Received:
18 November 2014

Revised:
24 March 2015

Accepted:
30 March 2015

doi: 10.1259/bjr.20140776

Cite this article as:

Brodin NP, Tang J, Skalina K, Quinn TJ, Basu I, Guha C, et al. Semi-automatic cone beam CT segmentation of *in vivo* pre-clinical subcutaneous tumours provides an efficient non-invasive alternative for tumour volume measurements. Br J Radiol 2015;88:20140776.

FULL PAPER

Semi-automatic cone beam CT segmentation of *in vivo* pre-clinical subcutaneous tumours provides an efficient non-invasive alternative for tumour volume measurements

^{1,2}N P BRODIN, PhD, ¹J TANG, MD, ³K SKALINA, MS, ¹TJ QUINN, BS, ^{1,2}I BASU, PhD, ^{1,2,3}C GUHA, MD, PhD and ^{1,2}W A TOMÉ, PhD

¹Department of Radiation Oncology, Montefiore Medical Center and Albert Einstein College of Medicine, Bronx, NY, USA

²Institute for Onco-Physics, Department of Radiation Oncology, Albert Einstein College of Medicine, Bronx, NY, USA

³Department of Pathology, Albert Einstein College of Medicine, Bronx, NY, USA

Address correspondence to: Professor Wolfgang Axel Tomé

E-mail: wolfgang.tome@einstein.yu.edu

Objective: To evaluate the feasibility and accuracy of using cone beam CT (CBCT) scans obtained in radiation studies using the small-animal radiation research platform to perform semi-automatic tumour segmentation of pre-clinical tumour volumes.

Methods: Volume measurements were evaluated for different anatomical tumour sites, the flank, thigh and dorsum of the hind foot, for a variety of tumour cell lines. The estimated tumour volumes from CBCT and manual calliper measurements using different volume equations were compared with the “gold standard”, measured by weighing the tumours following euthanasia and tumour resection. The correlation between tumour volumes estimated with the different methods, compared with the gold standard, was estimated by the Spearman’s rank correlation coefficient, root-mean-square deviation and the coefficient of determination.

Results: The semi-automatic CBCT volume segmentation performed favourably compared with manual calliper measures for flank tumours $\leq 2\text{ cm}^3$ and thigh tumours $\leq 1\text{ cm}^3$. For tumours $> 2\text{ cm}^3$ or foot tumours, the CBCT method was not able to accurately segment the tumour volumes and manual calliper measures were superior.

Conclusion: We demonstrated that tumour volumes of flank and thigh tumours, obtained as a part of radiation studies using image-guided small-animal irradiators, can be estimated more efficiently and accurately using semi-automatic segmentation from CBCT scans.

Advances in knowledge: This is the first study evaluating tumour volume assessment of pre-clinical subcutaneous tumours in different anatomical sites using on-board CBCT imaging. We also compared the accuracy of the CBCT method to manual calliper measures, using various volume calculation equations.

Accurate methods for assessing subcutaneous tumour volumes are vital components of pre-clinical cancer research. Longitudinal studies comparing different cancer treatment regimens in research animals (usually mice or rats) often use tumour volume assays as the main end point for evaluating treatment efficacy.¹ The current standard for tumour volume measurements for pre-clinical subcutaneous tumours consists of using manual callipers to determine the length, width and, in some cases, also depth of the tumour. Tumour volumes are then calculated based on a chosen mathematical formula, where a formula based on a modified ellipsoid has previously been shown to perform quite well.^{1,2} Calliper measures, although fast and convenient, are subject to several sources of uncertainty such as interobserver variability, differences in tumour shape, and amount of fatty tissue and fur surrounding the tumour.

Non-invasive imaging methods have become the standard for clinical tumour response assessment, and CT has been the main component for more than a decade.^{3,4} Previous studies have shown that small-animal ultrasound imaging or sequential micro-CT scans can be used to measure subcutaneous tumours in mice and rats more accurately than manual calliper measures.^{5–7} Improving the accuracy of tumour volume measurements will not only improve the quality of data in treatment efficacy studies, but it will also reduce the variability and thus reduce the number of animals required for tumour studies.

Taking micro-CT scans or ultrasound images of animals may, however, be quite a time consuming and potentially costly procedure, also requiring further anesthetizing of the animals. Here, we present a method for semi-automatic tumour volume determination based on cone beam CT

(CBCT) scans taken using the on-board imager of the small-animal radiation research platform (SARRP; XStrahl®, England, UK).^{8,9} With robotic-image-guided small-animal irradiators becoming increasingly available,¹⁰ this method provides a promising alternative for fast and less user-dependent tumour volume measurement using CBCT scans already obtained in the process of radiation therapy target localization. We compare the performance of CBCT volume segmentation to that of manual calliper measurements for different tumour sites and provide recommendations for pre-clinical tumour volume assessment based on these results.

