CHEST

Ying Wang
Geertruida H. de Bock
Rob J. van Klaveren
Peter van Ooyen
Wim Tukker
Yingru Zhao
Monique D. Dorrius
Rozemarijn Vliegenthart Proença
Wendy J. Post
Matthijs Oudkerk

Volumetric measurement of pulmonary nodules at low-dose chest CT: effect of reconstruction setting on measurement variability

Y. Wang · P. van Ooyen · W. Tukker ·
Y. Zhao · M. D. Dorrius ·
R. V. Proença · M. Oudkerk (✉)
Department of Radiology,
University Medical Center Groningen,
University of Groningen,
Hanzeplein 1,
9700 RB Groningen, The Netherlands
e-mail: m.oudkerk@rad.umcg.nl
Tel.: +31-503-614260
Fax: +31-503-611798

G. H. de Bock · W. J. Post
Department of Epidemiology,
University Medical Center Groningen,
University of Groningen,
Groningen, The Netherlands

R. J. van Klaveren
Department of Pulmonology,
Erasmus Medical Center,
Rotterdam, The Netherlands

Abstract *Objective:* To assess volumetric measurement variability in pulmonary nodules detected at low-dose chest CT with three reconstruction settings. *Methods:* The volume of 200 solid pulmonary nodules was measured three times using commercially available semi-automated software of low-dose chest CT data-sets reconstructed with 1 mm section thickness and a soft kernel (A), 2 mm and a soft kernel (B), and 2 mm and a sharp kernel (C), respectively. Repeatability coefficients of the three measurements within each setting were calculated by the Bland and Altman method. A three-level model was applied to test the impact of reconstruction setting on the measured volume. *Results:* The repeatability coefficients were 8.9, 22.5 and 37.5% for settings A, B and C. Three-level analysis showed that settings A and C yielded a 1.29 times higher estimate of nodule volume compared with setting B (*P*=0.03). The significant interaction among setting, nodule location and morphology demonstrated that the effect of the reconstruction setting was different for different types of nodules. Low-dose CT reconstructed with 1 mm section thickness and a soft kernel provided the most repeatable volume measurement. *Conclusion:* A wide, nodule-type-dependent range of agreement between volume measurements with different reconstruction settings suggests strict consistency is required for serial CT studies.

Keywords Multi-detector computed tomography · Pulmonary nodules · Volume measurement · Section thickness · Reconstruction kernel

Introduction

Currently, an increasing number of pulmonary nodules are being detected incidentally with the evolution of multi-detector computed tomography (CT) technology and the use of high resolution images [1]. Assessment of changes in nodule volume to determine the volume doubling time based on serial CT examinations has proven to be a valuable tool in the differentiation between benign and malignant pulmonary nodules as malignant nodules typically present with a volume doubling time between 30 and 400 days [2–7]. Three-dimensional (3-D) volumetric measurements are more accurate and reproducible than two-dimensional measurements and are increasingly used in lung cancer screening programs and tumour response evaluation after chemotherapy [8–13].

However, none of the available 3-D software tools appears to be completely reproducible and a certain degree of variability exists [14–18]. Potential sources of volumetric measurement error are, in addition to the type of software used, also dependent on the reader, nodule characteristics, imaging and reconstruction settings [15–17]. Several studies [19–22] explored the effect of reconstruction settings on volume measurement variability, but they are derived from

normal-dose chest CT, and no data are available from low-dose chest CT so far. Furthermore, all these studies focused on the agreement between the volumes measured with the different reconstruction settings, but none of them addressed the repeatability of each individual reconstruction setting. For reconstruction settings with a lower repeatability, agreement is also expected to be lower. The purpose of this study was to evaluate volume measurement variability in pulmonary nodules detected with low-dose chest CT with three reconstruction settings including the repeatability and to quantify the effect of reconstruction settings on the volume measurement.

Patients and methods

Nodule selection

This study is a side study of the Dutch-Belgian multi-centre randomised low-dose CT lung cancer screening trial NELSON [9]. The NELSON study has been approved by the Minister of Health of the Netherlands and the ethics committees of all four participating hospitals. Informed consent was obtained from all participants. The original approval and informed consent for the screening study included the ability to use data for future research, including the current side study.

