CI

e-med
www.e-med.org.uk

ORIGINAL ARTICLE

Assessing the effect of CT slice interval on unidimensional, bidimensional and volumetric measurements of solid tumours

Yongqiang Tan[a], Pingzhen Guo[a], Helen Mann[b], Sarah Elizabeth Marley[b],
Marietta Louise Juanita Scott[b], Lawrence H. Schwartz[a], Dana Cici Ghiorghiu[b], Binsheng Zhao[a]

[a]Department of Radiology, Columbia University Medical Center, New York, USA; [b]AstraZeneca UK, Alderley Park, Macclesfield, Cheshire, UK

Corresponding address: Binsheng Zhao, DSc, Department of Radiology, Columbia University Medical Center, 710 West 168th Street, New York, NY 10032, USA.
Email: bz2166@columbia.edu

Abstract

Objectives: To study the magnitude of differences in tumour unidimensional (1D), bidimensional (2D) and volumetric (VOL) measurements determined from computed tomography (CT) images reconstructed at 5, 2.5 and 1.25 mm slice intervals. **Materials and Methods:** A total of 118 lesions in lung, liver and lymph nodes were selected from 30 patients enrolled in early phase clinical trials. Each CT scan was reconstructed at 5, 2.5 and 1.25 mm slice intervals during the image acquisition. Lesions were semi-automatically segmented on each interval image series and supervised by a radiologist. 1D, 2D and VOL were computed based on the final segmentation results. Average measurement differences across different slice intervals were obtained using linear mixed-effects analysis of variance models. **Results:** Lesion diameters ranged from 6.1 to 80.1 mm (median 18.4 mm). The largest difference was seen between 1.25 and 5 mm (mean difference of 7.6% for 1D [$P < 0.0001$], 13.1% for 2D [$P < 0.0001$], -5.7% for VOL [$P = 0.0001$]). Mean differences between 1.25 and 2.5 mm were all within $\pm 3.5\%$ (within $\pm 6\%$ confidence interval). For VOL, there was a larger average difference between measurements on different slice intervals for the smaller lesions (<10 mm) compared with the larger lesions. **Conclusions:** Different slice intervals may give different 1D, 2D and VOL measurements. In clinical practice, it would be prudent to use the same slice interval for consecutive measurements.

Keywords: *Computed tomography (CT); response assessment; volume; measurement variability; slice interval.*

Introduction

Tumour diameter as measured on longitudinal computed tomography (CT) images is commonly used in oncology clinical trials and clinical practice to assess therapy response[1–3]. Over the past 3 decades, CT technology and computer hardware/software have improved significantly. With today's sub-millimetre isotropic spatial resolutions of CT images, plus advanced image analysis algorithms, accurate measurement of tumour volume is feasible. The advantages of volumetric CT over traditional diameter methods in the assessment of tumour response to therapy are currently under intensive investigation[4,5]. Tumour measurement can be affected by many factors arising from both the imaging acquisition and analysis procedures. Variability in tumour volumetric (VOL) measurements due to CT scanner, dose, slice interval and segmentation algorithms has previously been studied mainly using pulmonary nodules detected in CT lung cancer screening programs and in routine clinical practice for the purpose of non-invasive follow-up/diagnosis[6–9]. However, in those clinical settings, nodules were normally small (average approximately 10 mm) and the recommended image spatial resolutions were high (≤ 1 mm in the z-axial direction). The findings thus may not be applicable to the clinical trial setting where solid tumours are often large in size (average approximately 30–40 mm) and can originate from, and

spread to, any organ, especially the lungs, liver and lymph nodes.

When designing a clinical trial that incorporates volumetric CT for an investigational study of tumour response assessment, the first question is whether a slice interval of 5 mm, the most widely used imaging parameter in current cancer clinical trials and clinical practice, is appropriate for measuring tumour VOL. Although a few studies in lung cancer have reported the use of thinner slice intervals such as 2.5 mm or 1.25 mm (the potential slice intervals in future cancer clinical trials and clinical practice) when investigating the value of the volumetric response assessment[10–12], the magnitude of the difference in the VOL (as well as one-dimensional [1D] and two-dimensional [2D]) measurements of solid tumours across these different slice intervals has yet to be reported.

Understanding the magnitude of such differences is of vital importance, as it would help guide the design of the most appropriate CT imaging protocols for clinical trials and clinical care and improve the accuracy of interpreting tumour response to therapy. This is particularly important for phase I and II clinical trials in which a small number of patients are recruited, often from multiple centres. The small patient number and almost unavoidable differences in imaging equipment and acquisition/reconstruction techniques can make tumour measurement unreliable (noisy). Utilizing standardized imaging acquisition protocols and more accurate measurement metrics to quantify tumour changes is expected to lower measurement variability and thus require fewer number of patients, reduce trial duration and save costs for new drug development. The purpose of this study was to explore the differences in computer-aided 1D, 2D and VOL measurements of solid tumours found in lung, liver and lymph nodes (the 3 most metastatic sites of cancers) across the slice intervals of 5, 2.5 and 1.25 mm.