METHODS AND MATERIALS

Animals and tumour models

Traditionally, flank tumours are the most commonly used subcutaneous pre-clinical tumour models. In this study, we were interested in evaluating the efficacy of semi-automatic CBCT tumour volume segmentation for subcutaneous tumours in three different anatomical locations; the flank, thigh and dorsum of the hind foot. This study was performed on mice that were all part of on-going tumour studies with radiation therapy or focused ultrasound and the details regarding mouse strain, cancer cell lines and tumour location are presented in Table 1. We opted to include different strains of mice and tumours from different cancer cell lines to test the volume segmentation method on a data set that was representative of a broad variation of pre-clinical tumour models. All animal procedures were conducted in accordance with approved protocols from the Institutional Animal Care and Use Committee at the Albert Einstein College of Medicine.

Volume measurement by cone beam CT imaging

Animals were anesthetized during the entire procedure; whether it included imaging and therapy or only imaging, with a continuous flow of 1.5 l min^{-1} of 1.5% isoflurane in pure oxygen. The on-board CBCT imager on our SARRP consists of an amorphous Si flat-panel detector with 512×512 detector elements of pixel size 0.4 mm^2 that register radiographs, as the

stage on which the animal is placed rotates 360° . For all CBCT imaging, a 1-mm aluminium filter (provided by the SARRP manufacturer) was used along with 50-kV tube voltage, 0.7-mA tube current and 67-s acquisition time. The voxel resolution for the applied image reconstruction protocol was $0.275 \times 0.275 \times 0.275 \text{ mm}^3$, which produced high-resolution CBCT image series intended for radiation therapy target localization.

To perform the tumour volume segmentation, the digital imaging and communications in medicine (DICOM) files of the CBCT scan were exported to an in-house-developed MATLAB® program (MathWorks®, Natick, MA). The segmentation method was based on a simple Hounsfield unit (HU)-thresholding algorithm, with user input in the form of manually drawn regions of interest (ROIs) and defining the superior and inferior limits of the tumour. Because of differences in image acquisition between CBCT and helical CT, there will be differences in the estimated HUs. We calculated the CBCT HUs used in the segmentation algorithm relative to the pixel intensities of a CBCT-scanned water phantom according to:

$$\text{HU}_{\text{CBCT},j} = \frac{I_j - I_{\text{water}}}{I_{\text{water}}} \times 1000 \quad (1)$$

where I_j is the pixel intensity in the j th pixel, and I_{water} is the averaged pixel intensity obtained from a ROI in the water phantom. The HU thresholding for the segmentation algorithm was defined as to exclude pixels below a certain value (such as air and low-density fatty tissue) and above a certain value (such as the bone):

$$\text{Include} = \begin{cases} \text{Yes if } \text{HU}_{\text{CBCT},j} \geq \text{HU}_{\text{low threshold}} \\ \text{Yes if } \text{HU}_{\text{CBCT},j} < \text{HU}_{\text{high threshold}} \\ \text{No otherwise} \end{cases} \quad (2)$$

Using our imaging panel, there was some variation in the absolute pixel values between different experiments, depending on

Table 1. Detailed list of subcutaneous tumour models included in this study with the number of tumours included for primary correlation analysis, validation and accuracy estimated based on resected tumour weights

Tumour site	Mouse strain	Cell line	Injection
Flank ($n_{\text{primary}} = 103$) ($n_{\text{validation}} = 29$) ($n_{\text{resected}} = 27$)	C57BL/6	3LL Lewis lung cancer	2×10^5 cells in the left flank
	C57BL/6	RM1 prostate cancer	1×10^5 cells in the right flank
	Athymic nude	Huh7 $\times$ human hepatocellular cancer	1.3×10^6 cells in the right flank
	Athymic nude	HepG2 human hepatocellular cancer	3.5×10^6 cells in the right flank
	Rag2 ^{-/-} γ C ^{-/-}	HepG2 human hepatocellular cancer	5×10^6 cells in the right flank
Thigh ($n_{\text{primary}} = 29$) ($n_{\text{validation}} = 14$) ($n_{\text{resected}} = 32$)	C57BL/6	3LL Lewis lung cancer	3×10^5 or 2×10^5 cells in the right thigh
	C57BL/6	B16-F10 melanoma	2×10^5 cells in the right thigh
	C57BL/6	TPSA23 prostate cancer (engineered from TRAMP-C1)	5×10^5 cells in the right thigh
	Athymic nude	RM1 prostate cancer	5×10^4 cells in the right and left thighs
Dorsum of hind foot ($n_{\text{primary}} = 102$) ($n_{\text{validation}} = 32$)	C57BL/6	B16-F10 melanoma	2×10^5 cells in the right hind foot dorsum
	C57BL/6	RM1 prostate cancer	2.5×10^4 cells in the right hind foot dorsum
	C57BL/6	3LL Lewis lung cancer	1×10^5 cells in the right hind foot dorsum

when the CBCT scans had been performed. As such, we could not define absolute HU thresholds for all scans, and these were adjusted when necessary.

