学位論文

Doctoral Thesis

Radiological Analysis for the Detection and Qualitative Diagnosis of Small Pulmonary Nodules

（小型肺結節の検出および質的診断における放射線学的研究）

尾 田 済 太 郎

Seitaro Oda

熊本大学大学院医学教育部博士課程病態制御学専攻放射線診断学

指導教員

山下康行　教授

熊本大学大学院医学教育部博士課程医学専攻放射線診断学

２０１０年３月

学　位　論　文

Doctoral Thesis

論文題名　：　Radiological Analysis for the Detection and Qualitative Diagnosis of Small Pulmonary Nodules

（小型肺結節の検出および質的診断における放射線学的研究）

著 者 名　：　尾 田 済 太 郎

Seitaro Oda

指導教員名　：　熊本大学大学院医学教育部博士課程病態制御学専攻放射線診断学

山下康行　教授

審査委員名　：　放射線治療医学　教授　大屋夏生

呼吸器病態学　教授　興梠博次

機能病理学　教授　伊藤隆明

医用放射線技術学　教授　冨口静二

２０１０年３月

Contents

Abstract

Background and Purpose:

1) Chest radiography remains the most widely used imaging technique for the detection of chest diseases because of its low cost, simplicity, and low radiation dose. However, the false-negative rate for the detection of pulmonary nodules is relatively high. The dual-energy subtraction (DES) technique is now available in full-field digital flat-panel detector (FPD) radiography systems. We investigated the clinical efficacy of DES technique using FPD chest radiography systems in the detection of small pulmonary nodules. We also investigated the effect of virtual DES image using the massive-training artificial neural network (MTANN) which is a kind of artificial intelligence.

2) The introduction of CT lung cancer screening has increased the number of detected ground glass opacities (GGO) nodules. Monitoring the temporal change of the nodule by high-resolution CT can help in obtaining a differential diagnosis. However, conventional 2-dimensional CT measurements are not reliable. More reliable estimation of pulmonary nodule volume can now be obtained using 3-dimensional volumetry by automated computer software. We investigated the accuracy and reproducibility of computer-aided volumetry (CAV) software for GGO nodules.

Moreover, we evaluated the volume-doubling time (VDT) of histologically proved GGO nodules.

Materials and Methods:

1) Using an FPD radiography system we obtained 108 sets of chest radiographs of a chest phantom with- and without simulated nodules around 10mm in diameter. Each data set contained a standard- and a corresponding DES images. The attenuation of the simulated nodules was -450, -200 and 30 Hounsfield units (HU). Receiver-operating-characteristic (ROC) analysis was used for evaluating observer performance. We also evaluated 41 sets of chest radiographs from 26 patients with small pulmonary nodules and 15 normal subjects. Similarly, virtual DES image using MTANN was examined.

2) First we investigated the accuracy of CAV software by scanning a chest phantom that included simulated 3-, 5-, 8-, 10-, and 12-mm diameter GGO nodules with CT numbers of -800, -630, and -450 HU. Next, 2 radiologists performed 2 independent measurements of 59 human nodules. Intra- and inter-observer agreement was evaluated using Bland-Altman methods. Third, we retrospectively evaluated 47 GGO nodules (mixed GGO, n=28; pure GGO, n=19). They were histologically confirmed as atypical adenomatous hyperplasia (AAH, n=13), bronchioloalveolar carcinoma (BAC, n=22),

and adenocarcinoma (AC, n=12). Using CAV software, we calculated the VDT for each nodule.

Results:

1) For -450 HU nodules the mean area under the ROC curve (AUC) without and with DES images was 0.66 and 0.77, respectively; the difference was significant. For nodules with -200- and 30 HU, there was no significant difference in the AUC value. In clinical study, the mean AUC without and with DES images was 0.62 ± 0.05 and 0.68 ± 0.05, respectively; the difference was statistically significant. For part-solid nodules, the difference of the mean AUC value was statistically significant, for non-solid nodules it was not, and for solid nodules it was not. The virtual DES image using MTANN significantly improved the mean AUC value.

2) Measurement error for simulated GGO nodules measuring 3- and ≥5 mm in diameter ranged from 51.1 - 85.2% and from -4.1 - 7.1%, respectively. For intra-observer agreement, the 95% limits of agreement were -14.9 - 13.7% and -16.6 to 15.7% for observer A and B, respectively; for inter-observer agreement these values were -16.3 to 23.7% in nodules ≥ 8 mm in diameter.

The mean VDT of all GGO nodules was 486.4 ± 368.6 days. The mean VDT for AAH, BAC, and AC was 859.2 ± 428.9, 421.2 ± 228.4, and 202.1 ± 84.3 days, respectively; there were statistically significant differences for all comparative combinations of AAH, BAC, and AC.

Conclusions:

1) Use of DES technique at FPD chest radiography significantly improved the diagnostic performance of radiologists detecting small pulmonary nodules. The virtual DES image using MTANN also improved the diagnostic performance of radiologists

2) With CAV for GGO nodules, measurement error was small for nodules with a diameter of 5 mm or more. Intra- and inter-observer agreement was relatively high in nodules ≥ 8 mm in diameter.

The evaluation of VDT using 3-dimensional volumetry is a promising method for the differentiation of GGO nodules.

和文抄録

[目的]

① 胸部単純写真は簡便で低コスト、低被曝の点から胸部疾患の評価に最も広く用いられている。しかし、病変の検出における偽陰性率は高く、特に小型肺結節で 19-72%と報告されている。そこで、近年、flat-panel detector (FPD) X 線装置に応用された dual-energy subtraction (DES)技術を使用して、小型肺結節の検出能について検討した。また、人工知能の一種である massive training artificial neural network (MTANN)を応用した virtual DES 技術の有用性についても検討した。

② 肺癌 CT 検診の導入により、ground glass opacity (GGO)結節の検出が増えている。GGO 結節の診断では経過中の増大速度の評価が重要であるが、従来の 2 次元的測定では正確とは言い難い。近年、肺結節の growth rate の評価にコンピュータ支援3次元的体積測定が注目されている。そこで、GGO 結節におけるコンピュータ支援3次元的体積測定の精度を検証し、続いて GGO 結節の volume doubling time (VDT) について検討を行った。

[方法]

① FPD X 線装置を使用。濃度の異なる模擬 GGO 結節（−450, −200, 30 HU）を用いて標準画像と DES 画像を作成し、DES 画像の併用の有無における模擬 GGO 結節の検出能を receiver operating characteristic (ROC)解析にて検討した。続いて、小型肺結節を有する臨床症例における DES 画像での検出能を ROC 解析にて検討した。また、結節亜型（充実結節、mixed GGO 結節、pure GGO 結節）別の検出能についても検討した。続いて、MTANN を用いた virtual DES 画像における小型肺結節の検出能を臨床症例を用いて ROC 解析にて検討した。

② サイズ（直径 3−, 5−, 8−, 10−, 12mm）と濃度（−800, −630, −450 HU）の異なる模擬 GGO 結節を CT 撮像し、それぞれの GGO 結節の体積をコンピュータ支援ソフトウェアにて測定し、その精度を検証した。さらに 59 症例の GGO 結節の体積を 2 名の放射線科医が測定し、観察者内・観察者間誤差を検証した。

続いて、病理診断（atypical adenomatous hyperplasia [AAH], bronchioloalveolar carcinoma [BAC], adenocarcinoma [AC]）が得られた 47 GGO 結節における経時的な体積変化をコンピュータ支援ソフトウェアにて測定し、VDT を算出した。

[結果]

① −450HUの模擬結節におけるDES画像の有無におけるarea under the ROC curve (AUC)値は0.66と0.77であり検出能の有意な向上がみられた。−200HUと30HUの結節においてはDES画像を併用することによる有意な検出能の向上は認められなかった。臨床症例においてはDES画像を併用することで検出能の向上を認め、特にmixed GGO結節で有意な検出能の向上がみられた。

また、MTANNを応用したvirtual DESを併用することでも有意な検出能の向上が認められた。

② ソフトウェアの測定誤差は直径3mmの模擬GGO結節で51.1 − 85.2%、直径5mm以上で−4.1 − 7.1%だった。直径8mm以上のGGO結節における観察者内誤差の95%信頼区間は観察者Aで−14.9 − 13.7%、観察者Bで−16.6 − 15.7%であり観察者間誤差はで−16.3 − 23.7%であった。

全GGO結節の平均VDTは486.4±368.6日だった。疾患別の平均VDTでは、AAHで859.2±428.9日、BACで421.2±228.4日、ACで202.1±84.3日だった。AAH、BAC、ACそれぞれのVDTの間に有意差がみられた。

[結論]

① FPD X 線装置においてDES 画像を併用することで小型肺結節の検出能が有意に向上した。また、MTANN を応用した virtual DES 画像も小型肺結節の検出に有用である。

② GGO 結節におけるコンピュータ支援3次元的体積測定は直径 5mm 以上の GGO 結節で測定誤差が小さく、直径 8mm 以上の GGO 結節で測定者内・測定者間誤差が小さいことが判明した。

コンピュータ支援3次元的体積測定による VDT の評価は GGO の質的診断に有用と思われる。

Publication list

1. Oda S., Awai K., Suzuki K., Yanaga Y., Funama Y., MacMahon H., Yamashita Y. Detection of Small Pulmonary Nodules on Chest Radiographs: Effect of Rib Suppression by a Massive Training Artificial Neural Network (MTANN) Technique on the Performance of Radiologists. AJR Am J Roentgenol. 193: W397-402, 2009

2. Oda S., Awai K., Murao K., Ozawa A., Yanaga Y., Kawanaka K., Yamashita Y. Computer-Aided Volumetry of Pulmonary Nodules Showing Ground Glass Opacity at Multidetector Computed Tomography. AJR Am J Roentgenol. 194: 398-406, 2010

3. Oda S., Awai K., Funama Y., Utsunomiya D., Yanaga Y., Kawanaka K., Yamashita Y. Detection of Small Pulmonary Nodules with Various Attenuation by Using Dual-energy Subtraction Chest Radiography: Phantom Study with Receiver Operating Characteristics Analysis. Japanese Journal of Radiology (in-press)

4. Oda S, Awai K, Liu D, Nakaura T, Yanaga Y, Nomori H, Yamashita Y. Ground-glass opacities on thin-section helical CT: differentiation between bronchioloalveolar carcinoma and atypical adenomatous hyperplasia. .AJR Am J Roentgenol. 190: 1363-1368, 2008

5. Oda S, Awai K, Funama Y, Utsunomiya D, Yanaga Y, Kawanaka K, Nakaura T, Hirai T, Murakami R, Nomori H, Yamashita Y. Detection of Small Pulmonary Nodules on

Chest Radiographs: Efficacy of Dual-Energy Subtraction Technique Using Flat-Panel Detector Chest Radiography Clinical Radiology. (in-press)

6. Funama Y, Awai K, Liu D, Oda S, Yanaga Y, Nakaura T, Kawanaka K, Shimamura M, Yamashita Y. Detection of nodules showing ground-glass opacity in the lungs at low-dose multidetector computed tomography: phantom and clinical study. J Comput Assist Tomogr. 33: 49-53, 2009

7. Liu D, Awai K, Funama Y, Oda S, Nakaura T, Yanaga Y, Hatemura M, Kawanaka K, Yamashita Y. Identification and characterization of focal ground-glass opacity in the lungs by high-resolution CT using thin-section multidetector helical CT: experimental study using a chest CT phantom. Radiat Med. 26: 21-27, 2008

8. Yanaga Y, Awai K, Nakaura T, Oda S, Funama Y, Bae KT, Yamashita Y. Effect of contrast injection protocols with dose adjusted to the estimated lean patient body weight on aortic enhancement at CT angiography. AJR Am J Roentgenol. 192: 1071-1078, 2009

9. Nakaura T, Awai K, Yauaga Y, Nakayama Y, Oda S, Hatemura M, Nagayoshi Y, Ogawa H, Yamashita Y. Contrast injection protocols for coronary computed tomography angiography using a 64-detector scanner: comparison between patient weight-adjusted- and fixed iodine-dose protocols. Invest Radiol. 43: 512-519, 2008

10. Yanaga Y, Awai K, Nakaura T, Namimoto T, Oda S, Funama Y, Yamashita Y. Optimal contrast dose for depiction of hypervascular hepatocellular carcinoma at dynamic CT using 64-MDCT. AJR Am J Roentgenol. 190: 1003-1009, 2008

11. Nakaura T, Awai K, Yanaga Y, Nakayama Y, Oda S, Funama Y, Yamashita Y. Detection of early enhancement of hypervascular hepatocellular carcinoma using single breath-hold 3D pixel shift dynamic subtraction MDCT. AJR Am J Roentgenol. 190: W13-18, 2008

12. 彌永由美、粟井和夫、中浦猛、船間芳憲、尾田済太郎、中村信一、浪本智弘、山下康行. Multidetector CT urography (MDCTU). 映像情報メディカル 40:113-119;2008

13. 彌永由美、粟井和夫、中浦猛、伊牟田真功、浪本智弘、尾田済太郎、中村信一、山下康行. 心臓 CT における合理的な造影法. 臨床画像 24:26-35;2008

14. 彌永由美、粟井和夫、中浦猛、伊牟田真功、浪本智弘、尾田済太郎、中村信一、山下康行. 64 列 CT による肝パフージョン CT. Innervision 230(6), 47-51, 2008

Acknowledgements

These academic investigation took place during my postgraduate study period from 2007-2010, at the department of Diagnostic Radiology & Imaging, Kumamoto University School of Medicine.

I'd like to express my sincere thanks to Professor Yasuyuki Yamashita, Chairman of the Department of Diagnostic Radiology, Graduate School of Medical Sciences, Kumamoto University, his generous guidance and constructive instructions. Never would I have finished my postgraduate study without help in academic field.

I am deeply grateful Dr. Kazuo Awai, professor, Department of Diagnostic Image Analysis, Graduate School of Medical Sciences, Kumamoto University, who has been instructing me during the most part of my research with his profound insight into the Radiological Sciences shed on my whole study. I have learned much from his scientific way of thinking, serious working skill and style and honest personality.

I, of course, owe a great deal to Dr. Nakaura, Dr. Yanaga , Dr. Utsunomiya and Dr. Funama for their cooperation and helps in my work. My appreciation also goes to all the other members of the Department of Diagnostic Radiology, Graduate School of Medical Sciences, Kumamoto University.

Finally I want to say thank you to all people who have ever helped me and love me. I determine to do better for the sake of your love in the future.

Abbreviations

DES: dual-energy subtraction

FPD: flat-panel detector

MTANN: massive-training artificial neural network

GGO: ground glass opacity

CAV: computer-aided volumetry

VDT: volume-doubling time

HU: Hounsfield unit

ROC: receiver-operating-characteristic

AAH: atypical adenomatous hyperplasia

BAC: bronchioloalveolar carcinoma

AC: adenocarcinoma

AUC: area under the ROC curve

IAEA: International Atomic Energy Agency

PPV: positive predictive values

NPV: negative predictive values

DQE: detective quantum efficiency

CT: computed tomography

IRB: institutional review board

MTANN: massive-training artificial neural network

CAD: computer-aided diagnosis

RVME: relative volume measurement error

HRCT: high-resolution computed tomography

MDCT: multidetector computed tomography

ROI: region of interest

ER: estimated error

SD: standard deviation

CI: confidence interval

3D: 3 dimensional

Chapter 1. Background and objective

Contents

1. Clinical significance of small pulmonary nodules

Since the introduction of helical computed tomography (CT) in the early 1990s, the detection and characterization of small pulmonary nodules at CT has become routine in patient work-ups. Multi-detector row CT (MDCT), introduced in the late 1990s, offers high-volume coverage speed performance and is now one of the most robust tools for the diagnosis of pulmonary nodules.

At MDCT, many smokers were found to harbor small solid pulmonary nodules (1). The clinical importance such nodules measuring less than 1 cm in diameter differs substantially from that of larger nodules detected on chest radiographs, in that the vast majority of the small nodules are benign with an overall malignancy rate of less than 1 - 2% (2). The positive relationship between lesion size and likelihood of malignancy has been demonstrated in solid nodules (3).

Pulmonary nodules exhibiting ground glass opacity (GGO nodules) at MDCT have attracted attention. Among pulmonary nodules detected at lung cancer screening with low radiation-dose helical CT, 19 - 38% manifested small focal GGO nodules (4, 5) and GGO nodules have been suggested as a sign of early-stage adenocarcinoma or its precursor (4, 6-8). Henschke et al. (4) reported that at 18 - 63%, the malignancy rate of GGO nodules was much higher than of solid nodules. Therefore, the early detection

and accurate characterization of GGO nodules is clinically important. However, it is often difficult to determine the specific cause of GGO nodules because they may be attributable to various pathologic entities including focal inflammation, focal interstitial fibrosis (9), atypical adenomatous hyperplasia (AAH) (6, 7), bronchioloalveolar carcinoma (BAC) (6, 7, 10), and adenocarcinoma (6, 7, 10). Although monitoring the changes in the CT appearance of these lesions over time may help to obtain a differential diagnosis, a definitive diagnosis may require interventional procedures. As the objective assessment of temporal changes in the CT appearance of these lesions, such as their size or nature, is not necessarily easy, some special devices may be required.

2. Detection of small pulmonary nodules on chest radiographs

2.1 Current status of the detection of small pulmonary nodules on chest radiographs

Chest radiography remains the most widely used imaging technique for the detection of chest diseases because of its low cost, simplicity, and low radiation dose. However, the false-negative rate for the detection of pulmonary nodules is relatively high; it was reported to range from 19 - 72% (11, 12) and radiography is inferior to CT with respect to the detection of small pulmonary nodules (13). Failure to

detect pulmonary nodules on chest radiographs has been attributed to their size (14) and density (15) and obscuration by structures such as the ribs, clavicles, mediastinum, and pulmonary vessels (16). In fact, 82 - 95% of lung cancers missed by radiologists were partly obscured by overlying bones such as the ribs and/or clavicles (16). Therefore, more effective methods for detecting nodules on radiographs must be developed.

2.2 Dual-energy subtraction technique for chest radiography

The dual-energy subtraction (DES) technique for chest radiography can remove overlying bone structures to create soft-tissue-selective images (17, 18), thereby enhancing the visualization of pulmonary nodules overlaid by bones. Consequently, DES may improve the detectability of chest lesions, especially small pulmonary nodules.

Single- and double-exposure dual-energy systems are available. In the former system, DES images are obtained by exposing two storage phosphor plates separated by a copper filter. A disadvantage of the single-exposure system is the lower signal-to-noise ratio of the tissue-selective subtraction image. On the other hand, in double-exposure systems (18), two sequential radiographs are obtained at low- and high kV settings, respectively. Although the 200-millisecond delay between the two exposures may produce misregistration artifacts on the subtracted images, double-exposure

systems produce DES images with a better signal-to-noise ratio than single-exposure systems. The double-exposure DES technique is now available in full-field digital flat-panel detector (FPD) radiography systems that are widely used because of the rapid accessibility of images, improved image quality, and the possibility of reduced radiation exposure (19).

Earlier studies indicated the usefulness of the DES technique at computed radiography (CR) and conventional film-screen systems (20). Although a few earlier studies (21, 22) evaluated the performance of radiologists in the detection of pulmonary nodules using DES techniques with the FPD system, the usefulness of this technique with the FPD system has not been established.

2.3 Virtual dual-energy subtraction chest radiography using a massive training artificial neural network

The massive training artificial neural network (MTANN) technique is a highly nonlinear pattern-recognition technique that has been applied in computer-aided diagnosis for the detection of pulmonary nodules on CT scans (23, 24). We developed a novel method that achieves rib suppression on chest radiographs by using the MTANN (25). It can produce virtual DES images such as soft-tissue- and bone images from a

single standard chest radiograph obtained with a standard radiography system. For the generation of virtual DES images only software is needed; neither specialized equipment nor additional radiation dose exposure is required. Our trial described in ***chapter 4*** was the first clinical trial of the use of virtual DES images to detect pulmonary nodules.

3. Qualitative diagnosis of small pulmonary nodules on CT

3.1 Current status of the diagnosis of small pulmonary nodules on CT

Rapid technological advances in CT scanners, especially the development of multi-row detectors and sub-second gantry rotation times, greatly increased the speed of image acquisition. This facilitated a substantial decrease in the CT section thickness and interval. Thinner sections reduce the effect of volume averaging, thereby improving spatial resolution and allowing the detection of smaller lung nodules. Furthermore, pulmonary nodules can be evaluated in any direction by using multiplanar reconstruction images. This improved technology resulted in the incidental detection of large numbers of small pulmonary nodules, a fact noted in the literature, especially in CT studies for lung cancer screening (2). The frequency of detecting one or more small pulmonary nodules on a screening CT scan ranges from 5 - 60%, with

higher rates noted on thinner slices (11, 26-31). However, the majority of these small pulmonary nodules are benign (2). These false-positive CT findings may result in unnecessary intervention in patients harboring benign nodules.

In general, a differential diagnosis of solitary pulmonary nodules is based on morphologic findings on CT images, primarily the characteristics of the nodule margins, the internal nodular structure, and the relationship to surrounding structures. However, it is often difficult to differentiate between benign- and malignant nodules on a single CT study (32). Practically, the differentiation between a benign- and malignant nodular etiology involves the absence of growth on follow-up CT scans. Although the medical community recommended that the absence of nodular growth at 2-year follow-up be considered as proof of the benign nature of the nodule, the subjective evaluation of the growth rate by radiologists is unreliable with respect to slow-growing tumors such as BAC whose doubling time is very long (457 - 813 days on average) (33, 34). Therefore, accurate and robust methods to measure the size of pulmonary nodules must be developed.

3.2 Computer-aided volumetry of pulmonary nodules using MDCT

With MDCT scanners it is possible to scan a wide range, including areas containing pulmonary nodules, at a detector collimation of 0.500 - 0.625 mm under one breath-hold. This facilitates the 3-dimensional (3D) evaluation of pulmonary nodules. Although the size of pulmonary nodules is usually determined and compared using 2-dimensional (2D) CT evaluation as earlier described, such measurements are not reliable particularly with respect to small nodules (35). Yankelevitz et al. (36) reported that 3D volumetric measurements offered obvious advantages over conventional bilinear 2D measurements for the accurate representation of the nodular volume. Therefore, the pulmonary nodule volume can now be estimated reliably and accurately using 3D volumetric measurements and automated computer software. In studies on computer-aided volumetry (CAV) of pulmonary nodules using volumetric data obtained at MDCT (36-40), accuracy and reproducibility were sufficiently high (37). However, most of the software used in these studies only evaluated solid pulmonary nodules and CAV of GGO nodules was reported to be difficult (40). The development of CAV software that can be used to measure not only the volume of solid- but also of GGO nodules is a challenging issue.

3.3 Clinical application of volumetry of pulmonary nodules

One of the most important indicators of malignant pulmonary nodules is their growth rate; this is commonly expressed as the volume doubling time (VDT) (38). The VDT is an independent and significant prognostic factor in lung cancer patients (41). Shorter VDT may reflect greater histological tumor aggressiveness, suggesting that lung cancers with a short VDT are associated with a worse prognosis. Conventionally, the evaluation of tumor growth has been performed by bilinear 2D measurements on chest radiographs or CT images. The introduction of 3D CAV of pulmonary nodules may facilitate a more accurate and earlier evaluation of tumor growth using volumetric data obtained at MDCT. Most previously reported doubling times for pulmonary nodules were based on 2D measurements using a modified Schwartz equation (41, 42) that should be applied to round nodules. However, human pulmonary nodules are not necessarily round and their margins are not always smooth. Therefore, the VDT based on 2D measurements may be unreliable. To set a new standard, the VDT of pulmonary nodules should be re-evaluated with accurate 3D volumetry.

Accurate assessment of the therapeutic tumor response is clinically important. The "Response Evaluation Criteria in Solid Tumors (RECIST)" that use uni-

dimensional manual measurements of the sum of the longest diameter of the marker lesion (43) are considered to be the methodology of choice for assessing the tumor response to treatment (44). However, recent evidence suggests that the value of RECIST may be limited in situations where uni-dimensional measurements are highly variable. Moreover, in the evaluation of targeted drugs that do not yield frank disease remission, the RECIST classification, while compensating for errors of manual measurements, is insensitive to minimal tumor responses (45). The evaluation of pulmonary nodules with 3D CAV may decrease observer variation with very high confidence in the tumor response assessment.

4. Research Objectives

In ***chapter 2*** we described our phantom experiments using simulated pulmonary nodules with different X-ray attenuation in an effort to elucidate the relationship between the detectability of small pulmonary nodules with a dual-exposure DES technique and nodule density.

In ***chapter 3*** we reported our clinical study in which we evaluated the effectiveness of a dual-exposure DES technique in FPD chest radiography. In that

study we also analyzed the relationship between nodule detectability and nodule density in nodules located in the thorax.

As outlined in section 2.3, the MTANN technique can produce virtual DES images such as soft-tissue- and bone images from a single standard chest radiograph obtained with a standard radiography system. In ***chapter 4***, we assessed the effect of the MTANN technique on the performance of radiologists in the detection of pulmonary nodules on chest radiographs.

The introduction of 3D volumetry of pulmonary nodules using CAV software may facilitate a more accurate and earlier evaluation of tumor growth. Therefore, the value of 3D volumetry has been actively investigated in phantoms and in vivo. However, most of the software used in earlier studies only evaluated solid pulmonary nodules and CAV of GGO nodules can be difficult. In ***chapter 5***, we investigated the accuracy and reproducibility of results obtained on GGO nodules with newly developed CAV software that can also measure the volume of GGO nodules. Finally, in ***chapter 6***, we assessed the VDT of histologically proven GGO nodules using 3D CAV software and discussed whether VDT evaluation was useful for a differential diagnosis of GGO nodules.

Chapter 2. Detection of Small Pulmonary Nodules by Using Dual-energy Subtraction Chest Radiography: Phantom Study

Contents

ABSTRACT

Purpose: To investigate the detectability of simulated pulmonary nodules with different X-ray attenuation by flat-panel detector (FPD) chest radiography using a dual-exposure dual-energy subtraction (DES) technique.

Materials and Methods: Using a FPD radiography system we obtained 108 sets of chest radiographs of a chest phantom. They consisted of 54 sets each of chest radiographs with- and without simulated nodules. Each data set contained a standard- and a corresponding dual-energy subtracted chest radiograph (DES image). The diameter of the simulated nodules was 8-, 10-, and 12 mm; nodules of each size manifested attenuation of -450, -200, and 30 Hounsfield units (HU). We performed receiver operating characteristic (ROC) analysis to compare the observers' performance in detecting nodules.

Results: For -450 HU nodules the mean area under the ROC curve (AUC) without and with DES images was 0.66 and 0.77, respectively; the difference was significant (paired t-test, $p < 0.01$). For nodules with -200- and 30 HU, there was no significant difference in the AUC value (0.79 vs. 0.77, $p=0.13$; 0.92 vs. 0.94, $p=0.17$, respectively).

Conclusion: The addition of DES images to standard chest radiographs improved the performance of radiologists charged with detecting simulated nodules with an attenuation of -450 HU.

INTRODUCTION

Chest radiography still represents the most common tool in diagnostic radiology due to its low cost, low dose, and simple implementation. However, the false-negative rate of chest radiography for the detection of pulmonary nodules is relatively high; it was reported to range from 19 to 72 % (11, 12). The failure to detect pulmonary nodules on chest radiographs has been attributed to their size (14) and density (15) and obscuration by bony structures such as the ribs and/or clavicles (16). The dual-energy subtraction (DES) technique for chest radiography is one of the promising methods to reduce these anatomical noises. Using different energy X-ray beams, DES chest radiography can remove overlying bone structures and generate soft-tissue-selective images (17, 18). It enhances the visualization of pulmonary nodules overlaid by bones and may improve the detectability of pulmonary nodules. The DES technique is now available in full-field digital flat-panel detector (FPD) radiography systems that feature rapid accessibility of images, improved image quality, and the possibility of reduced radiation exposure (19).

The aim of this study was to investigate the detectability of simulated pulmonary nodules with different X-ray attenuation by FPD chest radiography using a dual-exposure DES technique.