Materials and methods

Imaging data

The image data were collected from a prospective exploratory study utilizing CT scan data acquired as per protocol for tumour assessment in several phase I and II cancer clinical trials testing systemic agents at a single cancer centre. Both the clinical trials and the standalone protocol of this study requiring CT images reconstructed at additional slice intervals were approved by the institutional review board.

The study used routine clinically acquired, contrast-enhanced, diagnostic multidetector row CT (MDCT) scans (LightSpeed 16 or 64 detector row VCT; GE Healthcare, Milwaukee, WI) with a slice collimation of 5 mm. During the data acquisition, images were reconstructed at a 5 mm slice interval using the standard reconstruction algorithm as per the defined clinical trial protocols. For the purposes of this study, data were also reconstructed at 2.5 mm and 1.25 mm slice intervals using the standard reconstruction algorithm. All CT images were transferred electronically from the CT consoles to a research server where patient identification information was removed. Following removal of patient identification information, the images were retrieved from the research workstations for further image analysis and tumour measurement.

Target lesion selection

In all, 118 lesions from 30 patients enrolled in these clinical trials were selected by a radiologist (L.H.S) according to RECIST 1.0 for the target lesion selection. The number of lesions per patient ranged from 1 to 10 (median 3.5), with a maximum of 5 lesions per organ (lung, liver and lymph nodes). Lesions were selected based on the measurability and reproducibility criteria, as described by RECIST[2]. Among the 118 lesions, there were 39 lung nodules, 39 liver metastases and 40 lymph nodes. Primary and metastatic tumours in the lung, liver and lymph nodes were chosen because these are the 3 most common sites of cancer metastases. Although a longest diameter of 10 mm on baseline scan is the smallest size meeting the criteria for target lesions according to RECIST 1.0, some responsive lesions may become smaller than 10 mm at follow-up scan time points. For this reason, the inclusion of a small proportion (19) of the lesions less than 10 mm in their longest diameter was permitted. Of these 19 lesions, 8 were lung, 7 liver and 4 lymph node nodules. The size of a lesion was determined by the average 1D size measured at the 3 slice intervals using the methods described in the following section. Table 1 shows the number of lesions according to anatomic site and size groups.

Computer-aided measurements

All lesions were automatically segmented by an operator (Y.T.) using in-house computer algorithms developed for lung, liver and lymph node lesions. The algorithms were mainly based on the marker-controlled watershed segmentation, followed by the boundary smoothing techniques[13]. The computer-generated lesion contours were then overlaid on the original images for review by an experienced radiologist (P.G. who has more than 20 years' experience in interpreting radiographic images).

The radiologist reviewed and, if necessary, used the computer mouse to adjust the boundaries of all lesions on the images at each given slice interval before proceeding to the next slice interval, starting from 5 mm images, followed by the 2.5 mm images and then the 1.25 mm images. In this way, the finer structures potentially depicted on the thinner slice images could not contribute to decisions made at a coarser scale. The time interval between working on any 2 different slice images was more than 1 week to reduce memory effects.

Table 1 **Lesion distribution per anatomic site and size groups**

Site of metastasis	Size group (1D)		Total
	<10 mm	≥10 mm	
Liver	7	32	39
Lung	8	31	39
Lymph	4	36	40
Total	19	99	118

$N = 118$ lesions.

Appropriate window/level settings of lung (1500/−500), liver (150/90) and chest/abdomen (340/60) were used for reviewing and editing each type of lesion, respectively. The final results were called computer-aided measurements (CAM).

Based on the final result of a segmented lesion, VOL, 1D and 2D measurements could be calculated automatically by computer. Lesion VOL was calculated by multiplying the number of segmented lesion voxels by the VOL of a voxel (= spacing in the *x*-axis multiplied by spacing in the *y*-axis multiplied by the slice interval). On each slice, the line (diameter) that had the maximum distance between any 2 pixels on the lesion contour and was completely inside the lesion was selected and the 1D measurement was taken as the largest in-slice diameter across all tumour slices. The maximum perpendicular diameter is a line perpendicular to the maximum diameter with the longest length inside the lesion. The 2D measurement of a tumour was the product of the maximum diameter and the maximum perpendicular diameter.

Statistical analysis

Linear mixed-effects analysis of variance (ANOVA) models[14] were used to assess the effect of CT slice intervals (5, 2.5 and 1.25 mm) on metastatic tumour measurements calculated with 1D, 2D and VOL methods. Analyses were performed with SAS software (version 8.1; SAS Institute, Cary, NC, USA).