When performing the segmentations, the program prompted the user to select the first and last CT slice containing the tumour tissue to define the superior and inferior borders of the tumour. The user was then asked to draw a ROI around the tumour in three different CT slices. Most of the tumours were considerably larger at the midpoint than at the edges, so only using a ROI at the midpoint would include a lot of normal tissue in the segmentation towards the end parts and only including ROIs at the edges would exclude a lot of tumour tissue, and as such, three manual ROIs were required. Empirically, we found that for the best performance, the locations of the ROIs were chosen so that one ROI was drawn close to the superior edge of the tumour, one close to the midpoint of the tumour and one close to the inferior tumour edge. With respect to the surrounding tissue and bone, the ROIs were drawn as exact tumour delineations (to the best of the observer's ability) to define the border between tumour and normal tissue, while generous margins were used with respect to air surrounding the tumour tissue, as air was automatically excluded based on the HU threshold. The ROIs were employed by the segmentation algorithm to exclude

anything outside the manually drawn ROIs as not being part of the tumour. As such, the program performed the automatic slice-by-slice tumour segmentation using the closest manually drawn ROI, as well as the HU-thresholding exclusion criteria.

As a final step, the volume segmentation was presented to the user as a three-dimensional image series, and the user was given the option to approve or redo the tumour segmentation. The tumour volume was then calculated as the number of voxels in the segmented volume multiplied by the volume of each voxel, that is, $0.275 \times 0.275 \times 0.275 = 0.0208 \text{ mm}^3$, and the data saved automatically.

Volume measurement by manual callipers

Manual calliper measurements were performed right before or after CBCT imaging, measuring length (l), width (w) and height (h) of the tumours. Calliper measurements were performed by different investigators (who routinely use manual callipers) to ensure that comparisons included the natural interobserver variation inherent to manual calliper measurements. Tumour volumes (V) were calculated using four different equations from previous publications^{1,2,5,7} to evaluate which equation was most accurate for determining tumour volumes in the different anatomical sites.

Figure 1. Examples of segmentations for three different tumours of the flank, thigh and dorsum of hind foot models are shown. The white lines represent the segmented tumour outline in CBCT slices other than those in which the manual regions of interest were drawn, thus representing automatically segmented tumour sections.

Figure 2. Comparison between gold standard and estimated flank tumour volumes. Data points on the left side of the identity line represent resected volumes being larger than the estimated volumes and *vice versa* on the right hand side. CBCT, cone beam CT; CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

$$V = \frac{(a \times b^2)}{2} \quad (3)$$

where a is the larger of the length and width and b is the shorter

$$V = \frac{\pi}{6} \times l \times w \times h \text{ based on the volume of an ellipsoid} \quad (4)$$

$$V = l \times w \times h \text{ based on the volume of a cuboid} \quad (5)$$

$$V = \frac{4\pi}{3} \times \frac{l \times w^2}{10} \text{ miscellaneous empirical volume formula} \quad (6)$$

Tumour volume comparison and validation

As the tumour-bearing animals were all part of on-going studies, tumours were resected and weighed following animal euthanasia for the animals where this was compliant with the main end point of the respective study. We then calculated the gold standard tumour volumes based on the resected tumour weights and a density of 1.030 g cm^{-3} based on the excellent correlation between subcutaneous tumour volume and weight previously published by Euhus *et al.*² There is of course a non-negligible uncertainty in these gold standard estimates based on potential difficulties in defining tumour tissue at the time of resection and the fact that the density may vary somewhat

between tumour models. To limit this uncertainty, the same experienced investigator performed the resections, and tumour tissue was consistently defined by its capsule in the flank and/or changes in tissue type or colour (specifically with regard to B16 melanoma tumours) in the thigh. The tumour volumes measured by CBCT or calculated based on manual calliper measurements were then compared with the gold standard tumour volumes to evaluate the accuracy of the different methods. Importantly, the different volume measurements were performed in a completely blinded set-up, that is, the observer taking the calliper measurements had no access to the CBCT volume data and *vice versa*. The different equations for calculating calliper-based tumour volumes were then compared directly to the CBCT volumes to evaluate the concordance between calliper measures and the semi-automatic CBCT segmentation. To ensure the robustness of these results across tumour models, we also performed the CBCT-to-calliper comparison in different subsets of tumours, used as independent validation cohorts.