MATERIALS AND METHODS

Dual-Energy Subtraction Chest Radiography

DES chest radiographs were obtained using a FPD digital chest system (Revolution XR/d, GE Healthcare, Milwaukee, WI). The detector featured an image size of 41 × 41 cm and a pixel dimension of 0.2 ×0.2 mm. DES images were acquired with a double-exposure technique with 200 ms between the high- (120-kV) and low-energy (60-kV) exposures; standard posteroanterior-, soft-tissue-, and bone images were generated. The imaging parameters included a 120-kV image at a speed equivalent to approximately 400, and a 60-kV image at a speed setting equivalent to approximately 1000. Our phantom study had shown that the entrance surface radiation dose for a standard posteroanterior chest radiograph with the FPD system was 0.097 mGy; it was 0.168 mGy for double-exposure DES imaging at the above speed setting (unpublished data). The radiation dose at double-exposure DES chest radiography with the FPD system is nearly equal to that of conventional computed- or film-screen radiography systems (19) and lower than that of the International Atomic Energy Agency (IAEA) guidance level for diagnostic chest radiography, i.e. 0.4 mGy at posteroanterior chest radiography (46).

Database

We used a chest phantom and simulated nodules (multi-purpose chest phantom N1, Kyoto Kagaku Co., Kyoto, Japan). In this phantom, simulated soft tissues such as pulmonary vessels, the chest wall, heart, diaphragm, and liver consist of polyurethane resin composites; simulated bone of an epoxide resin. The space between the pulmonary vessels, heart, and chest wall was filled with air. The chest wall can be removed from the other structures such as the simulated heart, pulmonary vessels, diaphragm, and liver. Simulated nodules were spheres made of urethane foam resin. We arbitrarily placed nodules on the bifurcations of the pulmonary vessels at any level of the lungs.

Using an FPD radiography system, we obtained 108 sets of chest radiographs of a chest phantom which consisted of 54 sets each of chest radiographs with- and without simulated pulmonary nodules. Each data set also contained a standard- and a corresponding dual-energy subtracted chest radiograph (DES image). The diameter of the simulated nodules was 8-, 10-, and 12 mm; nodules of each size manifested attenuation of -450, -200, and 30 Hounsfield units (HU, nominal values), the actual values measured at our CT scanners were -443, -205, 33 HU, respectively. The

simulated nodules were randomly inserted into the lung field of the chest phantom by one of the authors (Y. F.) who did not participate the observer performance study.

Observer Performance Study

We used a sequential-test method for receiver-operating-characteristic (ROC) analysis to evaluate the diagnostic performance of radiologists detecting pulmonary nodules on chest radiographs without and with DES images. Seven board-certified radiologists with 10 – 22 years of experience (mean 13.3 years) participated in this observer performance study. All specialized in body imaging and read chest radiographs regularly. They were allowed to change the level and width of the window on the monitor; reading time was not limited.

All observers recorded the location of detected simulated nodules on record sheets. They used a continuous rating scale and a line-marking method to rate their confidence level by placing marks on a 7-cm-long line on a recording form. The left end of the line indicated complete confidence that the chest radiographs without/with DES images did not, the right end indicated complete confidence that they did reveal a nodule. Intermediate levels of confidence were indicated by the position of the marks between the 2 line termini, where marks close to the right and left end indicated a greater and lesser degree of confidence, respectively. One author (Y.F.) then measured

the distance between the left end of the line and the mark and converted the distance to an ordinal confidence rating ranging from 0 - 100. A continuous rating scale containing a pair of horizontal lines was used in the sequential test. Observers first recorded their rating of chest radiographs without DES images on the upper line. Subsequently, they recorded their rating of DES images on the lower line. They entered their results for each case on a record form.

Statistical Analysis

We used ROC analysis to compare the radiologists' performance in detecting simulated pulmonary nodules with the different attenuation values on chest images acquired without and with the DES technique. A binormal ROC curve was fitted to each radiologist's confidence rating data acquired under the 2 reading conditions by applying quasi-maximum likelihood estimation (47). A computer program (ROCKIT; Charles E. Metz, University of Chicago, Chicago, IL) was used for obtaining binormal ROC curves from the ordinal-scale rating data (47). The area under the best-fit ROC curve (AUC) plotted in unit squares was calculated for each fitted curve.

The statistical significance of the difference in AUC values between the ROC curves obtained without and with DES images was tested. The paired t-test was

performed with a statistical software package (SPSS, version 15.0; SPSS, Chicago, IL), and p values less than 0.05 were considered to indicate statistically significant differences.

RESULTS

For all simulated nodules the AUC values for all 7 observers were higher with than without DES images (Table 1). Analysis of the overall performance of the 7 observers in the detection of simulated nodules indicated that the mean AUC values increased from 0.76 ± 0.08 (without DES images) to 0.79 ± 0.07 (with DES images); the difference was statistically significant ($p = 0.02$) (Fig. 1).

For simulated -450 HU nodules the mean AUC vales without and with DES images were 0.66 ± 0.06 and 0.77 ± 0.03, respectively; the difference was statistically significant ($p < 0.01$) (Fig. 2). On the other hand, for simulated nodules with -200- and 30 HU, there was no statistically significant difference in the AUC values without and with DES images (0.79 ± 0.07 vs. 0.77 ± 0.07, $p=0.13$; 0.92 ± 0.04 vs. 0.94 ± 0.03, $p=0.17$, respectively) (Figs. 3, 4).

When an observer correctly identified the location of the nodules and assigned a confidence level of 50 or more, we recorded the diagnosis as being correct and calculated mean sensitivity, specificity, accuracy, and positive- and negative predictive values (PPV, NPV) for all observers for all nodules. The sensitivity, specificity, accuracy, PPV, and NPV were 48.8%, 83.6%, 65.4%, 79.5%, and 58.9%, respectively, for standard images. These values were 63.2%, 83.6%, 73.4%, 82.6%, and 70.6%,

respectively, for DES images (Table 2). On DES images, sensitivity, accuracy and NPV were significantly improved.

A representative case is shown in Fig. 5.

DISCUSSION

Chest radiography is currently the most frequently used screening procedure for lung cancer because it is economical and easy to use. One of the main shortcomings of chest radiography is poor sensitivity for the detection of solitary pulmonary nodules that are smaller than 2 cm in diameter (48). Furthermore, the detection of small ground glass opacity (GGO) nodules is especially difficult on chest radiographs.

Results of previous studies indicated the usefulness of the DES technique at computed radiography or on film-screen systems (20). Few earlier studies evaluated the performance of radiologists in the detection of pulmonary nodules using DES techniques with the FPD system and the usefulness of this technique with the FPD system has not been established. Tagashira et al. (21) studied 50 patients with one or more nodules and 50 patients without nodules; they reported that the AUC values of 7 observers were significantly increased from 0.79 to 0.84 with DES images. Ricke et al. (22) evaluated 20 patients with a total of 59 pulmonary nodules; they found that the DES technique significantly improved sensitivity (from 33 to 42%), specificity (from 81 to 85%), and confidence in the detection of small pulmonary nodules. On the other hand, Rühl et al. (49), who studied 100 patients with a total of 149 pulmonary nodules, concluded that at FPD chest radiography the DES technique did not significantly improve the detection of pulmonary nodules. However, they did not analyze the

relationship between nodule detectability and nodule density. Our study is the first to investigate the detection of pulmonary nodules from the viewpoint of their relationship to nodule density using the DES technique with the FPD radiography system.

Our results suggested that the combined evaluation of DES images and original chest radiographs significantly improved the diagnostic performance of observers in the detection of simulated nodules with an attenuation of -450 HU but not of -200- and 30 HU. We posit that DES images enhance the visualization of pulmonary nodules with -450 HU, especially of nodules with overlapping bone shadows. Consequently, the detectability of nodules with -450 HU was improved.

Ide et al. (17) who studied 77 consecutive lung cancer patients and 77 healthy subjects reported that the DES technique at computed radiography failed to improve the detection of non-solid- and solid nodules, however, it significantly improved the detection of part-solid nodules, all of which were bronchioloalveolar carcinoma. We posit that -450 HU is equivalent to the density of part-solid nodules from bronchioloalveolar carcinoma (50).

The DES technique with the FPD system may also improve the detectability of part-solid nodules. On the other hand, the detectability of -200- and 30 HU nodules was equivalent on DES- and standard images.

FPD radiography systems are now widely used because of the rapid accessibility of images, improved image quality, and the possibility of reduced radiation exposure (19). The FDP radiography system yields higher detective quantum efficiency (DQE) than computed radiography and film-screen radiography systems (51). A higher DQE increases the ability to reveal objects in a noisy background (51) and reduces the radiation exposure without sacrificing image quality. At approximately 50% of the radiation dose, the quality of images obtained with the FPD radiography system was equal to that of computed- or film-screen radiography systems (52) and use of the double-exposure technique with FPD may not increase the radiation dose compared with these other systems.

Currently, single- and double-exposure dual-energy systems are available. In the former, DES images are obtained by exposing two storage phosphor plates separated by a copper filter. A disadvantage is the lower signal-to-noise ratio of the tissue-selective subtraction image. Also, theoretically, single-exposure systems cannot be used with FPD systems. On the other hand, in double-exposure systems (18), two sequential radiographs are obtained at low- and high kV settings, respectively. The 200-millisecond delay between the two exposures may produce misregistration artifacts

on the subtracted images. However, double-exposure systems produce DES images with a better signal-to-noise ratio than single-exposure systems.

There are several limitations in our study. First, we did not evaluate the effect of this technique on the efficacy of detecting pulmonary nodules with calcification. While such nodules may be subtracted on soft tissue images, they can be detected easily on standard- or bone images. Since most calcified nodules are benign, their detection may be of lower clinical significance. Second, we only evaluated the detectability of pulmonary nodules. We did not distinguish between nodules obscured by clavicles or ribs as compared to nodules not overlaid by structures that are subtracted by the DES technique. However, we analyzed the relationship between nodule detectability and bony structures in an unpublished clinical study. We found that diagnostic performance was significantly improved for nodules with overlapping bone shadows, for nodules without overlapping, it was not. Furthermore, Kido et al. (53, 54), who studied single-exposure DES images using computed radiography, revealed that DES images especially improved the detectability of nodules with overlapping bone shadows. Third, we conducted only phantom- but not clinical studies. The simulated nodules in the chest phantom were round and had clear margins and there were no misregistration artifacts such as may be attributable to cardiac-, respiratory-, and patient motion in the

clinical setting where pulmonary nodules are not necessarily round, their margins are not always clear, and there may be various lung diseases in the patient background. In addition, misregistration artifacts may be present on clinical images. Therefore, the detection of small pulmonary nodules on DES images may be more difficult in clinical- than phantom studies. The DES technique should be rigorously evaluated by large-scale clinical studies. Finally, this technique should also be evaluated with respect to the nodule size and the observers' experience with the interpretation of chest radiographs.

In conclusion, the addition of DES images to standard chest radiographs improved the performance of radiologists charged with detecting simulated nodules with an attenuation of -450 HU. The detectability of -200-, 0-, and 30 HU nodules was equivalent on DES- and standard images. We suggest that the DES technique with FPD chest radiography can reduce the number of GGO nodules that are missed in routine clinical practice.

Chapter 3. Detection of Small Pulmonary Nodules by Using Dual-energy Subtraction Chest Radiography: Clinical Study

Contents

ABSTRACT

Purpose: We investigated the effect of a double-exposure dual-energy subtraction (DES) technique on the performance of radiologists detecting small pulmonary nodules on flat-panel detector (FPD) chest radiographs.

Materials and Methods: Using on FPD radiography we obtained 41 sets of chest radiographs from 26 patients with pulmonary nodules measuring ≤ 20 mm and 15 normal subjects. Each data set included standard- and corresponding DES images. There were 6 non-solid-, 10 part-solid-, and 10 solid nodules. The mean size of the 26 nodules was 15.0 ± 4.8 mm. We performed receiver operating characteristic (ROC) analysis to compare the performance of 8 board-certified radiologists.

Results: For the 8 radiologists, the mean value of the area under the ROC curve (AUC) without and with DES images was 0.62 ± 0.05 and 0.68 ± 0.05, respectively; the difference was statistically significant ($p = 0.02$). For part-solid nodules, the difference of the mean AUC value was statistically significant (AUC = 0.61 ± 0.07 vs. 0.69 ± 0.05; $p < 0.01$); for non-solid nodules it was not (AUC = 0.62 ± 0.10 vs. 0.61 ± 0.09; $p = 0.73$), and for solid nodules it was not (AUC = 0.75 ± 0.10 vs. 0.78 ± 0.08; $p = 0.23$). For nodules with overlapping bone shadows, the difference of the mean AUC value was statistically significant ($p = 0.03$), for nodules without overlapping, it was not ($p = 0.26$).

Conclusion: Use of a double-exposure DES technique at FPD chest radiography significantly improved the diagnostic performance of radiologists detecting small pulmonary nodules.

INTRODUCTION

Chest radiography remains the most widely used imaging technique for the detection of chest diseases because of its low cost, simplicity, and low radiation dose. However, the false-negative rate for the detection of pulmonary nodules is relatively high; it was reported to range from 19 to 72%,(11, 12) and radiography is inferior to low-dose computed tomography (CT) with respect to the detection of small pulmonary nodules.(13) The failure to detect pulmonary nodules on chest radiographs has been attributed to their size(14) and density(15) and obscuration by structures such as the ribs, clavicles, mediastinum, and pulmonary vessels.(16) In fact, 82 - 95% of lung cancers missed by radiologists were partly obscured by overlying bones such as the ribs and/or clavicles.(16)

The dual-energy subtraction (DES) technique for chest radiography can remove overlying bone structures to create soft-tissue-selective images(17, 18) and it enhances the visualization of pulmonary nodules overlaid by bones. Consequently, its use may improve the detectability of chest lesions, especially small pulmonary nodules. The DES technique is now available in full-field digital flat-panel detector (FPD) radiography systems that are widely used because of the rapid accessibility of images, improved image quality, and the possibility of reduced radiation exposure.(19)

We investigated the clinical efficacy of a double-exposure DES technique using FPD chest radiography systems in the detection of pulmonary nodules smaller than 20 mm.

MATERIALS AND METHODS

Our study was approved by our institutional review board (IRB); informed consent was obtained from all patients who underwent double-exposure DES radiography. Our IRB also approved the participation of the 8 radiologists in the observer performance test. Informed consent for the observer performance study was obtained from all participants.

Database

Between August 2007 and February 2008, 136 patients underwent double-exposure DES radiography and chest CT for the evaluation of chest diseases in our department. One chest radiologist (K.A.) with 22 years of chest CT experience, who did not participate in the observer performance study, selected chest radiographs based on the following inclusion criteria: (1) DES radiography and chest CT were performed within 14 days, (2) no treatment or biopsy was performed between the 2 procedures, (3) the pulmonary nodules did not exceed 20 mm in the x-y (or transverse) plane on thin-section CT, (4) absence on CT scans of other pleural and parenchymal abnormalities such as consolidation, diffuse ground glass opacity (GGO), and pulmonary fibrosis. Based on these criteria we identified 26 patients, 11 men and 15 women ranging in age

from 40-85 years (mean 62.9 years). The same radiologist also selected 15 individuals, 11 men and 4 women ranging in age from 48-64 years (mean 55.5 years) without CT evidence of nodules or other chest diseases. Of the 26 patients with pulmonary nodules, 13 had bronchioloalveolar carcinoma, 9 adenocarcinoma, 1 squamous cell carcinoma, 2 solitary pulmonary metastasis, and 1 had a hamartoma. The histological diagnoses were confirmed at thoracic surgery. The mean size of the pulmonary nodules on CT was 15.0 ± 4.8 mm (range 7.1 - 20.0 mm). The chest radiologist classified the 26 nodules as non-solid- (pure GGO-, n=6), part-solid- (mixed GGO-, n=10) and solid nodules (n=10) according to their internal density.(4) He also classified the nodules into two types i.e. as nodules with and without overlapping bone shadows on chest radiographs. Overlapping was recorded when 50% or more of the nodular area overlapped with ribs and/or clavicles. Of the 26 nodules, 9 (34.6%) did and the other 17 did not overlap.

DES Chest Radiography and Chest CT

DES chest radiographs were obtained using an FPD digital chest system (Revolution XR/d, GE Healthcare, Milwaukee, WI). The detector featured an image size of 41 × 41 cm and a pixel dimension of 0.2 ×0.2 mm. DES images were acquired with a double-exposure technique with 200 ms between the high- (120-kV) and low-

energy (60-kV) exposures; standard posteroanterior-, soft-tissue-, and bone images were generated (Fig. 6). The imaging parameters included a 120-kV image at a speed equivalent of approximately 400, and a 60-kV image at a speed setting equivalent to approximately 1000. Our phantom study had shown that the radiation dose for a standard posteroanterior chest radiograph with the FPD system was 97 μGy; it was 168 μGy for double-exposure DES imaging at an above speed setting (unpublished data). The radiation dose at double-exposure DES chest radiography with the FPD system is nearly equal to that of conventional computed- or film-screen radiography systems.(19) CT imaging was with a 64-detector CT scanner (Brilliance-64, Philips Medical Systems, Cleveland, OH). The scanning parameters were: detector collimation 64 x 0.625 mm, helical pitch 0.673, rotation time 0.5 sec, tube voltage 120 kVp, tube current 250 mAs, and continuous reconstruction at 1.5 mm, and 5 mm slice thickness.

Observer Performance Study

We used a sequential-test method for receiver-operating-characteristic (ROC) analysis to evaluate the diagnostic performance of radiologists detecting small pulmonary nodules on chest images without and with DES images. They were 8 board-certified radiologists with 12 - 22 years of experience (mean 15.6 years). All

specialized in body imaging and read chest radiographs regularly. They were allowed to change the level and width of the window on the monitor; reading time was not limited.

All observers used a continuous rating scale and a line-marking method to rate their confidence level by placing marks on a 7-cm-long line on a recording form. The left end of the line indicated complete confidence that the chest radiographs without/with DES images did not, the right end indicated complete confidence that they did reveal a nodule. Intermediate levels of confidence were indicated by the position of the marks between the 2 line termini, where marks close to the right and left end indicated a greater and lesser degree of confidence, respectively. One author (Y.F.) then measured the distance between the left end of the line and the mark and converted the distance to an ordinal confidence rating ranging from 0 - 100. A continuous rating scale containing a pair of horizontal lines was used in the sequential test. Observers first recorded their rating of chest radiographs without DES images on the upper line. Subsequently, they recorded their rating of DES images on the lower line. They entered their results for each case on a record form.

First, to become familiarized with the observer study, each observer received training that involved reading images of 5 training cases that were not included in the 41

used in the observer performance study. The observers read the training cases during about 15 min just before the observer performance test. The 5 training cases consisted of 2 patients with- and 3 patients without nodules. The observers were instructed to use the rating scale consistently and uniformly. Before training and test-taking, they were informed that the purpose of the experiment was to evaluate whether the DES technique did, or did not enhance the detection of pulmonary nodules on chest radiographs.

Statistical Analysis

We used ROC analysis to compare the radiologists' performance in detecting pulmonary nodules on chest images acquired without and with the DES technique. A binormal ROC curve was fitted to each radiologist's confidence rating data acquired under the 2 reading conditions by applying quasi-maximum likelihood estimation.(47) A computer program (ROCKIT; Charles E. Metz, University of Chicago, Chicago, IL) was used for obtaining binormal ROC curves from the ordinal-scale rating data.(47) The area under the best-fit ROC curve (AUC) plotted in unit squares was calculated for each fitted curve.

The statistical significance of the difference in AUC values between the ROC curves obtained without and with DES images was tested. The paired t-test was

performed with a statistical software package (SPSS, version 15.0; SPSS, Chicago, IL), and p values less than 0.05 were considered to indicate statistically significant differences.

RESULTS

For all pulmonary nodules the AUC values for all 8 observers were higher with than without DES images (Table 3). Analysis of the overall performance of the 8 observers in the detection of pulmonary nodules indicated that the mean AUC values for all observers increased from 0.62 ± 0.05 (without DES images) to 0.68 ± 0.05 (with DES images); the difference was statistically significant ($p = 0.02$) (Fig. 7).

For part-solid nodules (n=10) the mean AUC values without and with DES images were 0.61 ± 0.07 and 0.69 ± 0.05, respectively; the difference was significant ($p < 0.01$). For non-solid nodules (n=6), these values were 0.62 ± 0.10 and 0.61 ± 0.09, respectively and the difference was not significant ($p = 0.73$). For solid nodules (n=10), they were 0.75 ± 0.10 and 0.78 ± 0.08, respectively and the difference was not significant ($p = 0.23$) (Table 4).

For nodules with overlapping bone shadows, the mean AUC values obtained without and with DES images were 0.66 ± 0.05 and 0.72 ± 0.05, respectively; the difference was significant ($p = 0.03$) (Table 5). For nodules without overlapping it was not significant (AUC = 0.58 ± 0.06 vs. 0.62 ± 0.06, respectively; $p = 0.26$).

When a threshold value of 50 was used for the confidence level, the mean of sensitivity, specificity, accuracy, and positive- and negative predictive value (PPV,

NPV) for all observers was 47.6%, 72.5%, 56.7%, 75.6%, and 44.4%, respectively, for standard images. These values were 63.0%, 72.5%, 66.5%, 80.1%, and 53.7%, respectively, for DES images (Table 6). On DES images, sensitivity, accuracy, PPV, and NPV were significantly improved.

Representative cases are shown in Figs. 8 and 9.

DISCUSSION

As lung cancer is now the primary cause of cancer-related deaths world-wide, its early detection and diagnosis are important for improving the survival rate. Chest radiography is currently the most frequently used screening procedure for lung cancer because it is economical and easy to use. However, the false-negative rate for the detection of pulmonary nodules is as high as 19 - 73% for conventional chest radiography(11, 12) and the detection of small lung cancers is especially difficult.(14) This is partly due to masking of the soft tissue lesion by the superposition of bony structures.

The significant improvement in the detection of pulmonary nodules by adding DES images in our study is supported by earlier investigations that evaluated DES methods at computed radiography or on film-screen systems.(20) Few earlier studies evaluated the performance of radiologists in the detection of pulmonary nodules using DES techniques with the FPD system and the usefulness of this technique with the FPD system has not been established. Tagashira et al. (21) studied 50 patients with one or more nodules and 50 patients without nodules; they reported that the AUC values of 7 observers were significantly increased from 0.79 to 0.84 with DES images. Ricke et al.(22) evaluated 20 patients with a total of 59 pulmonary nodules; they found that the

DES technique significantly improved the sensitivity (from 33 to 42%), specificity (from 81 to 85%), and confidence in the detection of small pulmonary nodules. On the other hand, Rühl et al. (49) who studied 100 patients with a total of 149 pulmonary nodules (3 - 45 mm, median 11 mm) concluded that at FPD chest radiography, the DES technique did not significantly improve the detection of pulmonary nodules. However, they did not analyze the relationship between nodule detectability and nodule location in the thorax. Our study is the first to investigate the detection of pulmonary nodules from the view point of their relationship to bony structures and nodule density using the DES technique with the FPD radiography system.

Our results suggested that the combined evaluation of DES images and original chest radiographs significantly improved the diagnostic performance of observers in the detection of small pulmonary nodules. The discrepancy between our findings and those of Rühl et al. may be explained by the inclusion of nodules measuring 20 mm or more in their study, suggesting that the DES technique may not improve the detection of large pulmonary nodules that are fully detectable on standard images only. The DES technique may be useful for the detection of small pulmonary nodules.

Our subgroup analysis showed that the improvement in detectability was statistically significant for nodules overlapping with bones; it was not significant for

non-overlapping nodules, suggesting that the DES technique may reduce the incidence of false-negatives by detecting pulmonary nodules that overlap with bones. Our results also suggest that the DES technique yields a statistically significant improvement in the detectability of part-solid nodules, but not of non-solid- and solid nodules. The subtlety of non-solid nodules on images acquired without and with the DES technique may render their detection difficult. For the detection of non-solid nodules, low-dose helical CT may be superior although the clinical significance of the detection of non-solid pulmonary nodules remains controversial. The DES technique was also not helpful for the detection of small solid nodules. Ide et al. (17) who studied 77 consecutive lung cancer patients and 77 healthy subjects reported that at computed radiography the DES technique failed to improve the detection of non-solid- and solid nodules, however, it significantly improved the detection of part-solid nodules. We posit that conventional chest radiography may yield sufficient detection performance for solid nodules and that the DES technique is useful for detecting part-solid nodules. The malignancy rate of part-solid- is much higher than of non-solid- and solid nodules; it was reported as 63% and 18% for part-solid- and non-solid nodules, respectively.(4) The prognosis is worse for lung cancer patients showing part-solid nodules on CT scans than for those with non-solid nodules,(8) indicating that an increase in the proportion of solid components

within a GGO nodule reduces the tumor doubling time, increases the rate of lymph node metastases and vascular invasion as well as the risk of recurrence.(8) Therefore, it is clinically valuable to detect part-solid nodules using the DES technique.

FPD radiography systems are now widely used because of the rapid accessibility of images, improved image quality, and the possibility of reduced radiation exposure.(19) The FDP radiography system yields higher detective quantum efficiency (DQE) than computed radiography and film-screen radiography systems.(51) A higher DQE increases the ability to reveal objects in a noisy background(51) and reduces the radiation exposure without sacrificing image quality. At approximately 50% of the radiation dose, the quality of images obtained with the FPD radiography system was equal to that of computed- or film-screen radiography systems (52) and use of the double-exposure technique with FPD may not increase the radiation dose compared with these other systems.

Currently, single- and double-exposure dual-energy systems are available. In the former, DES images are obtained by exposing two storage phosphor plates separated by a copper filter. A disadvantage is the lower signal-to-noise ratio of the tissue-selective subtraction image. Also, theoretically, single-exposure systems cannot be used with FPD systems. On the other hand, in double-exposure systems (18), two

sequential radiographs are obtained at low- and high kV settings, respectively. The 200-millisecond delay between the two exposures may produce misregistration artifacts on the subtracted images. However, double-exposure systems produce DES images with a better signal-to-noise ratio than single-exposure systems.

The massive-training artificial neural network (MTANN) technique is a nonlinear pattern-recognition technique. Suzuki et al. (25) developed a novel method that achieves rib suppression on chest radiographs by using a MTANN. It can produce virtual DES images such as soft-tissue- and bone images from a single chest image obtained with a standard radiography system. As only software is needed, it requires no specialized equipment and no additional radiation dose exposure. We reported that the MTANN technique significantly improved the diagnostic performance of radiologists in the detection of pulmonary nodules(55). We are in the process of comparing the efficacy of the MTANN- and the dual energy technique.

There are several limitations in our study. First, it included a relatively small number of patients; the DES technique should be rigorously evaluated by large-scale clinical studies. Second, we did not evaluate the observers' experience with the interpretation of chest radiographs. Chest radiography is the most common diagnostic tool of radiologists regardless of their years of experience. It may be necessary to

clarify the usefulness of the DES technique among observers with different levels of experience. Finally, we did not evaluate the effect of this technique on the efficacy of detecting pulmonary nodules with calcification. However, calcified nodules can be detected easily on standard- or bone images. Since most calcified nodules are benign, their detection may be of lower clinical significance.

In conclusion, the DES technique combined with the inspection of original chest radiographs significantly improved diagnostic performance in the detection of small pulmonary nodules. We suggest that the DES technique with FPD chest radiography can reduce the number of lung nodules that are missed in routine clinical practice. We recommend that this technique be used routinely to obtain maximum benefits.

Chapter 4. Detection of Small Pulmonary Nodules on Chest Radiographs: Clinical Utility of Virtual Dual-energy Subtraction technique by using a Massive Training Artificial Neural Network (MTANN)

Contents

ABSTRACT

Purpose: The massive-training artificial neural network (MTANN) is a nonlinear pattern-recognition technique that can suppress rib opacity on chest radiographs while maintaining soft-tissue contrast. It is called virtual dual-energy subtraction (DES) technique. We investigated the effect of virtual DES technique by MTANN on the performance of radiologists in the detection of pulmonary nodules on chest radiographs.

