EXPERIMENTAL

Marco Das
Julia Ley-Zaporozhan
H. A. Gietema
Andre Czech
Georg Mühlenbruch
Andreas H. Mahnken
Markus Katoh
Annemarie Bakai
Marcos Salganicoff
Stefan Diederich
Mathias Prokop
Hans-Ulrich Kauczor
Rolf W. Günther
Joachim E. Wildberger

Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners

M. Das (✉) · G. Mühlenbruch ·
A. H. Mahnken · M. Katoh ·
R. W. Günther · J. E. Wildberger
Department of Diagnostic Radiology,
RWTH Aachen University,
Pauwelsstrasse 30,
52072 Aachen, Germany
e-mail: das@rad.rwth-aachen.de
Tel.: +49-241-8035724
Fax: +49-241-8082411

J. Ley-Zaporozhan · H.-U. Kauczor
Department of Radiology,
German Cancer Research Center,
Heidelberg, Germany

H. A. Gietema · M. Prokop
Department of Diagnostic Radiology,
University Hospital Utrecht,
Utrecht, The Netherlands

A. Czech · S. Diederich
Department of Diagnostic Radiology,
Marienhospital Düsseldorf,
Düsseldorf, Germany

A. Bakai
CT Division,
Siemens Medical Solutions,
Forchheim, Germany

M. Salganicoff
CAD Applications,
Siemens Medical Solutions,
Malvern, PA, USA

Abstract The purpose of this study was to compare the accuracy of an automated volumetry software for phantom pulmonary nodules across various 16-slice multislice spiral CT (MSCT) scanners from different vendors. A lung phantom containing five different nodule categories (intraparenchymal, around a vessel, vessel attached, pleural, and attached to the pleura), with each category comprised of 7–9 nodules (total, n=40) of varying sizes (diameter 3–10 mm; volume 6.62 mm^3–525 mm^3), was scanned with four different 16-slice MSCT scanners (Siemens, GE, Philips, Toshiba). Routine and low-dose chest protocols with thin and thick collimations were applied. The data from all scanners were used for further analysis using a dedicated prototype volumetry software. Absolute percentage volume errors (APE) were calculated and compared. The mean APE for all nodules was 8.4% (±7.7%) for data acquired with the 16-slice Siemens scanner, 14.3% (±11.1%) for the GE scanner, 9.7% (±9.6%) for the Philips scanner and 7.5% (±7.2%) for the Toshiba scanner, respectively. The lowest APEs were found within the diameter size range of 5–10 mm and volumes >66 mm^3. Nodule volumetry is accurate with a reasonable volume error in data from different scanner vendors. This may have an important impact for intraindividual follow-up studies.

Keywords Multislice CT ·
Pulmonary nodule · Nodule
volumetry · Automated volumetry ·
Chest CT

Introduction

Pulmonary nodules are a common finding on chest CT examinations. Especially the increased use of multislice CT (MSCT) has led to an increased detection rate of pulmonary nodules [1, 2]. They can represent a wide variety of pulmonary pathologies [3, 4]. As shown in large screening trials, most of the small nodules (<5 mm) are benign [5–9],

and some of them disappear within follow-up [10]. However, with increasing nodule size (>5 mm), the likelihood of malignancy increases. Thus, it is important to differentiate between benign and malignant pulmonary nodules. Calcified pulmonary nodules are usually benign, while for solid (non-calcified) nodules guidelines suggest accurate size measurement, growth assessment or biopsy for nodule workup [11]. As a CT-guided biopsy or further

invasive workup is not practical for all detected pulmonary nodules, follow-up studies are performed to detect growth, usually based on manual axial measurements. However, these measurements are user dependent and may vary significantly between different radiologists (interreader variability) or even within the same radiologist (intrareader variability) [12]. Precise and reliable measurement of a pulmonary nodule is also crucial for therapy evaluation, e.g., in follow-up studies after chemotherapy. It has been shown that in terms of accuracy and reproducibility, the volumetric measurement of a nodule is more sensitive in the detection of growth and determination of tumor doubling time than the axial diameter measurement [13]. Algorithms have been developed and have been integrated into software tools for automated assessment of nodule volume. Initial studies have shown the value of the volumetry in different settings [14–24]. To our knowledge to date none of the algorithms has been tested on datasets acquired at different MSCT scanners from different vendors. This is of high importance as patients may undergo screening, staging or follow-up exams at different sites and scanners, and scanner-specific parameters may potentially influence volumetric measurements. The purpose of our study was to evaluate the accuracy of automated pulmonary nodule volumetry software using a phantom with artificial pulmonary nodules of known size and volume, which was scanned at four different widely used commercially available 16-slice MSCT scanners from different vendors.

