenters were similar between the TRMRA sequences. Lower signal thresholds of all observers were lower for IR than for MRA with GPD (202 ± 82 vs. 404 ± 136; p < 0.001) and with GFT (398 ± 177; p < 0.001).

There was mostly no correlation between each observer's chosen lower threshold and the resultant minimal RVOT diameter within IR datasets (A1: p = 0.198, A2: p = **0.009**, B: p = 0.056, C: p = 0.285), within GPD (A1: p = **0.042**, A2: p = 0.397, B: p = 0.897, C: p = 0.076), or within GFT datasets (A1: p = **0.042**, A2: p = 0.546, B: p = 0.088, C: p = 0.076).

Contribution of the image thresholds

As the 3D IR bSSFP sequence showed favorable intra- and interoperator reliability, DC were calculated for this sequence using different threshold conditions (Table 3).

The interoperator DC using the same threshold (objects A1 vs. C) was 0.96 ± 0.06, and 0.93 + 0.05 using different thresholds (objects A2 vs. C), which amounts to a difference in DC of 0.03 (p = 0.181).

Bland–Altman plots illustrate interoperator comparisons of minimal RVOT diameters, where the same threshold leads to a mean difference of 2.42% with a standard deviation of 7.65%, while different thresholds lead to a mean difference of − 9.12% with a standard deviation of 7.61% (p = 0.011) (Fig. 5).

Discussion

Factors such as imaging sequence, the brightness threshold chosen by the operator, and operator-dependent manual post-processing may cause variations in 3D reconstructions of CMR images. Knowledge of the potential biases and post-processing errors is required to make informed decisions before using patient-specific objects from those images for intervention or surgical planning. High geometric agreement coefficients, i.e. relative overlap of shapes between segmented objects, and not only of object volumes, are important for intervention planning, because adjacent structures (muscular RVOT, pulmonary artery branches) must be oriented correctly within the segmented model.

There is no gold standard for in vivo measurements. X-ray angiography, for example, which is used by interventionalists to guide their choice of equipment, provides a projectional summation image rather than a three-dimensional representation of the true shape.

Comparison of TRMRA with an extravascular (GPD) and an intravascular (GFT) contrast medium versus post-GFT contrast 3D IR bSSFP shows higher brightness thresholds chosen for both contrast sequences than for the latter, which is not surprising, given rapid TRMRA image acquisition after contrast infusion, so that no relevant diffusion of GPD into the extravascular space would have happened yet. This

Table 2 Excellent agreement, but large 95% confidence intervals between minimal right ventricular outflow tract diameters by acquisition sequence

	A1	A2	B	Intraoperator comparison (A1 vs. A2)	Interoperator comparison (A1 vs. B)
IR[a]	26.5 ± 4.5	24.4 (22.8–26.4)	23 ± 5.8	ICC[d] 0.978	ICC 0.995
				p < 0.001	p < 0.001
				MD[e] − 1.50	MD 0.36
				LoA[f] − 5.19 to 2.19	LoA − 1.19 to 1.91
				CoV[g] 7.8%	CoV 3.4%
GPD[b]	27 ± 4.7	26.1 ± 6.8	25.4 (24.4–27.7)	ICC 0.919	ICC 0.954
				p < 0.001	p < 0.001
				MD − 1.33	MD 1.39
				LoA − 7.74 to 5.07	LoA − 3.70 to 6.48
				CoV 12.3%	CoV 10.2%
GFT[c]	27.4 ± 7.3	27.3 (25.3–27.5)	25.2 (24.5–26.8)	ICC 0.938	ICC 0.991
				p = 0.001	p < 0.001
				MD − 1.94	MD 1.03
				LoA − 8.22 to 4.34	LoA − 0.94 to 3.00
				CoV 12.1%	CoV 4.0%
IR vs. GPD	p = **0.034**	p = **0.004**	p = **0.008**		
IR vs. GFT	p = 0.53	p = **0.021**	p = 0.086		
GPD vs. GFT	p = 0.35	p = **0.038**	p = **0.021**		