Statistical analysis

To estimate the correlation between tumour volumes determined by the different methods and compared with the gold standard, we calculated the Spearman's rank correlation coefficient for each comparison. This is a non-parametric test and does not depend on the assumption of a linear relation between compared variables. However, only measuring the correlation

Figure 3. Comparison between the gold standard and estimated flank tumour volumes $<2\text{cm}^3$. CBCT, cone beam CT; CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

coefficient can be misleading since a perfect correlation could exist even if tumour volumes were consistently under- or overestimated. To get a more complete measure of the accuracy of tumour volume segmentations, we also calculated the root-mean-square deviation (RMSD) of the different volume comparisons as well as the coefficient of determination (C_D). The better the agreement between volumes, the lower the RMSD and the closer the correlation coefficient, and the C_D will be equal to one. A C_D value outside of the range 0–1 indicates that the data are a poor fit to the assumed linear model; in this case that

would mean the tumour volumes estimated using a certain method is in very poor agreement to the comparative method.

We also investigated the interobserver variation of the CBCT segmentation method and manual calliper measurements, respectively. Three different observers, experienced in pre-clinical tumour studies, independently estimated the volumes of ten flank and ten thigh tumours using both methods, these tumours were then resected and weighed to compare the estimates to the gold standard. The volumes for the calliper measurements were

Figure 4. Comparison between the gold standard and estimated thigh tumour volumes. Data points on the left side of the identity line represent resected volumes being larger than the estimated volumes and *vice versa* on the right hand side. CBCT, cone beam CT; CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

Figure 5. Comparison between the gold standard and estimated thigh tumour volumes $<1\text{cm}^3$. CBCT, cone beam CT; CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

calculated using Equation (3), as this is most commonly used in the Radiation Oncology Laboratory of the Institute for Onco-Physics.

RESULTS

Semi-automatic cone beam CT tumour segmentation

All tumour volumes were efficiently segmented by loading the DICOM files from the CBCT scans into the MATLAB program and defining the required ROIs. Based on five randomly chosen CBCT scans, we measured the average segmentation time, including the time to load the DICOM data, to 98 ± 9 s. Figure 1 shows some examples of the resulting tumour segmentation for each of the three different sites; in a CT slice different from those in which the manual ROIs were drawn.

As shown, the segmentation of the flank tumours appeared to be the most robust, as there was often a clearly visible, less X-ray dense, subcutaneous fat layer separating the tumour from the normal tissue. For these tumours, the border between fatty tissue and tumour also helped the HU thresholding to better follow the tumour outline throughout the segmentation procedure. For the

thigh models, the tumour was clearly visible on the CBCT scan although there was rarely a clear border separating the tumour tissue from the thigh muscle, making it more difficult to define the extent of tumour tissue. For the dorsum of hind foot tumours, it was not possible to distinguish the tumour tissue from the rest of the foot, so the whole tumour-bearing foot was segmented and the final tumour volume was subsequently calculated by automatically removing HUs resembling those of the bone. This highlights the difficulty in accurately estimating tumour volumes in the foot model, which also affects calliper measurements, usually resulting in the measurement of the foot + tumour tissue as one.

Accuracy of tumour volume segmentation compared with the gold standard

There was quite good agreement between the estimated tumour volumes and the gold standard, for both the CBCT volumes and some of the equations using calliper measures, as can be seen in Figure 2. Calculations using Equations (3), (4) and (6) show good agreement with resected tumours although Equation (5), which estimates a cuboid volume, consistently overestimates the

Figure 6. Comparison between the gold standard and estimated flank and thigh tumour volumes. Data points on the left side of the identity line represent resected volumes being larger than the estimated volumes and *vice versa* on the right hand side. CBCT, cone beam CT; CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

Figure 7. Interobserver variations are shown for tumour volumes estimated using cone beam CT (CBCT) segmentation and manual callipers, respectively. Horizontal solid lines show the mean difference compared with the resected tumour volumes and uncertainty bars represent the standard deviation. Average diff., average difference; average std., average standard deviation; average IQR, average interquartile range; V_{Calliper} , tumour volumes estimated using callipers; V_{CBCT} , tumour volumes estimated using CBCT; V_{Resected} , resected tumour volume.

tumour volumes, also shown by a negative coefficient of determination. There is a clear trend suggesting that the performance is worse for large tumour volumes, especially for the CBCT-based volumes. We therefore analysed subsets of the tumour volumes divided into >2 or ≤ 2 cm³. In Figure 3, we can see that for tumours <2 cm³, the CBCT-generated tumour volumes are in far better agreement with the gold standard than calliper-based volumes, with lower RMSD, higher C_D and better correlation. The opposite holds true for tumours >2 cm³, where CBCT performs worse than calliper-based measures, where Equation (4) appeared to be the most accurate (Supplementary material Figure A).

