VOLUMETRY OF THE DOMINANT INTRAPROSTATIC TUMOUR LESION: INTERSEQUENCE AND INTEROBSERVER DIFFERENCES ON MULTIPARAMETRIC MRI

Hugh Harvey[1], Matthew R Orton[1], Veronica A Morgan[1], Chris Parker[2], David Dearnaley[2], Cyril Fisher[3], Nandita M deSouza[1]

[1] Cancer Research UK Imaging Centre and [2] Academic Urology Unit and [3] Department of Histopathology at The Institute of Cancer Research, London and Royal Marsden NHS Foundation Trust, UK

Short Title: PROSTATE TUMOUR VOLUMETRY: INTERSEQUENCE AND INTEROBSERVER DIFFERENCES

Type of Manuscript: Full Paper

Funding: Supported by Cancer Research UK (grant number CUK C1060/A808) and NHS funding to the NIHR Biomedical Research Centre and the Clinical Research Facility in Imaging.

Acknowledgements: Supported by Cancer Research UK (grant number CUK C1060/A808) and NHS funding to the NIHR Biomedical Research Centre and the Clinical Research Facility in Imaging. We would also like to thank Kathy May, DELINEATE trial co-ordinator (CCR 3635), Institute of Cancer Research and Royal Marsden NHS Foundation Trust; Karen Thomas, Research Data Management and Statistics Unit, The Royal Marsden NHS Foundation Trust; Alan Thompson, Urological Surgeon, The Royal Marsden NHS Foundation Trust.

Conflict of Interest Statement: The authors have no conflict of interest to declare.

Abstract:

OBJECTIVES: To establish interobserver reproducibility of tumour volumetry on individual multiparametric(mp) prostate MRI sequences, validate measurements with histology and determine whether functional to morphological volume ratios reflect Gleason score.

METHODS: Forty-one men with prostate cancer treated with prostatectomy (cohort 1) or radical radiotherapy (cohort 2) who had pre-treatment mp (T2-W, DW-, DCE-) MRI were studied retrospectively. Dominant intraprostatic lesions (DIPLs) were manually delineated on each sequence and volumes compared between observers (n=40 analysable) and with radical prostatectomy (n=20). Volume ratios of DW- and DCE- to T2-W MRI were documented and compared between Gleason grade 3+3, 3+4 and 4+3 or greater categories.

RESULTS: Limits of Agreement of DIPL volumes between observers were: T2-W 0.9, -1.1 cm^3, DW-MRI 1.3, -1.7 cm^3, DCE-MRI 0.74, -0.89 cm^3. In cohort 1, T2-W volumes overestimated fixed specimen histological volumes, (+33% observer 1, +16% observer 2); DW- and DCE-MRI underestimated histological volume, the latter markedly so (-32% observer 1, -79% observer 2). Differences between T2-W, DW- and DCE-MRI volumes were significant ($p<10^{-8}$). The proportional volume of DW- (73.9$\pm$18.1 % observer 1, 72.5$\pm$21.9% observer 2) and DCE-MRI volume (42.6$\pm$24.6% observer 1, 34.3$\pm$24.9% observer 2) to T2-W volume was significantly different ($p<10^{-8}$) but these volume ratios did not differ between Gleason grade.

CONCLUSIONS: The T2-W MRI DIPL volume variability between observers and with histology best reflects the GTV for radiotherapy planning. The volume of cellular tumour represented by DW-MRI is greater than the vascular (DCE) abnormality; ratios of both to T2-W volume are independent of Gleason score.

ADVANCES IN KNOWLEDGE:

1. Manual volume measurement of tumour is reproducible within $1cm^3$ between observers on all sequences, confirming suitability across observers for radiotherapy planning.

2. Volumes derived on T2-W MRI most accurately represent *in vivo* lesion volumes.

3. The proportion of cellular (DW-) or vascular (DCE-) volume to morphological (T2-W) volume is not affected by Gleason score.

Keywords:

multiparametric MRI; tumour volume; prostate cancer; tumour dominant intraprostatic lesion; histopathological validation, diffusion-weighted, dynamic contrast-enhanced, volumetry

Abbreviations:

MRI – Magnetic Resonance Imaging, CT – Computerized Tomography, DIPL – Dominant IntraProstatic Lesion, OAR – Organ(s)-At-Risk, mpMRI – multiparametric Magnetic Resonance Imaging, ADC – Apparent Diffusion Coefficient, DWI – Diffusion Weighted Imaging, DCE – Dynamic Contrast Enhanced sequence, GTV – Gross Tumour Volume, ROI – Region Of Interest, IRB - Institutional Review Board.

Introduction:

The soft tissue contrast on T2-W Magnetic Resonance Imaging (MRI) is preferred over x-ray computerized tomography (CT) for prostate tumour identification, staging [1-4] and defining the dominant intraprostatic lesion (DIPL) [5]. Furthermore, additional information available from Diffusion-Weighted Imaging (DW-) and Dynamic Contrast Enhanced (DCE-) MRI techniques, collectively termed multiparametric (mp)MRI, may be exploited to improve sensitivity and specificity for tumour identification over T2-W imaging alone [6]. Accurate definition of gross tumour volume (GTV) derived from these images is essential in planning radiation therapy [7], particularly when giving boost doses to the DIPL [8]: overestimation of the GTV increases the risk of radiation-induced complications to Organs-At-Risk (OARs) such as the rectal wall, and underestimation reduces the long-term efficacy of treatment [9]. However, as there is increasing evidence that the volumes defined on individual mpMRI sequences are significantly different from each other [10] and depend on underlying histology [11, 12], the optimal sequence on which to outline the GTV remains to be established.

Traditionally, tumour outlines are done on T2-W images for radiation therapy planning. Although this involves simultaneous viewing of all mpMRI images [13], the specific and independent influence of the DW- and DCE-MRI identified tumour, which may vary with Gleason grade, on the morphological (T2-W) outlines has not been documented. A recent large study showed that the maximum volume measured on mpMRI correlated best with histology [14]. The purpose of this study therefore was to establish the interobserver reproducibility of prostate tumour volumetry on individual sequences obtained from mpMRI, validate the measurements

against histology and determine whether the proportion of cellular (DW-) or vascular (DCE-) volume to morphological (T2-W) volume reflects Gleason score.

Methods:

Patients:

Imaging data was obtained from 41 men with prostate cancer (mean age 66.7 ± 7.6 years, PSA range 3.0 – 32.0 ng/mL, clinical grade T1-T3, Gleason grade 6-8) who had been enrolled consecutively in 2 unrelated prospective studies approved by the local Institutional Review Board (IRB) and had given written consent for use of their data. Acquired images were therefore analysed retrospectively. All patients had mpMRI with positive histology on a standardized 8-10 core randomly sampled transrectal ultrasound guided biopsy done between 4 and 12 weeks previously (median 85 days, range 8-231 days). All patients were treatment naïve at the time of scanning. The first 20 (Cohort 1) were treated with radical prostatectomy and the latter 21 (Cohort 2) underwent radiation therapy with dose-boosting to the DIPL. In cohort 1, mpMRI was performed a mean of 16.7 days (median 12 days, range 1–54) prior to prostatectomy. In Cohort 1 3 were Gleason grade 3+3, 12 were 3+4 and 5 were 4+3 or greater. In Cohort 2, 5 were Gleason grade 3+3, 10 were 3+4 and 6 were 4+3 or greater.

Image Acquisition:

All imaging was done with an endorectal coil. Cohort 1 was studied at 1.5-T and 55ml of room air was used for inflation of the balloon. Cohort 2 was studied at 3-T and the balloon was filled with 60ml of perfluorocarbon to reduce susceptibility artefact. Hyoscine butyl bromide 20 mg was administered intramuscularly in all cases. T2-W images were obtained in 3 planes

orthogonal to the prostate at both field strengths supplemented by position matched DW- and DCE-MRI sequences in the axial plane.

At 1.5-T (Magnetom Avanto, Siemens Medical Systems, Erlangen, Germany), a fast spin-echo T2-W sequence (FSE, TR/effective TE = 5500/96, echo-train length 16, FoV 140mm, matrix size 512 × 512, 20 contiguous slices, slice thickness 3.0 mm) was used together with single-shot echo-planar DW sequence (TR/TE = 2500/69, FoV 200mm, matrix size 128 × 128, 12 contiguous slices, slice thickness 4 mm, b values, 0, 300, 500, and 800 s/mm^2) with three orthogonal diffusion directions, resulting in a rotationally invariant trace image at each *b*-value. A gradient-echo sequence was used for DCE imaging (GRAPPA, TR/TE = 4.1 / 1.77 ms, FoV 300mm, flip angle 30°, matrix 128 x 128, 8 contiguous 5.0mm slices at 90 time points, temporal resolution 3.52 secs during intravenous administration of 0.2 mL/kg of gadopentetate dimeglumine [Magnevist, Bayer Schering Pharma] at 3.0 mL/s). Registration images with the same measurement parameters and positions were acquired with flip angles of 2°, 8°, 16°, 24°, and 30° to enable estimation of T1 before contrast administration.