For each measurement method (1D, 2D and VOL), a linear mixed-effects model was fitted to the natural logarithm of the tumour measurement. Data were analyzed on the logarithmic scale to satisfy the assumptions of the linear mixed-effects model. Patient and nodule (intra-patient) were included as random effects in the ANOVA model (to account for the correlated nature of the data due to multiple measurements from each patient and lesion, respectively) and the slice interval was included as a fixed effect. Comparisons were performed to provide the least squares mean (LSM) estimates and corresponding 95% confidence intervals (CI) for the pairwise comparisons of the tumour measurements on different slice intervals (1.25, 2.5 and 5 mm). These point and interval estimates were exponentially back transformed to provide estimates of the percentage relative difference.

For each measurement method, the average percentage difference between slice interval 1 versus slice interval 2 was defined as: (Geometric LSM [measurement at slice interval 1/measurement at slice interval 2]−1) × 100%.

In addition to the analyses of all the 118 lesions measured, the above analyses were repeated, stratified by size group (<10 mm versus ≥10 mm) and by site of metastasis (liver, lung or lymph node). The size of each lesion was determined by the average 1D measurement made across the 3 slice intervals.

Linear mixed-effects models were also fitted to the logarithm of the tumour burden for each patient, rather than the individual tumour measurements. The tumour burden was calculated for each patient as the sum of the measurements of their individual lesions. Patient was included as a random effect in the ANOVA model and slice interval was included as a fixed effect. Comparisons were made as described above for the analyses of the individual lesions.

The agreement between individual tumour measurements made using different slice intervals (analyses at both the lesion and patient level) was assessed using the statistical techniques of Bland and Altman[15]. The difference between the 2 measurements (1.25 versus 2.5 mm; 1.25 versus 5 mm or 2.5 versus 5 mm) was plotted against the mean of the measurements and a reference range (i.e. limits of agreement) of $d \pm 1.96\,s$, where d is the mean difference between the 2 measurements and s is the standard deviation of the difference between the 2 measurements. The reference range was obtained using log transformed measurements and back transformed to give limits for the percentage relative difference of the actual measurements.

Results

Computer-aided 2 maximal diameters and same-plane lesion contours superimposed on 1.25 mm, 2.5 mm and 5 mm slice interval images for examples of lesions in liver, lung and lymph nodes are given in Fig. 1. There were statistically significant differences between measurements (1D, 2D and VOL) made on CT images reconstructed at different slice intervals. However, average percentage relative differences were relatively small and all within ±10% (except for the 2D measurements made between 1.25 mm and 5 mm slice images) (Table 2). Across all 3 types of measurement, the largest average difference was seen between 1.25 and 5 mm, with a mean percentage difference (95% CI) of 7.6% (6.3%, 9.0%; $P<0.0001$); 13.1% (10.0%, 16.3%; $P<0.0001$) and −5.7% (−8.4%, −2.9%; $P=0.0001$) for 1D, 2D and VOL, respectively. For 1D and 2D measurements, the smallest bias was between 1.25 and 2.5 mm, with a mean difference (95% CI) of 2.2% (0.9%, 3.4%; $P=0.0009$) and 3.1% (0.3%, 6.0%; $P=0.0311$), respectively. For VOL measurements, the smallest mean
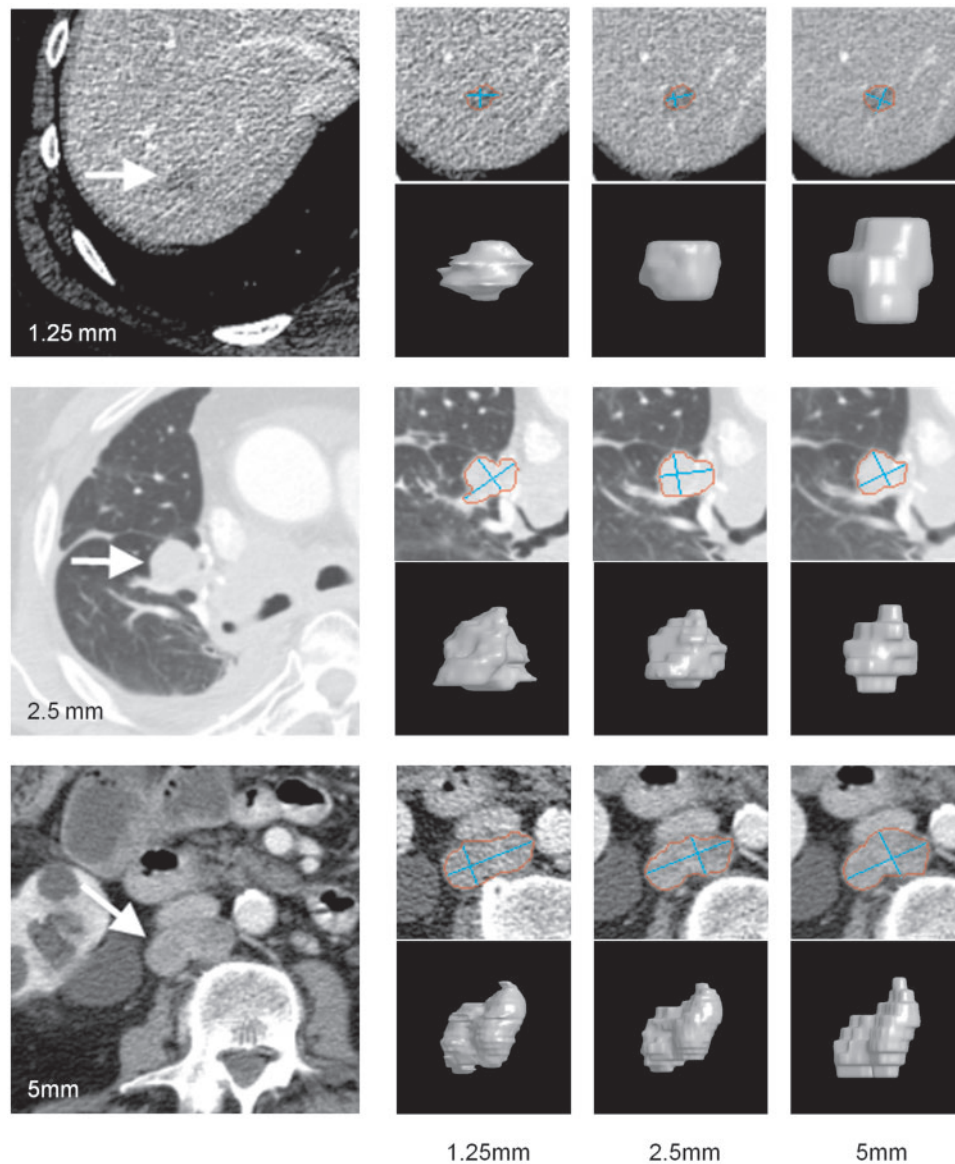
Figure 1   **Examples of lesions in liver, lung and lymph nodes on 1.25, 2.5 and 5 mm slice interval images. The left panel shows a close-up of (a) liver metastasis on a 1.25 mm slice interval image, (b) lung lesion on a 2.5 mm image, and (c) lymph node metastasis on a 5 mm image. The right panel shows the computer-aided contour, 2 maximal perpendicular diameters and three-dimensional views of each lesion on (from left to right, respectively) 1.25, 2.5 and 5 mm slice intervals.**