For the thigh tumours, the height was not measured for the tumours that were subsequently resected, and thus this comparison is limited to Equations (3) and (6). The results comparing estimated thigh tumour volumes to the gold standard are shown in Figure 4, where it is clear that manual calliper

measures are more accurate than the CBCT-based segmentation. Again, there is considerably worse agreement for large tumours, especially for the CBCT volumes. Here again, we performed a subset analysis dividing the tumours into >1 and ≤ 1 cm³, where the results for tumours <1 cm³ are shown in Figure 5. The performance of CBCT volume segmentation is dramatically improved for the smaller tumour volumes, comparable to the performance of manual callipers although still performing slightly worse with somewhat overestimated tumour volumes for CBCT and underestimated volumes for the calliper-based measures. The general performance of the CBCT tumour segmentation for both flank and thigh tumours is shown in Figure 6 for all tumours and in Supplementary material Figure B for tumours <1 cm³. Because of the poor performance for large tumours, the calliper measures performed generally better for all flank and thigh tumours, whereas the CBCT segmentation was superior or comparable for smaller tumours (≤ 1 cm³), depending on the chosen volume equation.

Figure 8. Comparison between cone beam CT (CBCT) and calliper-based flank tumour volumes. Data points on the left side of the identity line represent calliper-based volumes being larger than the estimated volumes and *vice versa* on the right hand side. CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

For the dorsum of hind foot tumours, there was no reliable way of estimating the tumour volume excluding the normal foot, so unfortunately gold standard tumour volumes could not be determined for those tumours.

Comparison of interobserver variation

The variation between three different observers was assessed using ten flank and ten thigh tumours, and the results are presented in Figure 7. There is considerable variation between observers using the calliper measurements, especially related to the flank tumour measurements. For the flank tumours, there was a consistent slight overestimation of the tumour volumes using the CBCT method, with little variation within the data for each observer. This was clearly shown by the substantial reduction in interquartile range of measurements for the CBCT method compared with manual callipers. For the thigh tumours, the variation among observers is comparable between CBCT and manual callipers, although with larger intraobserver variation for the calliper-based volumes.

Comparison of tumour volumes between cone beam CT and calliper measures

The tumour volumes estimated using the different methods were compared to determine which of the volume equations for the calliper measures showed the closest agreement to the CBCT

volumes and whether this was different between tumour sites. The comparison for flank tumours is shown in Figure 8, where Equation (3) shows excellent agreement with CBCT, whereas the other equations tend to over- or underestimate the volume compared with CBCT. To validate our findings, we performed the same comparison using an independent data set of flank tumours, which confirmed these results, although here Equations (3) and (4) performed equally well (see Supplementary material Figure C).

The CBCT tumour volumes compared with calliper-based tumour volumes for the thigh and dorsum of hind foot tumour models are shown in Figures 9 and 10, respectively. Thigh tumour volumes measured by callipers are systematically different when compared with CBCT-derived tumour volumes for all equations except Equation (5), which shows agreement for small tumours but considerable deviation for bigger tumours. This was, however, not confirmed in the validation data set where Equation (5) overestimated the volumes compared with CBCT, while the other equations compared quite well (see Supplementary material Figure D). The dorsum of hind foot tumour volumes based on calliper measures were all clearly different compared with the CBCT volumes, unsurprisingly, since reliable segmentation of those tumour volumes from CBCT was not possible. Precisely the same trend was seen in the dorsum of

Figure 9. Comparison between cone beam CT (CBCT) and calliper-based thigh tumour volumes. Data points on the left side of the identity line represent calliper-based volumes being larger than the estimated volumes and *vice versa* on the right hand side. CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

hind foot tumour validation data, given in Supplementary material Figure E.

To summarize the tumour volume comparisons between the different methods, all correlation estimates are presented in Table 2, for all three anatomical sites.

Recommendations

Based on these results we suggest the recommendations presented in Table 3 for estimating subcutaneous tumour volumes in small animals.

DISCUSSION

We showed in a completely blinded analysis that using a semi-automatic segmentation method, the volume of pre-clinical

subcutaneous flank, and if small enough (volume <1 cm³) also thigh tumours, can be efficiently estimated from CBCT scans already obtained in radiation studies using image-guided small-animal irradiators. This is to our knowledge, the first study to compare the efficacy of pre-clinical tumour volume estimation in different anatomical sites and to estimate the efficacy of using on-board CBCT imaging for tumour volume segmentation. The best tumour segmentation performance was seen for the flank tumours, which correlates well with the fact that the border between tumour and normal tissue was easier to distinguish in the flank model. This also explains why the reliability of CBCT tumour segmentation was poor for large flank and thigh tumours, as the larger tumours tended to be more invasive thus making it difficult to determine what was in fact tumour tissue. However, for most if not all pre-clinical studies in mice, tumour

Figure 10. Comparison between cone beam CT (CBCT) and calliper-based dorsum of hind foot tumour volumes. Data points on the left side of the identity line represent calliper-based volumes being larger than the estimated volumes and *vice versa* on the right hand side. CoeffDet, coefficient of determination; Eq, equation; RMSD, root-mean-square deviation.

volumes $<2\text{ cm}^3$ in the flank and 1 cm^3 in the thigh should cover the relevant range of appropriate tumour sizes.