At 3.0-T (Achieva, Philips, Best, Netherlands), a FSE T2-W sequence was also utilized (FSE, TR 2627 ms, TE 110 ms, FoV 120 mm, slice thickness 2.2 mm, matrix 220 × 184 extrapolated to 256 × 256) together with a single shot echo-planar DW sequence (TR 5000 ms, TE 54 ms, b = 0, 100, 300, 500, 800 s/mm^2, FOV 100 mm, slice thickness 2.2 mm, matrix 80 × 79 extrapolated to 176 × 176). A gradient-echo sequence was used for DCE imaging (3D FFE, SPAIR fat suppression, TR/TE = 4.4 / 2.1 ms, FoV 120mm, flip angle 16°, matrix 76 x 98 extrapolated to 224 x 244, 24 contiguous 2.3mm slices at 20 time points, temporal resolution 12 secs during

intravenous administration of 0.2 mL/kg of gadoterate meglumine [Dotarem, Guerbet, USA] at 2.0 mL/s). Registration images with the same measurement parameters and positions were acquired with a flip angle of 16° to enable estimation of T1 before contrast administration. The phase-encoding gradient was left to right in all cases to minimize motion artefacts in the prostate.

An external pelvic phased-array coil was used to acquire axial T1-W and T2-W images through the pelvis to assess lymph node status as part of the routine clinical examination at both 1.5-T and 3-T, but these images did not form part of the evaluation in this study.

Image Analysis:

Anonymised images were analysed on dedicated reporting workstations. Axial T2-W images, isotropic ADC maps calculated from monoexponential fit of the DW- data from all *b*-values, and grey-scale DCE images at peak contrast enhancement (range 58.5-62.5sec) had manual ROIs drawn around the DIPL on a 2D slice-by-slice basis. The DIPL was defined as the largest visible low signal intensity lesion on T2-W images with a corresponding subjective ADC reduction on DW-MRI from an octant with a positive biopsy. In cohort 1, the location of the DIPL was subsequently confirmed on the prostatectomy specimen. Smaller secondary lesions were ignored as these were not targets for dose boosting. Outlining was performed by free-hand drawing using a mouse-controlled cursor; margin recognition was based on the subjective assessment of the imaging features for each sequence according to current ESUR mpMRI guidelines [15]. T2-W images were assessed for regions of well-defined low T2-W signal, ADC maps for regions of

restricted diffusion, and DCE sequences for regions of brisk contrast uptake and early washout (Figure 1). ROI delineation on each sequence was done separately on a different occasion at least a week later to minimise possible memorisation of tumour margins.

GTVs were calculated by multiplying stacked ROI areas generated by the workstation software by sequence-specific slice thickness. A radiologist with 4 years prostate mpMRI experience performed all the DIPL ROI assessments. In addition, a second observer with 20 years' experience of prostate MRI, blinded to the first observer's ROIs, repeated identical assessments. Both observers were also blinded to the histopathological data.

Histopathological analysis:

All patients in cohort 1 underwent prostatectomy. The prostate was sectioned at 4 mm intervals in a plane perpendicular to the gland's posterior surface using a specially devised slicer to ensure accuracy of slicing [16]. Formalin fixed and paraffin wax embedded whole mount histopathological slides were prepared. The slicing axis matched the axial image acquisition angle so that stained sections from the embedded slices matched the imaging slices closely. Although the slice thickness did not match the imaging slice thickness, the segmentation of the whole tumour volume on both imaging and histology meant that slice by slice correlation of imaging with histology was not required. Tumour volumes of the DIPL were demarcated by a specialist histopathologist (Figure 1). The whole-mount slides were subsequently overlaid with a 1x1mm translucent grid sheet and photographed over a light source. Histopathological tumour volumes

of DIPLs were calculated by manual counting of overlying 1mm^2 grid squares, multiplied by the histological slice thickness.

Statistical Analysis:

Differences between the two observers for each sequence and histology were assessed using Bland-Altman plots and Limits of Agreement. The agreement was also assessed with a Pearson's correlation coefficient.

Analysis of Variance (ANOVA) was used to assess inter-sequence volume differences as well as differences in relative volumes between sequences across the 3 Gleason grade categories (3+3, 3+4, and 4+3 or greater). Paired-t-tests were used to detect significant differences between mpMRI volumes and histology for both observers.

A *p*-value of <0.05 was taken to be significant in all statistical tests. Analysis was performed in Microsoft Excel and SPSS v23. Bland-Altman plots were produced in GraphPad Prism v6.07.

Results:

One patient in cohort 2 had no DCE-MRI data, and artefacted T2-W data and was excluded from subsequent analysis.

Tumour Volume Variability

Differences between observers for T2-W, DW-MRI and DCE-MRI sequences:

GTVs drawn on T2-W images for both cohorts ranged from 0 to 7.0 cm^3 (mean 2.4 ± 1.93 cm^3) for observer 1 and from 0 to 7.2cm^3 (mean 2.29 ± 1.93 cm^3) for observer 2. Corresponding data for DW- and DCE-MRI are given in Table 1. Differences between volumes derived from all 3 sequences were significant for both observers (ANOVA, $p<10^{-8}$). Tumour volumes were smaller in cohort 1 than cohort 2 (Cohort 1: 1.5$\pm$ 1.6 cm^3 for observer 1 vs 1.8 $\pm$ 1.9 cm^3 for observer 2 and Cohort 2: 3.1 $\pm$ 2.0 cm^3 for observer 1 and 3.0 $\pm$1.9 cm^3 for observer 2)

Pearson's correlation tests demonstrated that interobserver GTVs for each sequence were significantly positively correlated at the 0.01 level (2-tailed), r =0.96 (T2-W), 0.94 (DW-MRI) and 0.92 (DCE-MRI). Limits of Agreement for interobserver variation in volumetry from each of the 3 sequences were: T2-W 0.9, -1.1 cm^3, DW-MRI 1.3, -1.7 cm^3, DCE-MRI 0.74, -0.89 cm^3; corresponding Bland-Altman plots are given in Figures 2 A, B and C respectively.

Differences between mpMRI sequences and histology (cohort 1):

Histological volumes ranged from 0.04 to 4.72cm^3 (mean 1.46 ± 1.50 cm^3). One patient's tumour was not detected on any mpMRI sequence but had a volume of 0.04cm^3 on histology. T2W-MRI GTVs overestimated histological volumes by 33 ± 76% (observer 1) and 16 ± 67% (observer 2) but had the highest correlation coefficient (r = 0.97 observer 1, 0.93 observer 2, p<0.0001). DW-MRI and DCE-MRI tended to underestimate histological volume (Table 2). Paired t-tests found that mean DCE-MRI GTVs were consistently and significantly different from histology (p=0.001 observer 1 and 0.0003 observer 2), whereas T2-W GTV differed from histology in observer 1 only (p=0.005) and DW-MRI differed from histology in observer 2 only (p=0.006). Bland-Altman plots for each sequence against histology with Limits of Agreement are exemplified for observer 1 in Figure 3.

Average volumes from all 3 sequences in cohort 1 were 1.33 ± 1.46 cm^3 for observer 1 and 1.05 ± 1.23 cm^3 for observer 2 (Table 2). A paired t-test showed no difference between this average volume and histology for observer 1 (p=0.2), although differences for observer 2 were significant (p=0.004).

Functional (DW- and DCE-) to morphological (T2-W) MRI tumour volume ratios and their relationship with Gleason score:

As assessed by cognitive fusion, there was a >90% overlap between ROIs from each sequence with each other. DW- and DCE-MRI derived volumes were consistently smaller than T2-W volumes: DW- to T2-W MRI ratios were 73.9 ± 18.1 % for observer 1 and 72.5 ± 21.9% for

observer 2. DCE- to T2-W MRI volume ratios were even lower (42.6 ± 24.6% for observer 1, 34.3 ± 24.9% for observer 2). . The proportion of the T2-W volume represented by the DW-MRI volume and the DCE-MRI volume was significantly different (ANOVA, $p<10^{-8}$).

Gleason grade was determined at prostatectomy in cohort 1 and pre-treatment in cohort 2. Eight patients had Gleason score 3+3 tumours, 21 were Gleason score 3+4 and 11 were Gleason score 4+3 or greater. DW-MRI to T2-W volume ratios in the 3 Gleason categories were 75.4 ± 15.9%, 72.6 ± 18.9% and 75.2 ± 19.4% respectively for observer 1 and 68.9 ± 17.8%, 67.7 ± 19.3% and 84.2 ± 26.1% for observer 2. DCE-MRI to T2-W volume ratios in the 3 Gleason categories were 44.9 ± 12.2%, 39.9 ± 27.4% and 45.7 ± 27.0% respectively for observer 1 and 33.5 ± 23.4%, 33.8 ± 24.8% and 35.8 ± 28.1% for observer 2. There were no significant differences in DW- and DCE- to T2-W MRI tumour volume ratios between the 3 Gleason grade categories for either observer (ANOVA, $p>0.05$), indicating no differences in the functional to morphological volumes with Gleason grade.