difference was between 2.5 mm and 5 mm, with a mean difference (95% CI) of -2.5% $(-5.3\%, 0.5\%;$ $P = 0.0969)$.

For VOL measurements, there was a bias towards larger measurements with a thicker slice interval (Table 2). For 1D and 2D measurements, there was a bias towards larger measurements with a smaller slice interval (Table 2). To understand how lesion size would affect the differences in measurements caused by the 3 slice intervals, the lesions were split into 2 subgroups based on a threshold of 10 mm (using the average 1D measurement across 3 slice intervals). Table 3 shows

that for VOL measurements, there was a much larger mean difference between the measurements made on different slice intervals for the small lesions compared with the larger lesions. As expected, the largest difference was found between 1.25 mm and 5 mm with a mean difference (95% CI) of -20.0% $(-30.3\%, -8.2\%; P = 0.0023)$ and -2.7% $(-4.8\%, -0.5\%; P = 0.0145)$ for small and large lesion VOL, respectively. There were differences in the observed bias between 1D and 2D measurements made on small and large lesions, but this was not consistently larger for either small or large lesion groups. The small lesions had similar low mean percent differences

Table 2 **Mean 95% confidence intervals and 95% reference ranges for the percentage relative differences between lesion measurements made on different slice intervals**

Measurement	Slice interval (mm)	% difference			P value
		Mean difference (%)	95% confidence interval (%)	95% reference range (%)	
1D	1.25 vs 5	7.6	(6.3, 9.0)	(−6.7, 24.1)	<0.0001
	2.5 vs 5	5.4	(4.0, 6.7)	(−7.7, 20.2)	<0.0001
	1.25 vs 2.5	2.2	(0.9, 3.4)	(−10.3, 16.3)	0.0009
2D	1.25 vs 5	13.1	(10.0, 16.3)	(−18.7, 57.4)	<0.0001
	2.5 vs 5	9.7	(6.7, 12.8)	(−20.3, 51.2)	<0.0001
	1.25 vs 2.5	3.1	(0.3, 6.0)	(−18.9, 31.0)	0.0311
VOL	1.25 vs 5	−5.7	(−8.4, −2.9)	(−35.5, 37.9)	0.0001
	2.5 vs 5	−2.5	(−5.3, 0.5)	(−27.1, 30.5)	0.0969
	1.25 vs 2.5	−3.3	(−6.1, −0.4)	(−26.3, 26.8)	0.0242

$N = 118$ lesions.