This side study concerned an analysis of category II and III solid nodules (15–500 mm^3, effective diameter 3.1–9.8 mm) according to the NELSON classification [9]. Participants screened in the Groningen NELSON study centre between September 2007 and February 2008 ($n=562$, 556 males, 6 females), were consecutively enrolled in this side study when they were diagnosed with non-calcified, solid nodules with a volume smaller than 500 mm^3. For every enrolled participant, all qualified nodules were included. Finally, 79 male and 3 female participants with 200 nodules were included in the analysis, having a mean ± SD age of 64 ±5 years.

Imaging

Low-dose chest CT examinations were performed by using 64 detector-row helical CT (Somatom-64, Siemens Medical systems, Forchheim, Germany) with the following parameters: 0.5 s tube rotation, 32×0.6 mm collimation with z-sample system and 28.8 mm table feed per rotation (pitch = 1.5). A caudal-cranial direction without contrast was used. Images were obtained from the level of the lung bases (posterior recesses) to the lung apex with the help of a scout view. Exposure settings were 20–30 mAs and 100–140 kVp depending on the weight of the participant. This corresponds to a dose-length-product <107 mGy·cm and an effective radiation dose <1.6 mSv. The fields-of-view of all three settings were the same and large enough to cover the complete lung cross-section.

Each low-dose chest CT examination was reconstructed using three different settings immediately after imaging with the following three combinations of section thickness/increment and kernel: (A) 1.0 mm/0.7 mm and a soft kernel (Siemens B30 filter), (B) 2.0 mm/1.4 mm and a soft kernel (Siemens B30) and (C) 2.0 mm/1.4 mm and a sharp kernel (Siemens B60). Setting A was the standard setting for the screening project [9], settings B and C were used for this particular study. Since a thicker section is less effective for nodule detection, only a 2 mm section thickness was chosen in our study [23]. The Siemens B30 kernel is the standard soft-tissue reconstruction kernel, and B60 is the bone reconstruction kernel, widely used in high resolution chest CT at normal dose levels.

Volumetric analysis

To evaluate the effect of reconstruction settings on volumetric measurement variability, each nodule was measured three consecutive times on every reconstructed CT image by one radiologist with 8 years of experience in chest CT. Digital workstations (Leonardo, Siemens Medical Solutions, Erlangen, Germany) were used for nodule volumetric analysis, with U.S. Food and Drug Administration (FDA) approved, commercially available software for semi-automated volume measurements (LungCare, Somaris/5, VB28B-W, Siemens Medical Solutions).

After the nodule was marked with a mouse click, the program automatically defined the volume of interest (VOI) around the nodule. A 3D template was generated, which represented the nodule in its optimal way, if possible without inclusion of surrounding structures. A second mouse click initiated the automated volume measurement, which included the following steps [14]: first, a fixed attenuation threshold of −400 HU was applied to extract a 3D-connected "structure of interest". This "structure of interest" can also include surrounding structures such as vessels or parts of the chest wall. Apart from this process, a small spherical 3D template originating from the click point was gradually expanded and its cross-correlation with the segmented nodule was computed for each step. The peak value of the cross-correlation curve was determined and an empirical cut-off value close to the peak value was used to separate the nodule from its potential adjacent structures. Finally, segmentation was completed by fusing the optimal 3D template and the "structure of interest". The segmented nodule was then shown in yellow on the volume-rendered display as the VOI. If the reader was not satisfied with the fit of the VOI, a manual modification of the segmentation by increasing or decreasing the volume could be performed. Manual adjustment of the VOI was not performed in this side study in order to avoid reader-induced volume measurement variability.

Nodule characteristics

The nodule characteristics (location and morphology) were determined by consensus double reading based on visual analysis of 1.0 mm thick axial images. Based on location, the nodules were classified into two subgroups: (1) purely intra-parenchymal, with little or no contact with the pleura, vessels or fissures or (2) attached, including pleural-based, fissure-attached and juxtavascular, with a pleura, fissure or vessel contact line larger than 50% of the nodule diameter.

Based on their morphology, nodules were classified into two groups: smooth-round and non-smooth-round (including polylobulated, spiculated and irregular nodules). Although attached and non-smooth-round nodules were reported to have high odds ratios of measurement variability, we decided to include them in order to be able to evaluate the effect of reconstruction settings on volume measurements for all nodule types encountered by CT screening [15, 16].

Statistical analysis

The measured volumes were logarithmically transformed (lnV) because of their skewed distribution. As the anti-log of differences on a log scale provides dimensionless ratios, the differences on the log scale were back-transformed to provide ratios among actual measurements. The ratios were further changed into relative volume difference (RVD, %) minus 100%.