Discussion:

We have established that manual ROI delineation of DIPL is reproducible for the purposes of radiotherapy planning between two observers with ~1 cm^3 limits of agreement on T2-W MRI. Both observers interpreted the mpMRI in accordance with ESUR guidelines. Additionally, they outlined in optimal ambient conditions for their individual preferences and had the ability to manipulate the window, brightness and magnification for decisions regarding tumour margins. Differences in observer perception of feature boundaries are likely to reflect the consistently lower measurements of observer 2. Although differences in lesion conspicuity due to different sequences and scanners will also be a factor, this study aimed to establish variability in the presence of these variations. Furthermore, as the diffusion-weighted image provided the most definitive contrast for lesion identification, its independence of field-strength re-inforces the validity of the findings across the 2 field-strengths used in this study. The concordance of each observer's measurements with histology however, remains the definitive test of the viability of the method. In practical terms, the measured tumour volume differed between observers by $1cm^3$ for T2-W images in the largest tumours in our cohort which should not cause differences in radiation therapy plans made on images outlined by different observers because of the addition of substantial additional margins when delineating a clinical target volume around the GTV.

We have demonstrated a correlation between mpMRI derived tumour volumes with histology that is similar to others [10, 14, 17-19]. In addition, we have shown that DIPL tumour volumes defined on the T2-W images were consistently larger than on DW- or DCE-MRI. Although they overestimate histological volumes, they are best suited to delineating the margins of the DIPL for radiation dose boosting, especially as a post-resection shrinkage factor of up to 1.15 in

histological samples [17] must be allowed for. Shrinkage is due to formalin fixation and was unavoidable in our study as the tissue was preserved immediately post-resection for optimal diagnostic purposes. In an early work Ponchetti et al showed that T2-W images overestimated small tumours by as much as 58%, however their MRI scans were done post-biopsy which may have confounded their MRI measurements [19]. In comparison, a study by Cornud et al [20] underestimated histology in nearly half the cases (49%) with a larger mean difference (-0.56cm^3) than we demonstrated (-0.08 to 0.30cm^3). A recent large study measuring all visible lesions in 202 patients also concluded that all sequences underestimated true volume and that the maximum volume from all sequences most closely matched histological volume. These results and those of others [21, 22] are likely to be influenced by the non-recognition on mpMRI of small, low Gleason score disease.

Estimation of tumour size has also been done on T2-W imaging using a maximal dimension approach utilising visual assessment of functional parameters to support the T2-W measurements [23]. Although these data correlate well with histological volumes, they also have been noted to underestimate them [24]. Other studies have used the functional information to define the T2-W ROIs, but have not interrogated the sequences individually [12]. Where individual sequences have been investigated, e.g. DCE-MRI comparison with histology [25], the focus has been on technical developments and comparison with histology, rather than on investigating the relationship of volumetry derived from individual sequences. The only other study comparing inter-sequence differences reported data from a small data set of 5 patients and, contrary to our findings, demonstrated no significant differences in GTVs between sequences as measured by 6 observers [26].

It is accepted that a combination of both T2-W and DW-MRI improves cancer detection and localisation [21, 27]. Use of a second additional functional technique such as DCE-MRI has been shown to further improve sensitivity [28] for tumour detection. In the assessment of volume on the other hand, the addition of DW- and DCE-MRI sequences to T2-W assessments has been reported to influence interobserver variability of tumour outlining [9]. The mean differences between observers on each of the 3 sequences in this study was <28%, smaller than the mean inter-sequence differences of up to 70%. T2-W sequences remain the preferred choice on which to delineate prostate tumour in current practice as their higher spatial resolution and low geometric distortion enables registration with CT images used for radiation therapy planning. As DW- techniques improve and thresholding of quantified ADC allows automated segmentation of tumour this may change, giving preference to semi-automated segmentation on ADC maps.

We have additionally shown that differences between T2-W and DW- or DCE-MRI derived GTVs scale consistently with tumour volume. These results suggest that the volume of neo-angiogenesis is smaller than the volume of abnormal cellular morphology demonstrated on T2-W or on DW-MRI respectively. The significant differences between GTVs derived from DCE-MRI compared with those from both other sequences and histology also may be in part due to the lower spatial resolution of this sequence.

Although whole mount histopathology is regarded as a gold standard for correlating image-derived tumour volume measurements, it should be noted that this technique also has innate margins of error and is subject to operator-dependent variation depending on experience and the equipment available. There is also documented variability in the interpretation and grading of Gleason grades [29], and substantial variability has been reported in the current clinical volume

estimation methods [30]. In our study, to minimize slice width variations we used a specially devised slicer to mitigate these effects. All samples were processed in the same manner and tumours demarcated by one histopathologist to reduce intra-operator variability.

All imaging in our study was performed with an endorectal coil which causes posterior deformation of the gland. Although this has potential for error when performing 2-D measurements, we would not expect an influence on volume measurements where tumour ROIs are defined on all slices with visible tumour. Histological assessments of tumour were limited by manual assessments of photographs with an overlain grid, but this has provided good correlation of imaging and pathological volumes in other tumour types [31, 32]. Digital analysis of histopathological volumes (planimetry) [33] is more robust where available.

A limitation of our data is the lack of information on spatial conformity of ROIs between sequences, which was assessed only by visual cognitive fusion to confirm concordance. In previous work aimed at identifying the index lesion, this has proved time-consuming with marginal improvements over cognitive fusion by an experienced observer [34]. In addition, field inhomogeneity at air-tissue interfaces can cause distortions and lead to errors in EPI based DWI, particularly at higher field strengths. However, the rectal balloon was filled with perfluorocarbon for our 3T data acquisition, minimising any such distortions. At 1.5T, being the most distorted, the volume measurements on DWI corresponded most closely with histology, indicating that distortion is not the key factor in measurement error of the DIPL. Another limitation of our study was the use of the peak enhancement DCE-MRI sequences for tumour delineation rather than the quantitative DCE parameter maps (K^{trans}, k_{ep}, v_e). It is unlikely however that this would have yielded GTVs from the DCE-MRI data that were significantly different.

In summary, we have established that mpMRI-derived GTV measurements of DIPLs derived from T2-W, DW-MRI and DCE sequences are reproducible between observers. GTV is largest on T2-W images and smallest on DCE-MRI images, and T2-W GTVs best approximate to *in vivo* tumour volume. Therefore, GTV should be delineated on T2-W images when defining the DIPL for radiation dose boosting. Differences in volumes derived from T2-W, DW- and DCE-MRI images are highly significant reflecting differences in cellular and vascular proportions; the proportion of the T2-W volume represented by the DW- and DCE-MRI volume in this sample were independent of Gleason grade.

Figure Legends:

Figure 1 Comparison of intersequence volumes with histology: Transverse T2-W (A), Diffusion-weighted (B) and Dynamic Contrast enhanced image at 30 secs post injection of gadoterate meglumine (C). Tumour outlines drawn on 3 separate occasions by observer 1 are overlaid. The volume in A was largest, and the volume in C smallest, although overlap between the outlines is noted in all cases. Whole-mount histological specimen at prostatectomy (D) confirms the presence of tumour at that location.

Figure 2: Bland-Altman plots showing differences in tumour volumetry between observers on T2-W (A), Diffusion Weighted (B) and Dynamic Contrast Enhanced (C) MRI in all patients (cohorts 1 and 2, n=40). The mean difference (solid line) and Limits of Agreement (dashed lines) representing $\pm$ 1.96 SD from the mean are given.

Figure 3: Bland-Altman plots showing differences in tumour volumetry with histology for observer 1 on T2-W (A), Diffusion Weighted (B) and Dynamic Contrast Enhanced (C) MRI in patients undergoing prostatectomy (cohort 1, n=20). The mean difference (solid line) and Limits of Agreement (dashed lines) representing $\pm$ 1.96 SD from the mean are given.

Table 1: Mean, standard deviation and median volume of dominant intraprostatic tumour lesion on each mpMRI sequence for all patients (cohorts 1 and 2, n=40) and for cohort 1 alone (n=20). Volume data derived from prostatectomy samples in cohort 1 is shown for comparison.