Table 3 **Mean and 95% confidence intervals for the percentage relative difference between lesion measurements made on different slice intervals for small and large lesion groups**

Measurement	Slice interval (mm)	Size group (1D):					
		<10 mm ($n = 19$), % difference			≥10 mm ($n = 99$), % difference		
		Mean difference (%)	95% confidence interval (%)	P value	Mean difference (%)	95% confidence interval (%)	P value
1D	1.25 vs 5	7.4	(3.3, 11.7)	0.0006	7.7	(6.3, 9.1)	<0.0001
	2.5 vs 5	9.4	(5.2, 13.7)	<0.0001	4.6	(3.3, 5.9)	<0.0001
	1.25 vs 2.5	−1.8	(−5.5, 2.1)	0.3572	2.9	(1.6, 4.3)	<0.0001
2D	1.25 vs 5	15.8	(7.1, 25.2)	0.0005	12.6	(9.4, 16.0)	<0.0001
	2.5 vs 5	20.5	(11.5, 30.3)	<0.0001	7.8	(4.7, 11.0)	<0.0001
	1.25 vs 2.5	−3.9	(−11.1, 3.9)	0.3046	4.5	(1.5, 7.6)	0.0033
VOL	1.25 vs 5	−20.0	(−30.3, −8.2)	0.0023	−2.7	(−4.8, −0.5)	0.0145
	2.5 vs 5	−7.5	(−19.5, 6.1)	0.2572	−1.4	(−3.6, 0.7)	0.1848
	1.25 vs 2.5	−13.5	(−24.7, −0.7)	0.0398	−1.2	(−3.4, 0.9)	0.2574

$N = 118$ lesions.

Table 4 **Mean and 95% confidence intervals for the percentage relative differences between lesion measurements made on different slice intervals by lesion type**

Measurement	Slice interval (mm)	Site of metastasis								
		Liver ($n = 39$), % difference			Lung ($n = 39$), % difference			Lymph ($n = 40$), % difference		
		Mean difference (%)	95% confidence interval (%)	P value	Mean difference (%)	95% confidence interval (%)	P value	Mean difference (%)	95% confidence interval (%)	P value
1D	1.25 vs 5	9.3	(6.6, 12.0)	<0.0001	5.8	(4.0, 7.6)	<0.0001	7.9	(5.4, 10.4)	<0.0001
	2.5 vs 5	5.6	(3.0, 8.2)	<0.0001	4.1	(2.3, 5.8)	<0.0001	6.4	(4.0, 8.9)	<0.0001
	1.25 vs 2.5	3.5	(1.0, 6.0)	0.0065	1.6	(−0.1, 3.4)	0.0583	1.4	(−0.9, 3.8)	0.2409
2D	1.25 vs 5	16.0	(10.6, 21.7)	<0.0001	10.5	(6.6, 14.6)	<0.0001	13.0	(6.5, 19.8)	<0.0001
	2.5 vs 5	11.6	(6.4, 17.1)	<0.0001	8.2	(4.3, 12.2)	<0.0001	9.5	(3.3, 16.1)	0.0028
	1.25 vs 2.5	4.0	(−0.9, 9.1)	0.1127	2.2	(−1.5, 6.0)	0.2431	3.2	(−2.7, 9.4)	0.2906
VOL	1.25 vs 5	−6.9	(−12.7, −0.7)	0.0297	−10.7	(−14.6, −6.7)	<0.0001	0.7	(−3.2, 4.8)	0.7207
	2.5 vs 5	−1.9	(−8.0, 4.6)	0.5603	−7.6	(−11.6, −3.4)	0.0007	2.2	(−1.8, 6.4)	0.2753
	1.25 vs 2.5	−5.1	(−11.0, 1.2)	0.1070	−3.4	(−7.6, 1.0)	0.1241	−1.5	(−5.3, 2.5)	0.4616

$N = 118$ lesions.

compared with those of the larger lesions between 1.25 mm and 2.5 mm for 1D and 2D measurements.

Table 4 shows measurement differences between different slice intervals when analyzed by site of metastasis. For the VOL measurement, the lung lesions showed the largest mean differences between the thinner (1.25 mm and 2.5 mm) and the thicker (5 mm) slice intervals compared with liver and lymph node metastases. This may

Table 5 **Mean and 95% confidence intervals for the percentage relative differences between lesion measurements made on different slice intervals at the total tumour burden (i.e. patient) level**

Measurement	Slice interval (mm)	% difference		
		Mean difference	95% confidence interval (%)	*P* value
1D	1.25 vs 5	6.3	(4.5, 8.1)	<0.0001
	2.5 vs 5	4.3	(2.5, 6.1)	<0.0001
	1.25 vs 2.5	1.9	(0.2, 3.7)	0.0314
2D	1.25 vs 5	9.2	(5.7, 12.9)	<0.0001
	2.5 vs 5	6.0	(2.6, 9.5)	0.0008
	1.25 vs 2.5	3.0	(−0.3, 6.5)	0.0731
VOL	1.25 vs 5	−3.2	(−5.5, −0.9)	0.0072
	2.5 vs 5	−1.9	(−4.2, 0.5)	0.1149
	1.25 vs 2.5	−1.4	(−3.7, 1.0)	0.2415