Materials and Methods: used 60 chest radiographs; 30 contained solitary pulmonary nodules and the other 30 contained no nodules. They were selected with a stratified random-sampling scheme from the Japanese standard digital image database developed by the Japanese Society of Radiological Technology. The mean size of the 30 pulmonary nodules was 14.7 ± 4.1 mm (standard deviation). Receiver-operating-characteristic (ROC) analysis was used for evaluating observer performance in detecting pulmonary nodules on the chest radiographs first without and then with virtual DES images. Seven board-certified radiologists and 5 radiology residents participated in this observer study.

Results: For all 12 observers, the mean values of the area under the best-fit ROC curve (AUC) for images without and with virtual DES were 0.816 ± 0.077 and 0.843 ± 0.074, respectively; the difference was statistically significant ($p = 0.019$). The mean AUC

values for images without and with virtual DES were 0.848 ± 0.059 and 0.883 ± 0.050, respectively, for the 7 board-certified radiologists ($p = 0.011$) and 0.770 ± 0.081 and 0.788 ± 0.074, respectively, for the 5 radiology residents ($p = 0.310$).

Conclusion: The combined evaluation of virtual DES images by MTANN technique and original chest radiographs significantly improved the diagnostic performance of radiologists in the detection of pulmonary nodules.

INTRODUCTION

Although chest radiography is widely used for the detection of pulmonary nodules (56, 57), the false-negative rate for pulmonary nodules on chest radiographs is relatively high (11, 16, 26, 58), and radiography is inferior to low-dose computed tomography (CT) with respect to the detectability of small nodules (13). Failure to detect pulmonary nodules has been attributed to their size (14) and density (15) and obscuration by structures such as the ribs, clavicles, mediastinum, and pulmonary vessels (16). Obscuration by bony structures can be improved by dual-energy subtraction chest radiography (17, 18, 59) and temporal subtraction techniques (60). However, dual-energy subtraction chest radiography requires special equipment, and temporal subtraction requires a previous radiograph of the same patient.

The massive-training artificial neural network (MTANN) technique is a highly nonlinear pattern-recognition technique that has been applied in computer-aided diagnosis for the detection of pulmonary nodules on chest radiographs (61) and CT scans (23, 24), for the differentiation between malignant and benign pulmonary nodules on CT images (62), and for the detection of colonic polyps on CT colonography images (63, 64). Rib suppression on chest radiographs by MTANN is a novel method (25). It can produce soft-tissue and bone images from a single chest image obtained by standard

radiographic methods. These processed images are called virtual DES images. Virtual DES image by MTANN is designed to suppress rib opacity on chest radiographs. However, the effect of rib suppression on the diagnostic performance of radiologists in detecting pulmonary nodules has not been elucidated. The purpose of this study was to investigate the clinical efficacy of virtual DES images by MTANN in the detection of nodules smaller than 20 mm on chest radiographs.

MATERIALS AND METHODS

Our institutional review board (IRB) approved the use of the database; informed consent for the inclusion of cases was waived because the database is open to public inspection (48)**.** Our IRB also approved the participation of the radiologists in the observer performance test. Informed consent for the observer performance study was obtained from all participants.

Virtual DES Chest Radiography with MTANN

Virtual DES chest radiographs by MTANN involve 2 steps, a training- (Fig. 10A) and an application step (Fig. 10B). In the training step, we used chest images obtained with a dual-energy radiography system. Dual-energy subtraction, a technique for separating bones from soft tissue on chest radiographs, can produce two tissue-selective images, i.e. a "bone" image and a "soft-tissue" image. Standard chest radiographs are used as input images to the MTANN, and the corresponding dual-energy bone images are used as the teaching images. We did not directly use dual-energy soft-tissue images as the teaching images because the MTANN trained with dual-energy soft-tissue images produced results that were slightly inferior to the MTANN trained with dual-energy bone images (25). In this study, we trained the

MTANN with 5 sets of chest radiographs obtained with a dual-energy radiography system. The training set consisted of 4 cases with nodules (average size 20 mm) and one case without any nodules (a nodule-free "normal" case). The nodule cases were selected using the criterion that at least 20% of the nodule area overlapped with ribs. The normal case was a "mean" or "typical" case in terms of rib contrast and size. An important property of the MTANN is its ability to be trained with a very small number of cases. Massive training with a large number of training samples (sub-regions) obtained from a small number of cases through a training-sample enrichment process allows the MTANN to avoid the "over-fitting" problem of an artificial neural network (25). In the application step, when an input chest image obtained with a standard radiography system is entered to the trained MTANN, it outputs a "bone-image-like" image where bones are isolated and enhanced. Then a "soft-tissue-image-like" image where ribs are suppressed is produced by subtraction of the "bone-image-like" image from the input chest image. Thus, our MTANN technique can produce a "bone-image-like" image and a "soft-tissue-image-like" image from a single chest image obtained with a standard radiography system. We detailed the technical aspects of image processing by MTANN elsewhere (25).

Database

To investigate the effect of rib suppression on the performance of radiologists in detecting pulmonary nodules on radiographs, we conducted an observer performance study. One chest radiologist (K.A.) with 22 years of experience in reading chest radiographs, who did not participate in the observer performance study, selected 60 chest radiographs from the Japanese Standard Digital Image Database developed by the Japanese Society of Radiological Technology (48). He then selected 30 patients with only one malignant solid pulmonary nodule not larger than 20 mm without calcification and 30 patients without nodules. Confirmation of the presence or absence of a pulmonary nodule was done by CT examination, the nodule diagnosis was on histologic and cytologic examination. All patients had no other pleural and parenchymal abnormalities such as consolidation, ground glass opacity and pulmonary fibrosis on CT.

The 60 patients were 36 men and 24 women ranging in age from 27 to 89 years (mean age 59.4 years); the mean age of patients with and without pulmonary nodules was 60.4 years (range 39 - 78 years) and 58.3 years (range 27 - 89 years), respectively. According to the two-tailed Student *t*-test, there was no significant difference in the age of patients with and without pulmonary nodules ($p = 0.555$). Of the 30 patients with pulmonary nodules, 27 were diagnosed with primary lung cancer and 3 with solitary pulmonary metastasis. The mean size of the pulmonary nodules was 14.7 ± 4.1 mm (standard deviation).

Observer Performance Study

We used a sequential-test method for receiver-operating-characteristic (ROC) analysis to evaluate the diagnostic performance of radiologists in detecting pulmonary nodules on chest images without and with virtual DES. The observers were 7 board-certified radiologists with 5 - 16 years of experience (mean 8.4 years) and 5 radiology residents with 2 - 3 years of experience (mean 2.4 years). All board-certified radiologists specialized in body imaging and read chest radiographs on a regular basis. They were allowed to change the level and width of the window on the monitor; reading time was not limited.

All observers used a continuous rating scale and a line-marking method to rate their confidence level by placing marks on a 7-cm-long line on a recording form. The left end of the line indicated complete confidence that the chest radiographs without/with rib suppression showed no nodule; the right end indicated complete confidence that they revealed a nodule. Intermediate levels of confidence were indicated by the position of the marks between the 2 line termini, where marks close to the right and left end indicated a greater and lesser degree of confidence, respectively. One author (Y.F.) then measured the distance between the left end of the line and the mark and converted the distance to an ordinal confidence rating ranging from 0 - 100.

A continuous rating scale containing a pair of horizontal lines was used in the sequential test. Observers first recorded their rating of chest radiographs without rib suppression on the upper line. Subsequently, they recorded their rating of rib-suppressed images on the lower line. They entered their results for each case on a record form.

First, to become familiarized with the observer study, each observer received training that involved reading images of 5 training cases that were not included in the 60 cases used in the observer performance study. The observers read the training cases during about 15 minutes just before the observer performance test. The 5 training cases consisted of 2 cases with- and 3 without a nodule. The observers were instructed to use the rating scale consistently and uniformly.

Before training and test-taking, the participating radiologists were informed that the purpose of the experiment was to evaluate whether virtual DES images did, or did not enhance the detection of pulmonary nodules on chest radiographs.

Statistical Analysis

We used ROC analysis to compare the radiologists' performance in detecting pulmonary nodules on chest images without and with rib suppression. A binormal ROC

curve was fitted to each radiologist's confidence rating data acquired under the 2 reading conditions by use of quasi-maximum likelihood estimation (47). A computer program (ROCKIT; Charles E. Metz, University of Chicago, Chicago, IL) was used for obtaining binormal ROC curves from the ordinal-scale rating data (47). The area under the best-fit ROC curve (AUC) plotted in the unit square was calculated for each fitted curve.

The statistical significance of the difference in AUC values between the ROC curves obtained without and with virtual DES was tested for all board-certified radiologists and all radiology residents. The paired t-test was performed with a statistical software package (SPSS, version 15.0; SPSS, Chicago, IL), and p values less than 0.05 were considered to indicate a significant difference.

RESULTS

Of the 30 nodules, 6 (20.0%) completely overlapped with bone shadows of the ribs and/or a clavicle, 17 (56.7%) overlapped partially, and 7 (23.3%) did not overlap on chest radiographs.

The AUC values for all 12 observers were significantly higher with than without virtual DES (Table 7). Analysis of the overall performance of the 12 observers in the detection of pulmonary nodules (Fig. 11) indicated that the mean AUC values for all observers increased from 0.816 ± 0.077 (without virtual DES) to 0.843 ± 0.074 (with virtual DES); the difference was statistically significant ($p = 0.019$).

For the board-certified radiologists, the mean AUC values obtained without and with virtual DES were 0.848 ± 0.059 and 0.883 ± 0.050, respectively; these values were significantly different ($p = 0.011$) (Table 7, Fig. 12).

On the other hand, for the group of residents, the mean AUC values obtained without and with virtual DES were 0.770 ± 0.081 and 0.788 ± 0.069, respectively; the difference was not significant ($p = 0.310$) (Table 7, Fig. 13).

Representative cases are shown in Figs. 14 and 15.

DISCUSSION

MTANN is a nonlinear pattern-recognition technique that can suppress rib opacity on chest radiographs while maintaining soft-tissue contrast (25). We found that the combined evaluation of virtual DES images and original chest radiographs statistically significantly improved the diagnostic performance of board-certified radiologists in the detection of pulmonary nodules. In the current study, we selected nodules of a size less than 20 mm for the observer performance study. Our results suggested that virtual DES images by MTANN technique helps skilled radiologists in the detection of relatively small pulmonary nodules. On the other hand, in the group of radiology residents the improvement was not statistically significant. Because the radiology residents had only 2 to 3 years of experience, we presume that their interpretation skills were still developing; thus, variations in their diagnostic performance were relatively large. A statistically significant difference was not obtained in this group, probably because the statistical power was not strong enough due to such variations and we expect that a statistically significant difference would be obtained by increasing the number of observers or cases. In addition, the lower performance of the residents may relate to the type of error seen in the 2 observer groups. Experience with the interpretation of chest radiographs appeared to decrease

the false-positive rate but had less effect on reducing the false-negative rate, suggesting that readers with less experience would report more false-positive findings. Virtual DES lowers the false-negative- rather than the false-positive rate by enhancing the visualization of pulmonary nodules on soft-tissue-selective images. Therefore, virtual DES may be more effective in experienced board-certified radiologists than in less experienced radiology residents.

Chest radiography remains the most widely used imaging technique for the detection of chest diseases because of its low cost, simplicity, and low radiation dose. However, the false-negative rate for the detection of pulmonary nodules is relatively high; it was reported to range from 12 to 90% (11, 16, 26, 58). Lung cancers missed on chest radiographs reportedly share the following characteristics: they are visually subtle (14, 65), small (14), and located predominantly in the upper lobes (65). Importantly, 82 - 95% of the lung cancers missed by radiologists were partly obscured by overlying bones such as the ribs and/or a clavicle (16). This suggests that the suppression of these structures on chest radiographs may improve the detection accuracy. Actually, in our database, more than 75% of the nodules overlapped with ribs or clavicles. However, in the present study we did not compare the detection of nodules overlaid by bone

structures and the identification of nodules without bone overlay on rib suppressed images, because our series contained an insufficient number of nodules in each category.

Because the MTANN technique suppresses bone attenuation, areas of calcification in a nodule or calcified nodules may become indistinct on virtual DES images. A solution consists of a careful comparison of the original chest radiographs and the virtual DES images. Bone-selective imaging is a counterpart of rib suppressed imaging, and the addition of bone-selective images in the interpretation of chest radiographs may help to avoid missing nodules with calcifications.

The dual-energy subtraction technique for chest radiography can remove overlying bone structures to create soft-tissue-selective images (17, 18) and it enhances the visualization of pulmonary nodules overlaid by bones. Single- (17, 59) and double-exposure dual energy systems (18, 66) are currently available. Chest radiographs acquired with either system yield improved nodule visibility and have been shown to improve the detection accuracy (66, 67). Despite this advantage, radiography systems with dual-energy subtraction are used in few institutions because they require specialized equipment and because they may deliver a higher radiation dose than standard chest radiography (66). A major advantage of the MTANN technique over the dual-energy subtraction technique is that the former requires no specialized equipment

for generating dual-energy x-ray exposures because only software is needed. In addition, there is no increase in the radiation dose exposure because chest radiographs acquired with standard radiography systems are used (25).

In the temporal subtraction technique, a previously acquired- is subtracted from a current chest radiograph; this enhances the detection of changes that occurred during the interval (60). This technique effectively improved the performance of radiologists in the detection of pulmonary nodules (60) and is commercially available in Japan (68). However, it cannot be applied unless a previously acquired radiograph is available. In contrast, the MTANN technique can be used even if there is only one radiograph.

Several previous studies have reported that computer-aided diagnosis (CAD) systems with automated detection methods for lung nodules significantly improved the diagnostic accuracy for identifying pulmonary nodules on chest radiographs (69). Kakeda et al. (69), who applied CAD to chest radiographs of 45 patients with solitary pulmonary nodules measuring up to 25 mm in diameter, reported a sensitivity of 73% with 3.15 false-positive findings per image. They also found that 75% of false-positive findings were attributable to the presence of normal anatomic structures such as ribs and clavicles. In a more recent study, Li et al. (65) also applied a commercially available CAD tool to chest radiographs; they reported an average of 5.9 false-positive findings

per image and a 96% false-positive rate attributable to interference by anatomic structures. We propose that adding CAD detection to the MTANN technique may yield a higher sensitivity and fewer false-positive results than the use of either method alone. Studies are underway in our laboratory to investigate whether the combined use of CAD and MTANN techniques improves the detection of pulmonary nodules.

According to Ide et al. (17) who evaluated the detectability of lung cancer on chest radiographs subjected to single-exposure dual-energy subtraction, this method did not improve the detectability of non-solid nodules also known as "pure ground glass opacity" nodules. These nodules have a high probability of being lung cancer, especially bronchioloalveolar carcinoma (BAC) (10). Because the MTANN technique is based on a dual-energy subtraction technique, it also may not be useful for the detection of non-solid nodules. Studies are underway to examine the detectability of non-solid nodules with the MTANN technique.

There are several limitations in our study. First, we extracted chest radiographs from the Japanese Standard Digital Image Database (48). Because images in that database were derived from many institutions and many were digitized from conventional chest radiographs, the image quality varied from image to image. We are in the process of investigating the nodule detectability on radiographs obtained with

recent digital radiography systems. Second, central hilar nodules were not included because their detection remains more difficult than that of peripheral nodules on standard chest radiographs. In future studies we will address the effect of our technique on the detection of central perihilar nodules. Third, we did not compare the effect of the MTANN technique with that of the clinically-used dual-energy subtraction technique. The efficacy of the MTANN- and the dual-energy technique should be tested and compared.

In conclusion, in combination, the MTANN technique for virtual DES and the inspection of original chest radiographs significantly improved the diagnostic performance of radiologists in the detection of small pulmonary nodules. We strongly suggest that this technique has the potential to reduce the number of lung nodules that are missed in routine clinical practice.

Chapter 5. Computer-Aided Volumetry of Pulmonary Nodules Showing Ground Glass Opacity at Multidetector CT: Feasibility Study Using Phantom and Clinical Cases

Contents

ABSTRACT

Purpose: To investigate the accuracy and reproducibility of results acquired with computer-aided volumetry (CAV) software for pulmonary nodules with ground glass opacity (GGO).

Materials and Methods: First we investigated the accuracy of CAV software by scanning a chest phantom that included simulated 3-, 5-, 8-, 10-, and 12-mm diameter GGO nodules with CT numbers of -800, -630, and -450 Hounsfield Units (HU) by thin-section helical computed tomography (CT). Three radiologists measured their volume and calculated the relative volume measurement error (RVME) defined as: (measured nodule volume minus assumed nodule volume) / assumed nodule volume x 100. Next, 2 radiologists performed 2 independent measurements of 59 human nodules. Intra- and inter-observer agreement was evaluated using Bland-Altman methods.

Results: RVME for simulated GGO nodules measuring 3- and $\geq$5 mm in diameter ranged from 51.1 - 85.2% and from -4.1 - 7.1%, respectively. In our clinical study, for intra-observer agreement, the 95% limits of agreement were -14.9 - 13.7% and -16.6 to 15.7% for observer A and B, respectively; for inter-observer agreement these values were -16.3 to 23.7% in nodules $\geq$ 8 mm in diameter.

Conclusion: With CAV for GGO nodules, RVME was small for nodules with a diameter of 5 mm or more. Intra- and inter-observer agreement was relatively high in nodules ≥ 8 mm in diameter.

INTRODUCTION

Among pulmonary nodules detected at lung cancer screening with low radiation-dose helical computed tomography (low-dose CT), 19 - 38% manifested focal ground-glass opacity (GGO) (4, 5). Nodules showing GGO (GGO nodules) may be attributable to focal inflammation, focal interstitial fibrosis (9), atypical adenomatous hyperplasia (AAH) (6, 7), bronchioloalveolar carcinoma (BAC) (6, 7, 10), or adenocarcinoma (6, 7, 10). While the differentiation of BAC and adenocarcinoma, which are malignant, from other diseases is important, it can be difficult on a single CT scan (32). Many inflammatory lesions resolve spontaneously or with antibiotic treatment (6), the size of GGO nodules attributable to BAC or adenocarcinoma, on the other hand, gradually increases (34, 70, 71). Focal interstitial fibrosis and AAH with pure GGO remain stable in size for months or years (9). Therefore, monitoring the nodule size for several months by high-resolution CT (HRCT) may help in obtaining a differential diagnosis. As the doubling time of BAC is very long (457 - 813 days on average) (33, 34), the subjective evaluation of the growth rate by radiologists is unreliable.

With multidetector CT (MDCT) scanners it is possible to scan a wide range, including areas containing pulmonary nodules, at a detector collimation of 0.500 - 0.625 mm under one breath-hold. This facilitates the three-dimensional evaluation of

pulmonary nodules. Previous studies assessed computer-aided volumetry (CAV) of pulmonary nodules using volumetric data obtained at MDCT (36-40); it reportedly yielded sufficiently high accuracy and reproducibility (37). However, most of the software used in these studies only evaluated solid pulmonary nodules and CAV of GGO nodules was reportedly difficult (40). We developed CAV software that can be used to measure not only the volume of solid- but also of GGO nodules. The purpose of this study was to investigate the accuracy and reproducibility of results obtained on GGO nodules with our CAV software.

MATERIALS AND METHODS

Computerized Volumetry of Pulmonary Nodules (Figs. 16, 17)

For volumetry of GGO nodules we used prototype software developed for the picture archiving and communication system (3D-analysis software for HOPE/DrABLE-EX, Fujitsu, Tokyo, Japan).

First, with a single click and drag maneuver, the radiologists roughly specified the region of interest (ROI) to include the target nodule on axial-, sagittal-, or coronal images. Then, the CAV software automatically segmented the nodule, using an improvement of our previously reported method (72). For the robust extraction of nodules from the lung parenchyma we adopted an adaptive threshold CT number based on the CT number of the target nodule and the lung parenchymal background. When the difference in the CT number of the target nodule and the lung parenchyma was relatively large, the CAV software chose the threshold CT number that approximated the mean value of the nodule and lung parenchyma. When the difference in the CT number of the nodule and lung parenchyma was relatively small, the threshold CT number approached non-linearly the average CT number of the lung parenchyma.

In the pulmonary nodule extraction process, structures connected to the nodule, e.g. vessels and bronchi, were roughly eliminated by the CAV software. Occasionally,

some structures remained around the nodules, or part of the nodule was not included in the extracted area. In such cases, the radiologists who performed volumetry edited the segmented area to include the nodule by changing the threshold level that separated the density of the nodule from the base density of the lung and by using the mathematical morphology "erosion/dilation" technique (73). "Erosion/dilation" is a useful digital image-processing technique to eliminate structures tangent with the nodule such as vessels and thoracic walls. Since this modification was based on the subjective judgment of each radiologist it resulted in volumetric intra- and inter-observer discrepancies. Although the prototype software permits free-hand shape-editing, the radiologists were not allowed to use this technique. Rather, they could use only the above image processing technique to assure the reproducibility of segmentation. The judgment of successful segmentation was based on the observers' visual assessment on axial CT images as well as sagittal- and coronal multiplanar reconstructed images. Finally, the volume of the segmented area was automatically measured by the software. The average calculation time required by the CAV system to analyze each nodule was less than 1.0 sec.

A computer workstation (FMV-H8230; Fujitsu, Tokyo, Japan) with dual 2.16-GHz processors (Core Duo; Intel, Santa Clara, Calif.) was used in this study.

Phantom Study

Chest phantom

To evaluate the accuracy of our CAV software we conducted a phantom study using simulated GGO nodules. Ours was a commercially available chest CT phantom and simulated GGO nodules (multi-purpose chest phantom N1, Kyoto Kagaku Co., Kyoto, Japan). In this phantom, simulated soft tissues such as pulmonary vessels, the chest wall, heart, diaphragm, and liver consist of polyurethane resin composites; simulated bone of an epoxide resin. The space between the pulmonary vessels, heart, and chest wall was filled with air. The chest wall can be removed from the other structures such as the simulated heart, pulmonary vessels, diaphragm, and liver. Simulated nodules were spheres made of urethane foam resin. The diameter of the simulated nodules was 3- (estimated error (ER) 20%), 5- (ER 6%), 8- (ER 4%), 10- (ER 3%), and 12 mm (ER 2.5%) and the CT number of the simulated nodules was -800, -630, and -450 HU.

In CT scans of the chest phantom harboring simulated nodules we placed a nodule on bifurcations of the pulmonary vessels in the right lung at the level of the carina.

CT scanning of the phantom study

CT imaging was with a 64-detector CT scanner (Brilliance-64, Philips Medical Systems, Cleveland, OH). The scanning parameters were: detector row width, 0.625 mm; helical pitch, 0.673; section thickness and section interval, 1 mm; rotation time, 0.5 sec; tube voltage, 120 kVp; tube current, 250 mAs.

Volumetry of simulated GGO nodules (Fig. 18)

Three radiologists with 4-, 10-, and 15 years of experience with chest CT subjected each simulated nodule to 3 volumetric measurements; the resultant values were averaged for each nodule. Each radiologist performed 3 volumetry sessions at 1-week or longer intervals.

We calculated the relative volume measurement error (RVME) for each nodule to evaluate the accuracy of our volumetry software. RVME was defined as: (measured nodule volume minus assumed nodule volume) / assumed nodule volume x 100. The assumed nodule volume is the calculated volume based on the diameter of a simulated spherical nodule. We calculated the mean RVME for each nodule by averaging the values reported by the 3 radiologists.

Clinical Study

To evaluate the reproducibility of our CAV software, represented by intra-and inter-observer agreement, we conducted a study using data derived from patients. We also investigated whether the necessity for editing the segment area or the edit time depends on the morphology of the GGO nodule. All patients who underwent CT examination at our institute gave their prior informed consent for the use of their CT images in future retrospective studies. Our institutional review board approved the use of the CT database. Informed specific study-related consent was waived.

Nodule selection

One chest radiologist (K.A.) with 21 years of chest CT experience reviewed the records of 211 consecutive patients suspected of harboring pulmonary nodules. The patients underwent thin-section helical CT of the chest at our institute during a 36-month period from January 2004 to December 2006**.** The radiologist, who did not participate in the volume measurement study, selected all patients who satisfied the following criteria: (a) GGO nodules that did not exceed 20 mm in the x-y (or transverse) plane, (b) a histological diagnosis based on findings at thoracic surgery or at CT-guided transcutaneous or bronchoscopic transbronchial biopsy, and (c) absence of

consolidation due either to the presence of organizing tissues after pneumonia or associated with idiopathic pulmonary fibrosis around the nodules. The latter criterion was applied because we found it difficult to define the nodule boundary. On the basis of these criteria, 59 nodules from 59 patients (18 men and 41 women ranging in age from 47-79 years [mean 65.4 years]) with 40 malignant- (BAC, n=38; adenocarcinoma, n=2) and 19 benign nodules (AAH, n=17; amyloidosis, n=1; goblet cell metaplasia, n=1) were identified.

The mean size in the x-y (or transverse) plane of the 59 nodules was 12.8 ± 4.7 mm (standard deviation) (range 4.0 - 20.0 mm); their mean density was -539.6 ± 96.3 HU (range -744 to -339 HU).

CT scanning

CT scans were obtained with a 4-detector scanner (LightSpeed QX/I, GE Medical Systems, Milwaukee, WI). After routine helical scanning of the whole thorax, thin-slice helical scans whose scan range of about 5 cm included the pulmonary nodules were obtained. The scan parameters for routine helical scans were: detector collimation, 4 x 2.5-mm; helical pitch (beam pitch), 1.5; slice thickness and interval, 5.0 mm; 0.8-sec rotation time, 120 kVp, and 250 mA. The scan parameters for thin-slice helical

scans were: detector row width, 4 x 1.25 mm; helical pitch, 0.75; slice thickness and interval, 1.25 mm; 0.8-sec rotation time, 120 kVp, and 160-200 mA. The reconstruction algorithm for thin-slice helical scans was "bone plus". Contrast enhancement was not used in any of the 59 nodules. Because the LightSpeed scanner was replaced with the Brilliance-64 instrument in January 2008, we used the LightSpeed scanner for the clinical- and the Brilliance-64 scanner for the phantom study.

Volumetry of GGO nodules (Fig. 19)

Two of the 3 radiologists who participated in the phantom study performed volumetry in the clinical study. They had 4 and 15 years of experience with chest CT, respectively.

Each radiologist performed 2 volumetry sessions at one month or longer intervals. Between the 2 sessions, the order of case presentation was changed to minimize the "learning effect". Before the first session, each radiologist underwent a training session that involved volumetry in 3 training cases to become familiar with the editing tools of the CAV software because the shape of some of the human nodules was more complex than the phantom nodules. The 3 training cases were different from the 59 patients included in our clinical study.

When the GGO nodules were not accurately extracted by the CAV software alone, the 2 radiologists subjectively edited the segmented area using techniques based on mathematical morphology. If an edit was required, we measured the edit time.

To investigate whether morphological nodule characteristics affected the necessity for editing, the chest radiologist (K.A.) who selected the 59 patients classified the GGO nodules into 3 categories based on the nature of the nodule margin. Thus, category 1 nodules (n=22) exhibited a well-defined, smooth margin; category 2 nodules (n=24) had a well-defined, irregular margin; and category 3 nodules (n=13) manifested an ill-defined margin. He also classified the 59 nodules as part-solid (mixed GGO, n=31) and non-solid (pure GGO, n=28) according to their internal density (4) . The readers recorded whether the GGO nodules were (n=12) or were not (n=47) adjacent to the pleural surface.

Statistical Analysis

All numerical values are reported as the mean ± standard deviation (SD). We used Bland-Altman (74) analysis to determine intra-and inter-observer agreement in the clinical study. We also assessed the intra-and inter-observer agreement for nodules < 8 mm and ≥ 8mm in diameter to evaluate the influence of the nodule size, and for nodules

that were or were not subjected to editing to reveal whether and how editing affected the reproducibility of CAV. On the Bland-Altman plots we used the percent difference of the average of 2 measured values (relative difference) rather than the absolute difference because there was an increase in the variability of the differences as the magnitude of the measurements increased.