		Observer 1	Observer 2	Histology
T2-W	Mean ± SD (cm^3) (All)	2.40 ± 1.93	2.29 ± 1.93	
	Median (cm^3) (All)	1.77	1.60	
	Mean ± SD (cm^3) (cohort 1)	1.84 ± 1.85	1.54 ± 1.60	1. 46 ± 1.5
	Median (cm^3) (cohort 1)	1.1	1.01	0.71
DW-MRI	Mean ± SD (cm^3) (All)	1.77 ± 1.58	1.62 ± 1.43	
	Median (cm^3) (All)	1.17	1.19	
	Mean ± SD (cm^3) (cohort 1)	1.46 ± 1.70	1.13 ± 1.35	
	Median (cm^3) (cohort 1)	0.79	0.61	
DCE-MRI	Mean ± SD (cm^3) (All)	1.01 ± 0.99	0.9 ± 1.1	
	Median (cm^3) (All)	0.82	0.71	
	Mean ± SD (cm^3) (cohort 1)	0.68 ± 0.98	0.48 ± 0.88	
	Median (cm^3) (cohort 1)	0.33	0.15	
Average of all sequences combined	Mean ± SD (cm^3) (All)	1.73 + 1.44	1.62 + 1.43	
	Median (cm^3) (All)	1.27	1.16	
	Mean ± SD (cm^3) (cohort 1)	1.33 + 1.46	1.05 + 1.23	
	Median (cm^3) (cohort 1)	0.71	0.56	

Table 2: Differences between volumes of the of dominant intraprostatic tumour lesion derived from each sequence for each observer and volumes derived from whole-mount prostatectomy specimens

	T2W vs. histology % and absolute (cm^3) difference	DWI vs. histology % and absolute (cm^3) difference	DCE vs. histology % and absolute (cm^3) difference	Average of 3 sequences vs. histology % and absolute (cm^3) difference
Observer 1	33.2 ± 76.3%	2.5 ± 57.8 %	-31.6 + 97.8 %	-7.5 + 49.1 %
Mean ± SD	0.38 ± 0.53 cm^3	0.0 ± 0.58 cm^3	-0.78 ± 0.91 cm^3	-0.13 + 0.44 cm^3
Observer 2	16.4 ± 67.5 %	-26.1 ± 36.4 %	-79.3 ± 28.3 %	-29.7 + 35.6 %
Mean ± SD	0.07 ± 0.57 cm^3	-0.33 ± 0.48 cm^3	-0.98 ± 0.98 cm^3	-0.41 + 0.56 cm^3

References:

1. Rasch C, Barillot I, Remeijer P, et al. Definition of the prostate in CT and MRI: a multi-observer study. International journal of radiation oncology, biology, physics. 1999;43(1):57-66.
2. Sefrova J, Paluska, Odrazka K, et al. Changes in target volumes definition by using MRI for prostate bed radiotherapy planning--preliminary results. Klin Onkol. 2010;23(4):256-63.
3. Parker CC, Damyanovich A, Haycocks T, et al. Magnetic resonance imaging in the radiation treatment planning of localized prostate cancer using intra-prostatic fiducial markers for computed tomography co-registration. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2003;66(2):217-24.
4. Khoo EL, Schick K, Plank AW, et al. Prostate contouring variation: can it be fixed? International journal of radiation oncology, biology, physics. 2012;82(5):1923-9.
5. Sefrova J, Odrazka K, Paluska P, et al. Magnetic resonance imaging in postprostatectomy radiotherapy planning. International journal of radiation oncology, biology, physics. 2012;82(2):911-8.
6. Scheenen TW, Rosenkrantz AB, Haider MA, Futterer JJ. Multiparametric Magnetic Resonance Imaging in Prostate Cancer Management: Current Status and Future Perspectives. Invest Radiol. 2015.
7. Pullini S, Signor MA, Pancot M, et al. Impact of multiparametric magnetic resonance imaging on risk group assessment of patients with prostate cancer addressed to external beam radiation therapy. European journal of radiology. 2016;85(4):764-70.
8. Croke J, Maclean J, Nyiri B, et al. Proposal of a post-prostatectomy clinical target volume based on pre-operative MRI: volumetric and dosimetric comparison to the RTOG guidelines. Radiation oncology. 2014;9:303.
9. Nyholm T, Jonsson J, Soderstrom K, et al. Variability in prostate and seminal vesicle delineations defined on magnetic resonance images, a multi-observer, -center and -sequence study. Radiation oncology. 2013;8:126.
10. Fedorov A, Penzkofer T, Hirsch MS, et al. The role of pathology correlation approach in prostate cancer index lesion detection and quantitative analysis with multiparametric MRI. Acad Radiol. 2015;22(5):548-55.
11. Borren A, Groenendaal G, Moman MR, et al. Accurate prostate tumour detection with multiparametric magnetic resonance imaging: dependence on histological properties. Acta Oncol. 2014;53(1):88-95.
12. Le Nobin J, Rosenkrantz AB, Villers A, et al. Image Guided Focal Therapy for Magnetic Resonance Imaging Visible Prostate Cancer: Defining a 3-Dimensional Treatment Margin Based on Magnetic Resonance Imaging Histology Co-Registration Analysis. The Journal of urology. 2015;194(2):364-70.
13. Engelhard K, Labanaris AP, Bogner K, et al. How good is post-biopsy multiparametric magnetic resonance imaging in detecting and characterising the index lesion of localised prostate cancer? Scandinavian journal of urology. 2014;48(6):499-505.
14. Bratan F, Melodelima C, Souchon R, et al. How accurate is multiparametric MR imaging in evaluation of prostate cancer volume? Radiology. 2015;275(1):144-54.
15. Barentsz JO, Richenberg J, Clements R, et al. ESUR prostate MR guidelines 2012. European Radiology. 2012;22(4):746-57.

16. Jhavar SG, Fisher C, Jackson A, et al. Processing of radical prostatectomy specimens for correlation of data from histopathological, molecular biological, and radiological studies: a new whole organ technique. Journal of Clinical Pathology. 2005;58(5):504-8.
17. Turkbey B, Mani H, Aras O, et al. Correlation of magnetic resonance imaging tumor volume with histopathology. The Journal of urology. 2012;188(4):1157-63.
18. Isebaert S, Van den Bergh L, Haustermans K, et al. Multiparametric MRI for prostate cancer localization in correlation to whole-mount histopathology. Journal of magnetic resonance imaging : JMRI. 2013;37(6):1392-401.
19. Ponchietti R, Di Loro F, Fanfani A, Amorosi A. Estimation of prostate cancer volume by endorectal coil magnetic resonance imaging vs. pathologic volume. European urology. 1999;35(1):32-5.
20. Cornud F, Khoury G, Bouazza N, et al. Tumor target volume for focal therapy of prostate cancer-does multiparametric magnetic resonance imaging allow for a reliable estimation? The Journal of urology. 2014;191(5):1272-9.
21. Mazaheri Y, Hricak H, Fine SW, et al. Prostate tumor volume measurement with combined T2-weighted imaging and diffusion-weighted MR: correlation with pathologic tumor volume. Radiology. 2009;252(2):449-57.
22. Le Nobin J, Orczyk C, Deng FM, et al. Prostate tumour volumes: evaluation of the agreement between magnetic resonance imaging and histology using novel co-registration software. BJU international. 2014;114(6b):E105-12.
23. Nakashima J, Tanimoto A, Imai Y, et al. Endorectal MRI for prediction of tumor site, tumor size, and local extension of prostate cancer. Urology. 2004;64(1):101-5.
24. Rud E, Klotz D, Rennesund K, et al. Detection of the index tumour and tumour volume in prostate cancer using T2-weighted and diffusion-weighted magnetic resonance imaging (MRI) alone. BJU international. 2014;114(6b):E32-E42.
25. Fennessy FM, Fedorov A, Penzkofer T, et al. Quantitative pharmacokinetic analysis of prostate cancer DCE-MRI at 3T: comparison of two arterial input functions on cancer detection with digitized whole mount histopathological validation. Magnetic resonance imaging. 2015;33(7):886-94.
26. Rischke HC, Nestle U, Fechter T, et al. 3 Tesla multiparametric MRI for GTV-definition of Dominant Intraprostatic Lesions in patients with Prostate Cancer--an interobserver variability study. Radiation oncology. 2013;8:183.
27. Haider MA, van der Kwast TH, Tanguay J, et al. Combined T2-weighted and diffusion-weighted MRI for localization of prostate cancer. AJR American journal of roentgenology. 2007;189(2):323-8.
28. Riches SF, Payne GS, Morgan VA, et al. MRI in the Detection of Prostate Cancer: Combined Apparent Diffusion Coefficient, Metabolite Ratio, and Vascular Parameters. American Journal of Roentgenology. 2009;193(6):1583-91.
29. Allsbrook WC, Jr., Mangold KA, Johnson MH, et al. Interobserver reproducibility of Gleason grading of prostatic carcinoma: general pathologist. Hum Pathol. 2001;32(1):81-8.
30. van der Kwast TH, Amin MB, Billis A, et al. International Society of Urological Pathology (ISUP) Consensus Conference on Handling and Staging of Radical Prostatectomy Specimens. Working group 2: T2 substaging and prostate cancer volume. Mod Pathol. 2011;24(1):16-25.