Materials and methods

Phantom specification

A chest phantom (Figs. 1 and 2, QRM, Möhrendorf, Germany) containing five different categories of pulmo-
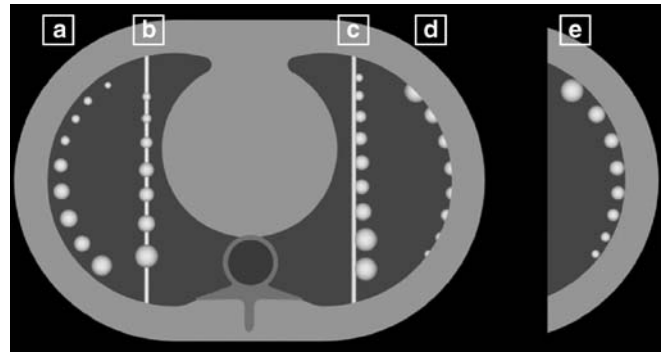

Fig. 1 Phantom in the Scanner gantry

Fig. 2 Scheme of the five different categories of nodules in the phantom

nary nodules (each category seven to nine nodules) was scanned. The phantom simulates an 11-cm-thick section of the chest with axial dimensions of 30×20 cm. The lung parenchyma is simulated with small cork pieces. The phantom contains a total of 40 nodules surrounded by these cork pieces. These nodules are manufactured in a round shape with precise knowledge of their volume. The nodules consist of a soft tissue equivalent material, with a physical density of 1.1 g/cm^3 and a CT density of 35 HU at 120 kV. Category A contains isolated intraparenchymal nodules, category B nodules around a vessel, category C nodules attached to a vessel, category D pleural nodules and category E nodules attached to the pleura. Each nodule category includes nodules with diameters from 3 mm to 10 mm and volumes between 13.24 mm^3 and 524.97 mm^3 (with the exception of category D in which the smallest volume is 6.62 mm^3).

Scan protocols

In this multicenter study the phantom was scanned at four different 16-slice MSCT scanners (SOMATOM Sensation 16, Siemens Medical Solutions, Forchheim, Germany; LightSpeed Pro 16, GE Healthcare Technologies, Waukesha, WI; Brilliance 16, Philips Medical Systems, Eindhoven, The Netherlands; Aquilion 16, Toshiba Medical Systems Corporation, Tokyo, Japan) with comparable standard-dose and low-dose protocols (Table 1). Thin- and thick-collimated examination protocols with comparable pitch settings with consecutive thin image reconstruction were applied. Standard sharp reconstruction kernels (as specified by each vendor) for lung window settings were used as recommended by each vendor. In total, 1,280 nodule measurements were obtained and compared to the true volume (four vendors $\times 2$ collimations $\times 2$ section thicknesses $\times 2$ dose settings $\times 40$ nodules).

Table 1 Overview of different scanners and scanner protocols

Scanner type	Vendor	Tube voltage [kV]	Collimation [mm]	Rotation time [s]	Tube current time product [mAs$_{eff.}$]	CTDIvol [mGy]
SOMATOM Sensation 16	Siemens	120	16×1.5	0.5	20	1.38
SOMATOM Sensation 16	Siemens	120	16×1.5	0.5	100	9.33
SOMATOM Sensation 16	Siemens	120	16×0.75	0.5	20	1.21
SOMATOM Sensation 16	Siemens	120	16×0.75	0.5	100	9.15
LightSpeed Pro 16	GE	120	16×1.25	0.5	20	2.31
LightSpeed Pro 16	GE	120	16×1.25	0.5	100	9.18
LightSpeed Pro 16	GE	120	16×0.625	0.5	20	1.43
LightSpeed Pro 16	GE	120	16×0.625	0.5	100	9.56
Brilliance 16	Philips	120	16×1.5	0.5	20	1.9
Brilliance 16	Philips	120	16×1.5	0.5	100	9.31
Brilliance 16	Philips	120	16×0.75	0.5	20	2.01
Brilliance 16	Philips	120	16×0.75	0.5	100	9.25
Aquilion 16	Toshiba	120	16×1	0.5	20	2.21
Aquilion 16	Toshiba	120	16×1	0.5	100	9.41
Aquilion 16	Toshiba	120	16×0.5	0.5	20	1.5
Aquilion 16	Toshiba	120	16×0.5	0.5	100	9.64