To determine the frequency of necessary edits (the edit rate) required by the morphology of the GGO nodules we performed the chi-square test. To assess the effect of nodule morphology on the edit time we used Student's t-test or Tukey multiple comparisons. Variables with a *p* value less than 0.05 were considered to indicate statistically significant differences.

Statistical analyses were performed with a software program (MedCalc Software; MedCalc, Mariakerke, Belgium; and SPSS, version 15.0; SPSS, Chicago, IL).

RESULT

Phantom Study (Fig. 20)

The average of the RVME for nodules with a diameter of 3 mm and a CT number of -800, -630, and -450 HU was 61.5, 51.1, and 85.2%, respectively. For all nodules measuring 5 mm or more, the average ranged from -4.1 to 7.1%. The standard deviation of the RVME for 3-mm diameter nodules with CT numbers of -800, -630, and -450 HU was 4.0, 9.8, and 13.5%, respectively; it was 0.9 to 6.2% for all nodules with a diameter of 5 mm or more.

Clinical Study

We were able to perform CAV in all 59 human GGO nodules.

Intra-observer agreement (Fig. 21 and Tables 8, 9)

The mean intra-observer agreement for all GGO nodules was 1.6 ± 11.1% (95% confidence interval [CI]: -1.3 to 4.5%) and 1.8 ± 11.5% (95% CI: -1.2 to 4.8%) for observer A and B, respectively. For intra-observer agreement a lower limit of agreement of -20.2% (95% CI: -25.2 to -15.2%) and -20.7% (95% CI: -25.8 to -15.5%) was obtained for observer A and B, respectively. The upper limits of agreement were

23.3% (95% CI: 18.4 to 28.3%) and 24.4% (95% CI: 19.2 to 29.5%) for observer A and B, respectively. The mean intra-observer agreement and lower and upper limit of agreement were 7.4 ± 16.7% (95% CI: -1.5 to 16.3%), -25.3% (95% CI: -40.8 to -9.8%), 40.1% (95% CI: 24.6 to 55.6%), respectively, in nodules < 8 mm in diameter, and -0.6 ± 7.3% (95% CI: -2.8 to 1.7%), -14.9% (95% CI: -18.7 to -11.0%), 13.7% (95% CI: 9.8 to 17.6%), respectively, in nodules ≥ 8 mm in diameter for observer A. Those for observer B were 7.0 ± 15.8% (95% CI: -0.9 to 14.9%), -24.0% (95% CI: -37.7 to -10.3%), 38.0% (95% CI: 24.3 to 51.7%), respectively, in nodules < 8 mm in diameter, and -0.4 ± 8.2% (95% CI: -3.0 to 2.2%), -16.6% (95% CI: -21.1 to -12.1%), 15.7% (95% CI: 11.2 to 20.2%), respectively, in nodules ≥ 8mm in diameter. The mean intra-observer agreement and lower and upper limit of agreement were 2.3 ± 8.1% (95% CI: -1.2 to 5.9%), -13.6% (95% CI: -19.7 to -7.5%), 18.3% (95% CI: 12.2 to 24.4%), respectively, in nodules without editing, and 1.1 ± 12.7% (95% CI: -3.2 to 5.4%), -23.9% (95% CI: -31.3 to -16.4%), 26.0% (95% CI: 18.6 to 33.5%), respectively, in nodules with editing for observer A. Those for observer B were 0.1 ± 8.8% (95% CI: -3.1 to 3.3%), -17.1% (95% CI: -22.6 to -11.5%), 17.3% (95% CI: 11.7 to 22.8%), respectively, in nodules without editing, and 3.8 ± 13.8% (95% CI: -1.6 to 9.1%), -23.3% (95% CI: -32.6 to -14.1%), 30.8% (95% CI: 21.6 to 40.1%), respectively, in nodules with editing.

Inter-observer agreement (Fig. 22 and Tables 8, 9)

The mean inter-observer agreement for all GGO nodules was 3.2 ± 11.8% (95% CI: 0.1 to 6.3%). For inter-observer agreement a lower limit of agreement of -20.0% (95% CI: -25.3 to -14.7%) and an upper limit of agreement of 26.4% (95% CI: 21.1 to 31.7%) was recorded. The mean inter-observer agreement and lower and upper limit of agreement were 1.9 ± 15.4% (95% CI: -6.0 to 9.8%), -28.4% (95% CI: -42.2 to -14.5%), 32.2% (95% CI: 18.3 to 46.0%), respectively, in nodules < 8 mm in diameter, and 3.7 ± 10.2% (95% CI: 0.5 to 6.9%), -16.3% (95% CI: -21.8 to -10.8%), 23.7% (95% CI: 18.3 to 29.2%), respectively, in nodules ≥ 8mm in diameter. The mean inter-observer agreement and lower and upper limit of agreement were 3.5 ± 8.5% (95% CI: -0.3 to 7.2%), -13.1% (95% CI: -19.6 to -6.6%), 20.1% (95% CI: 13.6 to 26.6%), respectively, in nodules without editing, and 2.7 ± 13.0% (95% CI: -2.8 to 8.2%), -22.9% (95% CI: -32.4 to -13.4%), 28.2% (95% CI: 18.7 to 37.8%), respectively, in nodules with editing.

Edit rate and edit time (Tables 10, 11)

Observers A and B applied editing to 36 (61.0%) and 28 nodules (47.5%), respectively. As shown in Table 2, the edit rate was 54.5% and 31.8% for category 1-, 50.0% and 45.8% for category 2-, and 92.3% and 76.9% for category 3 GGO nodules

for observer A and B, respectively. There was a statistically significant difference in the edit rate of the 3 categories (observer A: $p = 0.03$, observer B: $p = 0.04$; chi square test). For both observers, the edit rate was 53.6% for non-solid nodules; it was 67.7% (observer A) and 41.9% (observer B) for part-solid GGO nodules. There was no statistically significant difference in the edit rate between non-solid- and part-solid GGO nodules ($p = 0.40$, observer A; $p = 0.53$, observer B). The edit rate for nodules that were adjacent to the pleura was significantly higher than for nodules that were not (91.7% vs. 53.2%, $p = 0.04$, observer A; 83.3% vs. 38.3%, $p = 0.03$, observer B).

The mean edit time was 24.1 ± 10.0 and 35.3 ± 12.3 sec for observer A and B, respectively (Table 3). The mean edit time for category 3 nodules was statistically significantly longer than for category 1 and 2 ($p < 0.01$ and < 0.01 for observer A; $p < 0.01$ and < 0.01 for observer B). The difference in the edit time for both observers was not statistically significant for category 1 and 2 nodules ($p = 0.92$ and 0.61). For observer B the edit time for part-solid GGO was statistically significantly longer than for non-solid GGO ($p < 0.01$); for observer A there was no statistically significant difference ($p = 0.36$). There was no statistically significant difference in the edit time between nodules that were, or were not, adjacent to pleura (observer A: $p = 0.77$; observer B: $p = 0.35$).

DISCUSSION

Regardless of difficulty in the volumetry of GGO nodules, all GGO nodules in this study were successfully segmented with our CAV software. Our results were better than previously reported volumetry measurements of human solid nodules (75, 76); it was reported that 71 - 97% of nodules were successfully segmented.

Factors that affect CAV of pulmonary nodules are the algorithm of the volumetry software, the threshold values for nodule extraction, the nodule size and attenuation, the tube current, and the image reconstruction kernel (36, 40, 77, 78). The presence of pulmonary vessels, bronchi, or chest wall connecting with the nodule may require manual modification of nodule extraction and arbitrary manipulation by radiologists may be a factor affecting the accuracy of volumetry. de Hoop et al.(75) assessed the volumetry of solid nodules with 6 currently available semi-automated software packages; they found manual modification of nodule extraction improved the accuracy of volumetry without significantly affecting reproducibility.

In solid pulmonary nodules, CAV produced only minimal errors and yielded very high measurement reproducibility (38, 78-80). Kostis et al. (78) who evaluated CAV in human solid nodules reported that the overall standard measurement error in nodules 2 - 5 mm, 5 - 8 mm, and 8 - 10 mm was 18.5%, 10.6%, and 7.47%, respectively.

A study conducted by Das et al. (37) revealed that accurate volumetry was possible in simulated solid nodules with a diameter of 5 mm or larger and volumes greater than 66 mm^3. On the other hand, volumetry of GGO nodules is difficult because the attenuation difference between these nodules and normal lung parenchyma is small and other structures in the lung are visualized in the same attenuation range. Earlier volumetric studies on chest phantoms demonstrated a significantly higher volume measurement error for GGO- than solid nodules(40). In our phantom study we obtained accurate volumetric measurements for nodules larger than 5 mm in diameter; the mean measurement error for nodules 5 mm or larger was 2.35% (range -4.14 to 7.13%); our accuracy rate was equivalent to previous studies on solid nodules. In the clinical setting, however, pulmonary nodules are not necessarily round, their margins are not always clear, and they are frequently in contact with adjacent structures such as the chest wall, pulmonary vessels, and bronchi. In addition, artifacts due to respiratory and cardiac motion are often present on clinical images. On the other hand, the simulated nodules in the chest phantom were round and had clear margins and there were no motion artifacts. Therefore, volumetric measurement may be less accurate in clinical- than phantom studies. We observed measurement variability even in the volumetry of simulated GGO nodules in this study. As the software defines the optimal threshold CT

number within each ROI specified by the radiologist, the threshold value varies with the size or site of the ROI. In addition, in some instances editing was required even for simulated GGO nodules. This may account for the measurement variability encountered in our phantom study.

In our clinical investigation, intra- and inter-observer agreement was relatively high in nodules ≥ 8mm in diameter. For intra-observer agreement, the 95% limits of agreement were -14.9 to 13.7% and -16.6 to 15.7 % for observer A and B, respectively, and for inter-observer agreement they were -16.3 to 23.7 % in nodules ≥ 8 mm in diameter. Wormanns et al. (38) obtained excellent results; their intra- and inter-observer agreement was -3.9 to 5.7% and -5.5 to 6.6%, respectively, for automated volumetry of clinical solid nodules. Thus, volumetry measurements of GGO nodules exhibited larger variability than of solid nodules. Taking into account our intra- and inter-observer agreement results, the threshold for identifying an increase in the measured volume of a GGO nodule is a 30% increase, equivalent to a 9% increase in the nodule diameter. For example, if a nodule with a 10 mm diameter on the baseline scan has increased by 1 mm on the follow-up scan, it may be difficult to detect this change by visual assessment. Therefore, we concluded that our results on the intra- and

inter-observer agreement were clinically acceptable for the early detection of GGO nodule growth.

We also found that intra- and inter-observer agreement was slightly better for nodules without- than with editing. However, we posit that this difference is minimal and acceptable for clinical practice. Although editing based on the subjective judgment of each observer resulted in a few intra- and inter-observer discrepancies, appropriate editing may be indispensable for accurate volumetry of GGO nodules.

In our clinical study, edits were required in about half of the GGO nodules. For category 3 GGO nodules, the edit rate was significantly higher and the edit time was significantly longer than for category 1 and 2 nodules. However, for both observers, the mean edit time for category 3 GGO nodules was shorter than 60 sec. Furthermore, there was no statistically significant difference in the edit time between nodules that were, or were not, adjacent to the pleura and the mean edit time was approximately 30 sec for both pleura-adjacent and -non-adjacent GGO nodules, although the edit rate for nodules adjacent to the pleura was higher. Considering the acceptable length of the edit time, we suggest that our CAV software is a practical tool for the measurement of GGO nodules in humans.

One of the most important indicators of malignancy is the growth rate of the nodules; this is commonly expressed as the volume doubling time (38). CAV assessment of pulmonary nodules to estimate their growth rate has gained wider acceptance (36, 39). The reported average BAC doubling time, calculated from the maximal tumor diameter with the Schwartz equation (42) is 457 - 813 days (33, 34). In the interpretation of these data, it must be remembered that the reported doubling times for GGO nodules were not based on 3D data. According to Marten et al. (79), the relative mean error of volumetric measurements of pulmonary nodules was significantly higher for manual unidimensional measurements; intra- and inter-observer agreement for manual unidimensional measurements was lower than with automated volumetry. Yankelevitz et al. (36) also found that computer-aided 3D volumetry offered obvious advantages over conventional bilinear 2D measurements for the accurate representation of the nodular volume. We recommend that the doubling time of pulmonary nodules be re-evaluated with highly accurate CAV software that features high reproducibility.

There are some inherent problems with volumetry of GGO nodules. First, the size of some BACs with GGO decreased in the course of follow-up due to the collapse of alveolar spaces, fibrosis, or the severe narrowing of the alveolar space (71). Therefore, even if GGO nodules decrease in size, malignancy cannot be ruled out.

Second, since the cells of BAC with focal GGO manifest replacement growth without a marked reduction in the alveolar lumen and without marked thickening of alveolar septa (8), BAC with focal GGO is a relatively "soft tumor". Therefore, the volume of BAC with focal GGO may be affected by the lung volume. Furthermore, CAV of GGO nodules seems to be affected by a change in the attenuation of the lung surrounding parenchyma with inspiratory level. In volumetry of focal GGO, the lung volume may have to be taken into consideration.

There is ongoing debate concerning the management of patients with GGO nodules; no consensus has been reached on whether GGO nodules should be resected or followed up. At present there is no definite evidence concerning the natural history of GGO nodules.

Evaluating the volume-doubling time of GGO nodules with accurate CAV may yield information on tumor activity and facilitate the better management of GGO nodules. Routine CAV in the follow-up of GGO nodules may also be beneficial. Ikeda et al. (81) who performed 3D analysis of the CT number within GGO nodules using CAV software for differentiating among AAH, BAC, and adenocarcinoma found this method useful. However, even these 3D analyses require high-precision CAV.

During the follow-up of GGO nodules, solid components may appear or increase in size. According to Henschke et al.(4), the malignancy rate for part-solid GGO nodules was 63%. In adenocarcinoma with part-solid nodules, the ratio of solid- to non-solid parts is related to the prognosis; it is also useful for the differentiation of adenocarcinoma subtypes (10, 70). Therefore, a function for the identification and quantification of solid components within GGO nodules should be added to computer-aided diagnosis at CT follow-up of GGO nodules.

There are several potential limitations in our study. First, as we used a 64-detector CT scanner in the phantom- and a 4-detector scanner in the clinical study, the results of the phantom study may not be applicable to the clinical study. Das et al. (80, 82), who compared the accuracy of automated volumetry of solid phantom nodules using CT scanners from different vendors and scanners with different numbers of detectors, concluded that solid nodule volumetry was accurate and that the degree of volume error was acceptable considering that the data were acquired with different scanners. In GGO nodules, however, volumetry with different scanners may lead to a change of the cutoff value and turn out to be a source of variability. A technique that unifies and rectifies the image background between different types of scanner is required. Second, we used a section thickness and interval of 1 mm in the phantom

study; in the clinical study they were 1.25 mm because the clinical data were collected in a retrospective manner. Goo et al. (83) observed a tendency toward larger volume measurement errors with increasing section thickness. Thus, the measured volume in our clinical study may have been over-estimated. Third, we did not evaluate inter-scan variability. Wormanns et al. (38), who assessed solid pulmonary nodules on two consecutive CT scans obtained on the same day using CAV, reported that the 95% limit of agreement was -20.4 to 21.9% (standard error 1.5%). And Gietema et al.(84) also reported similar limits of agreement for inter-scan variability in volumetry of solid nodules. They suggested that most of the observed measurement variability was attributable to inter-scan differences. This indicates that the effect of inter-scan variability must be assessed. Lastly, there may have been a selection bias because we selected patients based on the strict criteria described above.

In conclusion, with our CAV software, volumetry was relatively accurate and intra- and inter-observer agreement may be clinically acceptable for the early detection of growth in GGO nodules ≥ 8mm in diameter.

Based on our findings we suggest that accurate CAV could play an important role in the follow-up and management of GGO nodules.

Chapter 6. Computer-Aided Volumetry of Pulmonary Nodules: Differential Diagnosis of Ground Glass Opacity Nodules by Volume-Doubling Time

Contents

ABSTRACT

Purpose: To investigate the volume-doubling time (VDT) of histologically proved pulmonary nodules showing ground glass opacity (GGO) at multidetector CT (MDCT) using computer-aided 3 dimensional (3D) volumetry (CAV) software.

Materials and Methods: We retrospectively evaluated 47 GGO nodules (mixed n=28, pure n=19) that had been examined by thin-section helical CT more than once. They were histologically confirmed as atypical adenomatous hyperplasia (AAH, n=13), bronchioloalveolar carcinoma (BAC, n=22), and adenocarcinoma (AC, n=12). Using semi-automatic CAV software, 2 radiologists independently performed 3D volumetry on the 47 GGO nodules. We averaged the measured values, calculated the VDT for each nodule using data acquired from the initial- and final CT study, and compared VDT among the 3 pathologies. We also compared the VDT of mixed- and pure GGO nodules.

Results: The mean VDT of all GGO nodules was 486.4 ± 368.6 days (range 89.0 - 1583.0 days). The mean VDT for AAH, BAC, and AC was 859.2 ± 428.9, 421.2 ± 228.4, and 202.1 ± 84.3 days, respectively; there were statistically significant differences for all comparative combinations of AAH, BAC, and AC (Steel-Dwass test, $p < 0.01$). The mean VDT for pure- and mixed GGO nodules was 628.5 ± 404.2 and

276.9 ± 155.9 days, respectively; it was significantly shorter for mixed- than pure GGO nodules (Mann-Whitney U-test, $p < 0.01$).

Conclusion: 3D volumetry of GGO nodules on CT using CAV software is a promising method for the differentiation of GGO nodules.

INTRODUCTION

The introduction of CT lung cancer screening has increased the number of detected ground glass opacity (GGO) nodules. Henschke et al. (4) reported that 44 (19%) of 233 positive results were lesions with GGO nodules; 15 (34%) of the 44 lesions were malignant. Among lesions with a solid component (mixed GGO nodules), 63% were malignant, nodules without solid components (pure GGO nodules) had a lower malignancy rate (18%). GGO nodules may be attributable to focal inflammation, focal interstitial fibrosis (9), atypical adenomatous hyperplasia (AAH) (6, 7), bronchioloalveolar carcinoma (BAC) (6, 7, 10), or adenocarcinoma (AC) (6, 7, 10). While the differentiation of BAC and AC, which are malignant, from other diseases is important, it can be difficult on a single CT scan (32). Many inflammatory lesions resolve spontaneously or with antibiotic treatment (6), on the other hand, the size of GGO nodules attributable to BAC or AC gradually increases (34, 70, 71). As focal interstitial fibrosis and AAH with pure GGO tend to remain stable in size for months or years (9), monitoring the nodule size for several months by high-resolution CT (HRCT) can help in obtaining a differential diagnosis. The reported average doubling time of BAC, calculated from the maximal tumor diameter with the Schwartz equation (42), is

very long (457 - 813 days on average) (33, 34), therefore, subjective evaluation of the growth rate may be unreliable.

Accurate 3 dimensional (3D) computer-aided volumetry (CAV) software to assess pulmonary nodules with volumetric data obtained at multidetector CT (MDCT) is available (36-40). However, most of the software used in earlier studies only evaluated solid pulmonary nodules and CAV of GGO nodules can be difficult (38, 40, 85). We developed enhanced CAV software to measure not only the volume of solid- but also of GGO nodules; it yielded sufficiently accurate and reproducible volume measurements (50).

The purpose of this study was to investigate the volume-doubling time (VDT) of histologically proved BAC, AC, and AAH that exhibited GGO nodules at MDCT. We used our improved 3D CAV software and assessed whether the VDT was useful to differentiate among BAC, AC, and AAH.

MATERIAS AND METHODS

All patients who underwent CT examination at our institute gave their prior informed consent for the use of their CT images in future retrospective studies. Our institutional review board approved this retrospective study and informed specific study-related consent was waived by the study subjects.

Patients and Nodule Selection

One chest radiologist (K.A.) with 22 years of chest CT experience reviewed the records of consecutive patients (April 2004 to April 2008) who underwent thin-section helical CT of the chest at our institute and were suspected of harboring pulmonary nodules. He did not participate in the volume measurement study. He selected patients who satisfied the following criteria: (a) GGO nodules that did not exceed 30 mm in the x-y (or transverse) plane, (b) thin-section helical CT scans were acquired more than once in the course of routine work-up, (c) no treatment or biopsy was performed between the these CT studies, (d) the histological diagnosis of AAH, BAC, or AC was based on findings at thoracic surgery or CT-guided transcutaneous- or bronchoscopic transbronchial biopsy, and (e) there was no consolidation due to either tissue organization after pneumonia or associated with idiopathic pulmonary fibrosis around

the nodules. The latter criterion was applied because we found it difficult to define the nodule boundary. On the basis of these criteria, 47 nodules from 39 patients (13 men and 26 women ranging in age from 31-79 years, mean 63.2 years) with 13 AAH, 22 BAC, 12 AC were identified. Their mean size in the x-y (or transverse) plane was 13.0 ± 4.9 mm (standard deviation) (range 5.5 - 27.2 mm); their mean density was -524.2 ± 166.3 Hounsfield units (HU) (range -786 - -135 HU). The chest radiologist classified the 47 GGO nodules as pure- (non-solid, n=28) and mixed GGO nodules (part-solid, n=19) according to their internal density (4). Table 12 shows the clinical and pathologic characteristics of the study population.

Image Acquisition

CT scans were obtained with a 4- (LightSpeed QX/I, GE Medical Systems, Milwaukee, WI) or a 64-detector CT scanner (Brilliance-64, Philips Medical Systems, Cleveland, OH). After routine helical scanning of the whole thorax, thin-slice helical scans whose scan range of about 5 cm included the pulmonary nodules were obtained on the LightSpeed scanner. The scan parameters for routine helical scans were: detector collimation, 4 x 2.5 mm; helical pitch (beam pitch), 1.5; slice thickness and interval, 5.0 mm; 0.8-sec rotation time, 120 kVp, and 250 mA. The parameters for thin-slice helical

scans were: detector row width, 4 x 1.25 mm; helical pitch, 0.75; slice thickness and interval, 1.25 mm; 0.8-sec rotation time, 120 kVp, and 160-200 mA. The reconstruction algorithm for thin-slice helical scans was "bone plus". At the Brilliance-64 scanner, the scanning parameters were: detector row width, 0.625 mm; helical pitch, 0.673; section thickness and section interval, 1 mm; rotation time, 0.5 sec; tube voltage, 120 kVp; tube current, 250 mAs. Contrast enhancement was not used in any of the patients.

Computerized Volumetry of Pulmonary Nodules (Fig. 23)

For volumetry of GGO nodules we used semi-automatic software developed for the picture-archiving and communication system (3D-analysis software, Fujitsu, Tokyo, Japan). It was an improved version of software we reported earlier (86).

First, with a single click and drag maneuver, the radiologists roughly specified the region of interest (ROI) to include the target nodule on axial-, sagittal-, or coronal images. Then, the CAV software automatically segmented the nodule, using an improvement of our previously reported method (72, 86). For the robust extraction of nodules from the lung parenchyma we adopted an adaptive threshold CT number based on the CT number of the target nodule and the lung parenchymal background (86). In this extraction process, structures connected to the nodule, e.g. vessels and bronchi,

were roughly eliminated by the CAV software. If some structures remained around the nodule, or part of the nodule was not included in the extracted area, the radiologists edited the segmented area to include the nodule by changing the threshold level that separated the density of the nodule from the base density of the lung and by using the mathematical morphology "erosion/dilation" technique (73). "Erosion/dilation" is a useful digital image-processing technique to eliminate structures tangent with the nodule such as vessels and thoracic walls. Although the software permits free-hand shape-editing, the radiologists were not allowed to use this technique. Rather, they could use only the above image processing technique to assure the reproducibility of segmentation. The judgment of successful segmentation was based on the observers' visual assessment on axial CT images as well as sagittal- and coronal multiplanar reconstructed images. Finally, the volume of the segmented area was automatically measured by the software. The average calculation time required by the CAV system to analyze each nodule was less than 1.0 sec. If an edit was required, the mean edit time was about 30 sec, even long edit times did not exceed 1 min.

A computer workstation (FMV-H8230; Fujitsu, Tokyo, Japan) with dual 2.16-GHz processors (Core Duo; Intel, Santa Clara, Calif.) was used in this study.

Evaluation of Volume Doubling Time

Using our CAV software, 2 radiologists independently performed volumetry of the pulmonary nodules. Nodule volumes were measured as described above on initial- and final CT scans for calculating VDT. The mean interval between the 2 CT scans was 188 days (range 34 - 1486 days). The VDT was calculated using the equation based on the modified Schwartz formula (42) of an exponential growth model.

$$VDT = [\log 2 \times t] / [\log (V_2/V_1)],$$

where V_1 and V_2 are the initial and final nodule volume and t is the interval between the 2 CT scans. The software calculates the VDT automatically by comparing the nodule volume on the 2 scans (Fig. 24). The average value of the 2 radiologists' measurements of VDT was adopted.

Statistical Analysis

Numerical data were expressed as the mean ± standard deviation (SD). Differences between the VDT among AAH, BAC, and AC were examined for statistical significance using the Steel-Dwass test. We also compared VDT between mixed- and

pure GGO nodules using the Mann-Whitney U-test. We applied these tests because the data sets of the VDT for BAC and pure GGO nodules were not normally distributed. The Spearman correlation coefficient was used to evaluate the strength of the correlation between the VDT and the maximum diameter of the GGO nodule on the initial CT scan and between the VDT and the mean internal density of the GGO nodule on the initial CT scan. We used this test because VDT data were not normally distributed.

Statistical analyses were performed with a software program (MedCalc Software; MedCalc, Mariakerke, Belgium; and R version 2.6.1, R Project for Statistical Computing). A p-value less than 0.05 denoted a statistically significant difference.

RESULTS

We were able to perform CAV and VDT calculations in all 47 GGO nodules. The mean VDT of all GGO nodules was 486.4 ± 368.6 days (range 89.0 - 1583.0 days). The mean VDT was 859.2 ± 428.9 (range 260.5 - 1583.0 days), 421.2 ± 228.4 (range 154.0 - 1182.0 days), and 202.1 ± 84.3 days (range 89.0 - 348.5 days) for AAH, BAC, and AC, respectively. There were statistically significant differences in the mean VDT for all combinations of AAH, BAC, and AC (Steel-Dwass test; AAH vs. BAC $p < 0.01$, AAH vs. AC $p < 0.01$, BAC vs. AC $p < 0.01$) (Fig. 25).

The mean VDT for pure- and mixed GGO nodules was 628.5 ± 404.2 days (range 154.0 - 1583.0 days) and 276.9 ± 155.9 days (range 89.0 - 724.5 days), respectively. It was significantly shorter for mixed- than pure GGO nodules (Mann-Whitney U-test, $p < 0.01$) (Fig. 26).

We obtained the best results when we used 460 days as the threshold for differentiating between malignant- (BAC and AC) and benign (AAH) GGO nodules; accuracy was 83.0%, sensitivity 82.4%, specificity 84.6%, and positive- and negative predictive values were 93.3% and 64.7%, respectively.

There was no correlation between the maximum GGO nodule diameter on the initial CT scan and the VDT ($r = -0.19$, $p = 0.19$) (Fig. 27). On the other hand, there

was a statistically significant correlation between the mean internal density of GGO nodules on the initial CT scan and the VDT ($r = -0.57$, $p < 0.01$) (Fig. 28).