Minimal right ventricular outflow tract diameters as expressed in mm. Normally distributed values are given as mean ± standard deviation (SD), non-normally distributed values as median (interquartile range). The lower part of the table shows comparisons between sequences for each operator

[a]IR—post-GFT inversion recovery sequence

[b]GPD—angiography with gadopentetate dimeglumine (Magnevist®)

[c]GFT—angiography with gadofosveset trisodium (Vasovist®)

[d]ICC—intraclass correlation coefficient using a two-way mixed effects model and p = error for intraclass correlation

[e]MD—mean difference in minimal object diameters (in mm)

[f]LoA—limits of agreement encompass the 95% confidence interval around the average difference between operators (in mm)

[g]CoV—coefficient of variation = SD of differences between two diameter measurements, divided by mean of two measurements

Statistically significant values are given in bold

suggestion is supported by similar first-pass characteristics of GFT versus the standard contrast agent gadobenate dimeglumine on TRMRA regarding signal-to-noise and contrast-to-noise ratios found by Frydrychowicz et al., even though their acquisitions were made at 3 T [12]. In fact, geometric agreement of objects made by different observers was similar with all sequences in our study (Table 1). Minimal RVOT diameters were measured smaller from the 3D IR bSSFP sequence than from GPD, and partly also from GFT (Table 2). This may be a consequence of ECG triggering, which reduces motion artifacts during image acquisition with the IR sequence. The 3D IR bSSFP sequence also showed superior performance regarding intra- and interoperator reliability for minimal object diameters (Table 2; Figs. 3, 4).

Besides assessing the fitness of imaging sequences for pre-interventional 3D segmentation, our comparisons provide a benchmark for intra- and interoperator variability in cardiovascular 3D blood pool modelling, since this has not been well studied so far despite increasing interest in pre-interventional 3D printing [10, 13]. We wanted to examine if different operators' manipulations while creating blood pool models resulted in significant differences in the final object. Interoperator geometric disagreement might then be overcome in part by automated segmentation, especially for contaminated datasets [10].

Geometric agreement between the objects of different operators was mostly in the 90% range. With dissimilarities of 3–19%, three-dimensional reconstructions can vary significantly depending on the circumstances of segmentation. Geometric agreement scores were not worse if models had been processed by different operators than by the same one, so different operators seem not to introduce too much variability. Other factors than the choice of operator