Nodule volume evaluation

The data from all scanners were stored on a hard drive and were transferred to a separate workstation for analysis using a commercially available lung analysis software (LungCARE; Siemens) with an implemented prototype algorithm for automated nodule volumetry. The user manually places a marker in each nodule. A 3D volume of interest (VOI) with a fixed volume of 30×30×30 mm^3 around the marked nodule is automatically generated. The next mouse click initiates the automated region growing segmentation of the nodule. Finally, nodule dimensions, density and volume are displayed. For this study, no further user interaction, although implemented in the software, was allowed. At the time of the measurements, the user performing the volumetry was blinded to the real volume of the nodules. Images were displayed in a lung window setting with a standardized window width of 1,200 HU and a center of −700 HU. Maximum intensity projections (MIP) were allowed; the automated nodule segmentation, however, is performed on the original axial sections. From the difference between the true nodule volume and the measured nodule volume, the absolute volume errors (APE) were calculated for each measurement. In addition, comparison of measurements from different scan protocols and different scanner types were made.

Data analysis

The absolute percentage error (APE) was determined as $APE = 100 \times |V_m - V_{rs}|/V_{rs}$ [22].

V_m represents the measured volume, and V_{rs} represents the true known reference volume of the phantom nodule. Linear correlation (Pearson r) was tested to evaluate the relationship between the measured volume and the reference volume. Mean APE values and the coefficient of determination (r^2) were calculated for all protocols and scanners. Three-way ANOVA was performed to detect any influence of the vendor, collimation and dose. All statistical analyses were performed using SAS statistical analysis software (The SAS System for Windows, JMP Version 5.0.1.2.; SAS Institute Inc., Cary, NC).

Results

The overall APE for SOMATOM Sensation 16 was 8.4% (±7.7%), for GE LightSpeed 16 14.3% (±11.1%), for Philips Brilliance 16 9.7% (±9.6%) and for Toshiba Aquilion 16 7.5% (±7.2%).

Especially the nodules of category D (pleural nodules) had very high APEs for all scanners and in all protocols. Looking at nodules with a volume larger than 66.29 mm^3 (diameter ≥5 mm), the APEs were smaller.

Siemens

The overall lowest APE (3.3%±2.4%) was found in the standard dose protocol with 1.5-mm collimation for nodule category B. The highest APE (19%±7.8%) was found in the low-dose protocol with 1.5-mm collimation for nodule category D (Table 2).

Table 2 Error in nodule volume measurements for different protocols and nodule categories (A = intraparenchymal, B = around a vessel, C = attached to a vessel, D = pleural and E = attached to the pleura). LD: low dose; SD: standard dose

SOMATOM Sensation 16	Mean APE (absolute percentage error)			
Collimation/dose	0.75 LD	1.5 LD	0.75 SD	1.5 SD
Category A	8.96±7.49	9.78±7.97	11.67±8.76	10.5±6.71
Category B	9.36±17.86	4.92±1.89	5.07±3.82	3.3±2.43
Category C	4.35±4.78	4.45±3.07	4.46±4.52	4.8±5.09
Category D	6.73±4.03	19.03±7.76	12.61±5.44	14.37±9.02
Category E	6.12±5.59	9.44±2.65	6.66±5.44	12.06±5.28
Size ≤33.86 mm^3	15.82±13.98	13.76±10.03	13.82±6.98	12.68±8.7
Size ≥66.29 mm^3	4.03±4.3	8.10±5.98	6.16±5.5	7.83±6.5

GE

The overall lowest APE (9.3%±7.3%) was found in the standard dose protocol with 0.625-mm collimation for nodule category B. The highest APE (21.2%±10.0%) was found in the standard dose protocol with 1.25-mm collimation for nodule category D (Table 3).