Repeatability of each setting

As three repeated volume measurements of pulmonary nodules were performed by the same reader and the same software for each reconstruction setting, the mean difference between replicates is expected to be zero, implying that the backward transformed mean difference has a ratio of 100%. Thus, we calculated the repeatability coefficients according to the Bland and Altman method, which defined the value below which 95% of the differences will lie and thus a larger coefficient corresponds to a lower repeatability [24]. Repeatability coefficients were calculated with one way analysis of variance, defined as 2.77 times the within-subject standard deviation.

Effect of setting and repeat measures on the measured volume

A multi-level model approach was performed to get unbiased effect estimates of reconstruction settings on the measured volume. We used nested data in the sense that some persons had more than one nodule and within each nodule we had several measurements. The multi-level model approach is particularly appropriate for these types of data where correlations between different measurements within participants could be present [25]. In detail, we used a three-level model accounting for correlations within participants, between nodules and repeat measures, and accounting for correlations within nodules between repeated measures. Consequently, the highest level consists of the participants, the second level of the nodules and the lowest level of the repeated measures.

The reconstruction settings, nodule location and morphology were used as independent variables. Both random and fixed effects were considered. In the modelling process, variables were included in the model sequentially. After each step the goodness of fit was determined by the difference in deviance (-2 × log likelihood) between the present and the previous model. The modelling started with an empty model, in which only a fixed intercept and three random terms were modelled. Thereafter, reconstruction setting, nodule characteristics and their interaction were subsequently added. For this multi-level analysis, the MLwiN software package for fitting multi-level models was used (version 2.02; Centre for Multilevel Modelling, Institute of Education, University of London, UK).

Results

Of 82 participants, the median number of nodules per participant was two and ranged from one to nine (Table 1). Of 200 nodules, 36% (71/200) were attached, and 33.5% (47/200) were non-smooth-round.

Repeatability of each setting

The plots of the standard deviation of three measurements for every nodule against their mean for three settings are illustrated in Fig. 1. The repeatability coefficients of lnV were 0.08, 0.20 and 0.32 for settings A, B and C respectively. After antilog transformation and presented as RVD, the coefficients were 8.9, 22.5 and 37.5% for the three settings respectively. Non-smooth-round, attached nodules had a lower repeatability than smooth-round and

Table 1 The distribution of 200 nodules in 82 participants

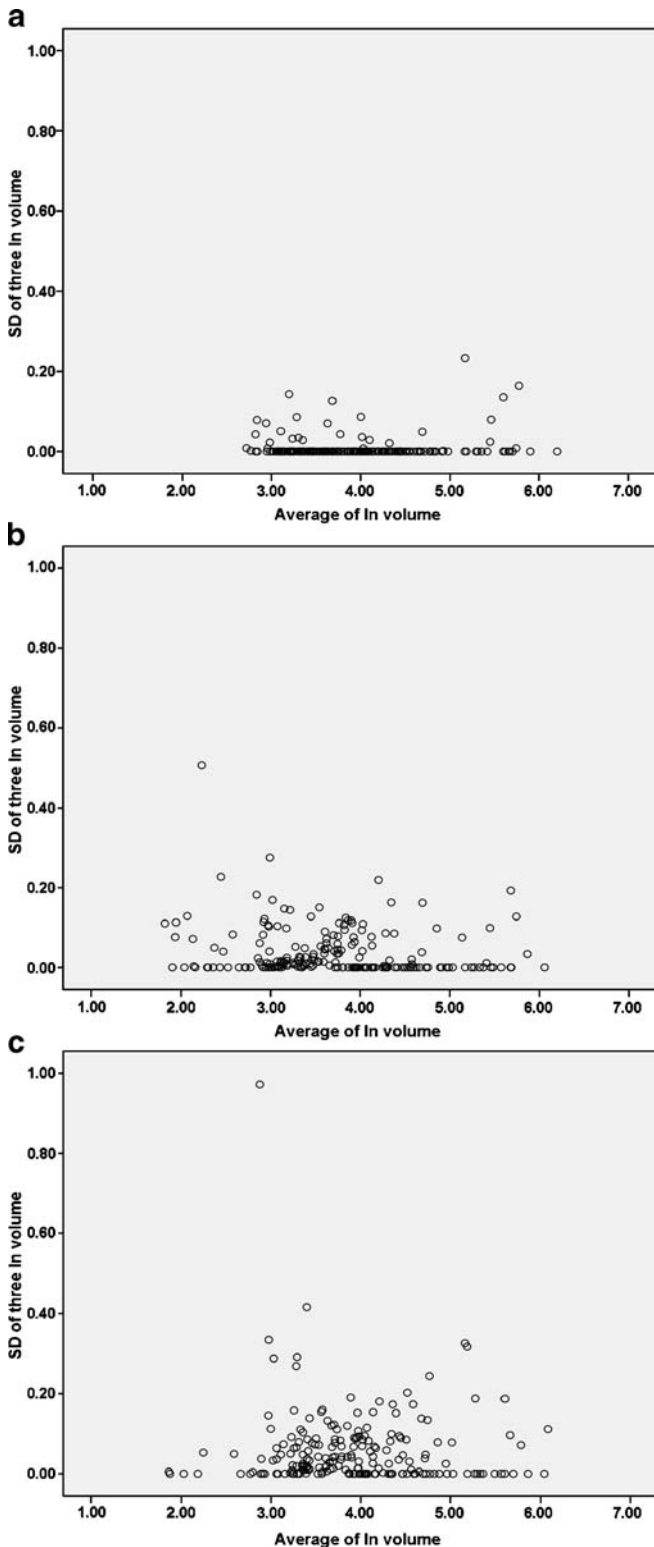| Number of nodules per participants	Number of participants (%)
1	34 (41)
2	18 (22)
3	11 (13)
4	9 (11)
5	4 (5)
6	3 (4)
7	2 (3)
9	1 (1)

