Artikel online veröffentlicht: 21.10.2021

Accuracy and Reproducibility of a Software Prototype for Semi-Automated Computer-Aided Volumetry of the solid and subsolid Components of part-solid Pulmonary Nodules

Genauigkeit und Reproduzierbarkeit eines Software-Prototypen zur halbautomatischen computergestützten Volumetrie der soliden und subsoliden Komponenten von teilsoliden Lungenrundherden

Autoren

Sebastian Werner[1], Regina Gast[1], Rainer Grimmer[2], Andreas Wimmer[3], Marius Horger[1]

Institute

1 Radiology, University Hospitals Tübingen, Tübingen, Germany
2 Medical Imaging, Siemens Healthineers AG, Erlangen, Germany
3 Siemens Healthcare GmbH, Forchheim, Germany

Key words

thorax, CT, CT-quantitative, adenocarcinoma, technical aspects

eingereicht 15.05.2021
akzeptiert 23.09.2021
online publiziert 21.10.2021

Bibliografie

Fortschr Röntgenstr 2022; 194: 296–305
DOI 10.1055/a-1656-9834
ISSN 1438-9029
© 2021. Thieme. All rights reserved.
Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Korrespondenzadresse

Dr. Sebastian Werner
Radiologie, Universitätsklinikum Tübingen, Hoppe-Seyler-Str. 3, 72076 Tubingen, Germany
Tel.: +49/70 71/2 96 18 76
sebastian.werner@med.uni-tuebingen.de

ABSTRACT

Purpose To test the accuracy and reproducibility of a software prototype for semi-automated computer-aided volumetry (CAV) of part-solid pulmonary nodules (PSN) with separate segmentation of the solid part.

Materials and Methods 66 PSNs were retrospectively identified in 34 thin-slice unenhanced chest CTs of 19 patients. CAV was performed by two medical students. Manual volumetry (MV) was carried out by two radiology residents. The reference standard was determined by an experienced radiologist in consensus with one of the residents. Visual assessment of CAV accuracy was performed. Measurement variability between CAV/MV and the reference standard as a measure of accuracy, CAV inter- and intra-rater variability as well as CAV intrascan variability between two recontruction kernels was determined via the Bland-Altman method and intraclass correlation coefficients (ICC).

Results Subjectively assessed accuracy of CAV/MV was 77 %/79 %–80 % for the solid part and 67 %/73 %–76 % for the entire nodule. Measurement variability between CAV and the reference standard ranged from −151–117 % for the solid part and −106–54 % for the entire nodule. Interrater variability was −16–16 % for the solid part (ICC 0.998) and −102–65 % for the entire nodule (ICC 0.880). Intra-rater variability was −70–49 % for the solid part (ICC 0.992) and −111–31 % for the entire nodule (ICC 0.929). Intrascan variability between the smooth and the sharp reconstruction kernel was −45–39 % for the solid part and −21–46 % for the entire nodule.

Conclusion Although the software prototype delivered satisfactory results when segmentation is evaluated subjectively, quantitative statistical analysis revealed room for improvement especially regarding the segmentation accuracy of the solid part and the reproducibility of measurements of the nodule's subsolid margins.

Key points:

- Assessed visually CAV delivers similar accuracy compared to manual volumetry
- Accuracy of CAV was higher for the entire nodule
- Reproducibility was better for the solid part
- Variability between the kernels was higher for the solid part

ZUSAMMENFASSUNG

Ziel Untersuchung der Genauigkeit und Reproduzierbarkeit eines Software-Prototypen zur semiautomatischen computergestützten Volumetrie (CAV) von teilsoliden Lungenrundherden (PSN) mit separater Segmentierung des soliden Anteils.

Material und Methoden 66 PSN wurden retrospektiv in 34 nativen Dünnschicht-Thorax-CTs von 19 Patienten identifiziert. Die CAV wurde von 2 Medizinstudenten durchgeführt, eine manuelle Volumetrie (MV) von 2 radiologischen Assistenzärzten. Der Referenzstandard wurde von einem erfahrenen Radiologen im Konsens mit einem der Assistenzärzte festgelegt. Die Genau-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

igkeit der CAV wurde visuell beurteilt. Die Messvariabilität zwischen CAV/MV und dem Referenzstandard als Maß für die Genauigkeit, die Inter- und Intrarater-Variabilität der CAV sowie die Intrascan-Variabilität der CAV zwischen 2 Rekonstruktionskernels wurden mittels der Bland-Altman-Methode und dem Intraclass-Korrelationskoeffizienten (ICC) bestimmt.

Ergebnisse Die subjektiv bewertete Genauigkeit der CAV/MV lag bei 77 %/79 %–80 % für den soliden Anteil und bei 67 %/73 %–76 % für den gesamten Rundherd. Die Messvariabilität zwischen CAV und dem Referenzstandard reichte von –151–117 % für den soliden Anteil und –106–54 % für den gesamten Rundherd. Die Interrater-Variabilität betrug –16–16 % für den soliden Anteil (ICC 0,998) und –102–65 % für den gesamten Rundherd (ICC 0,880). Die Intrarater-Variabilität betrug –70–49 % für den soliden Anteil (ICC 0,992) und –111–31 % für den gesamten Rundherd (ICC 0,929). Die Intrarater-Variabilität zwischen dem weichen und scharfen Rekonstruktionskernel betrug –45–39 % für den soliden Anteil und –21–46 % für den gesamten Rundherd.

Schlussfolgerung Obwohl der Software-Prototyp bei der subjektiven Bewertung der Segmentierung zufriedenstellende Ergebnisse lieferte, zeigte die quantitative statistische Analyse Potenzial für Verbesserungen, insbesondere hinsichtlich der Segmentierungsgenauigkeit des soliden Teils und der Reproduzierbarkeit der Messung der subsoliden Ränder des Rundherdes.

Kernaussagen:

- Visuell beurteilt liefert die CAV eine ähnliche Genauigkeit wie die manuelle Volumetrie.
- Die Genauigkeit der CAV war für den gesamten Knoten höher.
- Die Reproduzierbarkeit war für den soliden Anteil besser.
- Die Variabilität zwischen den Rekonstruktionskernels war für den soliden Anteil höher.