31. deSouza NM, Scoones D, Krausz T, Gilderdale DJ, Soutter WP. High-resolution MR imaging of stage I cervical neoplasia with a dedicated transvaginal coil: MR features and correlation of imaging and pathologic findings. American Journal of Roentgenology. 1996;166(3):553-9.
32. Yang DH, Kim JK, Kim KW, et al. MRI of Small Cell Carcinoma of the Uterine Cervix with Pathologic Correlation. American Journal of Roentgenology. 2004;182(5):1255-8.
33. Salarian M, Shahedi M, Gibson E, et al., editors. Toward quantitative digital histopathology for prostate cancer: comparison of inter-slide interpolation methods for tumour measurement2013.
34. Riches SF, Payne GS, Morgan VA, et al. Multivariate modelling of prostate cancer combining magnetic resonance derived T2, diffusion, dynamic contrast-enhanced and spectroscopic parameters. European Radiology. 2015:1-10.

Figure 1A

Click here to download Figure Figure 1A.tif

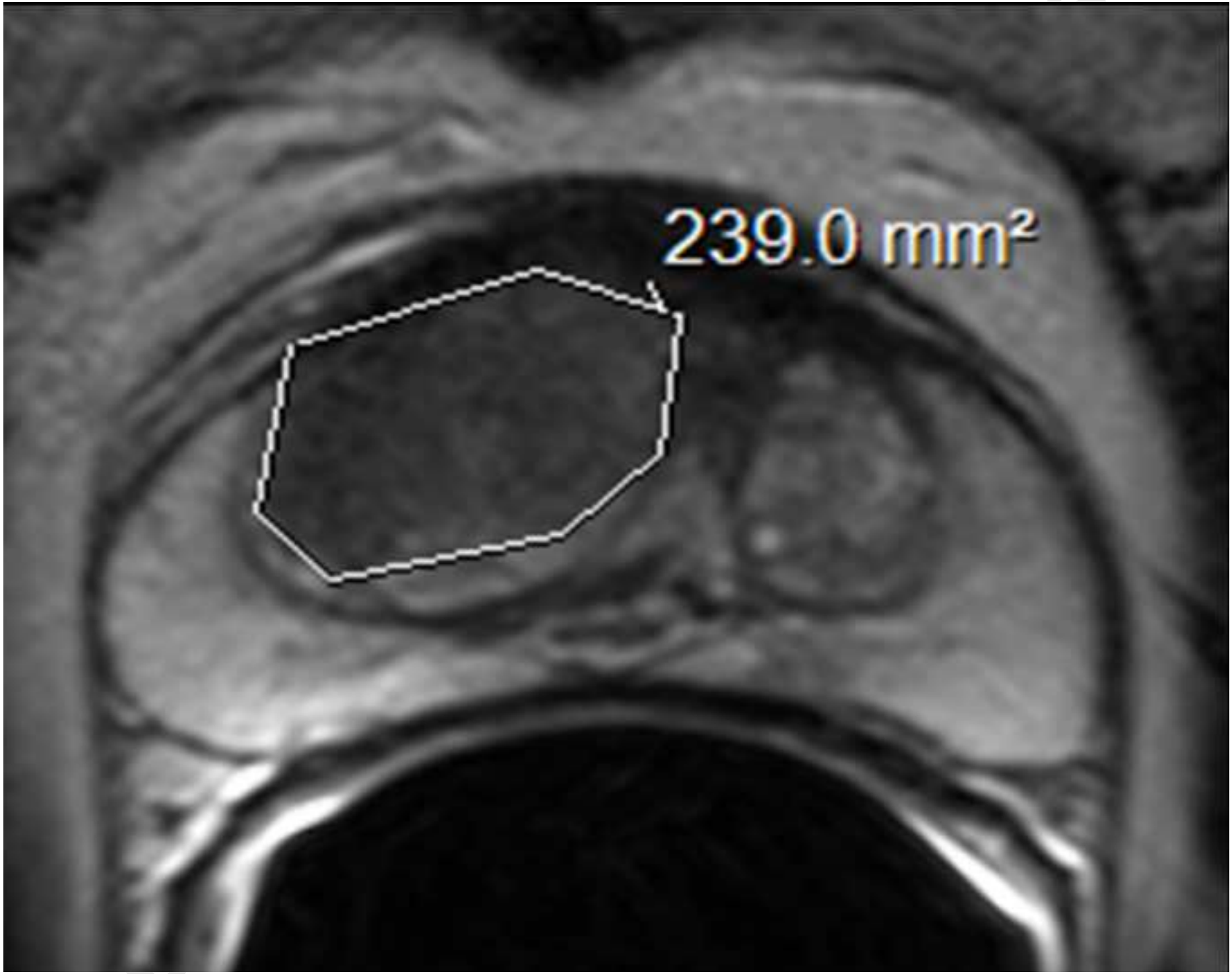

Figure 1B

Click here to download Figure Figure 1B.tif

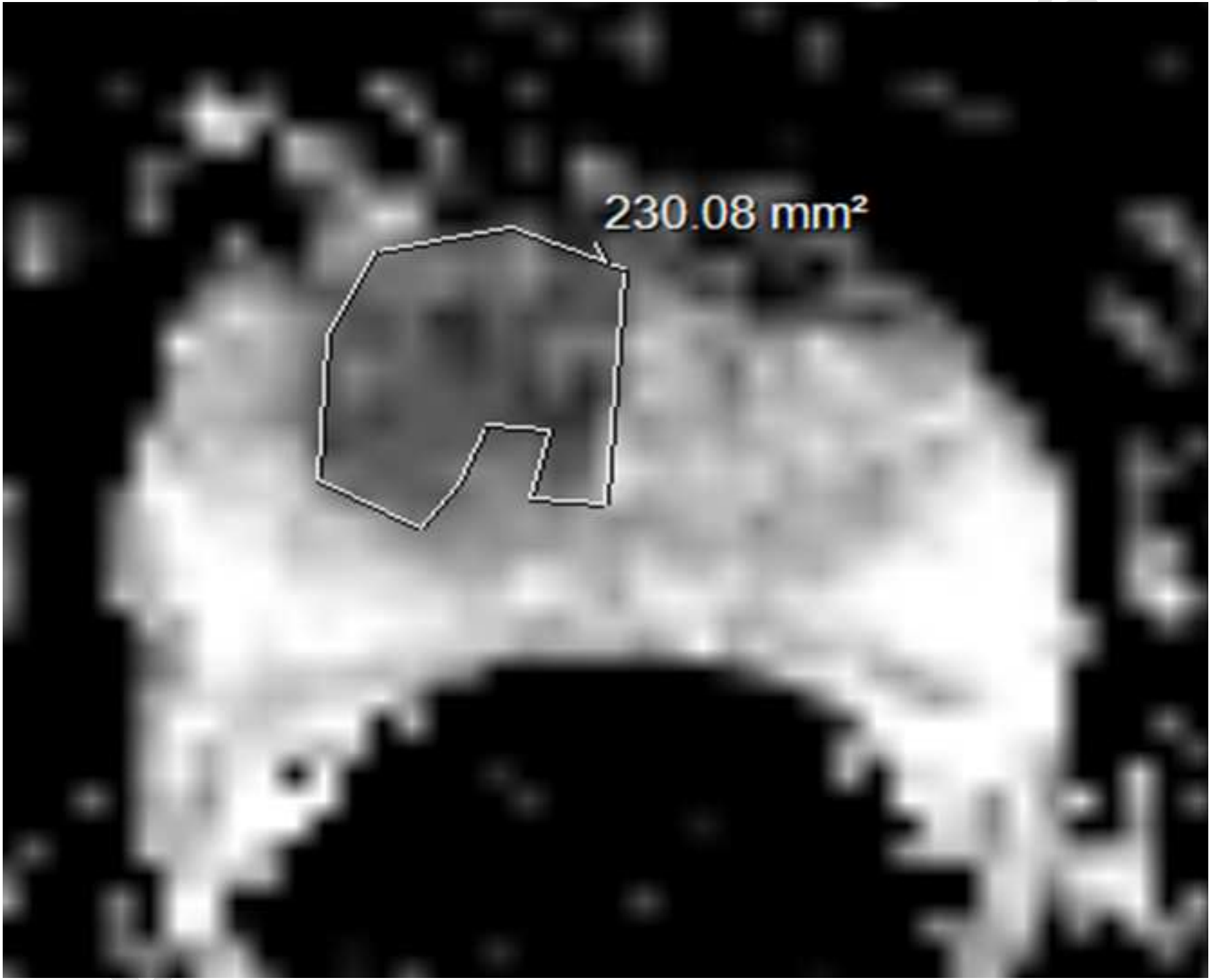

Figure 1C

Click here to download Figure Figure 1C.tif

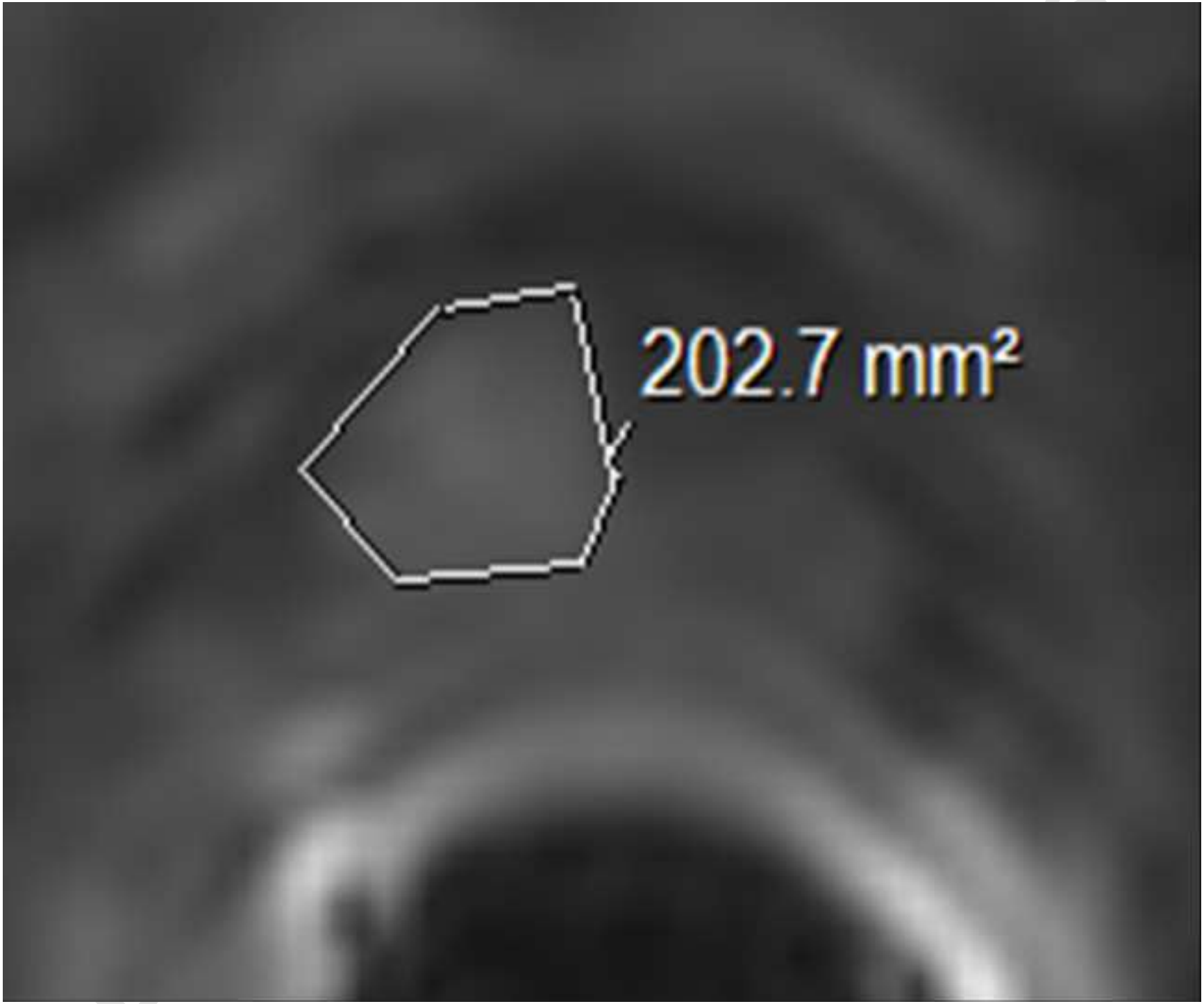

Figure 1D

Click here to download Figure Figure 1D.tif