Fig. 1 The plots of standard deviation of three measurements (logarithmically transformed volumes) for every nodule against their mean in each setting. **a** 1 mm section thickness and soft kernel, **b** 2 mm and soft kernel, **c** 2 mm and sharp kernel. Less variability was found in setting A than in B and C

purely intra-parenchymal nodules with settings A and C, but not with setting B (Table 2).

Effect of setting and repeat measures on the measured volume

The results of multi-level modelling are shown in Table 3. In this multi-level model, settings A and C both yielded a 1.29 [95% confidence interval (CI): 1.03–1.63] times higher estimate of the volume than setting B ($P=0.03$). Measuring time also yielded a significant though very small effect on the measured volume: the second and third measures yielded a 1.01 (95% CI: 1–1.02) times higher estimate for volume than the first measure, which indicated a 1% difference ($P=0.03$). It turned out that the differences between settings A and B, and between settings C and B were the same. The same held for the differences between time 2 and time 1, and for time 3 and time 1. Attached and non-smooth-round nodules yielded 1.74 (95% CI: 1.39–2.18) and 1.83 (95% CI: 1.42–2.37) times higher estimates of the volume than intra-parenchymal and smooth-round nodules ($P<0.01$), which indicated significant differences for the measured volumes of those types of nodules. In addition there was a significant interaction between setting and nodule characteristics (location and morphology) indicating that the effect of the reconstruction setting on the measured volumes was different for the different types of nodules ($P<0.05$).

Discussion

This study demonstrated that the repeatability of volume measurement of pulmonary nodules obtained at 1 mm section thickness combined with a soft kernel was almost two times better than that at 2 mm section thickness combined with a soft kernel and almost four times better than that obtained at 2 mm section thickness combined with a sharp kernel. Furthermore, the setting with 1 mm section thickness and a sharp kernel led to larger volumes, and the effect of setting on the measured volume was different for the different types of nodules.

The difference in repeatability among the three settings can serve as an indicator for choosing a certain reconstruction setting for low-dose chest CT, especially in lung cancer screening programs in which volumetric growth assessment is integrated into the nodule evaluation algorithm. Higher repeatability translates into more accurate growth assessment. Therefore, we recommend a reconstruction setting with 1 mm section thickness and a soft kernel as the standard reconstruction setting for low-dose CT lung cancer screening considering the volume measurement. Furthermore, if the reconstruction settings are different at follow-up, for example 2 mm at baseline and 1 mm at follow up, we should realise that a volume of 25% may not

Table 2 Repeatability: relative volume difference (%) among three assessments for each reconstructive setting in solid pulmonary nodules <500 mm^3 by location and morphology

Nodule characteristics		Number (%)	Relative volume difference (%)		
			1 mm and soft kernel (A)	2 mm and soft kernel (B)	2 mm and sharp kernel (C)
Location	Purely intra-parenchymal	129 (64.5)	6.0	23.6	26.2
	Attached	71 (35.5)	12.6	20.4	54.1
Morphology	Smooth-round	153 (76.5)	6.6	22.7	29.4
	Non-smooth-round	47 (23.5)	14.1	21.7	59.1
Total		200 (100)	8.9	22.5	37.5

represent real growth but could be caused by the use of different reconstruction settings. Therefore, consistency in the use of reconstruction settings is recommended during follow up in order to avoid unreliable growth assessments and false-positive test results in CT screening programs.