Citation Format

- Werner S, Gast R, Grimmer R et al. Genauigkeit und Reproduzierbarkeit eines Software-Prototypen zur halbautomatischen computergestützten Volumetrie der soliden und subsoliden Komponenten von teilsoliden Lungenrundherden. Fortschr Röntgenstr 2022; 194: 296–305

Introduction

Part-solid pulmonary nodules (PSN) are a frequent incidental finding in chest CTs, for example in the setting of a lung cancer screening program. Persistent PSN have been reported to be malignant more frequently than solid or pure ground glass nodules (PGGN) with rates as high as 93.3 % [1]. In case of malignancy they usually represent adenocarcinoma and its precursors [2]. On the other hand, they can also represent various benign entities like infection, inflammation, focal interstitial fibrosis, eosinophilic pneumonia, thoracic endometriosis, focal hemorrhage or organizing pneumonia and have been shown to be transient in up to 69.8 % of cases [3]. This variety can make diagnosis challenging and dependent on the experience of the radiologist who has to take into account multiple lesion features such as attenuation, location inside the lung, size, shape and whether singular vs. multiple. But even if one considers all these particularities, diagnosis is often not certain requiring follow-up examinations for documentation of resolution, persistence or growth [5]. Moreover, because the size of many of these nodules is small, how size is measured becomes especially important. Variations in CT-scanners, window settings, as well as inter- and intrarater performance are common and may have a critical impact on the assessment of size, especially in the follow-up [6]. Presently, manual uni- or bidirectional diameter measurements are the standard for lung cancer screening programs and day to day clinical care as reflected in current guidelines regarding the management of pulmonary nodules [5, 7]. In PSN both the diameter of the whole lesion and of the solid part should be measured while the focus should be on the solid component, because although the solid component does not always correlate with the pathologically determined invasive component, there is a general correlation between them. [4–6, 8].

There is accumulating evidence that semi-automated computer-aided volumetry (CAV) has several advantages compared to manual diameter measurements. The Dutch-Belgian lung cancer screening trial (NELSON), which was the first screening program to use semi-automated CAV instead of manual diameter measurements, achieved high negative predictive values and presumably fewer false-positive results compared to other lung cancer screening trials [9, 10]. Furthermore, the volume-based management protocol yielded high sensitivity and specificity for the 2-year lung cancer probability [11]. Heuvelmans et al. [12] concluded in their investigation of diameter and volume measurements for estimation of lung nodule size that the use of mean or maximum axial diameter to assess the size of intermediate-sized lung nodules leads to a substantial overestimation of nodule volume, compared with semi-automated volumetry and that median intranodular diameter variation exceeds the 1.5 mm growth cut-off advocated in screening guidelines such as LungRADS implying a significant potential for errors in nodule management. It is not trivial to measure the accuracy of semi-automated volumetry since the “true” size of any pulmonary nodule is in most cases unknown. The reference standard of volume measurement after nodule excision is not perfect due to factors like an inevitable bias toward larger nodules, differences in pathology handling techniques and variations in the degree of lung inflation [13]. Nevertheless, several phantom studies have delivered promising results [14–17]. Apart from accurate estimation of the nodule volume, in clinical practice it is arguably more important for the software to possess high levels of intra- and interrater reliability since many nodules require follow-up examinations. In the past the variability range of nodule volume has been reported to be in general approximately 25 % [18].

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

The objective of this study was to test the performance of a software prototype for semi-automated computer-aided volumetry of part-solid pulmonary nodules with separate segmentation of the whole lesion and the solid component and to compare results with those acquired by manual volumetry.

Material and Methods

Study population

This retrospective evaluation of CT image data was approved by the local institutional review board (registration number 187/2018BO2). A retrospective database search of the local radiology department identified 34 chest CT scans of 19 consecutive patients (median age 75 years; range, 55–91 years; 8 female) diagnosed with part-solid pulmonary nodules (n = 66) in the routine CT-work-up between February 2015 and February 2018.

CT examinational protocol

All chest CTs were obtained unenhanced at end-inspiratory phase. In total, 34 CT-image data sets with a mean of 2 (range, 1–10) follow-up examinations were evaluated. CT-examinations were performed using a multi-detector scanner (SOMATOM Definition Flash, Siemens Healthineers, Forchheim, Germany), a 300–350 mm field of view, a 512 × 512 reconstruction matrix, 120 kV, 100 effective mAs and a tube rotation time of 0.5 ms. In all patients a spiral acquisition was obtained from the apex to the base of the lungs. Patients were positioned supine with the arms stretched in elevation and abduction. Thin-slice CT scans (0.6 mm) were reconstructed using a smooth reconstruction kernel (filter, B31f). For 54 lesions both smooth and sharp kernel reconstructions (filter, B70f) were available. All chest CTs were analysed for the existence of additional pathologies e. g. pleural effusions, pulmonary oedema, haemorrhage or pneumonia, that could have impacted volumetry results in which case they were excluded from the final analysis.

Functioning of the software prototype

Implementation of manual segmentation

The complete chest CT image dataset is displayed in three planes. The reader identifies the PSN and uses a designated tool to manually draw the edges of the whole lesion and of the solid part on every image the nodule is visible on. The edges can be drawn and freely adjusted in all three planes. After finalizing the manual segmentation, the software automatically calculates the volume and longest axial diameter of both the entire nodule and the solid part without displaying results to the reader.

Implementation of semi-automated computer-aided segmentation

The complete chest CT image dataset is displayed in three planes. The reader identifies the PSN and subjectively selects the axial slice in which the lesion shows the longest diameter. A designated tool is used to draw a straight line *(seed line)* through the longest diameter. The software then immediately performs automatic segmentation separately for the entire nodule and the solid part and calculates the volumes and longest axial diameters. The reader is blinded to the segmentation results.

Technical description of semi-automated computer-aided segmentation

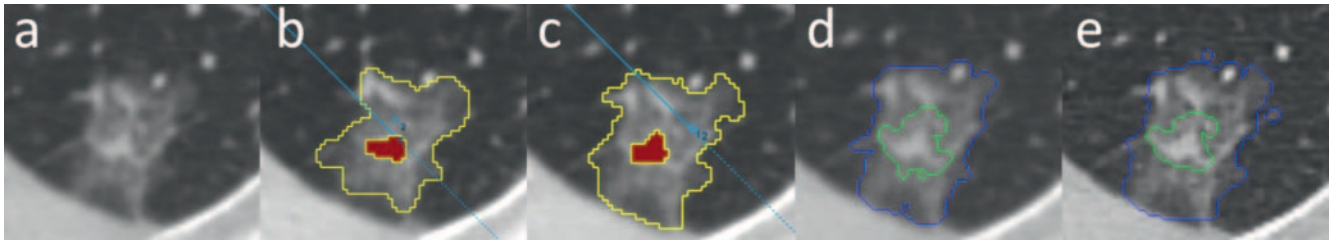
After initializing CAV via drawing, the seed line the algorithm then computes a histogram of the attenuation of the voxels marked by the seed line to differentiate between solid lesions (i. e. parenchymal consolidation obscuring pulmonary structures like vessels and bronchi) and part-solid lesions. If the 25 % quantile of the histogram exceeds a predefined attenuation threshold, a pure solid lesion is assumed. In this case the lesion is segmented through region growing followed by morphological operations. The algorithm determines whether the lesion shows direct pleural contact, in which case the nodule is separated from the pleura. A detailed description of the algorithm can be found in a study by Moltz et al. [19]. If the histogram analysis does not indicate a pure solid lesion (i. e. detects the presence of ground-glass opacification which shows higher attenuation than normal lung parenchyma and lower attenuation than the solid portion and pulmonary soft tissues such as vessels or bronchi), the entire part-solid lesion is segmented through region growing with boundaries determined via intensity analysis of the nodule region and surrounding parenchyma [20]. This is followed by morphological operations analogous to the ones performed for solid lesions. In part-solid lesions the denser structures belonging to the solid compartment of the lesion are identified via thresholding: the center of the largest solid structure is used as a seed point to segment a solid compartment with the same algorithm as for pure solid lesions described above. The solid compartment is restricted to the boundaries of the subsolid compartment. The algorithm accounts for partial volume effects when determining the volumes of the solid and the subsolid compartment [21]. The reported subsolid volume includes the volume of the solid compartment. Examples of a segmentation results are given in ▶ **Fig. 1, 2**.