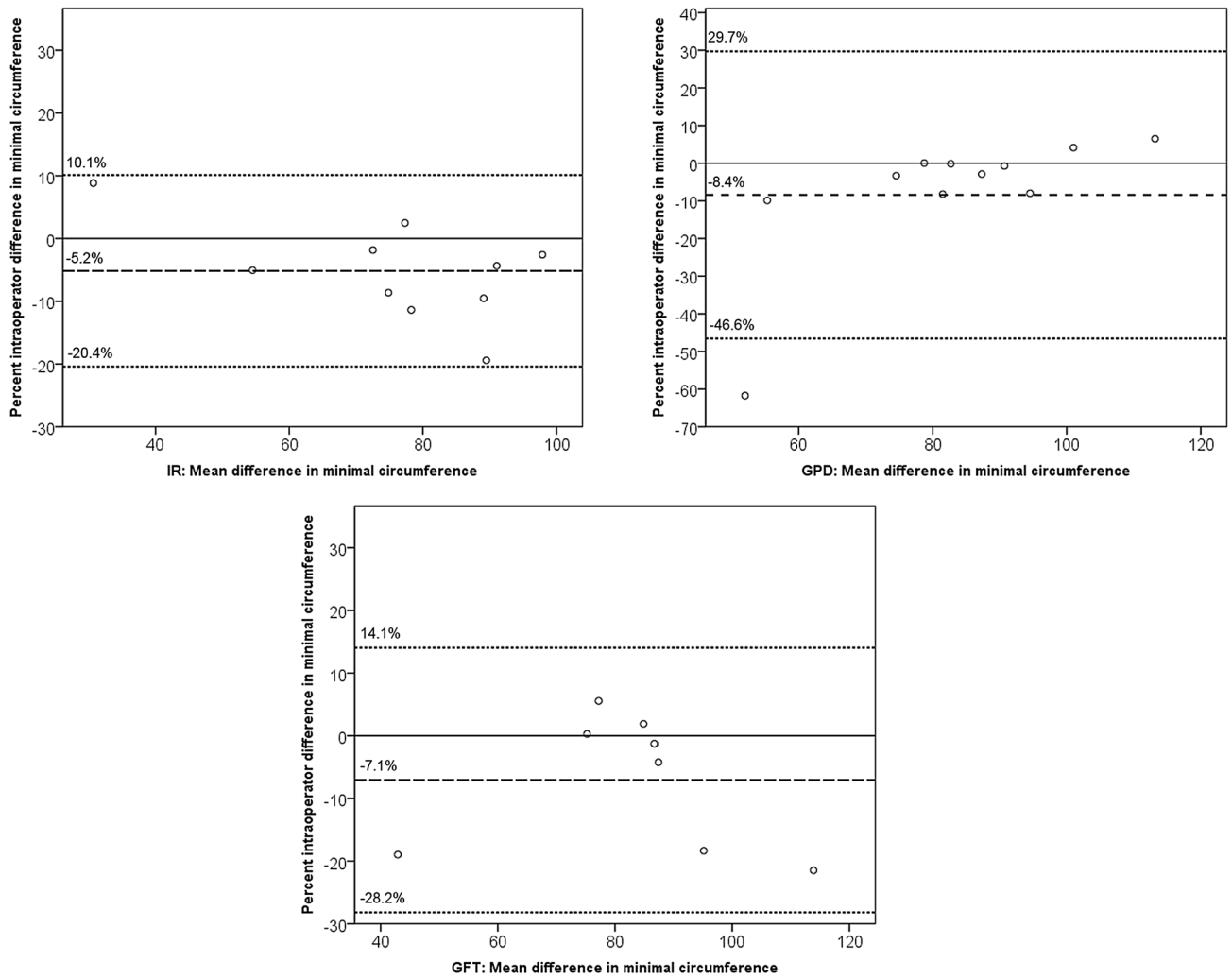

Fig. 3 Intraoperator variability in minimal right ventricular outflow tract circumference. Left panel: *IR* post-GFT inversion recovery sequence. Middle panel: *GPD* angiography with gadopentetate dimeglumine (Magnevist®). Right panel: *GFT* angiography with gadofosveset trisodium (Vasovist®). Dashed lines mark the mean difference as a percentage of object circumference, dotted lines mark its 95% confidence interval, and the solid line indicates perfect agreement

must play a role in variations in the patient-specific 3D models.

The question of thresholding is very important to developments of automated segmentation, with specific problems of fixed thresholding methods used to determine relevant brightness on MR images [14]. Therefore, some user interaction is typically needed to influence the choice of brightness threshold.

To determine the contribution of the chosen threshold to the total variation in DC, we performed interoperator comparisons using either a consensus threshold or individual thresholds between operators.

The difference in mean geometric agreement scores between the resulting objects was 0.03 according to Dice, whether the threshold was kept constant between observers or not, which corresponds to approximately half of the deviation from perfect overlap, but was not statistically significant. This means that the threshold used is not the most significant source of variation, but the same amount of variation is due to the individual's choices in the segmentation process. Thresholding is only the first step in the model creation workflow, and manual additions and subtractions may diversify the resulting objects again, whether made by different operators or by the same operator on a different day.

The reproducibility of minimal RVOT diameters does, however, vary with the chosen threshold, as was evident in our Bland–Altman analysis, showing that different thresholds lead to a higher mean interoperator difference (with a similar SD) than a constant threshold (Fig. 5).