Philips

The overall lowest APE (2.6%±2.6%) was found in the low-dose protocol with 1.5-mm collimation for nodule category C. The highest APE (15.8%±9.5%) was found in the low-dose protocol with 1.5-mm collimation for nodule category A (Table 4).

Toshiba

The overall lowest APE (3.3%±2.1%) was found in the low-dose protocol with 1-mm collimation for nodule category C. The highest APE (11.8%±16.5%) was found in the low-dose protocol with 0.5-mm collimation for nodule category D (Table 5).

Pearson correlation showed a very good correlation for all measured nodules for all protocols in all scanner types with the real volume. Overall correlation for Siemens, GE, Philips and Toshiba was r^2=0.99, r^2=0.98, r^2=0.99 and r^2=0.99.

A statistically significant influence of the vendor could be detected (P=0.0044). Additionally, a significant influence of the collimation was detected (P=0.0212) and consecutively of the slice thickness was found (P=0.0192), but interestingly no statistically significant influence of the dose settings could be detected (P=0.0992).

Discussion

Since computed tomography is the method of choice for imaging of pulmonary nodules, an increasing number of pulmonary nodules needs to be evaluated and followed over time [3, 5–9]. Thus, it is of increasing importance to use a precise and reliable method for quantitative assessment of pulmonary nodules in chest CT [21, 22, 25]. Especially in screening settings patients frequently undergo follow-up examinations for growth evaluation of pulmonary nodules. Sometimes follow-up is not performed on the same scanner, and the CT protocol may not be identical. Therefore, it is of great importance to achieve reliable objective volumetry results in all different scanner types from different vendors.

In a study performed by Goo et al. [22], simulated nodules of different sizes were used, and a significant influence of nodule size, section thickness and threshold settings was found. But compared to our study, only nodules surrounded by lung parenchyma with low CT attenuation were included; no nodules attached to the pleura or attached to a vessel were included, and all scans

Table 3 Error in nodule volume measurements for different protocols and nodule categories (A = intraparenchymal, B = around a vessel, C = attached to a vessel, D = pleural and E = attached to the pleura). LD: low dose; SD: standard dose

LightSpeed 16 GE	Mean APE (absolute percentage error)			
Collimation/dose	0.625 LD	1.25 LD	0.625 SD	1.25 SD
Category A	18.07±8.13	18.07±8.13	19.24±13.63	14.73±13.32
Category B	9.57±14.99	10.19±14.61	9.31±7.32	9.67±4.88
Category C	10.3±10.12	9.83±10.44	13.32±14.28	9.96±8.53
Category D	9.93±6.86	9.95±6.88	20.34±14.47	21.19±10.03
Category E	17.4±11.43	17.40±11.43	15.84±12.42	12.3±7.02
Size ≤33.86 mm^3	19.77±13.58	19.79±13.57	34.31±14.70	22.83±10.36
Size ≥66.29 mm^3	10.84±8.28	10.84±12.11	16.38±10.6	13.93±10.57

Table 4 Error in nodule volume measurements for different protocols and nodule categories (A = intraparenchymal, B = around a vessel, C = attached to a vessel, D = pleural and E = attached to the pleura). LD: low dose; SD: standard dose

Philips Brilliance 16	Mean APE (absolute percentage error)			
Collimation/dose	0.75 LD	1.5 LD	0.75 SD	1.5 SD
Category A	14.15±9.45	15.82±9.51	14.12±9.51	15.69±9.49
Category B	5.93±2.93	5.78±3.5	5.76±2.1	5.34±2.8
Category C	4.04±6.04	2.62±2.59	3.22±3.19	13.61±32.50
Category D	14.73±7.73	14.99±9.44	12.6±6.53	14.53±8.85
Category E	8.24±3.22	8.81±3.33	7.7±3.38	7.63±3.81
Size ≤33.86 mm^3	15.86±8.51	16.43±10.37	12.46±7.61	14.44±10.95
Size ≥66.29 mm^3	7.26±5.92	7.7±6.22	7.28±6.0	7.71±6.24

were conducted at only one single CT scanner. In a different study by Wiemker et al. [16], the influence of different parameters like the Hounsfield threshold, radiation dose, reconstruction filters and segmentation algorithms was tested. The authors showed an increasing connectivity of pulmonary nodules to surrounding structures on thinner sections, challenging the segmentation algorithm. On the other hand, thinner sections made the algorithm more independent of the Hounsfield threshold. The algorithm tested in this study stayed within a variance of 3%.