Manual and computer-assisted segmentation and volumetry

For each of the 66 part-solid nodules 4 sets of volume measurements (MV1, MV2, CAV1, CAV2) were produced by two radiology residents and two medical students, each set containing separate measurements of the entire PSN and the solid part: Manual volumetry performed by Radiologist 1 (MV1), manual volumetry performed by Radiologist 2 (MV2), CAV performed by medical student 1 (CAV1) and CAV performed by medical student 2 (CAV2). Radiologist 1 did not have any significant experience in reading chest CTs, Radiologist 2 had three years of experience.

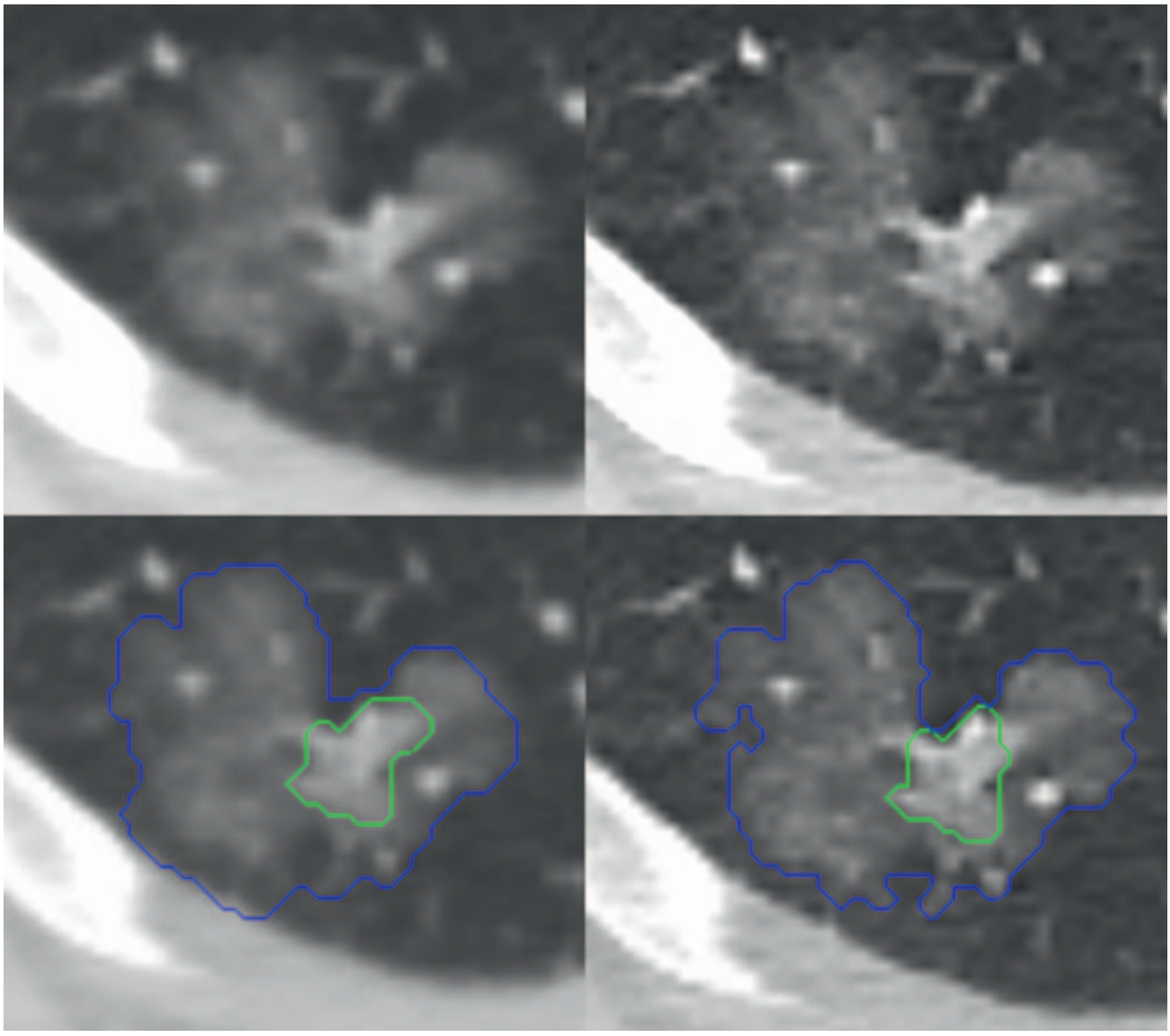
In a subset of 54 part-solid nodules CT datasets had been reconstructed with both the smooth and the sharp kernel. In this subset two additional sets of CAV measurements (CAVsmooth, CAVsharp) were produced by medical student 1, each set again containing separate measurements of the entire PSN and the solid part: CAV was performed three seperate times in each PSN using different variants of seed lines according to the following instructions: Seed line 1: “Draw a seed line through the longest diame-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.



▶ **Fig. 1** Sample images of a 78-year-old male patient with an adenocarcinoma in the right upper lobe. Images demonstrate results of manual segmentation using the smooth reconstruction kernel (**b**, **c**) and semi-automated computer-aided segmentation using the smooth **d** and sharp kernel **e**.

▶ **Abb. 1** Beispielbilder eines 78 Jahre alten männlichen Patienten mit einem Adenokarzinom im rechten Oberlappen. Die Bilder zeigen die Ergebnisse der manuellen Segmentierung mit dem weichen Rekonstruktionskernel (**b**, **c**) und der halbautomatischen computergestützten Segmentierung mit dem weichen **d** und harten Kernel **e**.