DISCUSSION

The tumor-doubling time was an independent and significant prognostic factor in lung cancer patients (41). Shorter VDT may reflect greater histological tumor aggressiveness, suggesting that lung cancers with a short VDT are associated with a worse prognosis. Conventionally, the evaluation of tumor growth has been performed by bilinear 2D measurements on chest radiographs or CT images. The introduction of 3D CAV of pulmonary nodules may facilitate a more accurate and earlier evaluation of tumor growth using volumetric data obtained at MDCT. Lillington (87) noted that the VDT of most benign pulmonary nodules was more than 450 days; for malignant lesions it was usually less than 400 days. Traditionally, the absence of lesion growth over a 2-year period was taken to indicate a benign status. On the other hand, 2-year stability does not guarantee a benign state because some malignant growths have a long VDT, for example, some studies have reported a VDT for AC of over 2 years (88). In our study, the mean VDT was 859.2 ± 428.9, 421.2 ± 228.4, and 202.1 ± 84.3 days for AAH, BAC, and AC, respectively; the mean VDT for pure- and mixed GGO nodules was 628.5 ± 404.2 and 276.9 ± 155.9 days, respectively. Based on our 3D CAV results, AC manifested a statistically significantly shorter mean VDT than BAC or AAH. AC was the most aggressive among the 3 tumor types we studied and mixed- were more

aggressive than pure GGO nodules. Lindell et al. (89) who used 2D volumetry to study pulmonary nodules reported that the mean VDT was 780 ± 1545 and 746 ± 1238 days for BAC and AC, respectively; it was 469 ± 452 and 568 ± 1222 days, respectively, for pure- and mixed GGO nodules. Their results with respect to the mean VDT differed from ours and their standard deviation was much longer than ours. We attribute this difference to the different volume measurement methods (2D- vs. 3D measurement) and the small number of lesions included in their investigation. Hasegawa et al. (33) evaluated the VDT of 61 primary lung cancers 2D volumetrically. Their mean VDT of well-differentiated AC, including BAC, was 533 ± 381 days; as in our study it was shorter for mixed- than pure GGO nodules (457 ± 260 vs. 813 ± 375 days). These findings suggest that tumors with mixed- are more aggressive than tumors with pure GGO nodules.

Yankelevitz et al. (90) considered lung cancers with VDT exceeding 400 day as overdiagnosed cancers. That is, therapeutic intervention in these cancers had no effect on mortality from lung cancer. Using data from the Mayo Lung- (91) and the Memorial Sloan-Kettering Cancer Center Project (92) they examined VDT of stage I lung cancers. They found that among these cancers, 2% reported in the former and 7% in the latter project could be considered as overdiagnosed. In our study, 22 (46.8%) of all 47 GGO

nodules manifested a VDT longer than 400 days. While no AC exhibited a VDT of 400 days or more, in 11 (84.6 %) of AAH and 10 (45.5 %) of BAC it exceeded 400 days; 18 (64.3 %) of pure- and 3 (15.8 %) of mixed GGO nodules manifested a VDT longer than 400 days. It must be noted that more than half of our pure GGO nodules, thought to have a good prognosis (71), had a VDT longer than 400 days. Our findings point to the clinical utility of 3D CAV. However, as the definition of overdiagnosis by Yankelevitz et al. [26] was based on 2D measurements on chest radiographs, the relationship between VDT and prognosis of tumors assessed with accurate 3D volumetry on HRCT scans must be re-evaluated.

We detected no correlation between the maximum diameter of GGO nodules and their VDT. On the other hand, there was a statistically significant correlation between the mean internal density of GGO nodules and VDT, indicating that GGO nodules with high internal density, most of them mixed GGO nodules, are found in more highly aggressive tumors and that they require careful observation and management.

Most previously reported doubling times for GGO nodules were based on 2D data and a modified Schwartz equation (41, 42) that should be applied to round nodules. However, human pulmonary nodules are not necessarily round and their margins are not

always smooth. Therefore, 2D- may be less accurate than 3D measurements. In fact, 3D CAV offered obvious advantages over conventional bilinear 2D measurements for the accurate representation of the nodular volume [16]. To set a new standard, we recommend that the doubling time of pulmonary nodules be re-evaluated with accurate 3D volumetry.

The adequate follow-up intervals for GGO nodules remain to be established. Our comparison of the VDT of pure- and mixed GGO nodules suggests that the interval should be shorter in patients with mixed- than pure-GGO nodules. In this study, the shortest VDT, indicative of high tumor aggressiveness, was 154 and 89 days for pure- and mixed GGO nodules, respectively. To detect GGO nodules with a short VDT, the first follow-up CT scan should be acquired 6- and 3 months, respectively, after the initial scan CT for pure- and mixed GGO nodules. If the first follow-up reveals no remarkable changes in the nodular volume, subsequent CT studies should be performed at one-year intervals, longer than the currently-applied term. We suggest that GGO nodules with a VDT exceeding 400 days should be monitored carefully without immediate surgical intervention. The adequate spacing of follow-up studies will avoid unnecessary radiation exposure of the patient as well as inappropriate intervention in overdiagnosed GGO nodules.

There are some inherent problems with volumetry of GGO nodules. First, the size of some BAC with GGO decreased in the course of follow-up due to the collapse of alveolar spaces, fibrosis, or the severe narrowing of the alveolar space (71). Therefore, even if GGO nodules decrease in size, malignancy cannot be ruled out. Second, since cells of BAC with focal GGO manifest replacement growth without a marked reduction in the alveolar lumen and without marked thickening of alveolar septa (8, 93), BAC exhibiting focal GGO tend to be relatively soft tumors. Therefore, the volume of BAC with focal GGO may be affected by the lung volume. At volumetry of focal GGO nodules, the lung volume may have to be taken into consideration.

Our study has several limitations. First, as the number of nodules available for study was small, large-scale clinical studies are needed for a vigorous investigation of VDT. Second, we used two types of CT scanner, a 64- and a 4-detector scanner, raising the possibility that the use of different instruments and scanning parameters may have affected our volumetric measurements, resulting in unreliable VDT results. Das et al. (80, 82) compared the accuracy of automated volumetry of phantom nodules using CT scanners from different vendors and scanners with different numbers of detectors. They concluded that nodule volumetry was accurate and that the degree of volume error was acceptable, findings that suggest that differences in the scanner type may have little

effect on CAV results. Nonetheless, a technique that unifies and rectifies the image background among different types of scanner is needed. We are engaged in developing new robust 3D CAV software that minimizes volumetric variability when image data acquired with CT scanners from different vendors are used. Lastly, there may have been a selection bias because we selected patients based on strict inclusion criteria.

In conclusion, evaluation of the VDT with 3D volumetry is a promising method for the differentiation of GGO nodules. It can play an important role in the follow-up and management of GGO nodules.

Figure legends and figures

Fig. 1 ROC curves of the diagnostic performance of all observers without and with DES images. The average AUC value is significantly improved with DES images ($p = 0.02$).

Fig. 2 ROC curves of the diagnostic performance of all observers for -450 HU nodules without and with DES images. The average AUC value is significantly improved with DES images ($p < 0.01$).

Fig. 3 ROC curves of the diagnostic performance of all observers for -200 HU nodules. There was no statistically significant difference in the AUC values without and with DES ($p=0.13$).

Fig. 4 ROC curves of the diagnostic performance of all observers for 30-HU nodules. There was no statistically significant difference in the AUC values without and with DES ($p=0.17$).

Fig. 5 Chest phantom with a simulated -450 HU nodule with a diameter of 10 mm.

A. Standard chest radiograph. There is a simulated nodule overlapped by the rib shadow in the right upper lung field (arrow).

B. Soft-tissue image produced with the DES technique. The nodule is seen more clearly on the DES image. The mean rating score of all observers was increased from 56.2 to 80.6 by using the DES image.

Fig. 6 Row images obtained at low- and high kVp processed with the DES technique generated the standard posteroanterior image (A), the soft-tissue image (B), and the bone image (C).

Fig. 7 ROC curves of the diagnostic performance of all observers without and with DES images. The average AUC value is significantly improved with DES images ($p = 0.02$).

Fig. 8 Representative case with a nodule.

A. Standard chest radiograph. There is a small nodule overlapping the rib shadow in the right upper lung field (arrow).

B. Soft-tissue image produced with the DES technique. The nodule is seen more clearly on the DES image.

C. CT scan shows a part-solid nodule that proved to be well-differentiated adenocarcinoma in the right upper lobe.

Fig. 9 Representative case with a nodule.

A. Standard chest radiograph. Note the subtle nodule overlapping the rib shadows in the right middle lung field (arrow).

B. Soft-tissue image produced with the DES technique. The visibility of the nodule is maintained on the processed image.

C. CT scan shows a non-solid nodule that proved to be bronchioloalveolar carcinoma in the right middle lobe.

Fig. 10 Computerized schema for rib-suppressed chest radiography with MTANN

A. Training step: In the training step, the MTANN can be trained with input images and the corresponding "teaching" images. Bone images obtained with a dual-energy radiography system were used as the teaching images for the MTANN.

B. Application step: The MTANN can produce soft-tissue and bone images from a single chest image obtained with a standard radiography system.

Fig. 11 ROC curves of the diagnostic performance of all radiologists observing images without and with rib suppression. The radiologists' accuracy was significantly higher on rib-suppressed- than non-suppressed images ($p = 0.019$).

Fig. 12 ROC curves of the diagnostic performance of board-certified radiologists inspecting images without and with rib suppression. The radiologists' accuracy was significantly higher on rib suppressed- than non-suppressed images ($p = 0.011$).

Fig. 13 ROC curves of the diagnostic performance of radiology residents observing images without and with rib suppression. There was no statistically significant difference in the AUC value between images without and with rib suppression ($p = 0.310$).

Fig. 14 Representative case with a nodule.

A. Original chest radiograph. Note the subcentimeter nodule in the right middle lung field (arrow).

B. Rib suppressed image produced with MTANN. The visibility of the nodule is maintained on the processed image.

Fig. 15 Representative case with a nodule.

A. Original chest radiograph. There is a nodule overlapping with the anterior rib in the right mid lung field (arrow).

B. Rib-suppressed image produced with MTANN. The nodule is seen more clearly on the rib-suppressed image.

Fig. 16 Diagram of the computerized scheme for volumetry of nodules on thin-section helical CT images.

Fig. 17 Computerized scheme for volumetry of nodules on thin-section helical CT images.

A. The target GGO nodule on HRCT

B. Radiologists manually specify the target GGO nodule and place a region of interest (ROI)

C. Software automatically analyzes the density of the nodule and surrounding lung parenchyma and estimates the nodule border. Occasionally, some structures such as vessels remained around the nodules (arrows)

D. Radiologists subjectively modify the nodule border with a semi-automatic edit tool based on the concept of mathematical morphology. Lastly, the software automatically calculates the nodule volume.

Fig. 18 Screen display of computer-aided volumetry software in the phantom study. Axial-, coronal-, and sagittal images and maximum intensity projections (MIPs) are displayed. The simulated GGO nodule (-630HU) can be accurately extracted by the software.

Fig. 19 Screen display of computer-aided volumetry software in the clinical study. The software accurately extracts the pure GGO nodule (Category 1).

Fig. 20 Average relative volume measurement error (RVME) for nodules measuring 3-, 5-, 8-, 10-, and 12 mm in diameter with a CT number of -450-, -630, -800 Houndsfield units (HU). In nodules with a CT number of -450 HU, the RVME was 85.2 ± 13.5% (standard deviation), -4.1 ± 3.4%, 4.8 ± 2.6%, 1.7 ± 3.4%, and -1.5 ± 4.2%, respectively. In nodules with a CT number of -630 HU, the RVME was 51.1±9.8%, 0.2 ± 6.2%, 3.3 ± 1.4%, 0.2 ± 3.3%, and 1.6 ± 3.7%, respectively. In nodules with a CT number of -800 HU, the RVME was 61.5 ± 4.0%, 7.1 ±2.1%, 4.8 ± 1.4%, 3.1 ± 1.2%, and 7.0 ±0.9%, respectively.

Fig. 21 Intra-observer agreement

A. Intra-observer agreement in nodules < 8mm in diameter for observer A. The mean intra-observer agreement was 7.4 % ± 16.7 % (95 % CI: -1.5 to 16.3 %). The 95 % limits of agreement were -25.3 % (95 % CI: -40.8 to -9.8 %) to 40.1 % (95 % CI: 24.6 to 55.6 %).

B. Intra-observer agreement in nodules ≥ 8mm in diameter for observer A. The mean intra-observer agreement was -0.6 % ± 7.3 % (95 % CI: -2.8 to 1.7 %). The 95 % limits of agreement were -14.9 % (95 % CI: -18.7 to -11.0) to 13.7 % (95 % CI: 9.8 to 17.6 %).

C. Intra-observer agreement in nodules < 8mm in diameter for observer B. The mean intra-observer agreement was 7.0 % ± 15.8 % (95 % CI: -0.9 to 14.9 %). The 95 % limits of agreement were -24.0 % (95 % CI: -37.7 to -10.3 %) to 38.0 % (95 % CI: 24.3 to 51.7 %).

D. Intra-observer agreement in nodules ≥ 8mm in diameter for observer B. The mean intra-observer agreement was -0.4 % ± 8.2 % (95 % CI: -3.0 to 2.2 %). The 95 % limits of agreement were -16.6 % (95 % CI: -21.1 to -12.1) to 15.7 % (95 % CI: 11.2 to 20.2 %).

Fig. 22 Inter-observer agreement

A. The mean inter-observer agreement in nodules < 8mm in diameter was 1.9 % ± 15.4 % (95 % CI: -6.0 to 9.8 %). The 95 % limits of agreement were -28.4 % (95 % CI: -42.2 to -14.5 %) to 32.2 % (95 % CI: 18.3 to 46.0 %).

B. The mean inter-observer agreement in nodules ≥ 8mm in diameter was 3.7 % ± 10.2 % (95 % CI: 0.5 to 6.9 %). The 95 % limits of agreement were -16.3 % (95 % CI: -21.8 to -10.8) to 23.7 % (95 % CI: 18.3 to 29.2 %).

Fig. 23 Example of nodule volumetry on thin-section helical CT images.

A. Target GGO nodule on HRCT.

B. Radiologists manually specify the target GGO nodule and place a region of interest (ROI).

C. The software automatically analyzes the density of the nodule and surrounding lung parenchyma and estimates the nodule border. Occasionally structures such as vessels remain around the nodule (arrow).

D. Using the concept of mathematical morphology, radiologists subjectively modify the nodule border with a semi-automatic edit tool. Lastly, the software automatically calculates the nodule volume.

Fig. 24 VDT measurements of GGO nodule with 3D CAV software.

Adenocarcinoma with a GGO nodule in a 63-year-old woman. The initial- and final CT images are displayed on the left and right, respectively. The interval between the two CT scans was 270 days. The software automatically calculates the VDT by comparing both scans. In this case, the VDT was 163 days.

Fig. 25 Graph showing the mean VDT for AAH, BAC and AC. There were statistically significant differences in the mean VDT for all combinations of AAH, BAC, and AC (Steel-Dwass test). The upper (lower) end of vertical lines, upper (lower) margin of boxes, horizontal lines in boxes, and circular symbols represent upper (lower) extremes, upper (lower) quartiles, medians, and outliers of data, respectively. *$p < 0.01$.

Fig. 26 Graph showing the mean VDT for pure- and mixed GGO nodules. VDT was significantly shorter for mixed- than pure GGO nodules (Mann-Whitney U-test). See Fig. 3 for explanation of symbols. *$p < 0.01$.

Fig. 27 Graph showing the correlation between the VDT and the maximum diameter- and between the VDT and the mean density of GGO nodules on the initial CT scan. No correlation was found ($r = -0.19$, $p = 0.19$).

Fig. 28 Graph showing the correlation between the VDT and the mean internal density of GGO nodules on the initial CT scan. There was a statistically significant correlation ($r = -0.57$, $p < 0.01$).

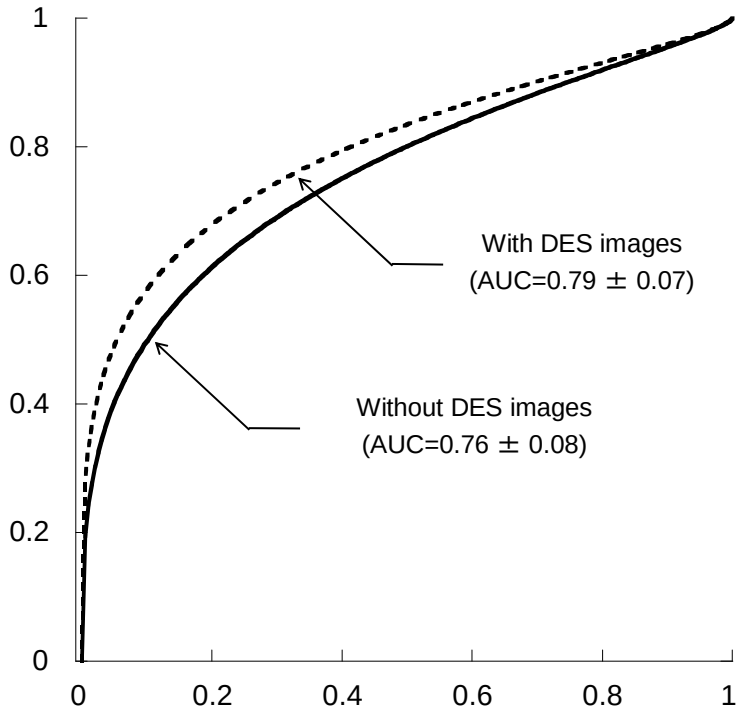


Fig. 1 ROC curves of the diagnostic performance of all observers without and with DES images. The average AUC value is significantly improved with DES images ($p = 0.02$).

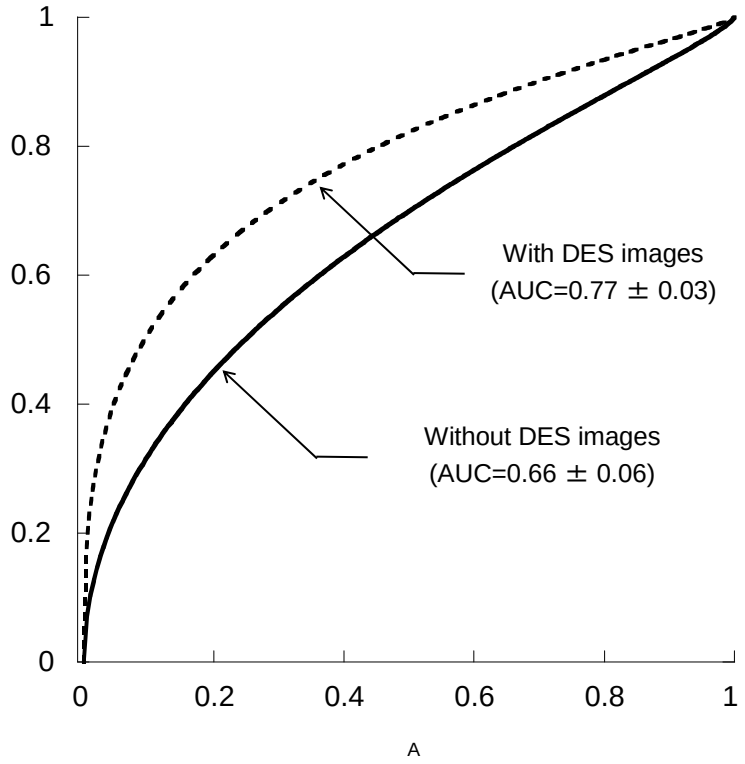


Fig. 2 ROC curves of the diagnostic performance of all observers for -450 HU nodules without and with DES images. The average AUC value is significantly improved with DES images ($p < 0.01$).

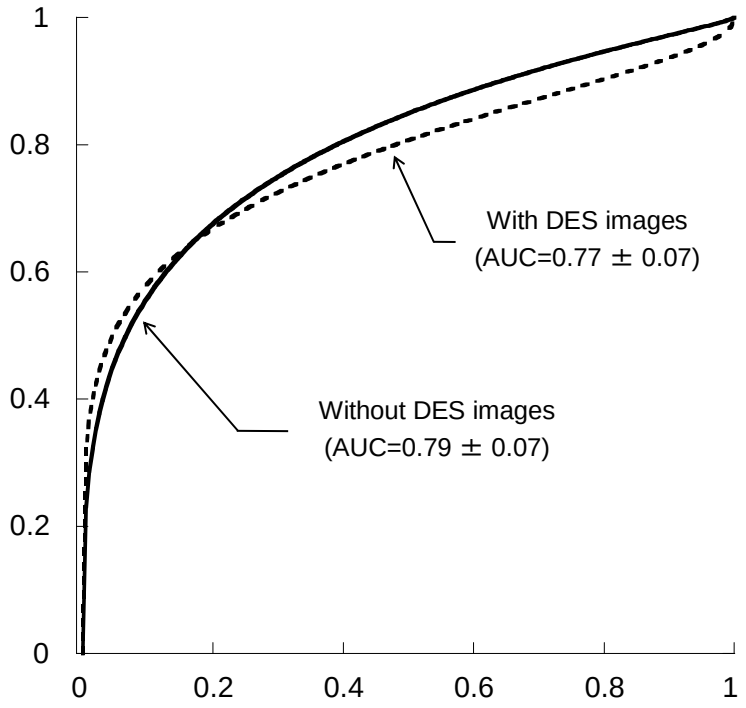


Fig. 3 ROC curves of the diagnostic performance of all observers for -200 HU nodules. There was no statistically significant difference in the AUC values without and with DES (p=0.13).

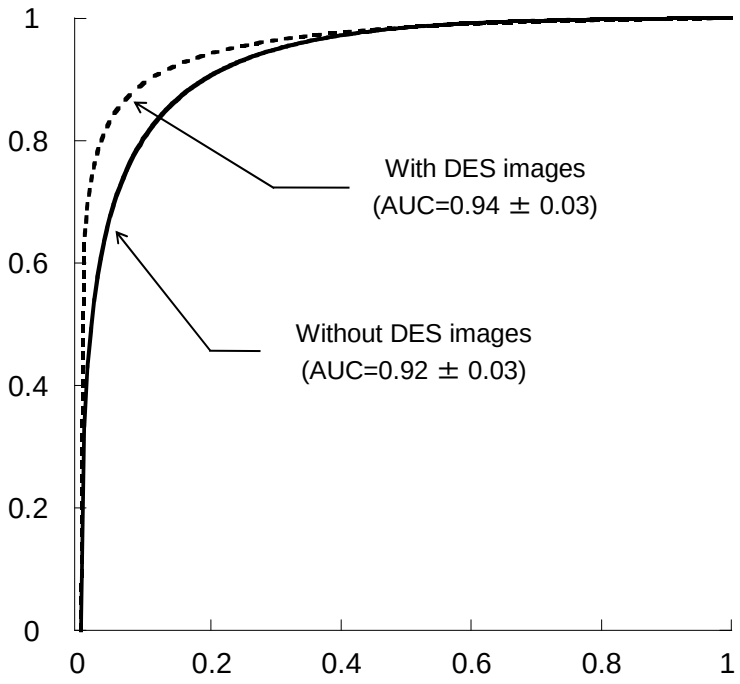
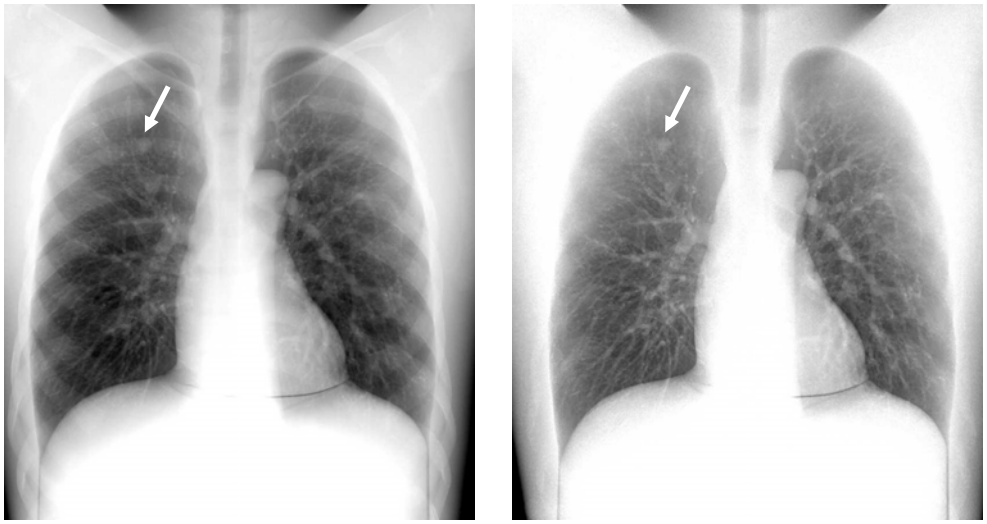


Fig. 4 ROC curves of the diagnostic performance of all observers for 30-HU nodules. There was no statistically significant difference in the AUC values without and with DES (p=0.17).

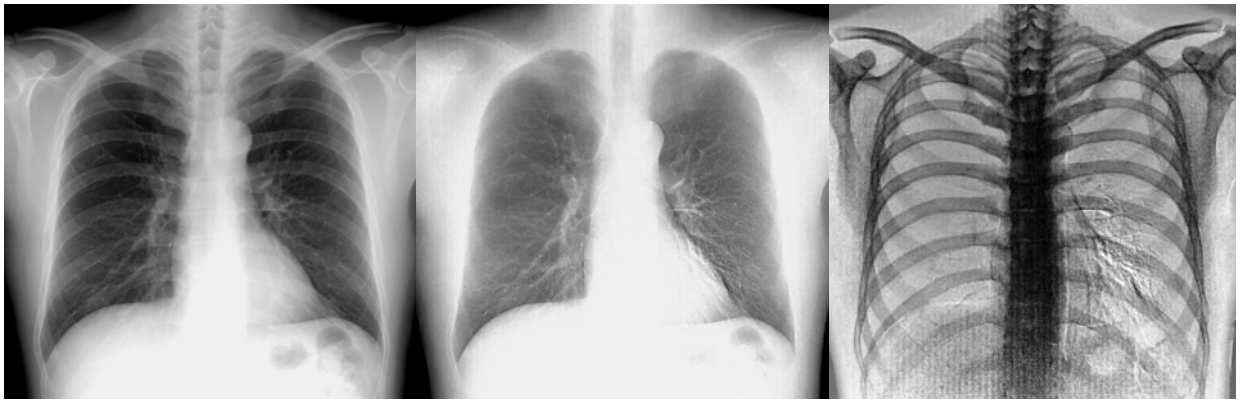


A **B**

Fig. 5 Chest phantom with a simulated -450 HU nodule with a diameter of 10 mm.

A. Standard chest radiograph. There is a simulated nodule overlapped by the rib shadow in the right upper lung field (arrow).

B. Soft-tissue image produced with the DES technique. The nodule is seen more clearly on the DES image. The mean rating score of all observers was increased from 56.2 to 80.6 by using the DES image.



A **B** **C**

Fig. 6 Row images obtained at low- and high kVp processed with the DES technique generated the standard posteroanterior image (A), the soft-tissue image (B), and the bone image (C).