In our study we demonstrated that there is poor agreement between the volume measurements on CT reconstructed with different section thicknesses and kernels, which has also been reported in previous studies. In two studies including 42 pulmonary metastases and 75 pulmonary nodules, an increase in section thickness led to an increase in nodule volume measured [20, 21]. However, we found that an increase in section thickness to 2 from 1 mm led to a decrease in the measured volume, which is contradictory to previous studies. This can be explained by an averaging effect of the CT density of adjoining voxels when the section thickness increases. For example, if there are two adjoining voxels along the longitudinal axis with CT values of -600 and -300 HU, only one voxel will be included in the segmentation by the volumetric software as the applied threshold was -400 HU. When the section thickness is doubled, these two voxels will form one voxel with an average of -450 HU. This voxel will then be excluded from segmentation as it is below the threshold and subsequently lead to a decrease in volume compared with the original one. When the two original voxels have CT values of -500 and -200 HU, then the new voxel will be included in the segmentation and lead to an increase in the measured volume. Therefore, the effect of section thickness on measured volume is dependent on the density of the nodules.

Smaller nodules (mean volume 71 mm^3) in our study were more inclined to be influenced by the partial-volume effect and had lower CT density than larger nodules (mean volume 348 and 1,259 mm^3) in the above-mentioned studies. Consequently, the increase in the section thickness caused a decrease in the measured volume in our study, contradictory to the mentioned studies (Fig. 2). With respect to the reconstruction kernel, Honda et al. have reported that using a high-spatial frequency reconstruction algorithm led to significantly larger volumes than using low-spatial frequency algorithms. A low-spatial algorithm

Table 3 Effects of reconstruction setting, repeated measures and nodule characteristics on the measured volume by multi-level modelling