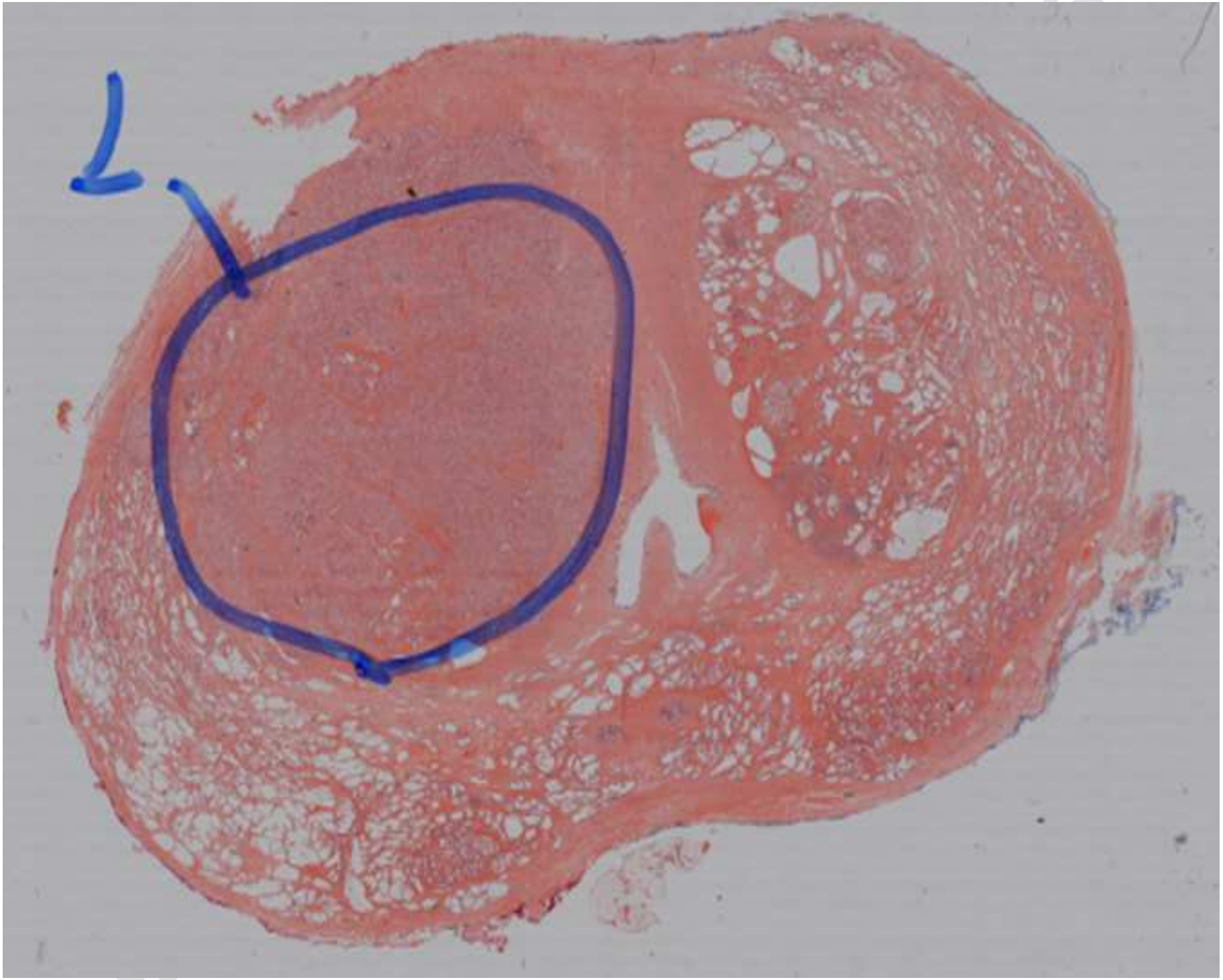

Figure 2A

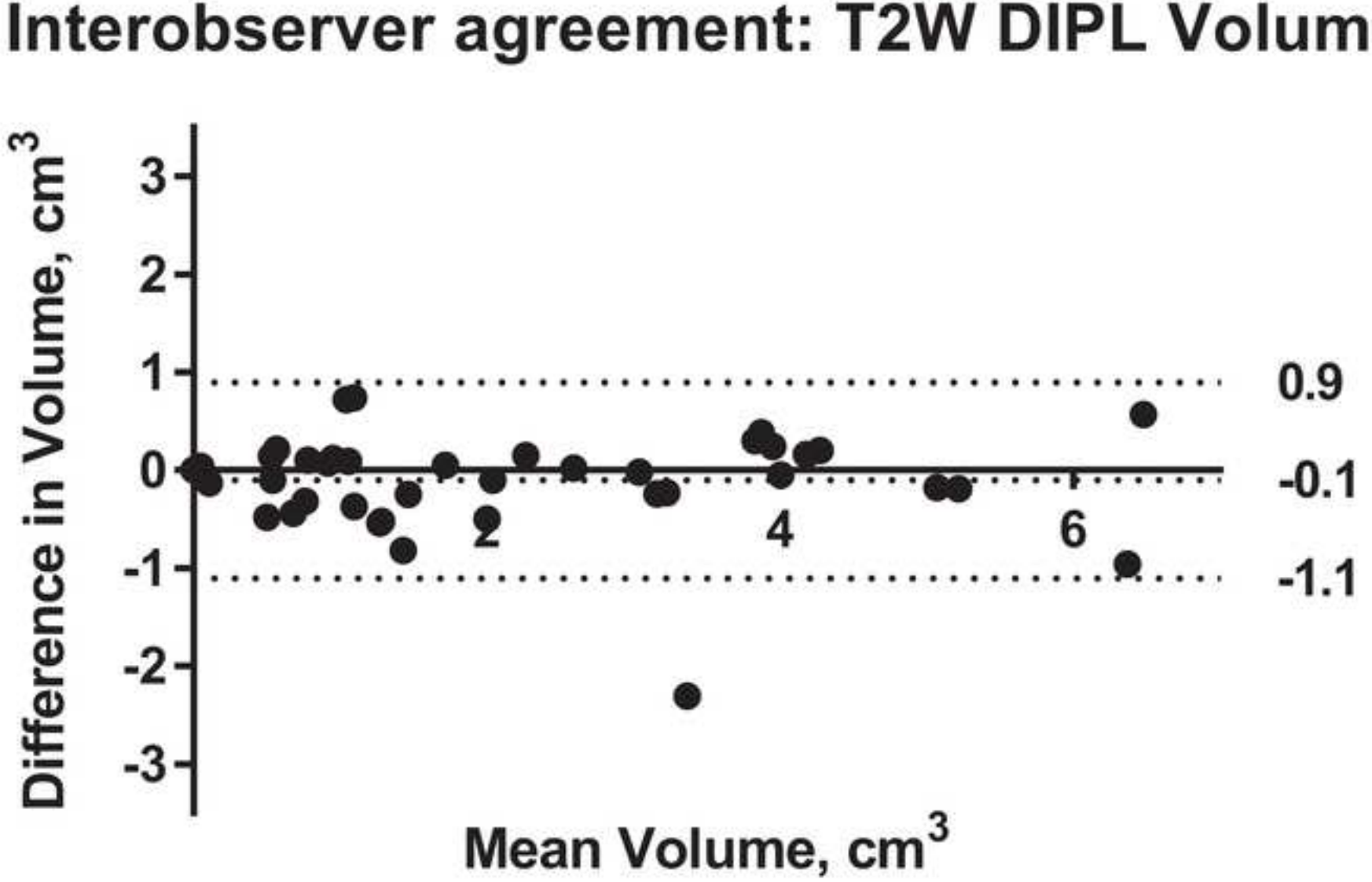

Figure 2B

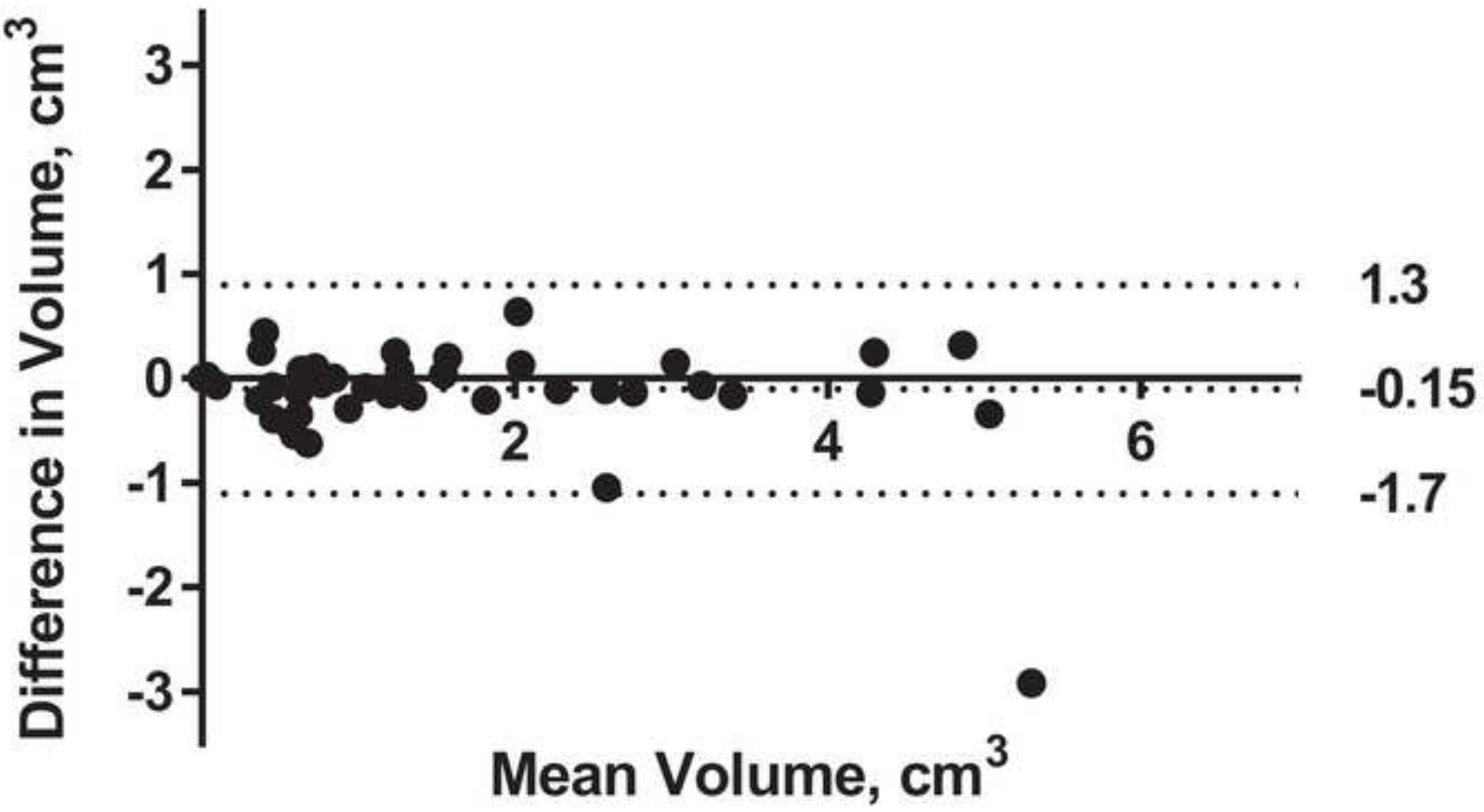

Figure 2C

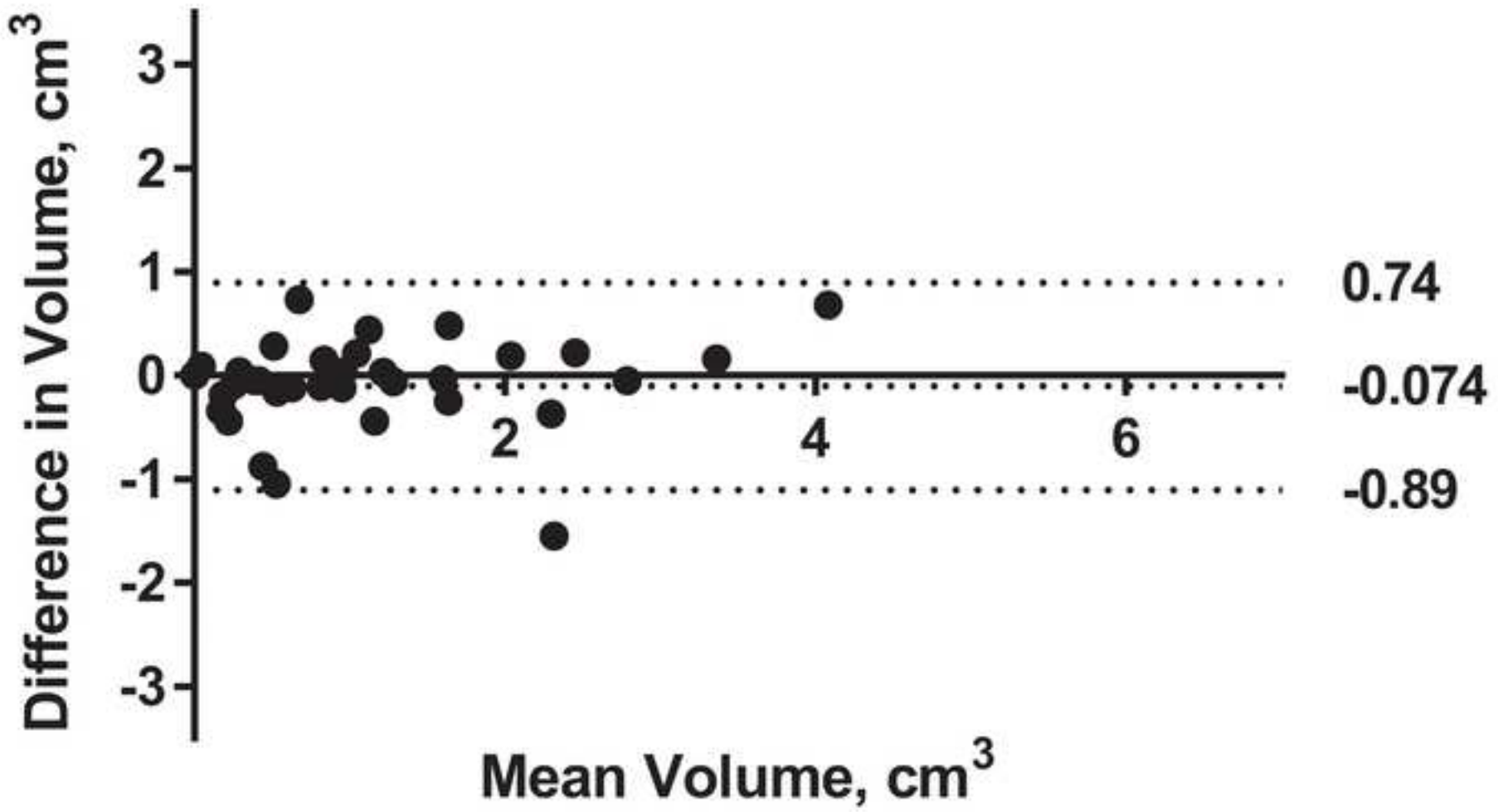

Figure 3A

Click here to download Figure Figure 3A.tif

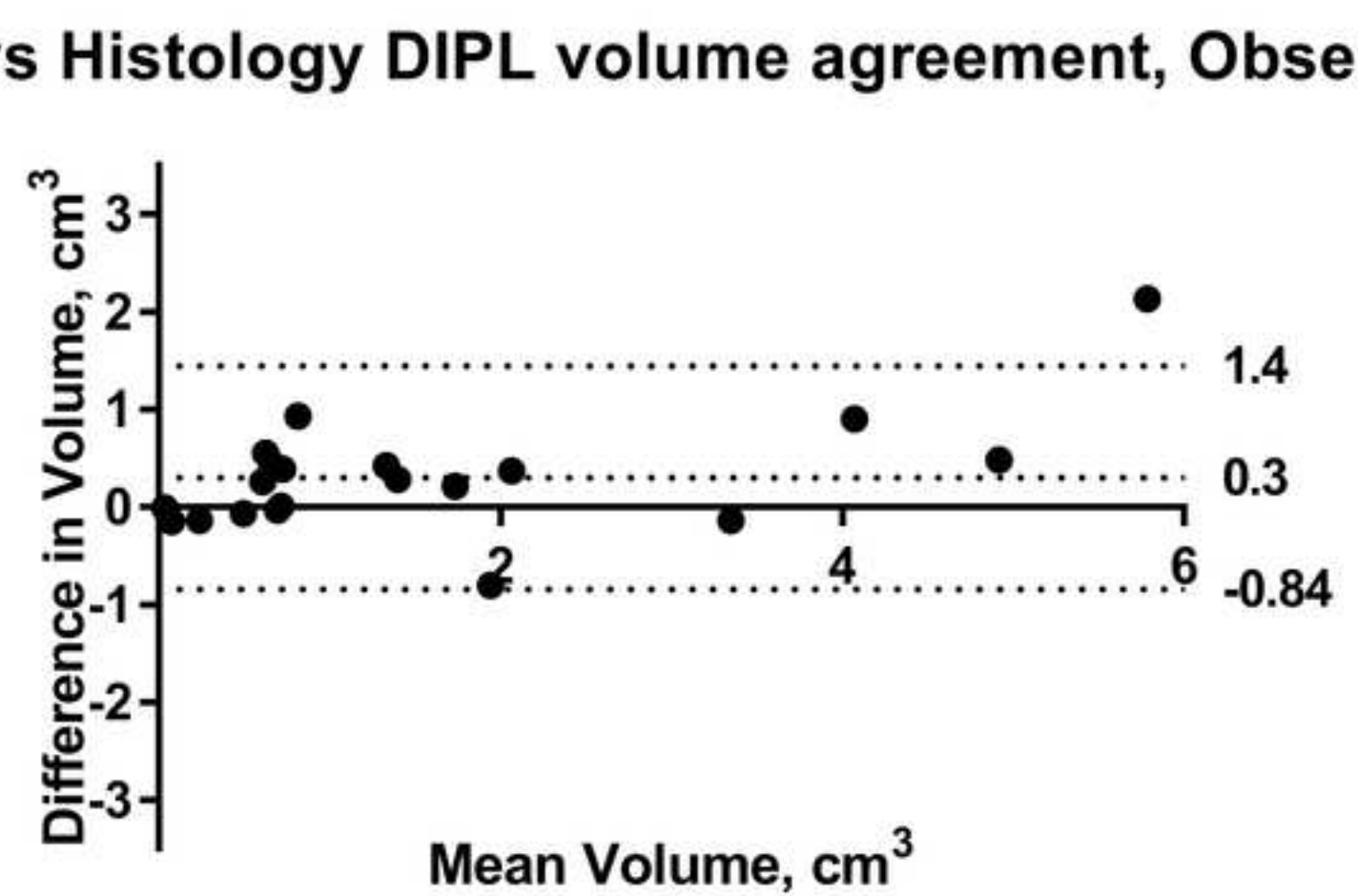

Figure 3B

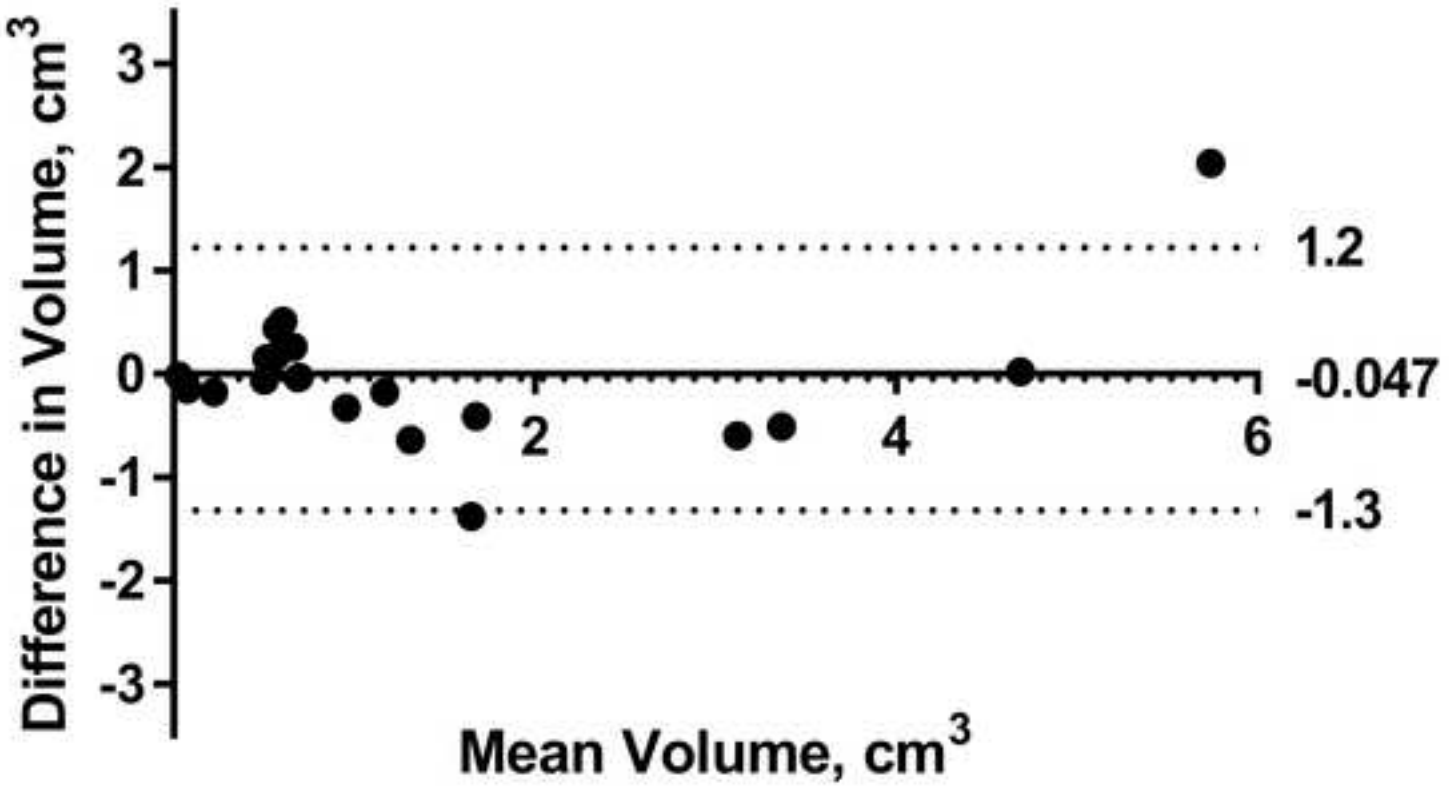

Figure 3C

Click here to download Figure Figure 3C.tif