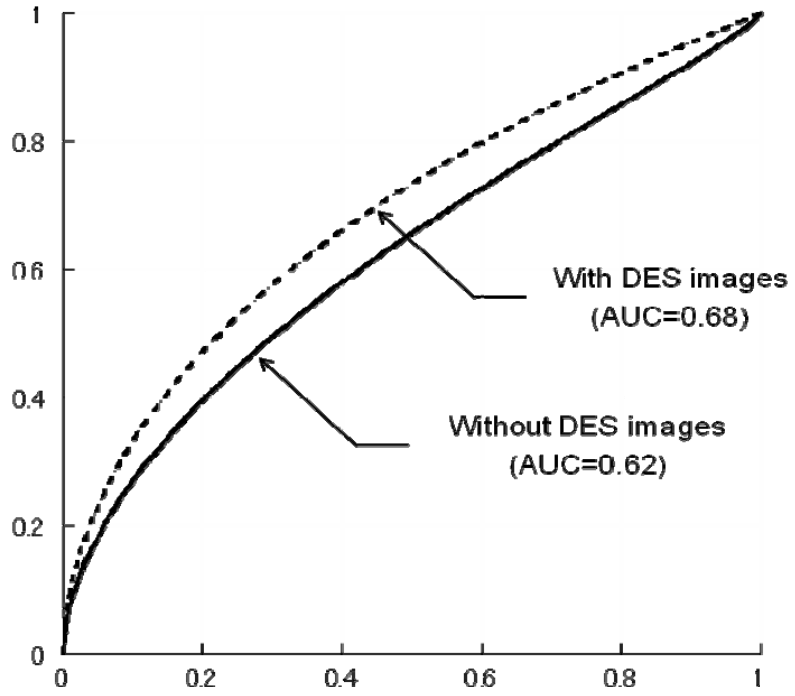
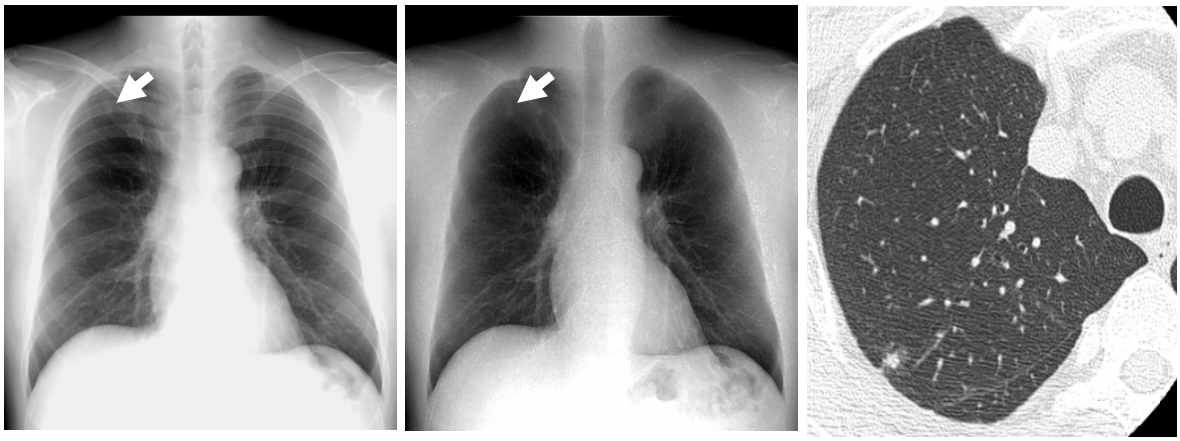


Fig. 7 ROC curves of the diagnostic performance of all observers without and with DES images. The average AUC value is significantly improved with DES images ($p = 0.02$).



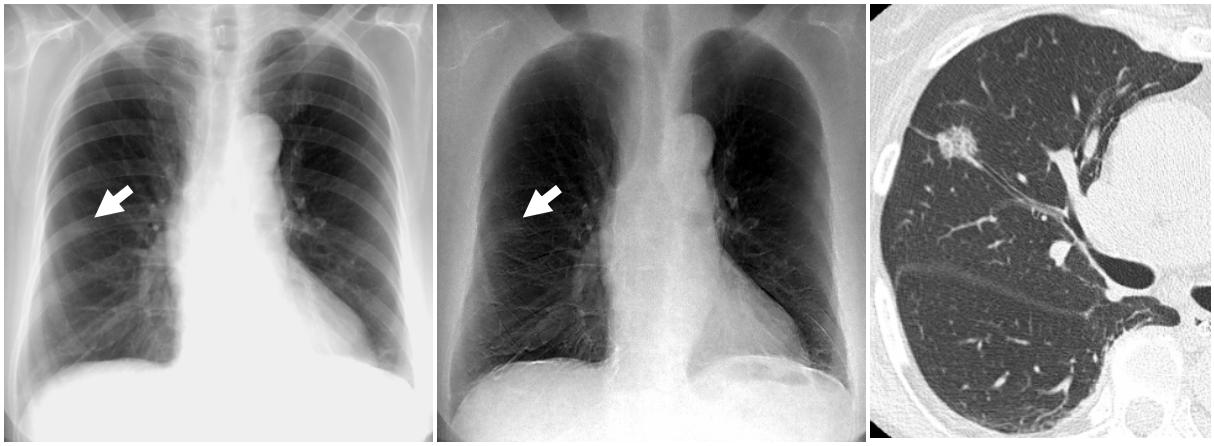
A **B** **C**

Fig. 8 Representative case with a nodule.

A. Standard chest radiograph. There is a small nodule overlapping the rib shadow in the right upper lung field (arrow).

B. Soft-tissue image produced with the DES technique. The nodule is seen more clearly on the DES image.

C. CT scan shows a part-solid nodule that proved to be well-differentiated adenocarcinoma in the right upper lobe.



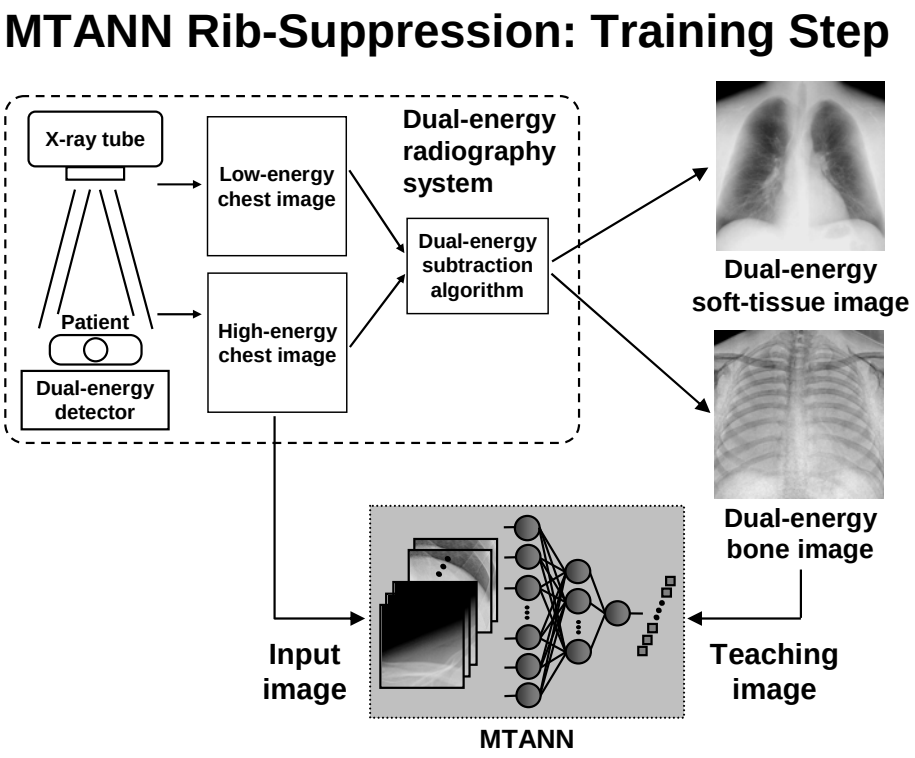
A **B** **C**

Fig. 9 Representative case with a nodule.

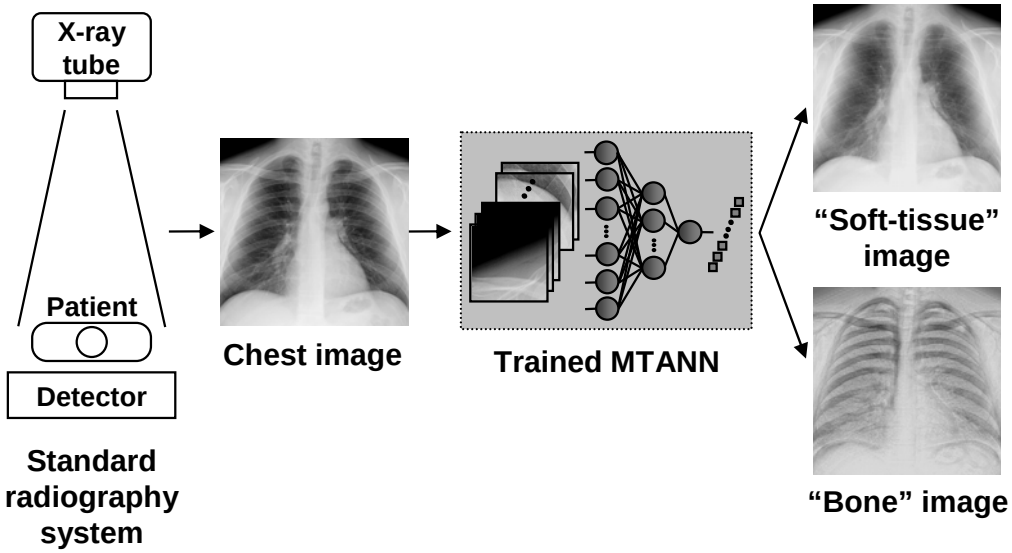
A. Standard chest radiograph. Note the subtle nodule overlapping the rib shadows in the right middle lung field (arrow).

B. Soft-tissue image produced with the DES technique. The visibility of the nodule is maintained on the processed image.

C. CT scan shows a non-solid nodule that proved to be bronchioloalveolar carcinoma in the right middle lobe.


MTANN Rib-Suppression: Training Step
X-ray tube
Low-energy chest image
Dual-energy radiography system
Dual-energy subtraction algorithm
Patient
High-energy chest image
Dual-energy detector
Dual-energy soft-tissue image
Dual-energy bone image
Input image
Teaching image
MTANN

A


MTANN Rib-Suppression: Application Step
X-ray tube
Patient
Detector
Standard radiography system
Chest image
Trained MTANN
"Soft-tissue" image
"Bone" image

B

Fig. 10 Computerized schema for rib-suppressed chest radiography with MTANN

A. Training step: In the training step, the MTANN can be trained with input images and the corresponding "teaching" images. Bone images obtained with a dual-energy radiography system were used as the teaching images for the MTANN.

B. Application step: The MTANN can produce soft-tissue and bone images from a single chest image obtained with a standard radiography system.

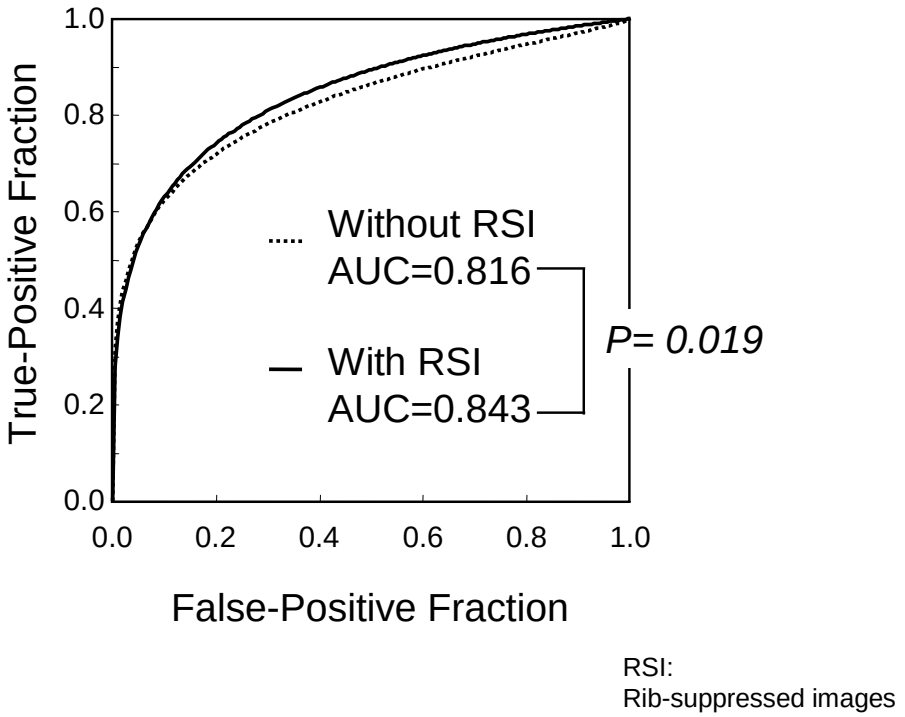


Fig. 11 ROC curves of the diagnostic performance of all radiologists observing images without and with rib suppression.

The radiologists' accuracy was significantly higher on rib-suppressed- than non-suppressed images ($p = 0.019$).

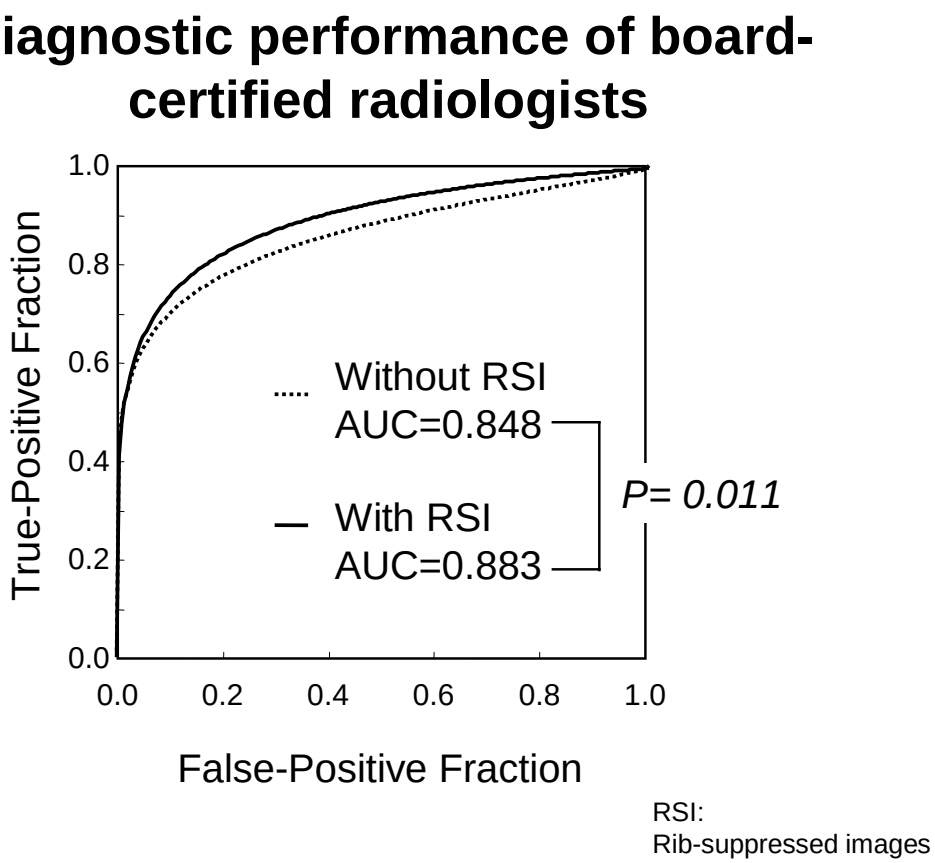


Fig. 12 ROC curves of the diagnostic performance of board-certified radiologists inspecting images without and with rib suppression.

The radiologists' accuracy was significantly higher on rib suppressed- than non-suppressed images (p = 0.011).

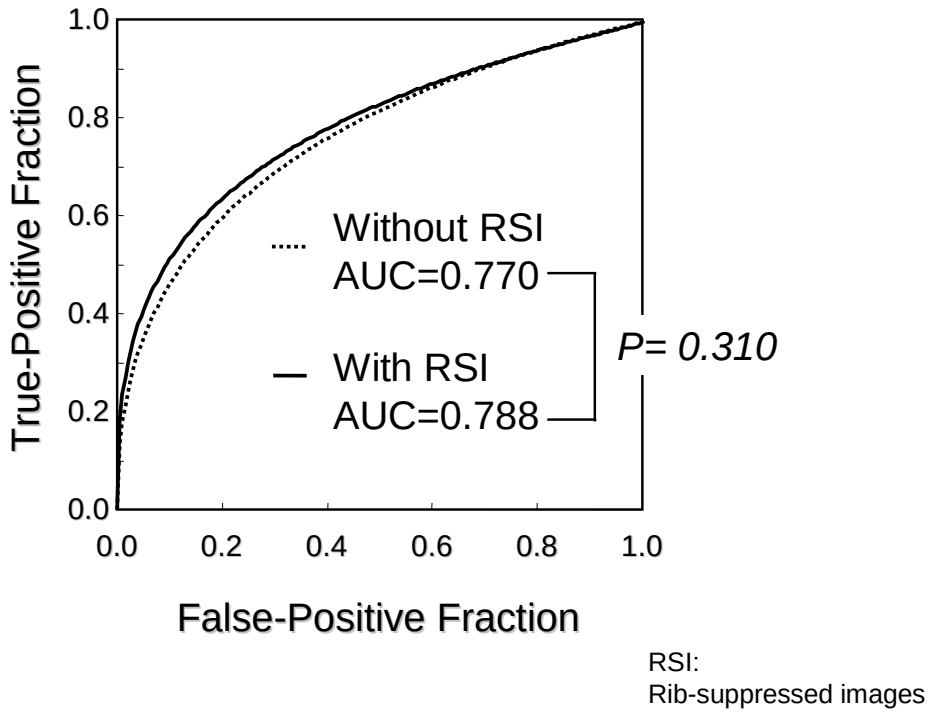
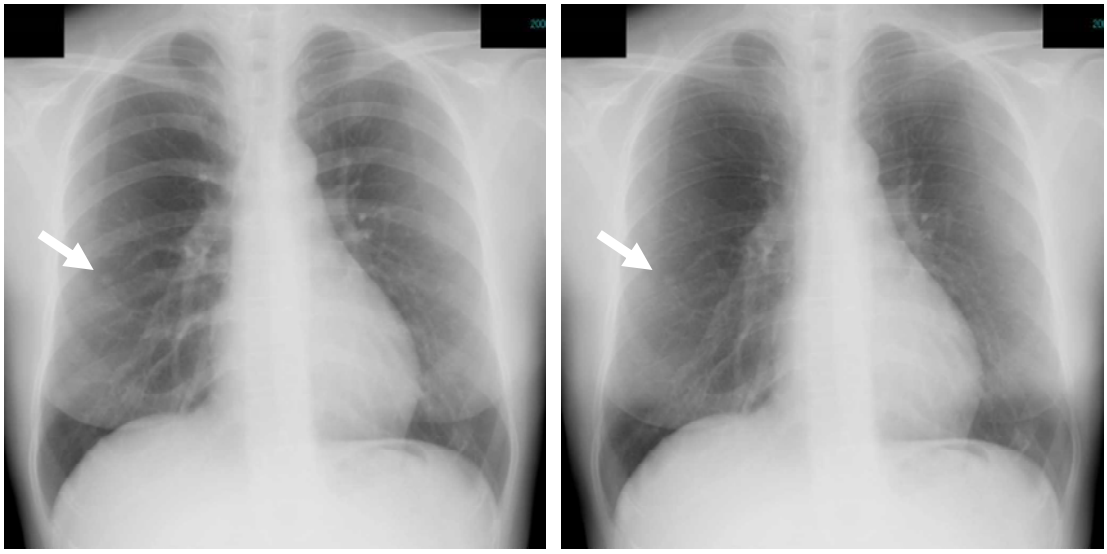


Fig. 13 ROC curves of the diagnostic performance of radiology residents observing images without and with rib suppression.

There was no statistically significant difference in the AUC value between images without and with rib suppression ($p = 0.310$).

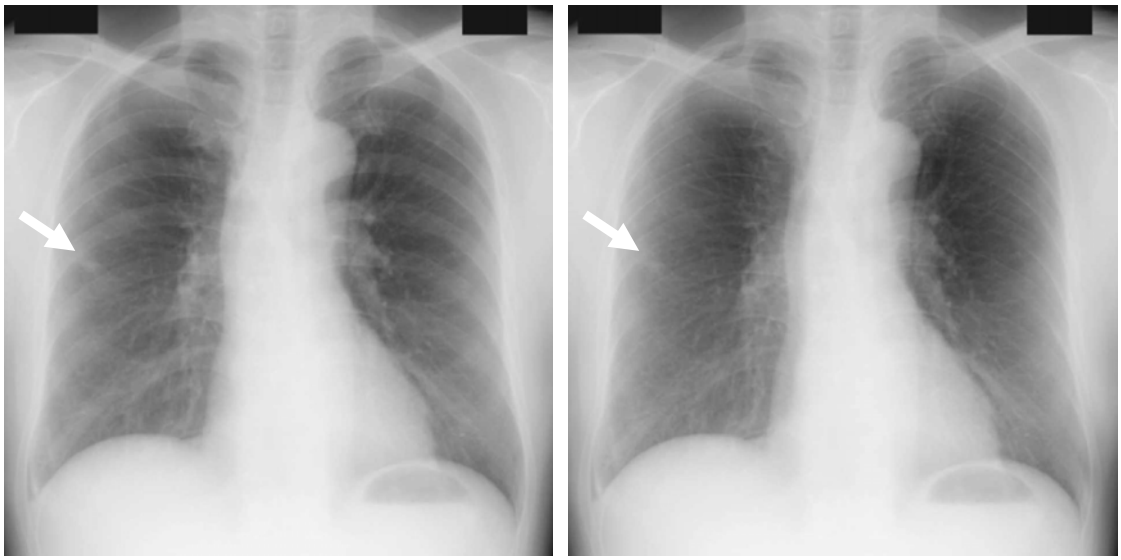


A **B**

Fig. 14 Representative case with a nodule.

A. Original chest radiograph. Note the subcentimeter nodule in the right middle lung field (arrow).

B. Rib suppressed image produced with MTANN. The visibility of the nodule is maintained on the processed image.



A **B**

Fig. 15 Representative case with a nodule.

A. Original chest radiograph. There is a nodule overlapping with the anterior rib in the right mid lung field (arrow).

B. Rib-suppressed image produced with MTANN. The nodule is seen more clearly on the rib-suppressed image.

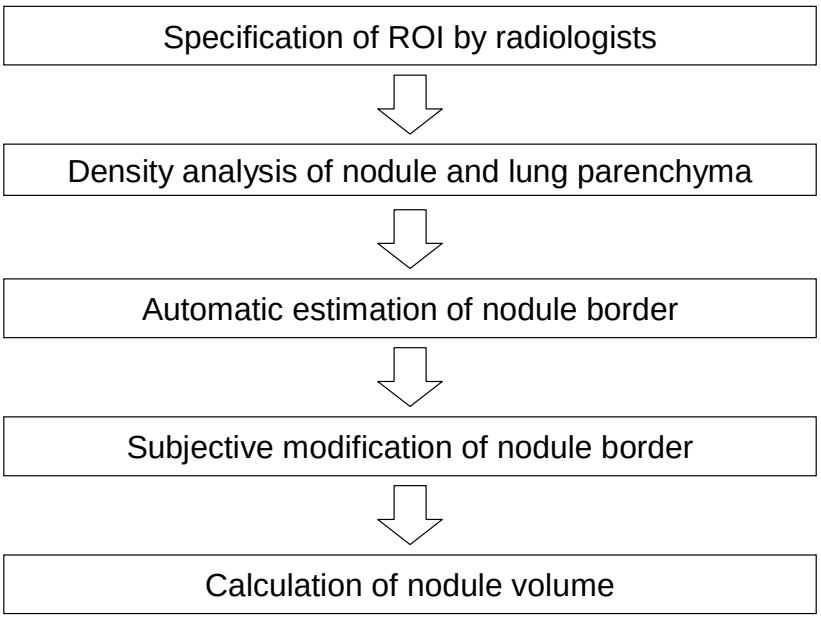


Fig. 16 Diagram of the computerized scheme for volumetry of nodules on thin-section helical CT images.

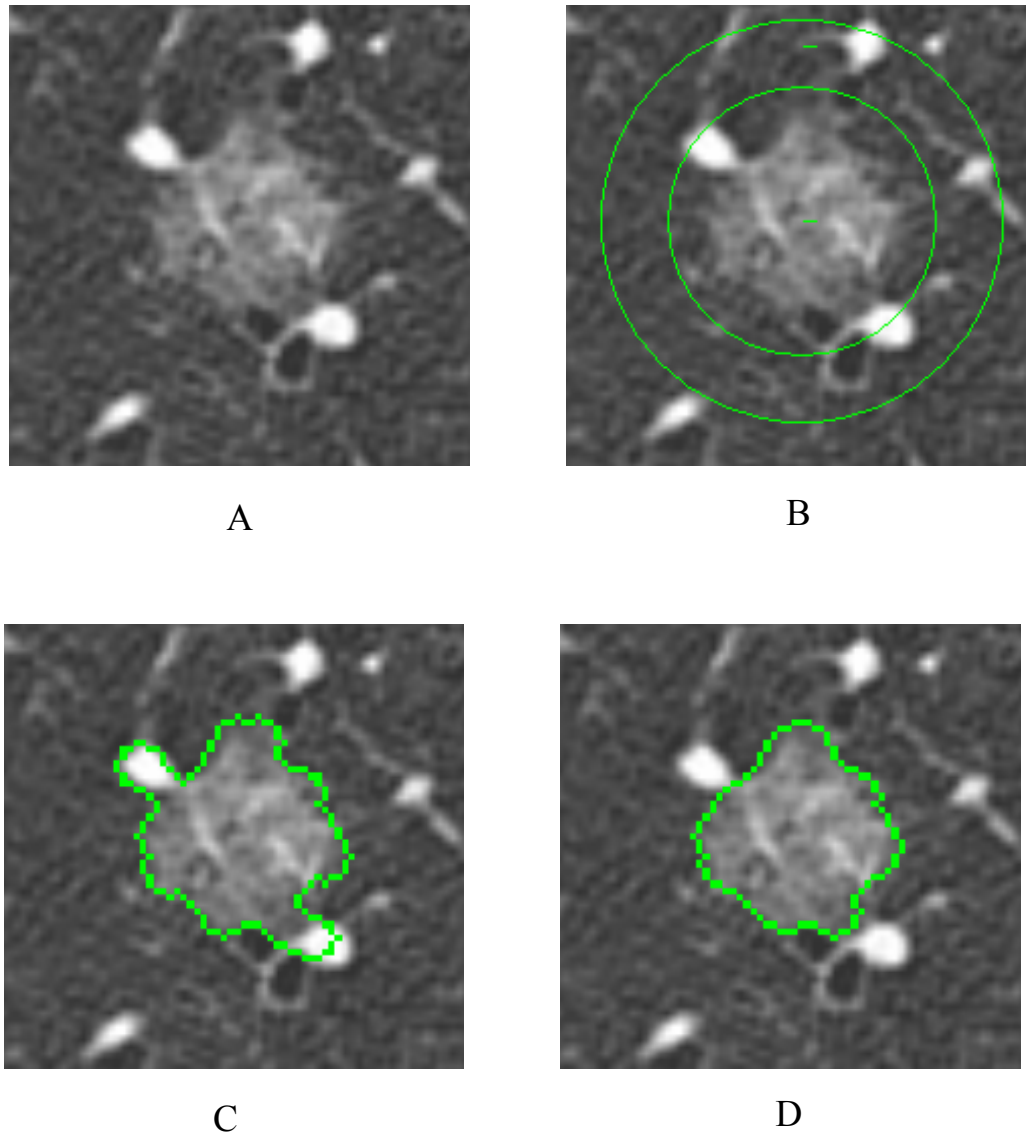


Fig. 17 Computerized scheme for volumetry of nodules on thin-section helical CT images.

A. The target GGO nodule on HRCT

B. Radiologists manually specify the target GGO nodule and place a region of interest (ROI)

C. Software automatically analyzes the density of the nodule and surrounding lung parenchyma and estimates the nodule border. Occasionally, some structures such as vessels remained around the nodules (arrows)

D. Radiologists subjectively modify the nodule border with a semi-automatic edit tool based on the concept of mathematical morphology. Lastly, the software automatically calculates the nodule volume.