▶ **Fig. 2** Sample images of a 76-year-old female patient with an adenocarcinoma in the right upper lobe. Images show the results of semi-automated computer-assisted segmentation using the smooth (left) and sharp reconstruction kernel (right).

▶ **Abb. 2** Beispielbilder einer 76 Jahre alten Patientin mit einem Adenokarzinom im rechten Oberlappen. Die Bilder zeigen die Ergebnisse der halbautomatischen computergestützten Segmentierung mit dem weichen (links) und dem harten Rekonstruktionskernel (rechts).

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

▶ **Table 1** Mean volumes [mm³] and longest axial diameters [mm] with standard deviations of the entire PSN and the solid lesion part acquired by manual volumetry and CAV (± standard deviation).

▶ **Tab. 1** Durchschnittliches Volumen [mm³] und größter axialer Durchmesser [mm] mit Standardabweichungen der gesamten PSN und der soliden Anteile, ermittelt mittels manueller Volumetrie und CAV (± Standardabweichung).

	manual volumetry (reader 1)	manual volumetry (reader 2)	CAV (student 1)
volume entire PSN	1401 (± 2929)	1607 (± 3420)	1213.0 (± 2706)
volume solid part	272 (± 500)	245 (± 376)	266 (± 440)
diameter PSN	15.0 (± 7.2)	15.8 (± 8.4)	12.3 (± 8.3)
diameter solid part	9.1 (± 3.7)	9.2 (± 3.6)	8.7 (± 4.2)

ter."; Seed line 2: "Draw a seed line through the longest diameter but be a little imprecise."; Seed line 3: "Draw a seed line through the approximate longest diameter and extend the seed line into the surrounding lung parenchyma." The reader was blinded to the segmentation results. The average of the three volume measurements was calculated (CAVsmooth). The manually-drawn seed lines were then transferred to the CT data sets that had been reconstructed with the sharp kernel to obtain CAV results for both kernels using the exact same seed lines and the average of the three volume measurements was calculated (CAVsharp).

Analysis

Subjective visual assessment

Four weeks after the production the blinded segmentation results MV1, MV2 and CAV1 were shown to a senior radiologist with 25 years of experience in reading chest CTs (Radiologist 3) and to Radiologist 2. They visually assessed the quality of the results in the following manner: A dedicated software program was used. The readers selected each of the segmented 66 PSNs from a list. The selected PSN is shown in the axial CT images with the segmentation results displayed as colored lines surrounding the edges of the entire nodule and the solid part, each color representing one of the three datasets. The readers were able to select which of the separate segmentation results were displayed at any time with the option to display any combination or no result at all. This ensured that the lesion itself could be examined well and that segmentation results could be compared directly. The readers visually evaluated the segmentation results, i. e. how exact the lines depicted the borders of the solid part and the entire nodule. Each single segmentation result was evaluated as either satisfactory or unsatisfactory via consensus reading.

Quantitative statistical analysis

The following parameters were evaluated:

1. CAV Accuracy

CAV accuracy was assessed via comparing semi-automated CAV (CAV1 and CAV2) to the calculated average of the two radiology residents' manual volume measurements (MV1 and MV2), which was defined as the reference standard, using the Bland-Altman method [22, 23].

2. CAV and manual volumetry interrater variability

The interrater variability of CAV and manual volumetry was assessed by comparing the results of semi-automated CAV performed by the two medical students (CAV1 and CAV2) and the results of manual volumetry performed by Radiologist 1 and 2 (MV1 and MV2) using the Bland-Altman method and calculating the intraclass correlation coefficient (ICC).

3. CAV intra-rater variability

The intra-rater variability of CAV was assessed by determining each minimum and maximum measurement out of the three separate measurements per PSN performed by medical student 1 in the CT datasets that had been reconstructed with the smooth kernel (CAVsmooth). These were then compared via the Bland-Altman method. Additionally, we calculated the ICC for the three separate measurements.

4. Variability between the smooth and sharp reconstruction kernel

Variability of CAV measurements between the smooth and the sharp reconstruction kernel was assessed via comparing the calculated average values of CAVsmooth with those of CAVsharp using the Bland-Altman method.

Bland-Altman analysis consists of calculation of the relative differences in volume measurements, i. e. the difference in two measurements divided by the mean volume. Volume measurement variability is defined as the 95 % confidence interval of these relative differences. ICC estimates and their 95 % confidence intervals were calculated based on a single rater, absolute-agreement, two-way random-effects model. A p-value of 0.05 was considered statistically significant. We used the computer software IBM SPSS Statistics 26 and GraphPad Prism 9.

Results

Mean values and standard deviations for volumes and diameters of the entire lesion and the solid part are presented in ▶ **Table 1**.

Subjective visual assessment

Manual segmentation of the solid part was rated as satisfactory in 79 %–80 %. Manual segmentation of the entire nodule was rated as satisfactory in 73 %–76 %. Semi-automated computer-assisted

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

▶ **Table 2** Results of subjective visual assessment. Percentage of segmentation results rated as satisfactory.

▶ **Tab. 2** Ergebnisse der visuellen Bewertung. Prozentsatz der Segmentierungsergebnisse, die als zufriedenstellend gewertet wurden.

solid part	Manual volumetry (Radiologist 1)	79 % (52/66)
	Manual volumetry (Radiologist 2)	80 % (53/66)
	CAV (Medical Student 1)	77 % (51/66)
entire PSN	Manual volumetry (Radiologist 1)	73 % (48/66)
	Manual volumetry (Radiologist 2)	76 % (50/66)
	CAV (Medical Student 1)	67 % (44/66)

segmentation delivered satisfactory results in 77 % for the solid part and 67 % for the entire nodule (▶ **Table 2**).