$N = 30$ patients.

have been due to the high contrast between the densities of the lung lesions and surrounding parenchyma, resulting in more partial volume artefacts on thicker than on thinner slice interval images. The lymph node metastases showed the smallest mean differences in VOL measurements across the 3 slice intervals (apart from a similar mean difference between 2.5 mm and 5 mm for the liver metastases). This may have been because the greatest extent of a lymph node metastasis is normally towards the *z*-direction, possibly reducing the effect of the slice interval on the VOL measurement. The liver lesions showed the largest differences in VOL measurements between 1.25 mm and 2.5 mm compared with the lymph node metastases and lung lesions. This may have been due to the high noise level of the abdominal images reconstructed at the 1.25 mm slice interval. Possibly for the same reason, liver lesions showed the largest mean differences across all slice interval images for 1D and 2D measurements (apart from the 1D measurements between 2.5 mm and 5 mm).

Each patient's tumour level data were summarized to provide a measure of tumour burden for 30 patients measured on 3 different slice intervals (Table 5). Mean percentage differences at the tumour burden (patient) level (Table 5) were smaller for all measurement types when compared with those from the tumour level analysis (Table 2). The largest average differences were between 1.25 mm and 5 mm slice intervals, while the smallest average differences were between 1.25 mm and 2.5 mm for all measurements.

Bland–Altman plots were used to examine the agreement between measurements on different slice intervals (Fig. 2). The plots show the greater agreement between measurements made on different slice interval images for larger lesions compared with the smaller lesions for VOL measurements. Across all 3 measurements (1D, 2D and VOL), there was most agreement between measurements made on 1.25 mm and 2.5 mm slice intervals, and least agreement between measurements made on 1.25 mm and 5 mm slice intervals (Table 2).

Discussion

To date, only a few studies have attempted to address the effects of the CT slice interval on tumour VOL measurements for the purpose of therapy response assessment[12,16]. This may be, in part, due to the lack of automated/semi-automated segmentation algorithms that can efficiently and accurately assist the measurement of lesion VOL. With the help of an elliptical approximation and a perimeter method, Winer-Muram et al.[16] measured the VOL of both phantom and lung lesions (13–38 mm) acquired on different slice intervals (2–10 mm). They proposed 2 different partial volume compensatory equations, based on the phantom data, for tumour VOL measured on different slice intervals, one for each of the 2 methods. The reported results, however, were controversial. Although smaller differences between tumour VOL obtained on thick and thin section images were observed in most of the lesions, the differences increased for approximately one-quarter to one-third of the lesions segmented with the different methods[16]. Zhao et al.[12] published a result reporting the effect of the slice interval on 1D, 2D and VOL measurements of lung metastasis. However, due to the technical limitations of the CT scanners (up to 4 detector row) at the time their study was conducted, without changing the then standard clinical imaging protocol of 7.5 mm slice collimation, they were only able to reconstruct the slice intervals down to 5 mm and 3.75 mm. They found that there was a significant difference in VOL measurements as slice intervals decreased from 7.5 to 5 mm and from 7.5 to 3.75 mm. But there were no significant differences between 5 mm and 7.5 mm for 1D, 2D and area measurements. This finding supported a smooth transition of the CT slice interval from 7.5 mm to 5 mm, the current standard, for assessing unidimensional and bidimensional measurements.