Some of these results could be observed in our study as well, as the algorithm had a higher APE in nodules attached to the pleura and was dependent of the anatomical location. One of the most interesting and important findings was the fact that no statistically significant difference could be detected for different radiation dose settings. This indicates robust results for the volumetry algorithm either on low-dose or standard-dose settings, which may yield an important impact, especially in follow-up examinations in oncology and in screening patients. On the other hand, a statistically significant difference was found between thin and thicker collimations, indicating a preferable use of thin collimation in comparison to thicker collimation. As standard vendor-specific kernel and comparable pitch were used for all acquistion protocols, further investigations may evaluate if differences in vendor-specific acquisition in this regard has an impact on the volumetry results.

In a third study by Ko et al. [23], who used phantom nodules to evaluate volumetric measurement methods for plastic nodules, an influence of radiation dose, nodule size and nodule attenuation was reported. We found lower APEs for bigger nodules, too. This is important for follow-up studies. Measurements for nodules smaller than 4 mm in diameter can be very difficult, but the measurement of nodules larger than 4 mm was more reliable. The published volumetric measurement results for different algorithms vary significantly. For example, the algorithm tested by Wiemker et al. [16] had a variance of only up to 3%, as did the one tested in a study by Yankelevitz et al. [18]. In our study, the measurement errors were higher; however, none of the above investigators tested nodules with such a large variety of anatomical locations, with diameters/volumes starting at 3 mm/6.67 mm^3. Finally, Marten et al. [14, 15] evaluated the accuracy of volumetric measurements using MSCT and flat panel volumetric CT [VCT] for phantom nodules before and after deformation. They showed that VCT was superior to MSCT regarding the accuracy of phantom nodule volumetry. However, scanners compared in our study are commonly used worldwide, and VCT is not yet applicable in clinical practice. If the technical limitations of VCT are to be overcome, it might serve as an alternative in the future.

Limitations

As for any phantom evaluation for the accuracy of intrapulmonary nodules, the nodules used in the phantom were perfectly roundly shaped. Therefore, only solid non-

Table 5 Error in nodule volume measurements for different protocols and nodule categories (A = intraparenchymal, B = around a vessel, C = attached to a vessel, D = pleural and E = attached to the pleura). LD: low dose; SD: standard dose

Toshiba Aquilion 16	Mean APE (absolute percentage error)			
Collimation/dose	0.5 LD	1.0 LD	0.5 SD	1.0 SD
Category A	9.24±7.76	11.19±7.25	9.56±7.18	11.32±7.5
Category B	4.39±3.66	4.82±6.73	4.49±4.24	4.43±3.29
Category C	4.7±4.63	3.29±2.14	4.14±3.87	8.54±13.16
Category D	11.76±16.46	8.21±6.21	9.67±10.29	10.3±7.3
Category E	5.19±1.81	6.23±1.71	7.73±3.63	9.79±4.99
Size ≤33.86 mm^3	14.35±13.89	11.28±7.72	13.35±8.17	16.74±11.15
Size ≥66.29 mm^3	4.63±4.09	5.62±4.57	5.03±4.25	6.39±4.81

deformable nodules were studied, and no ground-glass opacities were evaluated. Secondly, no motion deformation from breathing or heart motion could be included in this phantom study. Interaction for nodule volumetry might have reduced the APE, but in terms of reproducibility, no user interaction was allowed within this study. Comparing data from different vendors, one has to take vendor-specific image acquisition and image reconstruction parameters into account that might influence volumetry results. The volumetry software that was used in this study was developed by Siemens. Thus, during development it has been tested mostly with Siemens data.

In summary, nodule volumetry as was shown in this study is accurate and feasible for different datasets from different scanner vendors and different scan protocols within a reasonable range of volumetric error for most of the nodules. Low-dose settings can be used with the same results, while thin collimation should be preferred. Further improvement of volumetry algorithms is needed for very small nodules and nodules attached to the pleura.