| Variable		Regression coefficient (95% confidence interval)	P value
Settings	A: 1 mm thickness and soft kernel	1.29 (1.03–1.63)	0.03
	B: 2 mm thickness and soft kernel	1	
	C: 2 mm thickness and sharp kernel	1.29 (1.03–1.63)	0.03
Measuring times	1	1	
	2	1.01 (1–1.02)	0.03
	3	1.01 (1–1.02)	0.03
Location	Intra-parenchymal	1	
	Attached	1.74 (1.39–2.18)	<0.01
Morphology	Smooth-round	1	
	Non-smooth-round	1.83 (1.42–2.37)	<0.01
Setting × location	B × intra-parenchymal	1	
	A × attached	0.93 (0.88–0.99)	0.03
	C × attached	0.90 (0.84–0.95)	0.01
Setting × morphology	B × smooth-round	1	
	A × non-smooth-round	0.97 (0.90–1.03)	0.36
	C × non-smooth-round	0.91 (0.85–0.98)	0.01
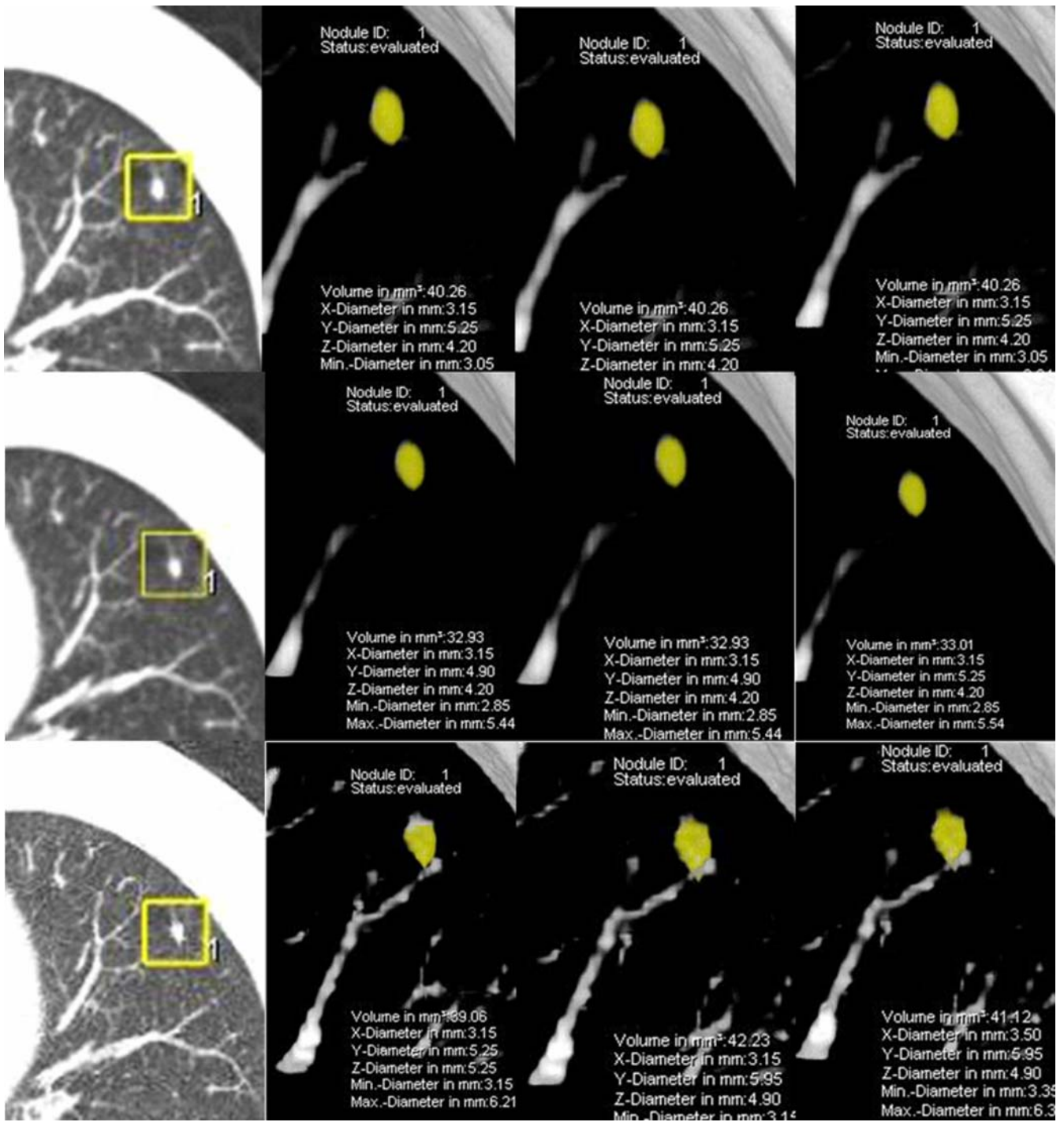
Fig. 2 Volume-rendered display showing three repeated measurements of a sub-centimetre intra-parenchymal pulmonary nodule with reconstruction settings A (*top*), B (*middle*) and C (*bottom*). Nodule measurements are completely identical with setting A (40.3 mm^3), slightly different with setting B (32.9, 32.9, 33.0 mm^3) and moderately different with setting C (39.1, 42.2, 41.1 mm^3). The means of measured volume with settings A and C were larger than with setting B (40.3 vs 32.9 vs 40.8 mm^3)

is equivalent to the soft kernel in our study [22]. The results of Honda et al. are in accordance with our results and demonstrate that a sharp kernel leads to a larger mean measured volume than a soft kernel. This can be explained by the fact that the spatial resolution in the *X-Y* plane increases when a sharp kernel is used and mild irregularities are included in the segmentation, which increases the measured volume (Fig. 2). The effect of reconstruction

settings on volume measurement was quantified by multi-level modelling and demonstrated that 1 mm section thickness and a sharp kernel led to larger volumes and that the effect was different for the different nodule types (Table 3). Therefore, consistency in the use of reconstruction settings is recommended for follow-up CT.

A limitation of our study is that we used only three reconstruction settings. Therefore, it is difficult to generalise the conclusions on the effect of section thickness and kernel. We decided to choose these three particular settings because they are most commonly used in low-dose chest CT. Furthermore, the accuracy of segmentation was not visually evaluated. It should be mentioned that manual adjustment of segmented volume will influence the standardisation of volume measurement in large populations and therefore was not applied in our study. Finally, the results of our study are only applicable for the specific software used. Different software or software versions may show different effects of reconstruction setting on measured volume.