In the present study, we analyzed lesions not only in lungs, but also in liver and lymph nodes, as these are the 3 most frequent sites where tumours originate and spread. To study the effects of the slice interval on the
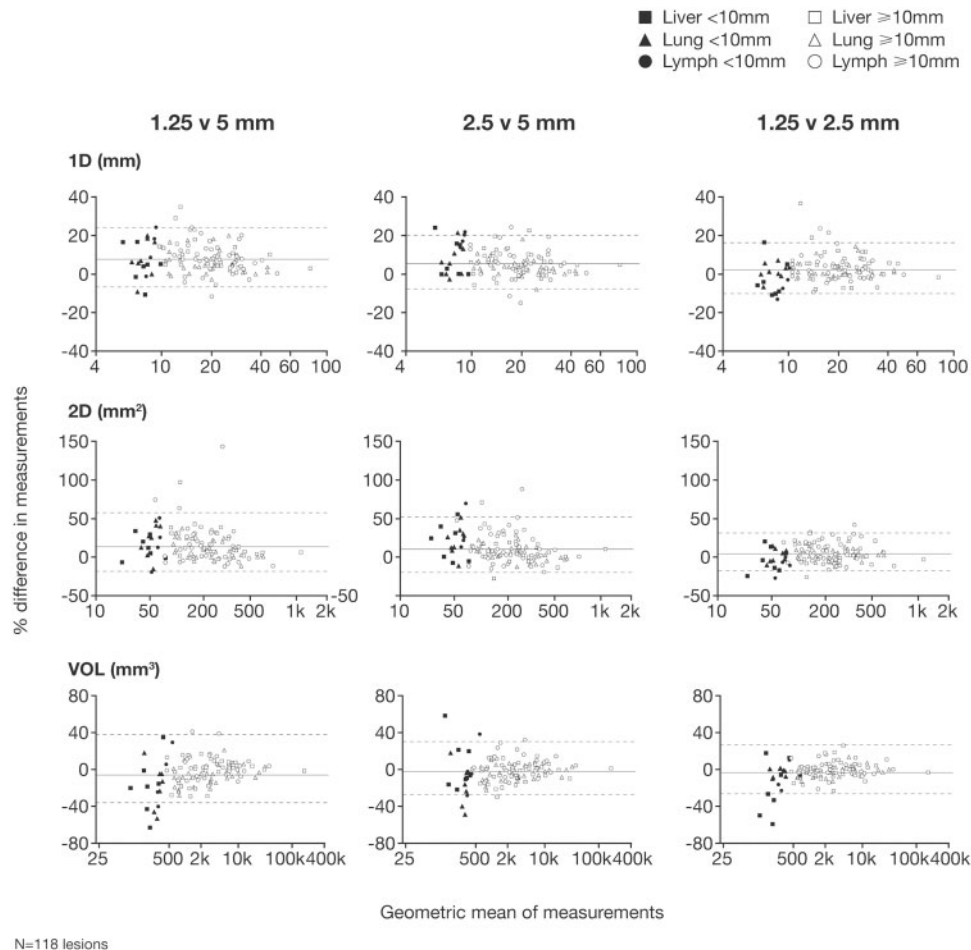
Figure 2 **Bland–Altman plots to assess the agreement between tumour measurements (1D, 2D and VOL) made on different slice intervals. The solid line shows the mean percentage difference in measurements and the dashed lines show the 95% reference range (data presented in Table 2).**

lesion size measurements, we chose the 3 slice intervals of 5 mm, 2.5 mm and 1.25 mm, which are either commonly used in today's clinical practice and clinical trials (e.g. 5 mm) or likely cover the range of slice intervals that may be used in the future (e.g. 2.5 mm and 1.25 mm).

This study found that the impact of slice interval was relatively small, although there were differences in the 3 measurements across the slice intervals. For all measurements, the largest biases were seen between 1.25 mm and 5 mm, whereas the smallest average differences were seen between 1.25 mm and 2.5 mm (except for the VOL, a slightly smaller average difference was found between 2.5 mm and 5 mm). The particularly small mean differences between 1.25 mm and 2.5 mm slice intervals for all 3 measurements (all within ±3.5%) indicate that 1.25 mm and 2.5 mm slice intervals may be interchangeable in the measurement of lesion size.

Our data also showed that tumour VOL measurements were biased towards larger results with a thicker slice

interval. This confirmed previous findings reported for both large lung lesions[12,16] and small lung nodules[8,9]. The reasons may be due to more obvious partial volume effects at thicker than at thinner slice intervals and because the computer-aided methods appeared to favour the incorporation of less dense areas on the top and bottom slices of a lesion (likely caused by the partial volume artefact) into the lesion VOL. For 1D and 2D measurements, however, there was a bias towards larger measurements with a thinner slice interval. This may be because there are more sampling slices along the z-axial direction for a thinner slice reconstruction. As a result, a larger in-plane extent (diameter) of the lesion has a higher likelihood of being captured by one of the thinner slices. Such trends were also noticed in Zhao et al.'s[12] previous work, between 3.75 mm and 5 mm, and between 3.75 mm and 7.5 mm slice intervals.

We analyzed the effects of the slice intervals on the tumour measurements by lesion size, categorizing the lesions as either small (<10 mm) or large (≥10 mm).

The mean differences in the lesion VOL between the different slice intervals were considerably larger in the small lesion group compared with those in the large lesion group. For 1D and 2D measurements, the mean differences across the 3 slice intervals remained almost unchanged, apart from 2D measurements in the small lesion group. Larger biases for 2D measurements in the small lesion group indicate that the slice interval of 5 mm may not be suitable for 2D measurements for lesions smaller than 10 mm.

The effects of the slice intervals on the tumour measurements by lesion type were also analyzed. The above observations were largely true for lung, liver and lymph nodes and therefore can be considered an artefact of data acquisition (slice interval) rather than tumour type/location or segmentation algorithm.

The mean percentage differences in tumour burden (i.e. patient level) measurements on different slice intervals were also analyzed and found to be consistently smaller when compared with those at the individual tumour level. One explanation for the VOL measurements, at least, could be that the larger biases were generally found for the smaller lesions, which had less influence on the average differences at the tumour burden level.