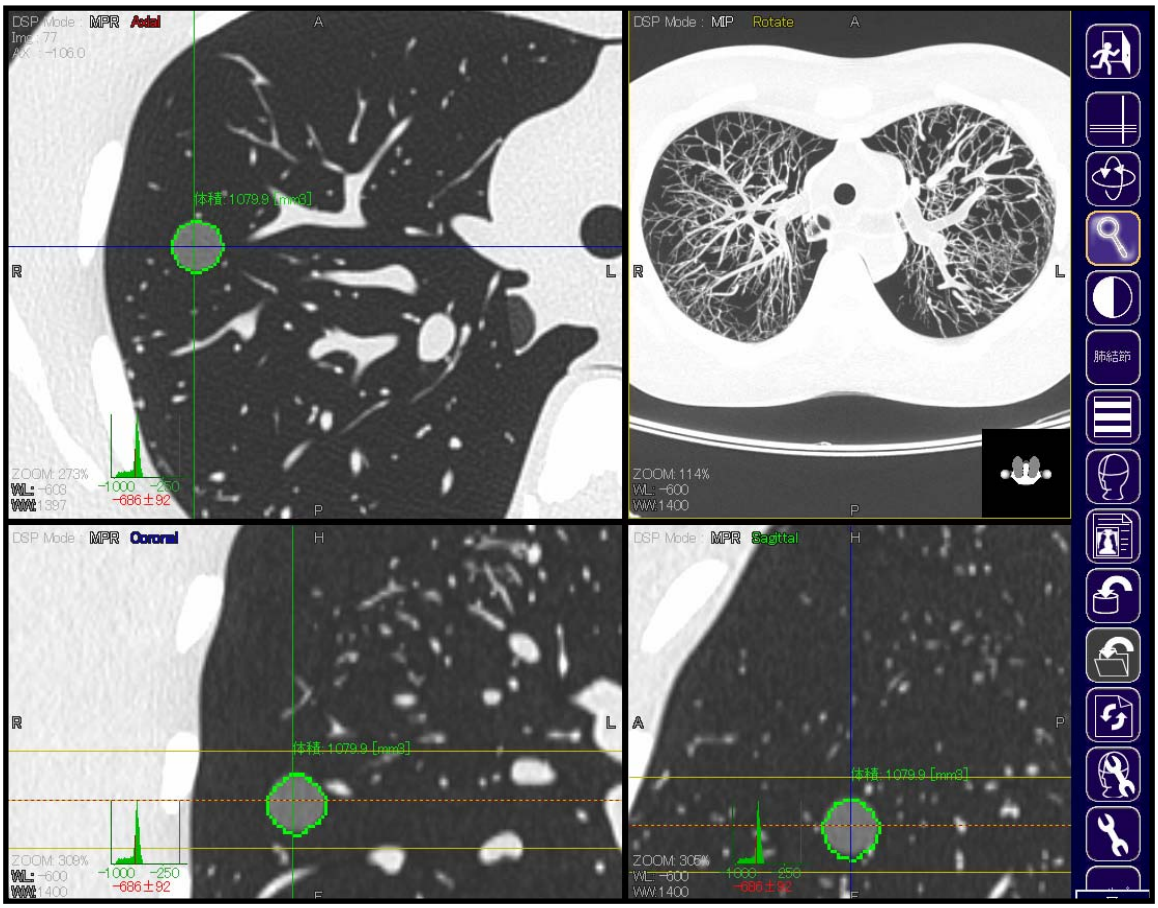


Fig. 18 Screen display of computer-aided volumetry software in the phantom study. Axial-, coronal-, and sagittal images and maximum intensity projections (MIPs) are displayed. The simulated GGO nodule (-630HU) can be accurately extracted by the software.

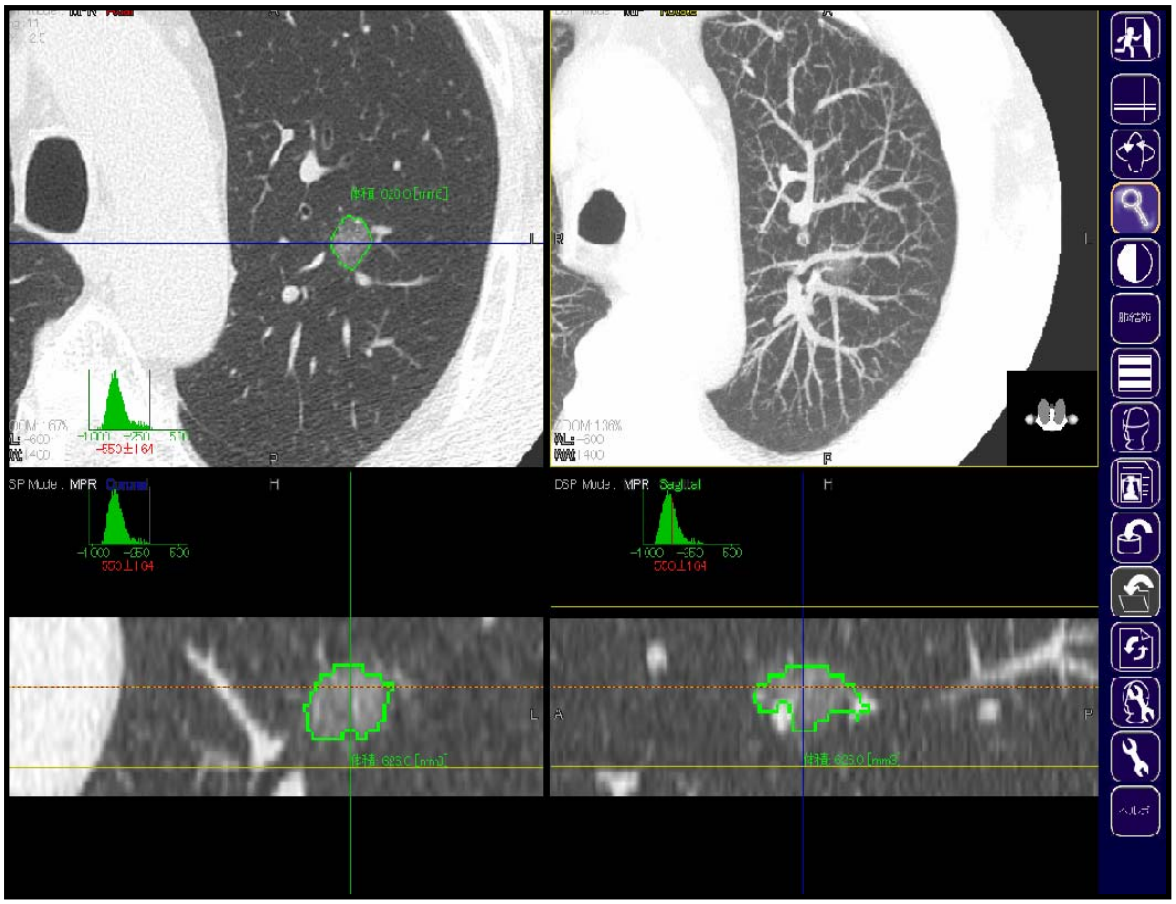


Fig. 19 Screen display of computer-aided volumetry software in the clinical study.

The software accurately extracts the pure GGO nodule (Category 1).

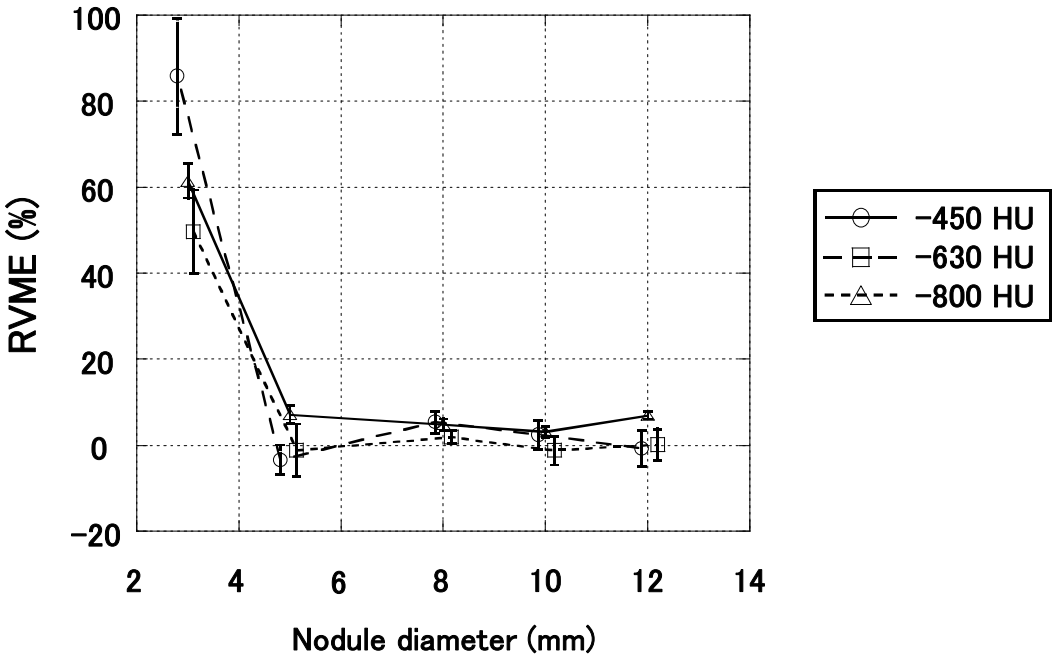
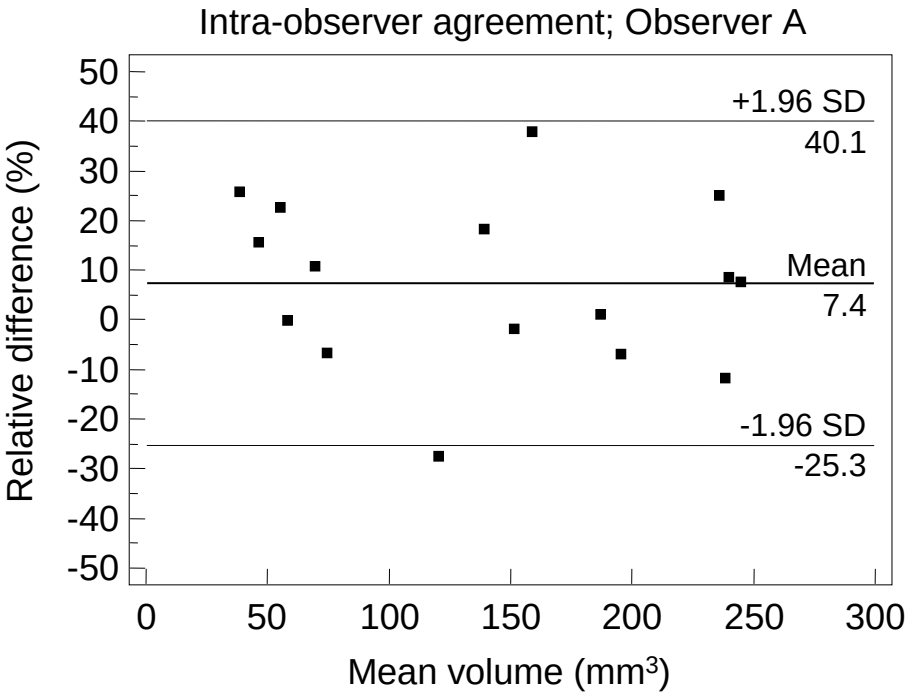
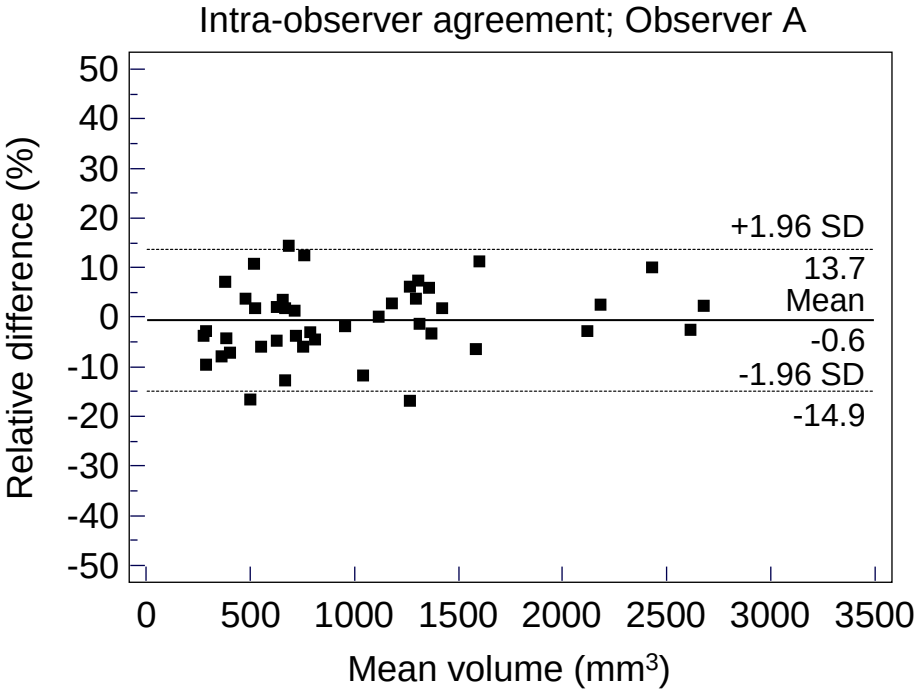


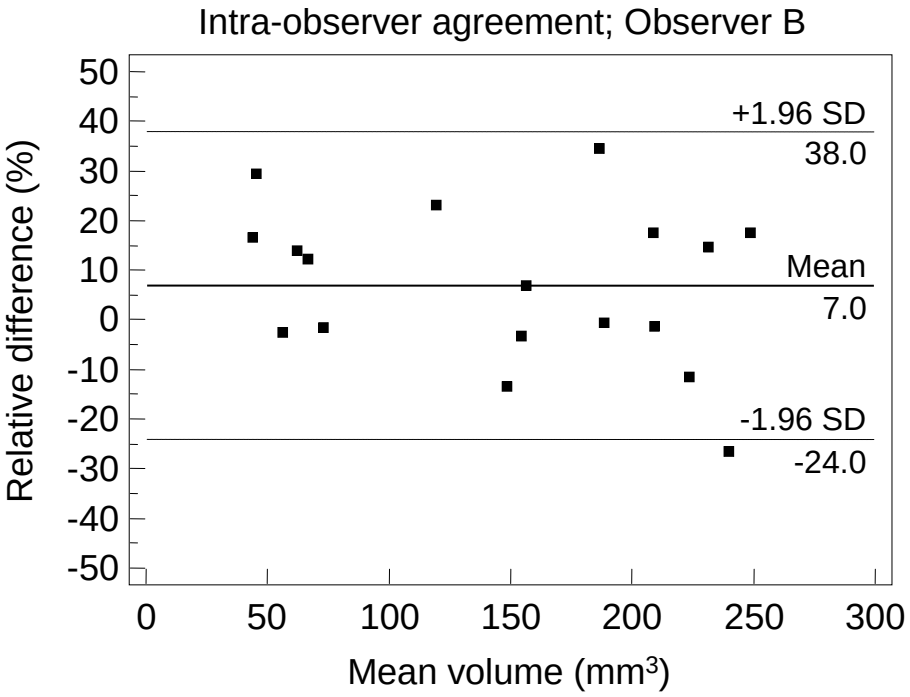
Fig. 20 Average relative volume measurement error (RVME) for nodules measuring 3-, 5-, 8-, 10-, and 12 mm in diameter with a CT number of -450-, -630, -800 Houndsfield units (HU). In nodules with a CT number of -450 HU, the RVME was 85.2 ± 13.5% (standard deviation), -4.1 ± 3.4%, 4.8 ± 2.6%, 1.7 ± 3.4%, and -1.5 ± 4.2%, respectively. In nodules with a CT number of -630 HU, the RVME was 51.1±9.8%, 0.2 ± 6.2%, 3.3 ± 1.4%, 0.2 ± 3.3%, and 1.6 ± 3.7%, respectively. In nodules with a CT number of -800 HU, the RVME was 61.5 ± 4.0%, 7.1 ±2.1%, 4.8 ± 1.4%, 3.1 ± 1.2%, and 7.0 ±0.9%, respectively.


Intra-observer agreement; Observer A
Relative difference (%)
50
40
30
20
10
0
-10
-20
-30
-40
-50
+1.96 SD
40.1
Mean
7.4
-1.96 SD
-25.3
0
50
100
150
200
250
300
Mean volume (mm3)

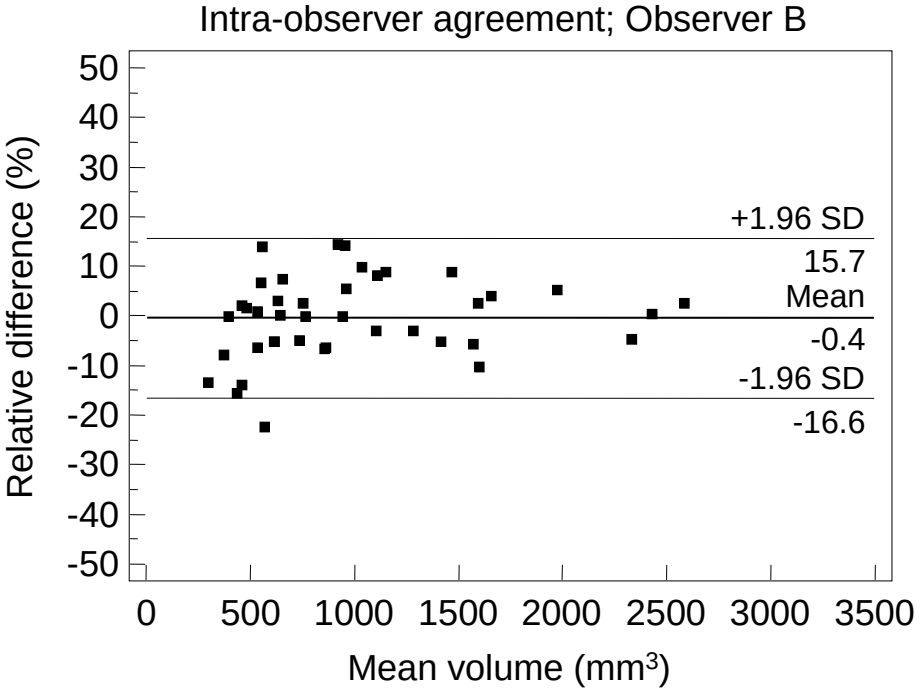
A


Intra-observer agreement; Observer A
Relative difference (%)
50
40
30
20
10
0
-10
-20
-30
-40
-50
+1.96 SD
13.7
Mean
-0.6
-1.96 SD
-14.9
0
500
1000
1500
2000
2500
3000
3500
Mean volume (mm3)

B


Intra-observer agreement; Observer B
Relative difference (%)
50
40
30
20
10
0
-10
-20
-30
-40
-50
+1.96 SD
38.0
Mean
7.0
-1.96 SD
-24.0
0
50
100
150
200
250
300
Mean volume (mm3)

C


Intra-observer agreement; Observer B
Relative difference (%)
50
40
30
20
10
0
-10
-20
-30
-40
-50
+1.96 SD
15.7
Mean
-0.4
-1.96 SD
-16.6
0
500
1000
1500
2000
2500
3000
3500
Mean volume (mm3)

D

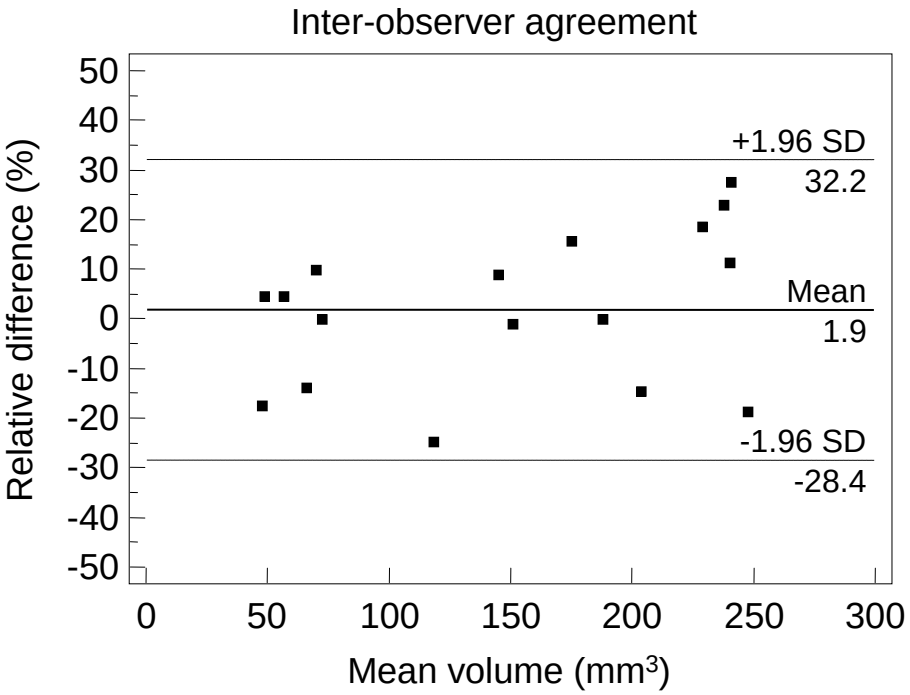
Fig. 21 Intra-observer agreement

A. Intra-observer agreement in nodules < 8mm in diameter for observer A. The mean intra-observer agreement was 7.4 % ± 16.7 % (95 % CI: -1.5 to 16.3 %). The 95 % limits of agreement were -25.3 % (95 % CI: -40.8 to -9.8 %) to 40.1 % (95 % CI: 24.6 to 55.6 %).

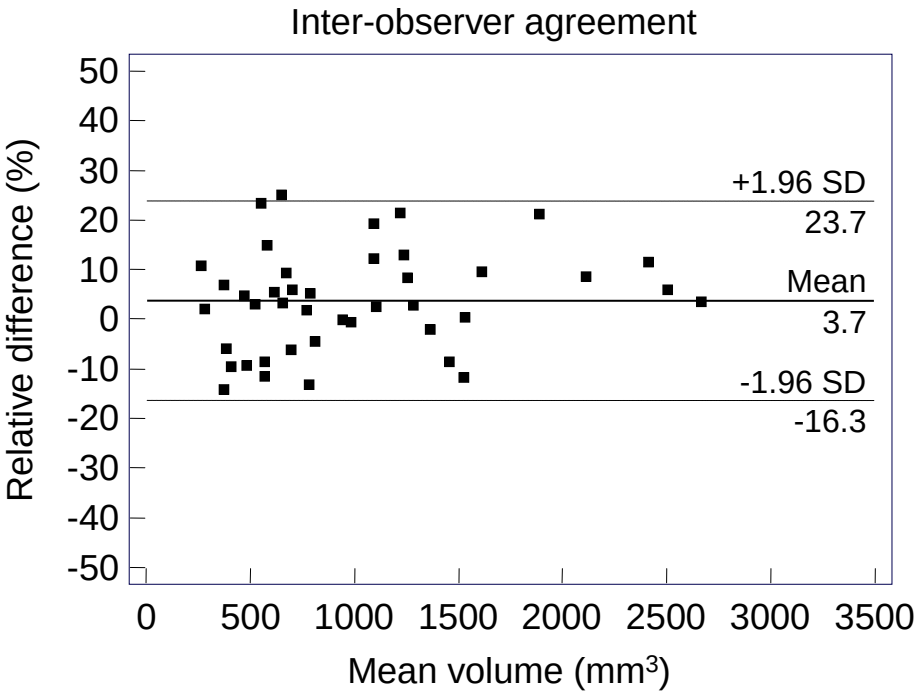
B. Intra-observer agreement in nodules ≥ 8mm in diameter for observer A. The mean intra-observer agreement was -0.6 % ± 7.3 % (95 % CI: -2.8 to 1.7 %). The 95 % limits of agreement were -14.9 % (95 % CI: -18.7 to -11.0) to 13.7 % (95 % CI: 9.8 to 17.6 %).

C. Intra-observer agreement in nodules < 8mm in diameter for observer B. The mean intra-observer agreement was 7.0 % ± 15.8 % (95 % CI: -0.9 to 14.9 %). The 95 % limits of agreement were -24.0 % (95 % CI: -37.7 to -10.3 %) to 38.0 % (95 % CI: 24.3 to 51.7 %).

D. Intra-observer agreement in nodules ≥ 8mm in diameter for observer B. The mean intra-observer agreement was -0.4 % ± 8.2 % (95 % CI: -3.0 to 2.2 %). The 95 % limits of agreement were -16.6 % (95 % CI: -21.1 to -12.1) to 15.7 % (95 % CI: 11.2 to 20.2 %).


Inter-observer agreement
Relative difference (%)
50
40
30
20
10
0
-10
-20
-30
-40
-50
+1.96 SD
32.2
Mean
1.9
-1.96 SD
-28.4
0
50
100
150
200
250
300
Mean volume (mm3)

A


Inter-observer agreement
Relative difference (%)
50
40
30
20
10
0
-10
-20
-30
-40
-50
+1.96 SD
23.7
Mean
3.7
-1.96 SD
-16.3
0
500
1000
1500
2000
2500
3000
3500
Mean volume (mm3)

B

Fig. 22 Inter-observer agreement

A. The mean inter-observer agreement in nodules < 8mm in diameter was 1.9 % ± 15.4 % (95 % CI: -6.0 to 9.8 %). The 95 % limits of agreement were -28.4 % (95 % CI: -42.2 to -14.5 %) to 32.2 % (95 % CI: 18.3 to 46.0 %).

B. The mean inter-observer agreement in nodules ≥ 8mm in diameter was 3.7 % ± 10.2 % (95 % CI: 0.5 to 6.9 %). The 95 % limits of agreement were -16.3 % (95 % CI: -21.8 to -10.8) to 23.7 % (95 % CI: 18.3 to 29.2 %).

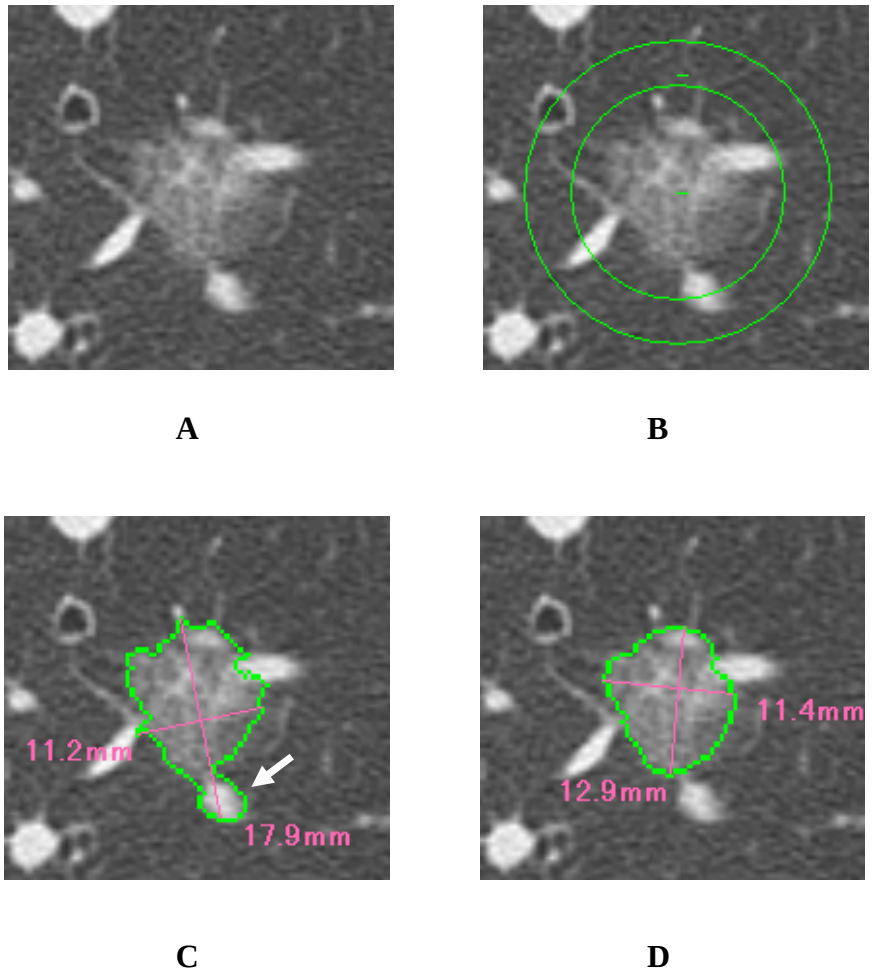


Fig. 23 Example of nodule volumetry on thin-section helical CT images.

A Target GGO nodule on HRCT.

B Radiologists manually specify the target GGO nodule and place a region of interest (ROI).

C The software automatically analyzes the density of the nodule and surrounding lung parenchyma and estimates the nodule border. Occasionally structures such as vessels remain around the nodule (arrow).