An idealized diameter was calculated from the minimal circumferences measured, because standard prosthetic valves or stents are designed for a uniform diameter and will

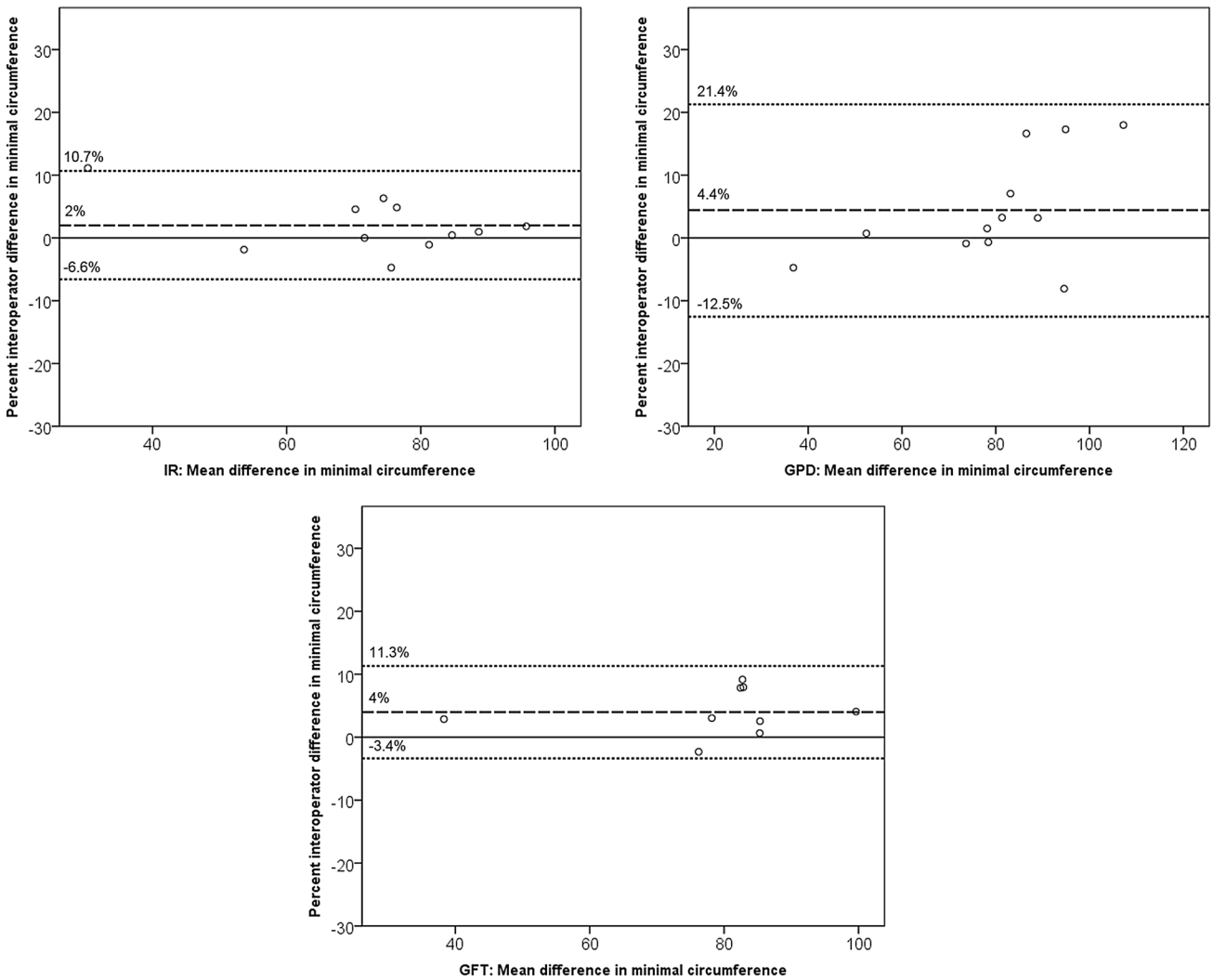

Fig. 4 Interoperator variability in minimal right ventricular outflow tract circumference. Left panel: *IR* post-GFT inversion recovery sequence. Middle panel: *GPD* angiography with gadopentetate dimeglumine (Magnevist®). Right panel: *GFT* angiography with gadofosveset trisodium (Vasovist®). Dashed lines mark the mean difference as a percentage of object circumference, dotted lines mark its 95% confidence interval, and the solid line indicates perfect agreement

Table 3 Influence of thresholds on geometric agreement and minimal diameters of 3D RVOT models, based on the 3D IR bSSFP sequence