As shown by Jensen *et al*,⁷ tumour volumes estimated using a micro-CT scanner and by delineating the tumour in each slice would likely lead to a slightly more accurate volume measurement, although it would be considerably less cost- and time-effective than our semi-automatic CBCT method. Another interesting approach for assessing pre-clinical tumour volumes was presented by Ayers *et al*⁵ using small-animal ultrasound imaging. Here, the authors showed a clear advantage compared with calliper measures by manually delineating the tumour in the ultrasound images, although this requires access to a small-animal ultrasound system. Both of these studies were based exclusively on flank tumours, so the role of ultrasound or micro-CT imaging for the thigh or dorsum of hind foot tumours remains unclear.

Several other methods for semi-automatic tumour segmentation based on CT scans are available in the literature for clinical,^{11–13} as well as pre-clinical investigations.^{14,15} The segmentation method presented in the present study is simple and efficient for pre-clinical subcutaneous tumours, whilst it is dependent on some manual input. More sophisticated methods employing features such as phase and gradient maps, landmark-based lesion tracking or region-growing methods based on non-Euclidian radial basis functions could be applied to further decrease the user dependence.^{11,13–15} Here, we find the manually drawn ROIs in our method to be a crucial component that provides the necessary flexibility needed to handle segmentation of tumours with considerable variation in shape and size. Irregular tumours with large variations between different CT slices will, however, be difficult to accurately segment using the presented semi-automatic method.

Table 2. Correlation results for all tumour volume comparisons presented as the Spearman's rank correlation (S_R), root-mean-square deviation (RMSD) and coefficient of determination (C_D)

Tumour site		CBCT	Equation (3)	Equation (4)	Equation (5)	Equation (6)
General (flank + thigh)	Resected tumour	$S_R = 0.91$	$S_R = 0.95$			$S_R = 0.96$
		RMSD = 0.58	RMSD = 0.35	–	–	RMSD = 0.36
		$C_D = 0.71$	$C_D = 0.87$			$C_D = 0.86$
General (<1 cm ³)	Resected tumour	$S_R = 0.95$	$S_R = 0.86$			$S_R = 0.87$
		RMSD = 0.13	RMSD = 0.13	–	–	RMSD = 0.11
		$C_D = 0.80$	$C_D = 0.75$			$C_D = 0.84$
Flank	Resected tumour	$S_R = 0.97$	$S_R = 0.95$	$S_R = 0.96$	$S_R = 0.96$	$S_R = 0.95$
		RMSD = 0.47	RMSD = 0.47	RMSD = 0.36	RMSD = 1.55	RMSD = 0.47
		$C_D = 0.88$	$C_D = 0.85$	$C_D = 0.91$	$C_D = -0.66$	$C_D = 0.85$
	CBCT	–	$S_R = 0.82$	$S_R = 0.86$	$S_R = 0.86$	$S_R = 0.78$
			RMSD = 0.22	RMSD = 0.27	RMSD = 0.40	RMSD = 0.24
			$C_D = 0.54$	$C_D = 0.33$	$C_D = -0.45$	$C_D = 0.46$
Flank (<2 cm ³)	Resected tumour	$S_R = 0.97$	$S_R = 0.90$	$S_R = 0.91$	$S_R = 0.91$	$S_R = 0.91$
		RMSD = 0.14	RMSD = 0.31	RMSD = 0.24	RMSD = 0.87	RMSD = 0.24
		$C_D = 0.95$	$C_D = 0.79$	$C_D = 0.84$	$C_D = 0.41$	$C_D = 0.84$
Flank (>2 cm ³)	Resected tumour	$S_R = 0.23$	$S_R = 0.70$	$S_R = 0.70$	$S_R = 0.70$	$S_R = 0.38$
		RMSD = 0.93	RMSD = 0.80	RMSD = 0.62	RMSD = 2.86	RMSD = 0.90
		$C_D = 0.18$	$C_D = 0.29$	$C_D = 0.39$	$C_D = -2.54$	$C_D = -0.42$
Thigh	Resected tumour	$S_R = 0.85$	$S_R = 0.95$	–	–	$S_R = 0.95$
		RMSD = 0.67	RMSD = 0.20			RMSD = 0.24
		$C_D = 0.04$	$C_D = 0.89$			$C_D = 0.84$
	CBCT	–	$S_R = 0.80$	$S_R = 0.82$	$S_R = 0.82$	$S_R = 0.79$
			RMSD = 0.15	RMSD = 0.15	RMSD = 0.09	RMSD = 0.16
			$C_D = 0.36$	$C_D = 0.35$	$C_D = 0.75$	$C_D = 0.26$
Thigh (<1 cm ³)	Resected tumour	$S_R = 0.96$	$S_R = 0.85$	–	–	$S_R = 0.87$
		RMSD = 0.14	RMSD = 0.10			RMSD = 0.12
		$C_D = 0.76$	$C_D = 0.81$			$C_D = 0.62$
Thigh (>1 cm ³)	Resected tumour	$S_R = -0.44$	$S_R = 0.90$	–	–	$S_R = 0.68$
		RMSD = 0.86	RMSD = 0.36			RMSD = 0.44
		$C_D = -1.86$	$C_D = 0.20$			$C_D = -0.10$
Foot	CBCT	–	$S_R = 0.67$	$S_R = 0.67$	$S_R = 0.67$	$S_R = 0.67$
			RMSD = 0.10	RMSD = 0.10	RMSD = 0.08	RMSD = 0.11
			$C_D = -1.27$	$C_D = -0.34$	$C_D = 0.54$	$C_D = -0.72$