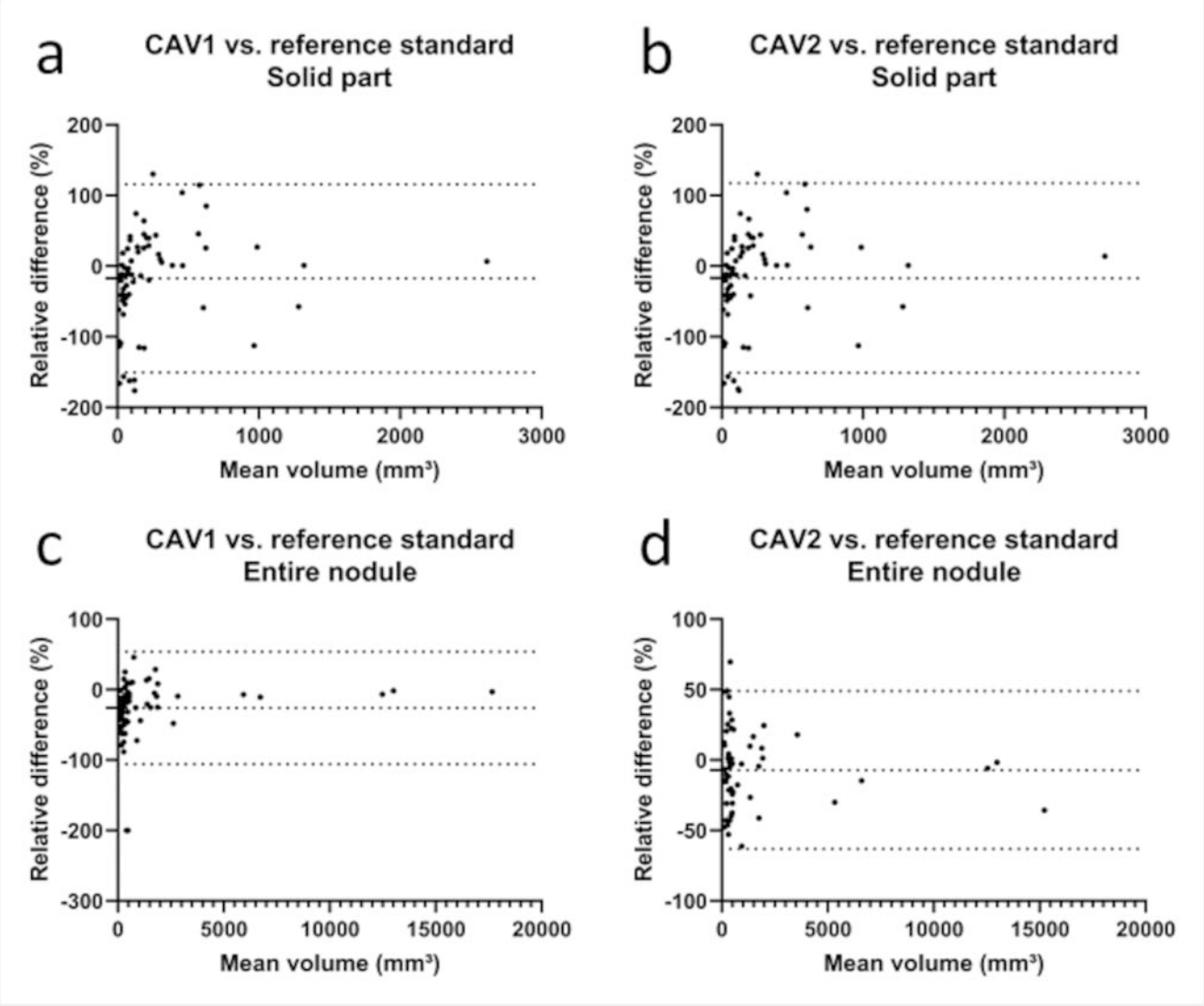
Statistical analysis of volumetry

Numbers in brackets following ICC values indicate the lower and upper bounds of their 95 % confidence intervals.

1. CAV Accuracy

For the solid part relative variability between CAV1 / CAV2 and the reference standard was −150–116 %/−151–117 % with a mean relative difference of −17 %/−17 %. For the entire nodule relative variability was −106–54 %/−63–49 % with a mean relative difference of −26 %/−7 %. The respective Bland-Altman plots are shown in ▶ **Fig. 3**.