	Interoperator comparison with same thresholds (A1 vs. C)	Interoperator comparison with different thresholds (A2 vs. C)	p value
Dice coefficient	0.96 ± 0.06	0.93 + 0.05	0.181
Mean difference in minimal RVOT diameter	2.4 (− 0.9 to 4)%	− 9.1 ± 7.6%	**0.011**

Dice coefficients and differences in minimal RVOT diameter for interoperator comparisons using the same versus different thresholds. Normally distributed values are expressed as mean ± standard deviation, values with other distribution as mean (interquartile range)

RVOT right ventricular outflow tract, *3D IR bSSFP* post-gadofosveset inversion recovery balanced steady state free precession sequence

Statistically significant value is given in bold

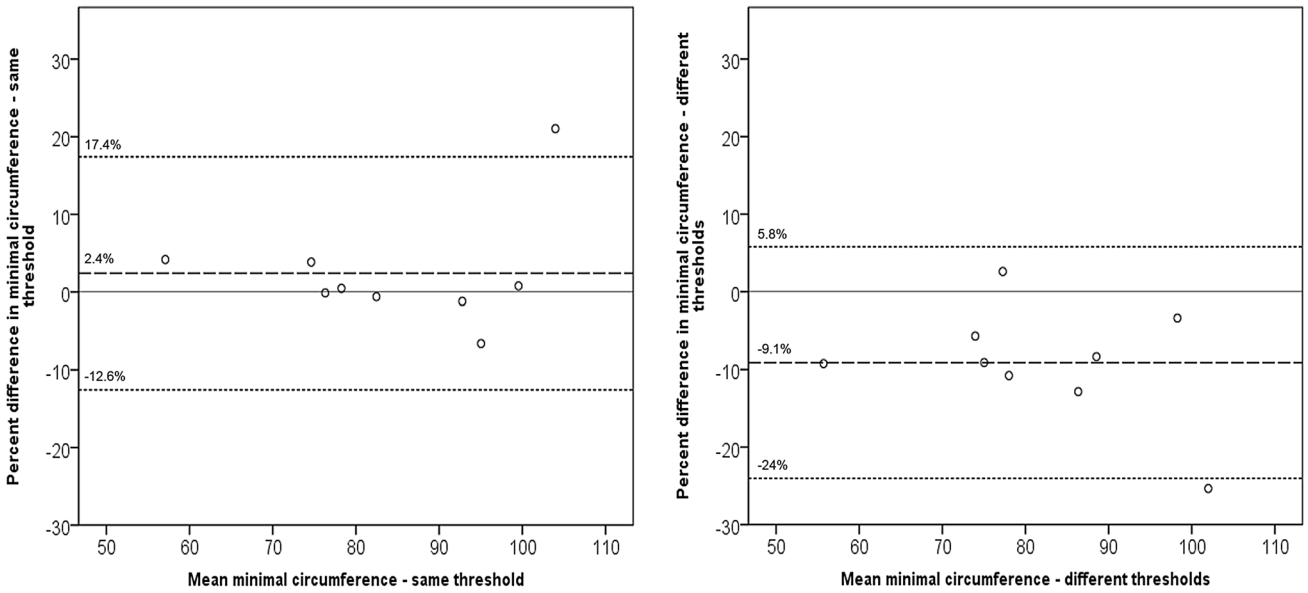

Fig. 5 Interoperator variability in minimal right ventricular outflow tract circumferences with the post-GFT inversion recovery sequence depending on threshold. Left panel: same threshold (objects A1 vs. C); right panel: different thresholds (objects A2 vs. C). Dashed lines mark the mean difference as a percentage of object circumference, dotted lines mark its 95% confidence interval, and the solid line indicates perfect agreement

extend an oval vessel into a circular shape as far as tissue properties allow.

Minimal RVOT diameters differed on average by − 1.9 to − 1.3 mm between objects from the same operator and by 0.4 to 1.4 mm between operators. These differences in minimal diameters between 3D models were similar to reported reproducibilities of aortic diameters in two-dimensional measurements of MRA and 3D SSFP images [15]. An et al. [16] reported the variability of inner diameter measurements of iliac and femoral arteries from post-mortem computed tomography datasets and found low variability (standard deviations of 3–14% of mean inner vessel diameter) in segmentations of these completely immobile arteries with no difference between the vendor software and the software which we also used, comparable to macroscopical direct measurements (standard deviations of 8–9% of mean diameter).