CBCT, cone beam CT.

The volume estimation method(s) with superior performance are indicated in bold font.

For orthotopic tumour models such as tumours within the lung or liver, our simple segmentation method will likely not be sufficient, and the more sophisticated methods developed for these purposes will have to be employed.^{12,13,15} Although beyond the scope of this study, directly comparing the performance and speed of available segmentation methods for various subcutaneous and orthotopic pre-clinical tumour models may provide useful guidelines for

tumour volume measurements. Such an investigation would likely depend both on achievable image resolution and the potential use of contrast agents to highlight tumour tissue.

We also investigated whether there was a difference in tumour segmentation accuracy based on which tumour cell line was used, by colour-coding the volume comparisons accordingly.

Table 3. Recommended methods for estimating pre-clinical subcutaneous tumour volumes in longitudinal studies

Tumour site	Recommended method(s)	
Flank	For volumes $\leq 2 \text{ cm}^3$ CBCT or manual calliper estimates using: $V = \frac{\pi}{6} \times l \times w \times h$ [Equation (4)]	For volumes $> 2 \text{ cm}^3$ Manual calliper estimates using: $V = \frac{\pi}{6} \times l \times w \times h$ [Equation (4)]
Thigh	For volumes $\leq 1 \text{ cm}^3$ CBCT or manual calliper estimates using: $V = \frac{(a \times b^2)}{2}$ [Equation (3)]	For volumes $> 1 \text{ cm}^3$ manual calliper estimates using: $V = \frac{(a \times b^2)}{2}$ [Equation (3)]
Dorsum of hind foot	CBCT not recommended. Use manual callipers and equation based on the apparent geometric shape of the dorsum of hind foot tumours	

CBCT, cone beam CT.

There was no noticeable difference between the different cell lines, for flank or thigh tumours (data not shown), although we cannot exclude that such a difference exists as this study was not powered to show such a subgroup difference.

Previous studies have recommended using Equation (3) for estimating tumour volumes based on manual calliper measurements.^{2,5-7} Interestingly, we found that while Equation (3) was the most accurate for small thigh tumours and correlated best with CBCT in flank tumours, Equation (4) was shown to be the most accurate for flank tumours when compared with the gold standard. Since we use many different tumour cell lines to simulate the typical variation of tumour types in a pre-clinical setting, it seems that different subcutaneous tumour types grow in slightly different geometrical patterns. Both Equation (3) and (4) are, however, based on estimating ellipsoidal volumes suggesting that this is indeed the most appropriate geometric shape for approximating subcutaneous tumour volumes.

For tumours implanted in the dorsum of the hind foot, it is very difficult to distinguish the tumour tissue from the rest of the foot, and CBCT is not recommended for this tumour site. Using manual callipers, the investigator should take note of the shape of the tumour growth and apply a suitable geometrical formula when calculating the volume.

As the use of small-animal irradiators is becoming more common and widespread in pre-clinical radiation oncology

research, the method presented here can provide an accurate non-invasive alternative to manual calliper measurements, especially in flank tumours $< 2 \text{ cm}^3$. Importantly, the results presented here should be validated in a large independent data set of pre-clinical subcutaneous tumours to ensure the general usefulness of the volume segmentation method. The individual investigator can always adjust the trade-off between time and accuracy by changing the number of manually drawn ROIs that are required, should there be a need for more accurate measurements.

SUPPLEMENTARY INFORMATION

Supplementary material showing the comparison for flank tumour volumes $> 2 \text{ cm}^3$, flank and thigh tumour volumes $\leq 1 \text{ cm}^3$ and validation comparison data for flank, thigh and foot tumours is available online. Also included in the Supplementary material is the MATLAB code for the semi-automatic tumour volume segmentation program.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

FUNDING

This work has been supported in part by R01-EB009040 from the US National Institutes of Health.