In conclusion, low-dose CT reconstructed with 1 mm section thickness and a soft kernel provided the most repeatable volume measurement with special software used. A wide, nodule-type-dependent range of agreement between volume measurements at low-dose CT reconstructed with different settings suggested consistency in the use of reconstruction settings for serial CT.

References

1. Fischbach F, Knollmann F, Griesshaber V, Freund T, Akkol E, Felix R (2003) Detection of pulmonary nodules by multislice computed tomography: improved detection rate with reduced slice thickness. Eur Radiol 13:2378–2383
2. Hasegawa M, Sone S, Takashima S et al (2000) Growth rate of small lung cancers detected on mass CT screening. Br J Radiol 73:1252–1259
3. Lillington GA (1991) Management of solitary pulmonary nodules. Dis Mon 37:271–318
4. Revel MP, Merlin A, Peyrard S et al (2006) Software volumetric evaluation of doubling times for differentiating benign versus malignant pulmonary nodules. Am J Roentgenol 187:135–142
5. Aoki T, Nakata H, Watanabe H et al (2000) Evolution of peripheral lung adenocarcinomas: CT findings correlated with histology and tumor doubling time. Am J Roentgenol 174:763–768
6. Yankelevitz DF, Reeves AP, Kostis WJ (2000) Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. Radiology 217:251–256
7. Yankelevitz DF, Gupta R, Zhao B, Henschke CI (1999) Small pulmonary nodules: evaluation with repeat CT–preliminary experience. Radiology 212:561–566
8. Revel MP, Bissery A, Bienvenu M, Aycard L, Lefort C, Frija G (2004) Are two-dimensional CT measurements of small noncalcified pulmonary nodules reliable? Radiology 231:453–458
9. Xu DM, Gietema H, de Koning H et al (2006) Nodule management protocol of the NELSON randomised lung cancer screening trial. Lung Cancer 54:177–184
10. Erasmus JJ, Gladish GW, Broemeling L et al (2003) Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. J Clin Oncol 21:2574–2582
11. Kostis WJ, Yankelevitz DF, Reeves AP, Fluture SC, Henschke CI (2004) Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. Radiology 231:446–452
12. Erasmus JJ, Gladish GW, Broemeling L et al (2003) Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. J Clin Oncol 21:2574–2582
13. Tran LN, Brown MS, Goldin JG et al (2004) Comparison of treatment response classifications between unidimensional, bidimensional, and volumetric measurements of metastatic lung lesions on chest computed tomography. Acad Radiol 11:1355–1360
14. Sohaib SA, Turner B, Hanson JA, Farquharson M, Oliver RT, Reznek RH (2000) CT assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. Br J Radiol 73:1178–1184
15. Wormanns D, Kohl G, Klotz E et al (2004) Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. Eur Radiol 14:86–92
16. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ et al (2008) Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung cancer screening program. Radiology 248:625–631
17. Gietema HA, Wang Y, Xu D et al (2006) Pulmonary nodules detected at lung cancer screening: interobserver variability of semiautomated volume measurements. Radiology 241:251–257
18. Ko JP, Rusinek H, Jacobs EL et al (2003) Small pulmonary nodules: volume measurement at chest CT—phantom study. Radiology 228:864–870
19. Revel MP, Lefort C, Bissery A et al (2004) Pulmonary nodules: preliminary experience with three-dimensional evaluation. Radiology 231:459–466
20. Goo JM, Tongdee T, Tongdee R, Yeo K, Hildebolt CF, Bae KT (2005) Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy. Radiology 235:850–856

21. Petrou M, Quint LE, Nan B, Baker LH (2007) Pulmonary nodule volumetric measurement variability as a function of CT slice thickness and nodule morphology. Am J Roentgenol 188:306–312

22. Honda O, Sumikawa H, Johkoh T et al (2007) Computer-assisted lung nodule volumetry from multi-detector row CT: influence of image reconstruction parameters. Eur J Radiol 62:106–113

23. Kim JS, Kim JH, Cho G, Bae KT (2005) Automated detection of pulmonary nodules on CT images: effect of section thickness and reconstruction interval—initial results. Radiology 236:295–299

24. Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8:135–160

25. Snijders TAB, Bosker RJ (1999) Multilevel analysis: an introduction to basic and advanced multilevel modelling. Sage, Wiltshire, UK