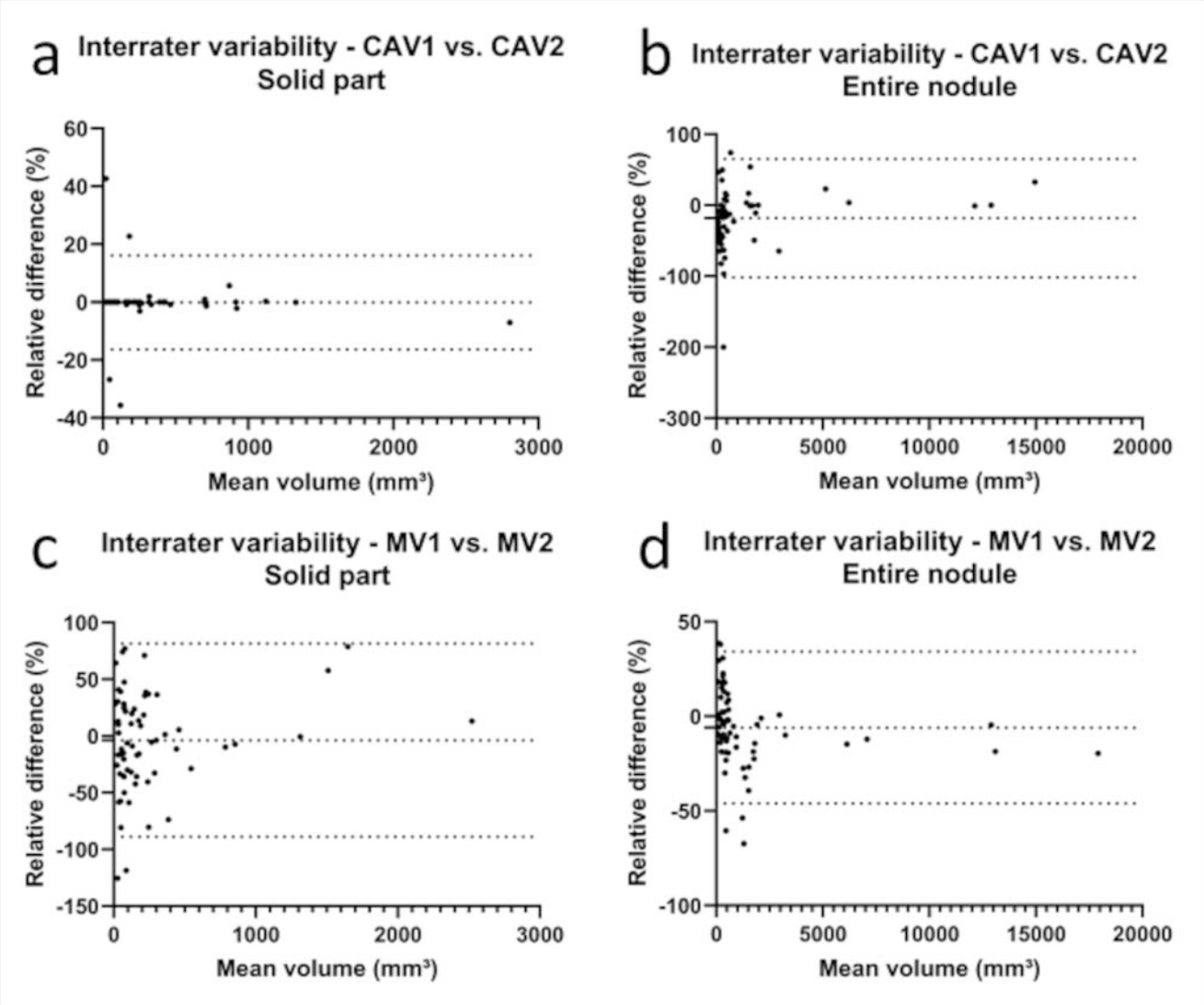
2. CAV interrater variability



▶ **Fig. 3** CAV Accuracy. Bland-Altman plots depicting variability between CAV1 and CAV2 and the reference standard. The mean differences (middle dotted line) and the upper and lower 95 % limits of agreement (upper and lower dotted lines) were as follows (limits of agreement in parenthesis): **a**: −17 (−150–116), **b**: −17 (−151–117), **c**: −26 (−106–54), **d**: −7 (−63–49).

▶ **Abb. 3** CAV Accuracy. Bland-Altman-Plots zur Darstellung der Variabilität zwischen CAV1 und CAV2 und dem Referenzstandard. Die Mittelwerte der Differenz (mittlere gestrichelte Linie) und die oberen und unteren 95 % Limits of Agreement (obere und untere gestrichelte Linien) waren wie folgt (Limits of Agreement in Klammern): **a**: −17 (−150–116), **b**: −17 (−151–117), **c**: −26 (−106–54), **d**: −7 (−63–49).

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.



▶ **Fig. 4** Interrater variability. Bland-Altman plots depicting interrater variability between CAV1 and CAV2 and between MV1 and MV2. The mean difference (middle dotted line) and the upper and lower 95 % limits of agreement (upper and lower dotted lines) were as follows (limits of agreement in parenthesis): **a**: –0.075 (–16–16), **b**: –18 (–102–65), **c**: –3.6 (–89–82), **d**: –5.9 (–46–34).

▶ **Abb. 4** Interrater-Variabilität. Bland-Altman-Plots zur Darstellung der Interrater-Variabilität zwischen CAV1 und CAV2 und zwischen MV1 und MV2. Die Mittelwerte der Differenz (mittlere gestrichelte Linie) und die oberen und unteren 95 % Limits of Agreement (obere und untere gestrichelte Linien) waren wie folgt (Limits of Agreement in Klammern): **a**: –0.075 (–16–16), **b**: –18 (–102–65), **c**: –3.6 (–89–82), **d**: –5.9 (–46–34).

For the solid part relative variability between CAV1 and CAV2 was –16–16 % with a mean relative difference of –0.075 %. For the entire nodule relative variability was –102–65 % with a mean relative difference of –18 %. The respective Bland-Altman plots are shown in ▶ **Fig. 4**. Regarding the solid part the ICC was 0.998 (0.997, 0.999). For the entire lesion the ICC was 0.880 (0.806, 0.926).

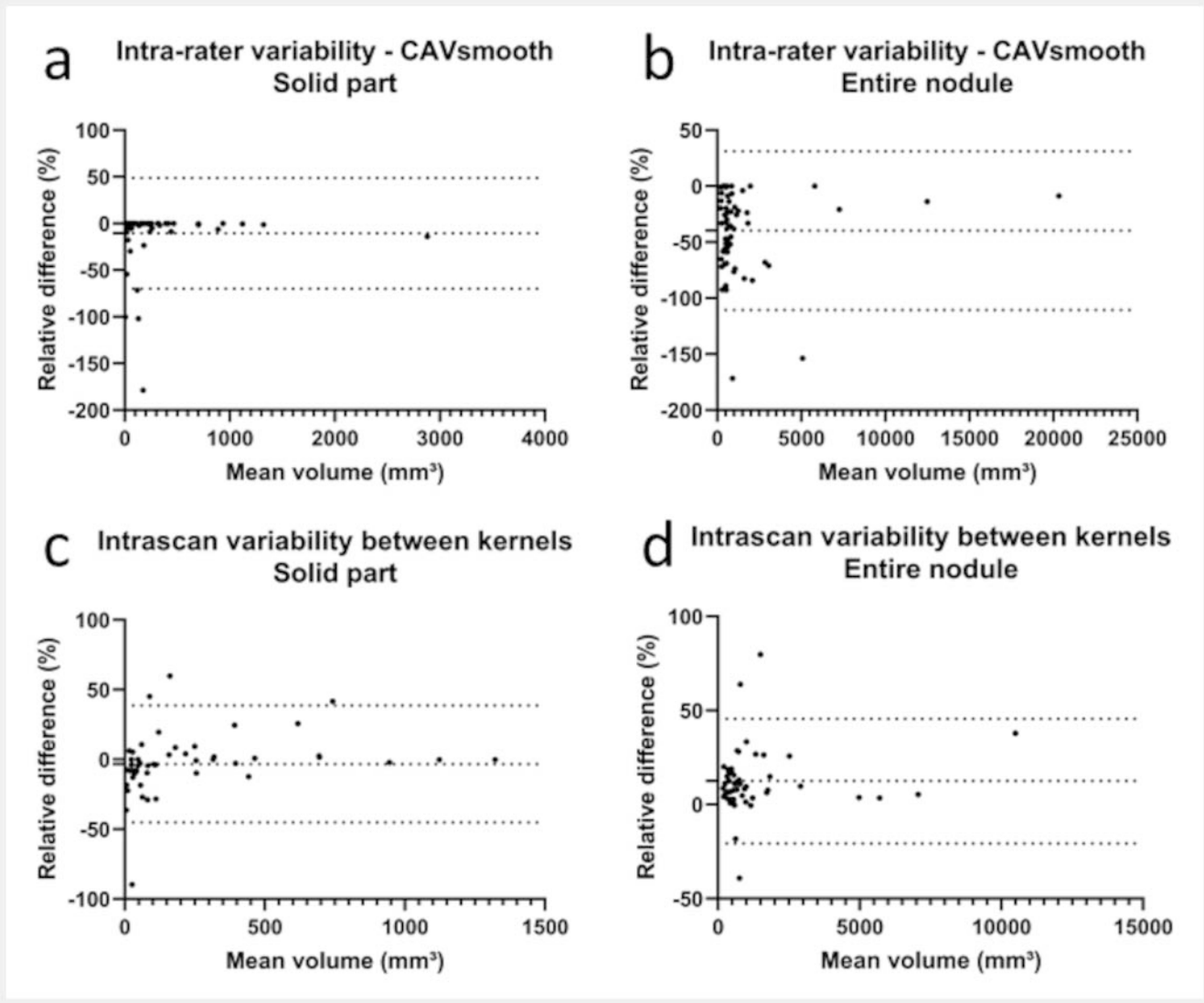
3. CAV intra-rater variability

For the solid part relative intra-rater variability was –70–49 % with a mean relative difference of –10 %. For the entire nodule variability was –111–31 % with a mean relative difference of –40 %. The respective Bland-Altman plots are shown in ▶ **Fig. 5a, b**. The ICC of the three separate measurements per PSN performed by medical student 1 was 0.992 (0.988, 0.995) for the solid part and 0.929 (0.883, 0.958) for the entire nodule.