There were several limitations in this study. When comparing lesion measurements between longitudinal CT scan images reconstructed with inconsistent slice intervals, there are other variables associated with the re-scan, e.g. patient reposition and organ movement. However, it was not possible to consider these variables in this study because the 3 slice intervals were reconstructed from the same scan. The true VOL of the lesions were unknown, therefore we were unable to investigate which slice interval would produce the most accurate and reproducible VOL measurement. The measurements were performed by one radiologist using 3 in-house computer segmentation algorithms developed for lesions in lung, liver and lymph nodes. The results warrant further validation.

In summary, significant differences between lesion 1D, 2D and VOL measurements were found using different slice intervals, although the mean differences across the different slice intervals were generally low. 1D and 2D measurements were biased towards larger measurements with a thinner slice interval, whereas VOL measurements were biased towards larger measurements with a thicker slice interval. Across all 3 measurements, there was best agreement between the measurements made on 1.25 mm and 2.5 mm slice intervals, followed by 2.5 mm and 5 mm, and then 1.25 mm and 5 mm. With the VOL measurements, there was a larger average difference between measurements on different slice intervals for the small lesions (<10 mm) compared with the larger lesions (≥10 mm). We therefore suggest that the same imaging acquisition parameters be used to follow up the same patient's lesion changes during the course of therapy.

References

[1] Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer 1981; 47: 207–214. doi:10.1002/1097-0142(19810101)47:1<207::AID-CNCR2820470134>3.0.CO;2-6. PMid:7459811.

[2] Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst 2000; 92: 205–216. doi:10.1093/jnci/92.3.205. PMid:10655437.

[3] Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009; 45: 228–247. doi:10.1016/j.ejca.2008.10.026. PMid:19097774.

[4] Mozley PD, Schwartz LH, Bendtsen C, Zhao B, Petrick N, Buckler AJ. Change in lung tumor volume as a biomarker of treatment response: a critical review of the evidence. Ann Oncol 2010; 21: 1751–1755. doi:10.1093/annonc/mdq051. PMid: 20332135.

[5] Mozley PD, Bendtsen C, Zhao B, et al. Measurement of tumor volumes improves RECIST-based response assessments in advanced lung cancer. Transl Oncol 2012; 5: 19–25. PMid:22348172.

[6] Gavrielides MA, Kinnard LM, Myers KJ, Petrick N. Noncalcified lung nodules: volumetric assessment with thoracic CT. Radiology 2009; 251: 26–37. doi:10.1148/radiol.2511071897. PMid: 19332844.

[7] Wormanns D, Kohl G, Klotz E, et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. Eur Radiol 2004; 14: 86–92. doi:10.1007/s00330-003-2132-0. PMid:14615902.

[8] Petrou M, Quint LE, Nan B, Baker LH. Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. Am J Roentgenol 2007; 188: 306–312. doi:10.2214/AJR.05.1063.

[9] Way TW, Chan HP, Goodsitt MM, et al. Effect of CT scanning parameters on volumetric measurements of pulmonary nodules by 3D active contour segmentation: a phantom study. Phys Med Biol 2008; 53: 1295–1312. doi:10.1088/0031-9155/53/5/009. PMid:18296763.

[10] Riely GJ, Kris MG, Zhao B, et al. Prospective assessment of discontinuation and re-initiation of erlotinib or gefitinib in patients with acquired resistance to erlotinib or gefitinib followed by the addition of everolimus. Clin Cancer Res 2007; 13: 5150–5155. doi:10.1158/1078-0432.CCR-07-0560. PMid: 17785570.

[11] Zhao B, Oxnard GR, Moskowitz CS, Kris MG, Pao W, Guo P. A pilot study of volume measurement as a method of tumor response evaluation to aid biomarker development. Clin Cancer Res 2010; 16: 4647–4653. doi:10.1158/1078-0432.CCR-10-0125. PMid:20534736.

[12] Zhao B, Schwartz LH, Moskowitz CS, et al. Pulmonary metastases: effect of CT section thickness on measurement—initial experience. Radiology 2005; 234: 934–939. doi:10.1148/radiol.2343040020. PMid:15681690.

[13] Yan J, Zhao B, Wang L, Zelenetz A, Schwartz LH. Marker-controlled watershed for lymphoma segmentation in sequential CT

images. Med Phys 2006; 33: 2452–2460. doi:10.1118/1.2207133. PMid:16898448.

[14] Burton P, Gurrin L, Sly P. Extending the simple linear regression model to account for correlated responses: an introduction to generalized estimating equations and multi-level mixed modelling. Stat Med 1998; 17: 1261–1291. doi:10.1002/(SICI)1097-0258(19980615)17:11<1261::AID-SIM846>3.0.CO;2-Z. PMid:9670414.

[15] Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999; 8: 135–160. doi:10.1191/096228099673819272. PMid:10501650.

[16] Winer-Muram HT, Jennings SG, Meyer CA, et al. Effects of varying CT section width on volumetric measurement of lung tumors and application of compensatory equations. Radiology 2003; 229: 184–194. doi:10.1148/radiol.2291020859. PMid: 14519875.