References

1. Wormanns D, Ludwig K, Beyer F, Heindel W, Diederich S (2005) Detection of pulmonary nodules at multirow-detector CT: effectiveness of double reading to improve sensitivity at standard-dose and low-dose chest CT. Eur Radiol 15:14–22
2. Valencia R, Denecke T, Lehmkuhl L, Fischbach F, Felix R, Knollmann F (2006) Value of axial and coronal maximum intensity projection (MIP) images in the detection of pulmonary nodules by multislice spiral CT: comparison with axial 1-mm and 5-mm slices. Eur Radiol 16:325–332
3. Takashima S, Sone S, Li F, Maruyama Y, Hasegawa M, Kadoya M (2003) Indeterminate solitary pulmonary nodules revealed at population-based CT screening of the lung: using first follow-up diagnostic CT to differentiate benign and malignant lesions. AJR Am J Roentgenol 180:1255–1263
4. Tan BB, Flaherty KR, Kazerooni EA, Iannettoni MD (2003) The solitary pulmonary nodule. Chest 123:89S–96S
5. Diederich S, Wormanns D, Semik M et al (2002) Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. Radiology 222:773–781
6. Henschke CI, McCauley DI, Yankelevitz DF et al (1999) Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 354:99–105
7. Henschke CI, Naidich DP, Yankelevitz DF et al (2001) Early lung cancer action project: initial findings on repeat screenings. Cancer 92:153–159
8. Sone S, Takashima S, Li F et al (1998) Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 351:1242–1245

9. Swensen SJ, Jett JR, Hartman TE et al (2003) Lung cancer screening with CT: Mayo Clinic experience. Radiology 226:756–761
10. Diederich S, Hansen J, Wormanns D (2005) Resolving small pulmonary nodules: CT features. Eur Radiol 15:2064–2069
11. MacMahon H, Austin JH, Gamsu G et al (2005) Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. Radiology 237:395–400
12. Bogot NR, Kazerooni EA, Kelly AM, Quint LE, Desjardins B, Nan B (2005) Interobserver and intraobserver variability in the assessment of pulmonary nodule size on CT using film and computer display methods. Acad Radiol 12:948–956
13. Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C (2006) Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. Eur Radiol 16:781–790
14. Marten K, Engelke C, Grabbe E, Rummeny EJ (2004) [Flat-panel detector-based computed tomography: accuracy of experimental growth rate assessment in pulmonary nodules]. Rofo 176:752–757
15. Marten K, Funke M, Engelke C (2004) Flat panel detector-based volumetric CT: prototype evaluation with volumetry of small artificial nodules in a pulmonary phantom. J Thorac Imaging 19:156–163
16. Wiemker R, Rogalla P, Blaffert T et al (2005) Aspects of computer-aided detection (CAD) and volumetry of pulmonary nodules using multislice CT. Br J Radiol 78 Spec No:S46–S56
17. 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

18. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI (2000) Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. Radiology 217:251–256
19. Boll DT, Gilkeson RC, Fleiter TR, Blackham KA, Duerk JL, Lewin JS (2004) Volumetric assessment of pulmonary nodules with ECG-gated MDCT. AJR Am J Roentgenol 183:1217–1223
20. Mullally W, Betke M, Wang J, Ko JP (2004) Segmentation of nodules on chest computed tomography for growth assessment. Med Phys 31:839–848
21. Bolte H, Riedel C, Jahnke T et al (2006) Reproducibility of computer-aided volumetry of artificial small pulmonary nodules in ex vivo porcine lungs. Invest Radiol 41:28–35
22. 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
23. Ko JP, Rusinek H, Jacobs EL et al (2003) Small pulmonary nodules: volume measurement at chest CT-phantom study. Radiology 228:864–870
24. Shin HO, Blietz M, Frericks B, Baus S, Savellano D, Galanski M (2006) Insertion of virtual pulmonary nodules in CT data of the chest: development of a software tool. Eur Radiol 16:2567–2574 Nov
25. Benjamin MS, Drucker EA, McLoud TC, Shepard JA (2003) Small pulmonary nodules: detection at chest CT and outcome. Radiology 226:489–493