4. Variability between the smooth and sharp reconstruction kernel

For the solid part relative variability of CAV measurements between the smooth and the sharp reconstruction kernel was –45–39 % with a mean relative difference of –3.2 %. For the entire nodule variability was –21–46 % with a mean relative difference of 13 %. The respective Bland-Altman plots are shown in ▶ **Fig. 5c, d**.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.



▶ **Fig. 5** CAV intra-rater variability; intrascan variability between reconstruction kernels. Bland-Altman plots depicting intra-rater variability for CAVsmooth and the intrascan variability between the smooth and sharp reconstruction kernels. The mean difference (middle dotted line) and the upper and lower 95 % limits of agreement (upper and lower dotted lines) were as follows (limits of agreement in parenthesis): **a**: –10 (–70–49), **b**: –40 (–111–31), **c**: –3.2 (–45–39), **d**: 13 (–21–46).

▶ **Abb. 5** Intrarater-Variabilität. Bland-Altman-Plots zur Darstellung der Intrarater-Variabilität für CAVsmooth und der Intrascan-Variabilität zwischen dem weichen und harten Rekonstruktionskernel. Die Mittelwerte der Differenz (mittlere gestrichelte Linie) und die oberen und unteren 95 % Limits of Agreement (obere und untere gestrichelte Linien) waren wie folgt (Limits of Agreement in Klammern): **a**: –10 (–70–49), **b**: –40 (–111–31), **c**: –3.2 (–45–39), **d**: 13 (–21–46)

Discussion

Overall the software prototype showed mixed results. Subjective assessment of CAV yielded satisfactory results with a somewhat higher rate of satisfactory segmentation results for the solid part. On the other hand, Bland-Altman analysis showed comparatively lower accuracy and interestingly better results for the entire nodule compared to the solid part. Since both, the subjective assessment of results as well as the establishment of the reference standard were based on subjective visual delineation of the solid and subsolid part's edges, this could be a result of relatively high intra- and interrater variability regarding this task. The reduced difference in attenuation between the ground-glass component of a subsolid nodule and the surrounding lung parenchyma is a known segmentation problem [13]. The subjective impression of the authors is that when performing manual segmentation, the edges of the solid lesion parts often can be more easily and confidently identified than those of the ground glass part because they are more sharply delineated. The volumes measured by CAV were lower compared to the manually derived reference standard. In clinical practice, rather than measuring the true size of a PSN – which is not known – it is more important to detect size changes during follow-up, which requires high intra- and interrater reliability. Bland-Altman analysis showed low interrater variability for

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

the solid part but relatively high variability for the entire nodule. Expressed as ICCs the agreement was high for both. Interestingly, the interrater variability of manual segmentation was lower for the entire nodule compared to the solid part. Intra-rater variability of CAV was relatively high overall with lower values for the solid part compared to the entire nodule. Expressed as ICCs the agreement was high. Regarding differences between the two reconstruction kernels we found that with the smooth kernel the volume of the solid part was measured slightly lower and the volume of the entire nodule somewhat higher. Overall, variability between the kernels was higher for the solid part compared to the entire nodule.

These findings are important because accurate and especially precise size measurement of PSNs, a task that can be difficult to accomplish adequately when performed manually, is vital for the estimation of their malignant potential in the initial assessment and in a follow-up scenario. Additionally, valid quantification is particularly important for the solid part of malignant nodules due to its known general correlation to the invasive component [6].

There are not many publications examining semi-automated volumetry of part-solid nodules. Most publications examine subsolid nodules in general, of which part-solid nodules are a subset. Moreover, the studies including part-solid nodules did not for the most part perform separate segmentation for the solid part.

In regard to the subjective evaluation of the segmentations' quality, Benzakoun et al. [24] examined 47 PGGNs and 50 PSNs and found satisfactory results in 81 %. Charbonnier et al. [25] found satisfactory results in 80.6 % for the solid parts of 170 subsolid nodules. These values are slightly better, but similar to ours. Intra-rater variability for the entire nodule in other studies was lower than our own. Kim et al. [26] analyzed 72 PGGNs and 22 PSNs and found a variability of –7.6 % to 8.5 %. Park et al. [27] examined 30 PGGNs and found a maximum variability of –9.1 % to 10.1 % with a sharp reconstruction kernel and of –11.6 % to 11.8 % with a medium sharp reconstruction kernel. Higher variability in our study might be a result of the deliberate manipulation of the seed lines in repeated measurements and using a smooth kernel. Expressed as an ICC Scholten et al. [29] found an agreement of 0.92 which almost equals our own results. Regarding interrater variability for the entire lesion Kim et al. [26] found a variability of –11.7 % to 18.1 % and Park et al. [27] of –15.8 % to 13.4 % with a sharp reconstruction kernel and –11.1 % to 6.2 % with a the medium sharp kernel. Those values also are lower than our own. However, expressed as ICCs we found comparatively lower values as for example Scholten et al. [28, 29] or Kamiya et al. [30]. In those two studies which included 24 PGGNs and 20 PSNs and 19 PGGNs and 14 PSNs respectively, Scholten et al. found ICCs between 0.920 and 0.957. Kamiya et al. found an ICC of 0.940 in an analysis of 4 PGGNs and 92 PSNs. Expressed as relative volume deviation other authors found values between –1.2 % and 18.1 % [26, 27]. With respect to volume measurements of the solid nodule part, Kamiya et al. in the study cited above, found ICCs between 0.994–0.996, which are similar to our own results.

Regarding differences in volume between manual and semi-automated measurements the study by Scholten et al. demonstrated that the average volume was 24.3 %–26.5 % smaller when measured manually [29]. This stands in contrast to our results which showed the reverse.

Our study is limited by its retrospective design with typical drawbacks such as the fact that sharp kernel reconstructions were not available for all nodules. The number of nodules is rather low in absolute terms but similar to other studies on this topic. We did not have a histological gold standard to determine the accuracy of volume and diameter measurements, but this is a common problem concerning publications on this issue.