While all limits of agreement left a wide span of possible error between observers, the average span was smallest with the IR sequence, which was ECG triggered. This has similarly been observed with aortic measurements and is most likely due to improved vessel wall sharpness from reduced motion of cardiovascular structures during the cardiac cycle [17].

These differences can be clinically relevant. Therefore, no matter which type of sequence is post-processed, correlation of quantitative information by addition of source images into patient evaluation is imperative when planning interventions.

The segmentations in this study were made from datasets of patients with congenital heart disease. While one could argue that this might not result in normal reference values for segmentation variability, the population is representative of patients who require intervention planning because of their structural cardiovascular abnormalities. One limitation is that gadofosveset, the contrast agent used for two of the imaging sequences in this paper, is no longer available on the market. Finally, this study was performed retrospectively, and therefore some patients were missing some sequences.

In conclusion, 3D models made from TRMRA with intra- and extravascular contrast agents or, especially, from 3D IR bSSFP datasets can be constructed by different observers, with different thresholds, with very good geometric agreement. Differences in minimal RVOT diameters can, however, be clinically relevant when used for intervention planning. These data can also provide insights for the development of automated segmentation algorithms. Similar ideas may be applicable for other interventions, such as transcatheter aortic valve replacement or mitral valve surgeries, where 3D printed models are also sometimes used.

Compliance with ethical standards

Conflict of interest TH and AT have significant ownership interest in VARYFII Imaging, LLC. Imaging, LLC. All the other authors were not consultants or employees for VARYFII Imaging, LLC and had control of inclusion of any data and information that might present a conflict of interest for TH or AT, and had no other conflicts of interest.