REFERENCES

- Tomayko MM, Reynolds CP. Determination of subcutaneous tumor size in athymic (nude) mice. *Cancer Chemother Pharmacol* 1989; **24**: 148–54. doi: [10.1007/bf00300234](https://doi.org/10.1007/bf00300234)
- Euhus DM, Hudd C, LaRegina MC, Johnson FE. Tumor measurement in the nude mouse. *J Surg Oncol* 1986; **31**: 229–34. doi: [10.1002/jso.2930310402](https://doi.org/10.1002/jso.2930310402)
- Prasad SR, Jhaveri KS, Saini S, Hahn PF, Halpern EF, Sumner JE. CT tumor measurement for therapeutic response assessment: comparison of unidimensional, bidimensional, and volumetric techniques initial observations. *Radiology* 2002; **225**: 416–19. doi: [10.1148/radiol.2252011604](https://doi.org/10.1148/radiol.2252011604)
- Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000; **92**: 205–16. doi: [10.1093/jnci/92.3.205](https://doi.org/10.1093/jnci/92.3.205)
- Ayers GD, McKinley ET, Zhao P, Fritz JM, Metry RE, Deal BC, et al. Volume of preclinical xenograft tumors is more accurately assessed by ultrasound imaging than manual caliper measurements. *J Ultrasound Med* 2010; **29**: 891–901.
- Ishimori T, Tatsumi M, Wahl RL. Tumor response assessment is more robust with sequential CT scanning than external caliper measurements. *Acad Radiol* 2005; **12**: 776–81. doi: [10.1016/j.acra.2005.03.062](https://doi.org/10.1016/j.acra.2005.03.062)
- Jensen MM, Jørgensen JT, Binderup T, Kjaer A. Tumor volume in subcutaneous mouse xenografts measured by microCT is more accurate and reproducible than determined

- by 18F-FDG-microPET or external caliper. *BMC Med Imaging* 2008; **8**: 16. doi: [10.1186/1471-2342-8-16](https://doi.org/10.1186/1471-2342-8-16)
8. Matinfar M, Ford E, Iordachita I, Wong J, Kazanzides P. Image-guided small animal radiation research platform: calibration of treatment beam alignment. *Phys Med Biol* 2009; **54**: 891–905. doi: [10.1088/0031-9155/54/4/005](https://doi.org/10.1088/0031-9155/54/4/005)
 9. Wong J, Armour E, Kazanzides P, Iordachita I, Tryggestad E, Deng H, et al. High-resolution, small animal radiation research platform with x-ray tomographic guidance capabilities. *Int J Radiat Oncol Biol Phys* 2008; **71**: 1591–9. doi: [10.1016/j.ijrobp.2008.04.025](https://doi.org/10.1016/j.ijrobp.2008.04.025)
 10. Bazalova M, Nelson G, Noll JM, Graves EE. Modality comparison for small animal radiotherapy: a simulation study. *Med Phys* 2014; **41**: 011710. doi: [10.1118/1.4842415](https://doi.org/10.1118/1.4842415)
 11. Dankerl P, Cavallaro A, Dietzel M, Tsymbal A, Kramer M, Seifert S, et al. Clinical evaluation of semi-automatic landmark-based lesion tracking software for CT-scans. *Cancer Imaging* 2014; **14**: 6. doi: [10.1186/1470-7330-14-6](https://doi.org/10.1186/1470-7330-14-6)
 12. Tacher V, Lin M, Chao M, Gjestebj L, Bhagat N, Mahammedi A, et al. Semiautomatic volumetric tumor segmentation for hepatocellular carcinoma: comparison between C-arm cone beam computed tomography and MRI. *Acad Radiol* 2013; **20**: 446–52. doi: [10.1016/j.acra.2012.11.009](https://doi.org/10.1016/j.acra.2012.11.009)
 13. Velazquez ER, Parmar C, Jermoumi M, Mak RH, van Baardwijk A, Fennessy FM, et al. Volumetric CT-based segmentation of NSCLC using 3D-Slicer. *Sci Rep* 2013; **3**: 3529. doi: [10.1038/srep03529](https://doi.org/10.1038/srep03529)
 14. Ali R, Gunduz-Demir C, Szilágyi T, Durkee B, Graves EE. Semi-automatic segmentation of subcutaneous tumours from micro-computed tomography images. *Phys Med Biol* 2013; **58**: 8007–19. doi: [10.1088/0031-9155/58/22/8007](https://doi.org/10.1088/0031-9155/58/22/8007)
 15. Pellerin O, Lin M, Bhagat N, Ardon R, Mory B, Geschwind JF. Comparison of semi-automatic volumetric VX2 hepatic tumor segmentation from cone beam CT and multi-detector CT with histology in rabbit models. *Acad Radiol* 2013; **20**: 115–21. doi: [10.1016/j.acra.2012.07.011](https://doi.org/10.1016/j.acra.2012.07.011)