In conclusion, although the software prototype delivers satisfactory results when segmentation is evaluated subjectively, quantitative statistical analysis revealed room for improvement especially regarding the segmentation accuracy of the solid part and the reproducibility of measurements of the nodule's subsolid margins.

CLINICAL RELEVANCE

- Accurate and reliable size measurement plays an important role in the management of PSNs, which possess relatively high malignant potential
- The workload regarding PSN management is going to increase with the implementation of lung cancer screening programs
- CAV has the potential to make nodule size quantification easier and faster if the software's accuracy and especially the reproducibility can reach the level of manual size measurement or even surpass it

Interessenkonflikt

Rainer Grimmer is an employee of Medical Imaging, Siemens Healthineers AG, Erlangen, Germany.
Andreas Wimmer is an employee of Siemens Healthcare GmbH, Forchheim, Germany
Die Autorinnen/Autoren geben an, dass kein Interessenkonflikt besteht.

References

[1] Nakata M, Saeki H, Takata I et al. Focal ground-glass opacity detected by low-dose helical CT. Chest 2002; 121: 1464–1467. doi:10.1378/chest.121.5.1464

[2] Seidelman JL, Myers JL, Quint LE. Incidental, subsolid pulmonary nodules at CT: etiology and management. Cancer Imaging 2013; 13: 365–373

[3] Lee SM, Park CM, Goo JM et al. Transient part-solid nodules detected at screening thin-section CT for lung cancer: comparison with persistent part-solid nodules. Radiology 2010; 255: 242–251

[4] Kim H, Park CM, Koh JM et al. Pulmonary subsolid nodules: what radiologists need to know about the imaging features and management strategy. Diagn Interv Radiol 2014; 20: 47–57

[5] MacMahon H, Naidich DP, Goo JM et al. Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. Radiology 2017; 284: 228–243

[6] Travis WD, Brambilla E, Noguchi M et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011; 6: 244–285

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

[7] American College of Radiology Committee on Lung-RADS®. Lung-RADS Assessment Categories version 1.1.. Available at (Accessed on September 2nd, 2021) https://www.acr.org/-/media/ACR/Files/RADS/Lung-RADS/LungRADSAssessmentCategoriesv1-1.pdf

[8] Travis WD, Asamura H, Bankier AA et al. The IASLC Lung Cancer Staging Project: Proposals for Coding T Categories for Subsolid Nodules and Assessment of Tumor Size in Part-Solid Tumors in the Forthcoming Eighth Edition of the TNM Classification of Lung Cancer. J Thorac Oncol 2016; 11: 1204–1223

[9] van Klaveren RJ, Oudkerk M, Prokop M et al. Management of lung nodules detected by volume CT scanning. N Engl J Med 2009; 361: 2221–2229

[10] Lee H, Um SW. Probability of lung cancer based on the size threshold and volume-doubling time for lung nodules detected in low-dose CT screening. Ann Transl Med 2015; 3: 21

[11] Horeweg N, van Rosmalen J, Heuvelmans MA et al. Lung cancer probability in patients with CT-detected pulmonary nodules: a prespecified analysis of data from the NELSON trial of low-dose CT screening. Lancet Oncol 2014; 15: 1332–1341

[12] Heuvelmans MA, Walter JE, Vliegenthart R et al. Disagreement of diameter and volume measurements for pulmonary nodule size estimation in CT lung cancer screening. Thorax 2018; 73: 779–781

[13] Devaraj A, van Ginneken B, Nair A et al. Use of Volumetry for Lung Nodule Management: Theory and Practice. Radiology 2017; 284: 630–644

[14] Siegelman JW, Supanich MP, Gavrielides MA. Pulmonary nodules with ground-glass opacity can be reliably measured with low-dose techniques regardless of iterative reconstruction: results of a phantom study. Am J Roentgenol 2015; 204: 1242–1247

[15] Kim H, Park CM, Lee SM et al. A comparison of two commercial volumetry software programs in the analysis of pulmonary ground-glass nodules: segmentation capability and measurement accuracy. Korean J Radiol 2013; 14: 683–691

[16] Doo KW, Kang EY, Yong HS et al. Accuracy of lung nodule volumetry in low-dose CT with iterative reconstruction: an anthropomorphic thoracic phantom study. Br J Radiol 2014; 87: 20130644

[17] Xie X, Zhao Y, Snijder RA et al. Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. Eur Radiol 2013; 23: 139–147

[18] Kim H, Park CM. Current perspectives for the size measurement of screening-detected lung nodules. J Thorac Dis 2018; 10: 1242–1244

[19] Moltz JH, Bornemann L, Kuhnigk J-M et al. Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans. IEEE Journal of selected topics in signal processing 2009; 3: 122–134

[20] Lassen B, Jacobs C, Kuhnigk J et al. Robust semi-automatic segmentation of pulmonary subsolid nodules in chest computed tomography scans. Physics in Medicine & Biology 2015; 60: 1307

[21] Heckel F, Meine H, Moltz JH et al. Segmentation-based partial volume correction for volume estimation of solid lesions in CT. IEEE Trans Med Imaging 2014; 33: 462–480

[22] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307–310

[23] Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999; 8: 135–160

[24] Benzakoun J, Bommart S, Coste J et al. Computer-aided diagnosis (CAD) of subsolid nodules: Evaluation of a commercial CAD system. Eur J Radiol 2016; 85: 1728–1734

[25] Charbonnier JP, Chung K, Scholten ET et al. Automatic segmentation of the solid core and enclosed vessels in subsolid pulmonary nodules. Sci Rep 2018; 8: 646

[26] Kim H, Park CM, Woo S et al. Pure and part-solid pulmonary ground-glass nodules: measurement variability of volume and mass in nodules with a solid portion less than or equal to 5 mm. Radiology 2013; 269: 585–593

[27] Park CM, Goo JM, Lee HJ et al. Persistent pure ground-glass nodules in the lung: interscan variability of semiautomated volume and attenuation measurements. Am J Roentgenol 2010; 195: W408–W414

[28] Scholten ET, de Jong PA, Jacobs C et al. Interscan variation of semi-automated volumetry of subsolid pulmonary nodules. Eur Radiol 2015; 25: 1040–1047

[29] Scholten ET, de Hoop B, Jacobs C et al. Semi-automatic quantification of subsolid pulmonary nodules: comparison with manual measurements. PLoS One 2013; 8: e80249

[30] Kamiya S, Iwano S, Umakoshi H et al. Computer-aided Volumetry of Part-Solid Lung Cancers by Using CT: Solid Component Size Predicts Prognosis. Radiology 2018; 287: 1030–1040

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.