Research involving human participants This human study has been approved by the appropriate institutional ethics committee and has therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Farooqi KM, Uppu SC, Nguyen K, Srivastava S, Ko HH, Choueiter N, Wollstein A, Parness IA, Narula J, Sanz J, Nielsen JC (2016) Application of virtual three-dimensional models for simultaneous visualization of intracardiac anatomic relationships in double outlet right ventricle. Pediatr Cardiol 37(1):90–98. https://doi.org/10.1007/s00246-015-1244-z
2. Yoo SJ, van Arsdell GS (2017) 3D printing in surgical management of double outlet right ventricle. Front Pediatr 5:289. https://doi.org/10.3389/fped.2017.00289
3. Valverde I, Sarnago F, Prieto R, Zunzunegui JL (2017) Three-dimensional printing in vitro simulation of percutaneous pulmonary valve implantation in large right ventricular outflow tract. Eur Heart J 38(16):1262–1263. https://doi.org/10.1093/eurheartj/ehw546
4. Byrne N, Velasco Forte M, Tandon A, Valverde I, Hussain T (2016) 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 5:2048004016645467. https://doi.org/10.1177/2048004016645467
5. Taha AA, Hanbury A (2015) Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging 15:29. https://doi.org/10.1186/s12880-015-0068-x
6. Burkhardt BE, Byrne N, Velasco Forte MN, Iannaccone F, De Beule M, Morgan GJ, Hussain T (2018) Evaluation of a modified Cheatham-Platinum stent for the treatment of aortic coarctation by finite element modelling. JRSM Cardiovasc Dis 7:2048004018773958. https://doi.org/10.1177/2048004018773958
7. Valsangiacomo Buechel ER, Grosse-Wortmann L, Fratz S, Eichhorn J, Sarikouch S, Greil GF, Beerbaum P, Bucciarelli-Ducci C, Bonello B, Sieverding L, Schwitter J, Helbing WA, Eacvi Galderisi M, Miller O, Sicari R, Rosa J, Thaulow E, Edvardsen T, Brockmeier K, Qureshi S, Stein J (2015) Indications for cardiovascular magnetic resonance in children with congenital and acquired heart disease: an expert consensus paper of the Imaging Working Group of the AEPC and the Cardiovascular Magnetic Resonance Section of the EACVI. Eur Heart J Cardiovasc Imaging 16(3):281–297. https://doi.org/10.1093/ehjci/jeu129
8. Baumgartner H, Bonhoeffer P, De Groot NM, de Haan F, Deanfield JE, Galie N, Gatzoulis MA, Gohlke-Baerwolf C, Kaemmerer H, Kilner P, Meijboom F, Mulder BJ, Oechslin E, Oliver JM, Serraf A, Szatmari A, Thaulow E, Vouhe PR, Walma E; Task Force on the Management of Grown-up Congenital Heart Disease of the European Society of C, Association for European Paediatric C, Guidelines ESCCfP (2010) ESC Guidelines for the management of grown-up congenital heart disease (new version 2010). Eur Heart J 31(23):2915–2957. https://doi.org/10.1093/eurheartj/ehq249
9. Makowski MR, Wiethoff AJ, Uribe S, Parish V, Botnar RM, Bell A, Kiesewetter C, Beerbaum P, Jansen CH, Razavi R, Schaeffter T, Greil GF (2011) Congenital heart disease: cardiovascular MR imaging by using an intravascular blood pool contrast agent. Radiology 260(3):680–688. https://doi.org/10.1148/radiol.11102327
10. Tandon A, Byrne N, Velasco MLN, Zhang S, Dyer AK, Dillenbeck JM, Greil GF, Hussain T (2016) Use of a semi-automated cardiac segmentation tool improves reproducibility and speed of segmentation of contaminated right heart magnetic resonance angiography. Int J Cardiovasc Imaging 32(8):1273–1279. https://doi.org/10.1007/s10554-016-0906-0
11. Dice LR (1945) Measures of the amount of ecologic association between species. Ecology 26(3):297–302
12. Frydrychowicz A, Russe MF, Bock J, Stalder AF, Bley TA, Harloff A, Markl M (2010) Comparison of gadofosveset trisodium and gadobenate dimeglumine during time-resolved thoracic MR angiography at 3T. Acad Radiol 17(11):1394–1400. https://doi.org/10.1016/j.acra.2010.05.022
13. Meier LM, Meineri M, Qua Hiansen J, Horlick EM (2017) Structural and congenital heart disease interventions: the role of three-dimensional printing. Neth Heart J 25(2):65–75. https://doi.org/10.1007/s12471-016-0942-3
14. Karim R, Bhagirath P, Claus P, James Housden R, Chen Z, Karimaghaloo Z, Sohn HM, Lara Rodriguez L, Vera S, Alba X, Hennemuth A, Peitgen HO, Arbel T, Gonzalez Ballester MA, Frangi AF, Gotte M, Razavi R, Schaeffter T, Rhode K (2016) Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images. Med Image Anal 30:95–107. https://doi.org/10.1016/j.media.2016.01.004
15. Veldhoen S, Behzadi C, Derlin T, Rybczinsky M, von Kodolitsch Y, Sheikhzadeh S, Henes FO, Bley TA, Adam G, Bannas P (2015) Exact monitoring of aortic diameters in Marfan patients without gadolinium contrast: intraindividual comparison of 2D SSFP imaging with 3D CE-MRA and echocardiography. Eur Radiol 25(3):872–882. https://doi.org/10.1007/s00330-014-3457-6
16. An G, Hong L, Zhou XB, Yang Q, Li MQ, Tang XY (2017) Accuracy and efficiency of computer-aided anatomical analysis using 3D visualization software based on semi-automated and automated segmentations. Ann Anat 210:76–83. https://doi.org/10.1016/j.aanat.2016.11.009
17. Potthast S, Mitsumori L, Stanescu LA, Richardson ML, Branch K, Dubinsky TJ, Maki JH (2010) Measuring aortic diameter with different MR techniques: comparison of three-dimensional (3D) navigated steady-state free-precession (SSFP), 3D contrast-enhanced magnetic resonance angiography (CE-MRA), 2D T2 black blood, and 2D cine SSFP. J Magn Reson Imaging 31(1):177–184. https://doi.org/10.1002/jmri.22016