D Using the concept of mathematical morphology, radiologists subjectively modify the nodule border with a semi-automatic edit tool. Lastly, the software automatically calculates the nodule volume.

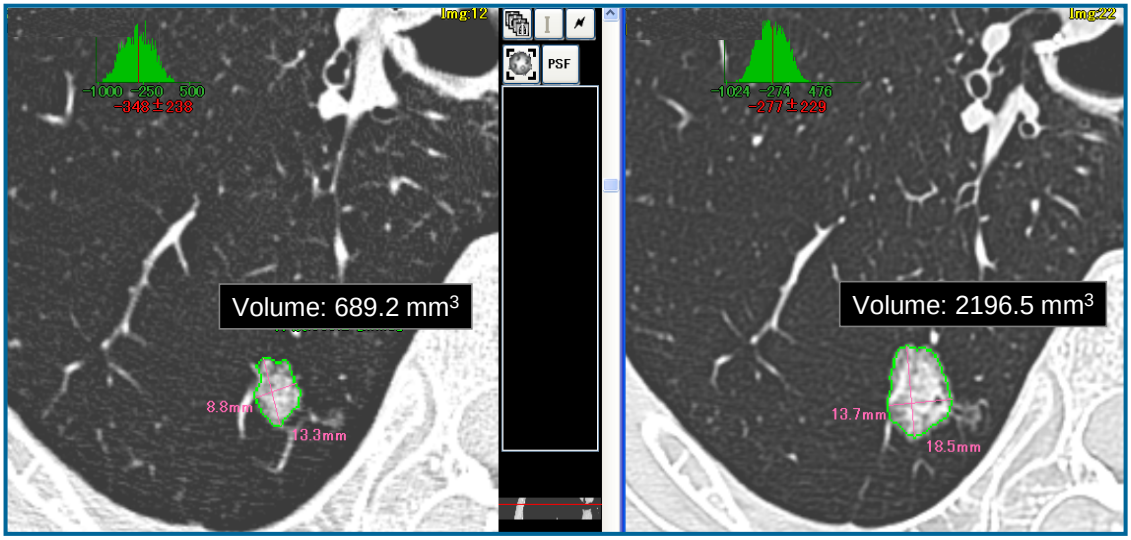


Fig. 24 VDT measurements of GGO nodule with 3D CAV software.

Adenocarcinoma with a GGO nodule in a 63-year-old woman. The initial- and final CT images are displayed on the left and right, respectively. The interval between the two CT scans was 270 days. The software automatically calculates the VDT by comparing both scans. In this case, the VDT was 163 days.

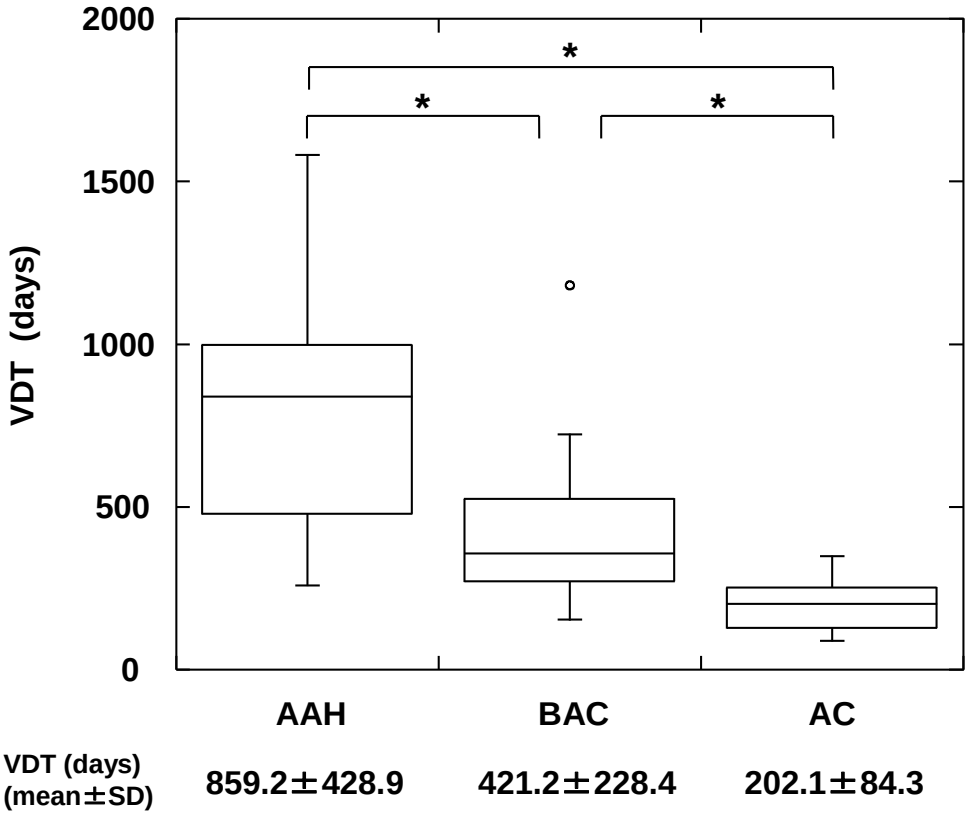


Fig. 25 Graph showing the mean VDT for AAH, BAC and AC. There were statistically significant differences in the mean VDT for all combinations of AAH, BAC, and AC (Steel-Dwass test). The upper (lower) end of vertical lines, upper (lower) margin of boxes, horizontal lines in boxes, and circular symbols represent upper (lower) extremes, upper (lower) quartiles, medians, and outliers of data, respectively. *$p < 0.01$.

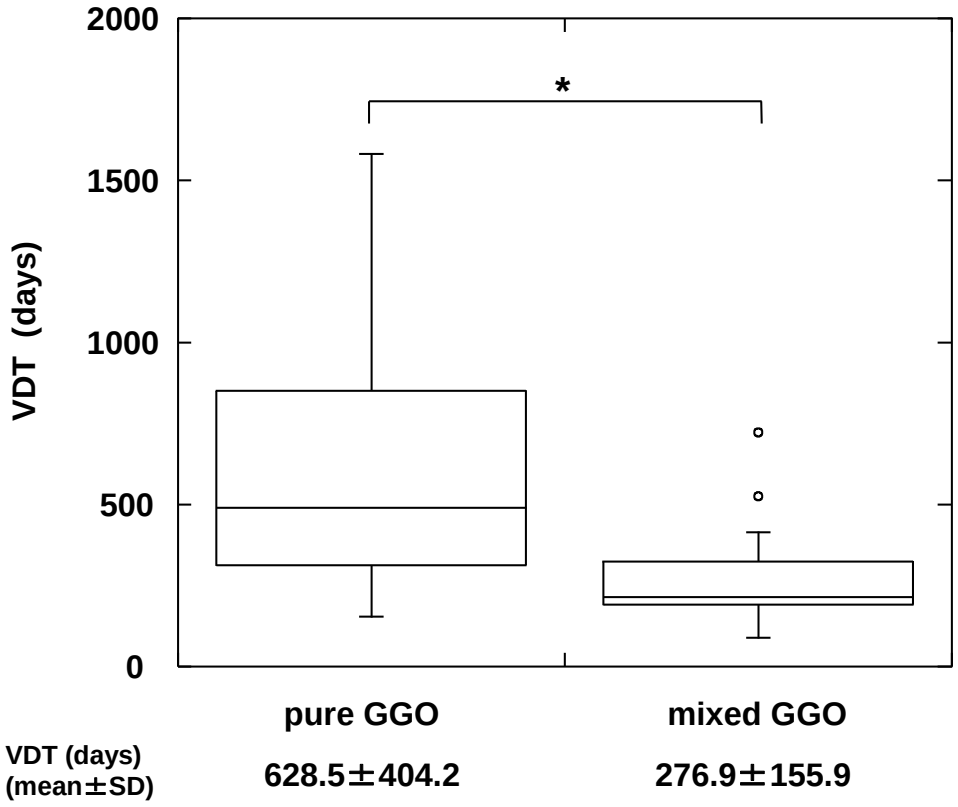


Fig. 26 Graph showing the mean VDT for pure- and mixed GGO nodules. VDT was significantly shorter for mixed- than pure GGO nodules (Mann-Whitney U-test). See Fig. 3 for explanation of symbols. *$p < 0.01$.

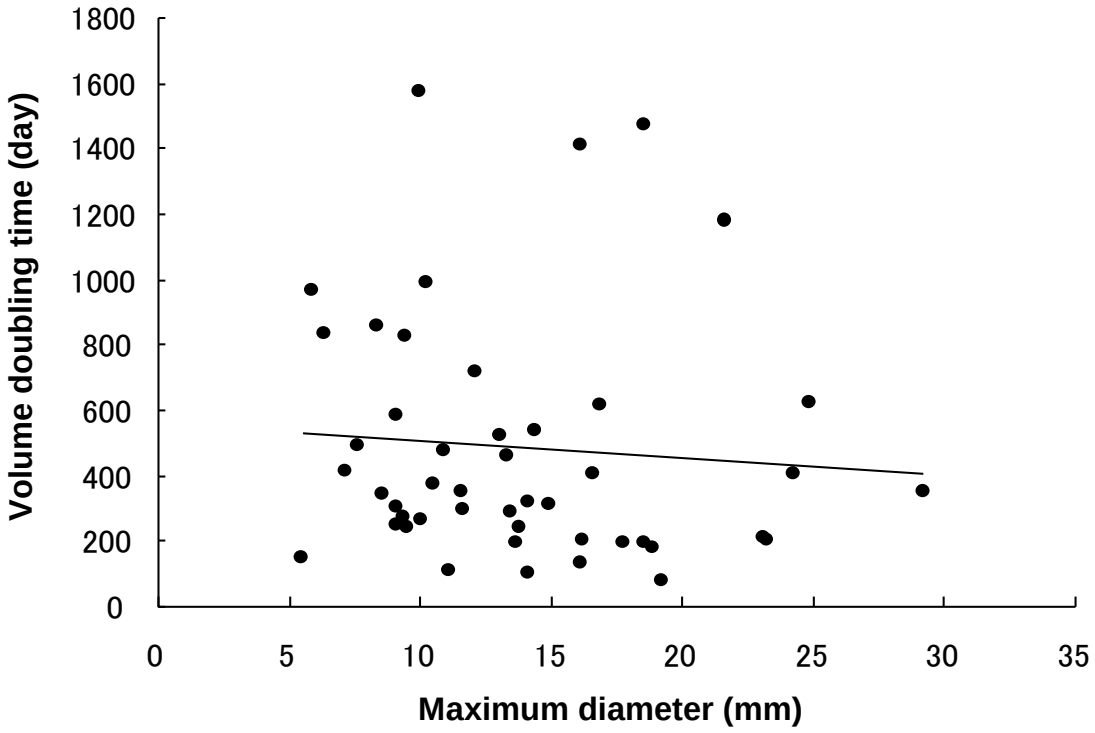


Fig. 27 Graph showing the correlation between the VDT and the maximum diameter- and between the VDT and the mean density of GGO nodules on the initial CT scan. No correlation was found ($r = -0.19$, $p = 0.19$).

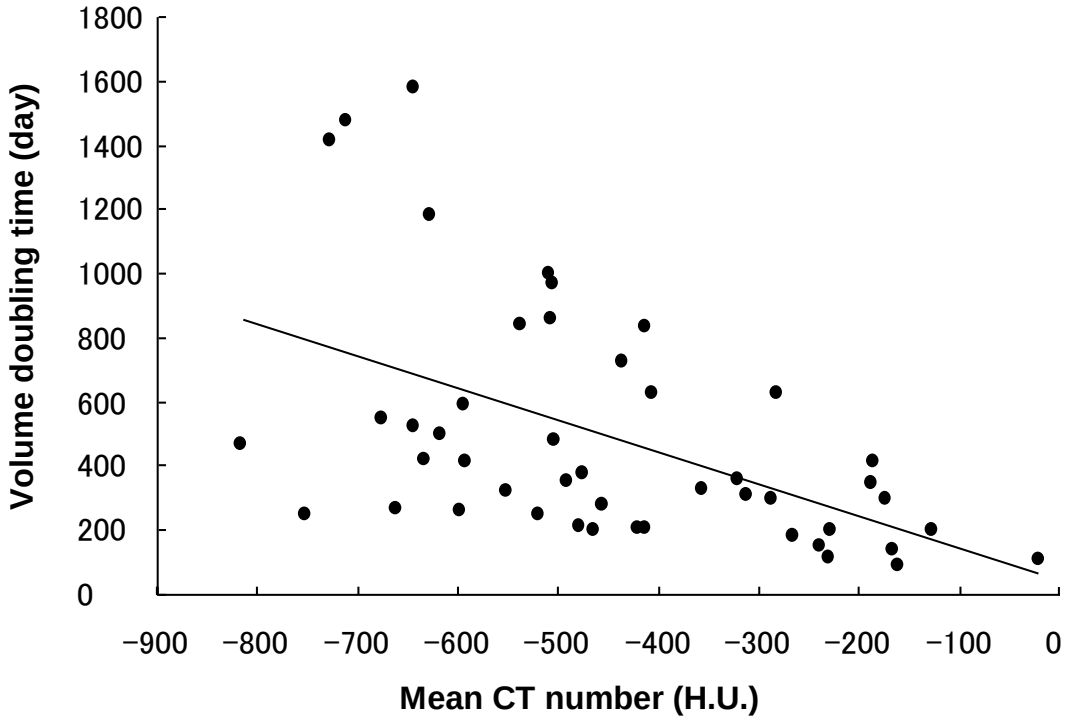


Fig. 28 Graph showing the correlation between the VDT and the mean internal density of GGO nodules on the initial CT scan. There was a statistically significant correlation ($r = -0.57$, $p < 0.01$).

Table lists and tables

Table 11. Edit time for GGO nodules according to morphological subtypes

Table 12. Clinical and pathologic Characteristics of All 47 GGO Nodules

Table 1. AUC values for all observers for the detection of simulated nodules without and with DES images

	AUC value	
Observer	Without DES images	With DES images
1	0.74	0.77
2	0.71	0.82
3	0.79	0.82
4	0.86	0.87
5	0.83	0.84
6	0.74	0.80
7	0.63	0.66
Mean ± SD	0.76 ± 0.08	0.79 ± 0.07

AUC, The area under the best-fit ROC curve; DES, dual-energy subtraction; SD, standard deviation

Table 2. Results of nodule detection for all observers without and with DES images

	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
All nodules					
Standard image	48.8	83.6	65.4	79.5	58.9
DES image	63.2*	83.6	73.4*	82.6	70.6*
Nodules with -450 HU					
Standard image	39.7	81.7	60.7	71.2	57.8
DES image	52.4	84.1	68.3*	81.1	65.4
Nodules with -200 HU					
Standard image	51.6	85.7	68.7	84.0	63.5
DES image	55.6	85.7	70.6	88.4	67.1
Nodules with 30 HU					
Standard image	76.2	83.3	79.8	81.6	79.2
DES image	80.2	81.0	80.6	81.8	82.1

*Statistically significant difference compared with the value of the corresponding standard image

PPV, positive predictive value; NPV, negative predictive value; DES, dual-energy subtraction

Table 3. AUC values for all observers for the detection of pulmonary nodules without and with DES images

	AUC value	
Observer	Without DES images	With DES images
1	0.591	0.625
2	0.633	0.641
3	0.546	0.598
4	0.589	0.734
5	0.68	0.685
6	0.690	0.709
7	0.625	0.739
8	0.605	0.698
Mean ± SD	0.620 ± 0.048	0.679 ± 0.052

Table 4. AUC values for each nodule type without and with DES images

Nodule type	AUC value		p value
	Without DES images	With DES images	
Non-solid nodule	0.617 ± 0.101	0.608 ± 0.093	0.725
Part-solid nodule	0.611 ± 0.068	0.689 ± 0.053	< 0.001
Solid nodule	0.750 ± 0.098	0.781 ± 0.078	0.225

Table 5. AUC values for nodule location without and with DES images

Nodule location	AUC value		p value
	Without DES images	With DES images	
With overlapping by bone shadows	0.664 ± 0.051	0.721 ± 0.054	0.025
Without overlapping	0.584 ± 0.055	0.617 ± 0.055	0.255

Table 6. Results for nodule detection for all observers without and with DES images

	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)
Standard image	47.6	72.5	56.7	75.6	44.4
DES image	63.0	72.5	66.5	80.1	53.7
	$p < 0.01$	$p = 0.89$	$p = 0.02$	$p = 0.12$	$p = 0.02$

Table 7. AUC values for board-certified radiologists and radiology residents for the detection of pulmonary nodules without and with rib-suppressed images

Observer	AUC Value	
	Without RSI	With RSI
Board-cerified Radiologists		
1	0.859	0.962
2	0.936	0.880
3	0.856	0.91
4	0.896	0.821
5	0.773	0.820
6	0.841	0.913
7	0.775	0.890
Mean±SD	0.848±0.059	0.883±0.050
Radiology Residents		
11	0.834	0.810
12	0.669	0.686
13	0.872	0.848
14	0.741	0.844
15	0.735	0.751
Mean±SD	0.770±0.081	0.788±0.069
All	0.816±0.077	0.843±0.074

RSI: rib-suppressed image

Table 8. Comparison of intra- and inter-observer agreement

	Mean (95%CI)	Upper limit (95%CI)	Lower limit (95%CI)
All nodules			
Intra-observer agreement for observer A	1.6% (-1.3 – 4.5%)	23.3% (18.4 – 28.3%)	-20.2% (-25.2 – -15.2%)
Intra-observer agreement for observer B	1.8% (-1.2 – 4.8%)	24.4% (19.2 – 29.5%)	-20.7% (-25.8 – -15.5%)
Inter-observer agreement	3.2% (0.1 – 6.3%)	26.4% (21.1 – 31.7%)	-20.0% (-25.3 – -14.7%)
Nodules < 8mm in diameter			
Intra-observer agreement for observer A	7.4% (-1.5 – 16.3%)	40.1% (24.6 – 55.6%)	-25.3% (-40.8. – -9.8%)
Intra-observer agreement for observer B	7.0% (-0.9 – 14.9%)	38.0% (24.3 – 51.7%)	-24.0% (-37.7 – -10.3%)
Inter-observer agreement	1.9% (-6.0 – 9.8%)	32.2% (18.3 – 46.0%)	-28.4% (-42.2 – -14.5%)
Nodules ≧ 8mm in diameter			
Intra-observer agreement for observer A	-0.6% (-2.8 – 1.7%)	13.7% (9.8 – 17.6%)	-14.9% (-18.7 – -11.0%)
Intra-observer agreement for observer B	-0.4% (-3.0 – 2.2%)	15.7% (11.2 – 20.2%)	-16.6% (-21.1 – -12.1%)
Inter-observer agreement	3.7% (0.5 – 6.9%)	23.7% (18.3 – 29.2%)	-16.3% (-21.8 – -10.8%)

Table 9. Influence of editing on intra- and inter-observer agreement

	Mean (95%CI)	Upper limit (95%CI)	Lower limit (95%CI)
Nodules without editing			
Intra-observer agreement for observer A	2.3% (-1.2 – 5.9%)	18.3% (12.2 – 24.4%)	-13.6% (-19.7 – -7.5%)
Intra-observer agreement for observer B	0.1% (-3.1 – 3.3%)	17.3% (11.7 – 22.8%)	-17.1% (-22.6 – -11.5%)
Inter-observer agreement	3.5% (-0.3 – 7.2%)	20.1% (13.6 – 26.6%)	-13.1% (-19.6 – -6.6%)
Nodules with editing			
Intra-observer agreement for observer A	1.1% (-3.2 – 5.4%)	26.0% (18.6 – 33.5%)	-23.9% (-31.3 – -16.4%)
Intra-observer agreement for observer B	3.8% (-1.6 – 9.1%)	30.8% (21.6 – 40.1%)	-23.3% (-32.6 – -14.1%)
Inter-observer agreement	2.7% (-2.8 – 8.2%)	28.2% (18.7 – 37.8%)	-22.9% (-32.4 – -13.4%)

Table 10. Edit rate for GGO nodules according to morphological subtypes

GGO subtype	Observer A		Observer B	
	Edit rate	*p* value	Edit rate	*p* value
Margin characteristics				
Category 1 (22)	54.5% (12)	*p*=0.03	31.8% (7)	*p*=0.04
Category 2 (24)	50.0% (12)		45.8% (11)	
Category 3 (13)	92.3% (12)		76.9% (10)	
Internal density				
Non-solid (28)	53.6% (15)	*p*=0.40	53.6% (15)	*p*=0.53
Part-solid (31)	67.7% (21)		41.9% (13)	
Relationship with pleura				
Adjacent to pleura (12)	91.7% (11)	*p*=0.04	83.3% (9)	*p*=0.03
Not adjacent to pleura (47)	53.2% (25)		38.3% (18)	
Total (n=59)	61.0% (36)		47.5% (28)	

Note: numbers in parentheses indicate number of patients.

Table 11. Edit time for GGO nodules according to morphological subtypes

	Observer A		Observer B	
GGO subtype	Edit time (sec)	*p* value	Edit time (sec)	*p* value
Margin characteristics				
Category 1 (22)	17.5±5.2	*p*=0.92	29.6±9.9	*p*=0.61
Category 2 (24)	20.4±9.9	*p*<0.01	31.4±11.6	*p*<0.01
Category 3 (13)	33.0±7.4	p<0.01	43.6±10.7	
Internal density				
Non-solid (28)	22.2±8.8	*p*=0.36	29.5±8.8	*p*=0.01
Part-solid (31)	25.4±10.9		41.9±12.7	
Relationship with pleura				
Adjacent to pleura (12)	24.8±8.0	*p*=0.77	32.3±9.9	*p*=0.35
Not adjacent to pleura (47)	23.7±11.0		36.3±13.4	
Total (n=59)	24.1±10.0		35.3±12.3	

Table 12. Clinical and pathologic Characteristics of All 47 GGO Nodules

	AAH	BAC	AC
Nodules（Patients）	13 (9)	22 (20)	12 (10)
Patient sex			
Male	6	6	3
Female	7	16	9
Mean age (y)	58.8 ± 13.1	63.9 ± 9.0	67.4 ± 6.1
GGO subtype			
pure-GGO	13	14	1
mixed-GGO	0	8	11
Mean maximum diameter (mm)	10.1 ± 3.2	13.3 ± 5.6	13.4 ± 4.6
Mean CT attenuation (HU)	-614.5 ± 98.4	-611.4 ± 123.0	-323.3 ± 210.7
Mean interval between two CT scans (days)	331.6 ± 436.9	123.9 ± 118.5	85.9 ± 39.8

Note- AAH = atypical adenomatous hyperplasia, BAC = bronchioloalveolar carcinoma

AC = adenocarcinoma

References

1. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med 1997; 157:849-855.

2. McWilliams AM, Mayo JR, Ahn MI, MacDonald SL, Lam SC. Lung cancer screening using multi-slice thin-section computed tomography and autofluorescence bronchoscopy. J Thorac Oncol 2006; 1:61-68.

3. Henschke CI, Yankelevitz DF, Naidich DP, et al. CT screening for lung cancer: suspiciousness of nodules according to size on baseline scans. Radiology 2004; 231:164-168.

4. Henschke CI, Yankelevitz DF, Mirtcheva R, McGuinness G, McCauley D, Miettinen OS. CT screening for lung cancer: frequency and significance of part-solid and nonsolid nodules. AJR Am J Roentgenol 2002; 178:1053-1057.

5. Li F, Sone S, Abe H, MacMahon H, Armato SG, 3rd, Doi K. Lung cancers missed at low-dose helical CT screening in a general population: comparison of clinical, histopathologic, and imaging findings. Radiology 2002; 225:673-683.

6. Nakata M, Saeki H, Takata I, et al. Focal ground-glass opacity detected by low-dose helical CT. Chest 2002; 121:1464-1467.

7. Nakajima R, Yokose T, Kakinuma R, Nagai K, Nishiwaki Y, Ochiai A. Localized pure ground-glass opacity on high-resolution CT: histologic characteristics. J Comput Assist Tomogr 2002; 26:323-329.

8. Noguchi M, Morikawa A, Kawasaki M, et al. Small adenocarcinoma of the lung. Histologic characteristics and prognosis. Cancer 1995; 75:2844-2852.

9. Park CM, Goo JM, Lee HJ, et al. Focal interstitial fibrosis manifesting as nodular ground-glass opacity: thin-section CT findings. Eur Radiol 2007; 17:2325-2331.

10. Kuriyama K, Seto M, Kasugai T, et al. Ground-glass opacity on thin-section CT: value in differentiating subtypes of adenocarcinoma of the lung. AJR Am J Roentgenol 1999; 173:465-469.

11. Kaneko M, Eguchi K, Ohmatsu H, et al. Peripheral lung cancer: screening and detection with low-dose spiral CT versus radiography. Radiology 1996; 201:798-802.

12. Quekel LG, Kessels AG, Goei R, van Engelshoven JM. Miss rate of lung cancer on the chest radiograph in clinical practice. Chest 1999; 115:720-724.

13. Sone S, Takashima S, Li F, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. Lancet 1998; 351:1242-1245.

14. Woodring JH. Pitfalls in the radiologic diagnosis of lung cancer. AJR Am J Roentgenol 1990; 154:1165-1175.

15. Tsubamoto M, Kuriyama K, Kido S, et al. Detection of lung cancer on chest radiographs: analysis on the basis of size and extent of ground-glass opacity at thin-section CT. Radiology 2002; 224:139-144.

16. Shah PK, Austin JH, White CS, et al. Missed non-small cell lung cancer: radiographic findings of potentially resectable lesions evident only in retrospect. Radiology 2003; 226:235-241.

17. Ide K, Mogami H, Murakami T, Yasuhara Y, Miyagawa M, Mochizuki T. Detection of lung cancer using single-exposure dual-energy subtraction chest radiography. Radiat Med 2007; 25:195-201.

18. Uemura M, Miyagawa M, Yasuhara Y, et al. Clinical evaluation of pulmonary nodules with dual-exposure dual-energy subtraction chest radiography. Radiat Med 2005; 23:391-397.

19. Bacher K, Smeets P, Bonnarens K, De Hauwere A, Verstraete K, Thierens H. Dose reduction in patients undergoing chest imaging: digital amorphous silicon flat-panel detector radiography versus conventional film-screen radiography and phosphor-based computed radiography. AJR Am J Roentgenol 2003; 181:923-

929.

20. Kelcz F, Zink FE, Peppler WW, Kruger DG, Ergun DL, Mistretta CA. Conventional chest radiography vs dual-energy computed radiography in the detection and characterization of pulmonary nodules. AJR Am J Roentgenol 1994; 162:271-278.

21. Tagashira H, Arakawa K, Yoshimoto M, Mochizuki T, Murase K. Detectability of lung nodules using flat panel detector with dual energy subtraction by two shot method: evaluation by ROC method. Eur J Radiol 2007; 64:279-284.

22. Ricke J, Fischbach F, Freund T, et al. Clinical results of CsI-detector-based dual-exposure dual energy in chest radiography. Eur Radiol 2003; 13:2577-2582.

23. Suzuki K, Armato SG, 3rd, Li F, Sone S, Doi K. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. Med Phys 2003; 30:1602-1617.

24. Arimura H, Katsuragawa S, Suzuki K, et al. Computerized scheme for automated detection of lung nodules in low-dose computed tomography images for lung cancer screening. Acad Radiol 2004; 11:617-629.

25. Suzuki K, Abe H, MacMahon H, Doi K. Image-processing technique for

suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). IEEE Trans Med Imaging 2006; 25:406-416.

26. Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999; 354:99-105.

27. Henschke CI, Naidich DP, Yankelevitz DF, et al. Early lung cancer action project: initial findings on repeat screenings. Cancer 2001; 92:153-159.

28. Sobue T, Moriyama N, Kaneko M, et al. Screening for lung cancer with low-dose helical computed tomography: anti-lung cancer association project. J Clin Oncol 2002; 20:911-920.

29. Li F, Sone S, Abe H, Macmahon H, Doi K. Malignant versus benign nodules at CT screening for lung cancer: comparison of thin-section CT findings. Radiology 2004; 233:793-798.

30. Swensen SJ. CT screening for lung cancer. AJR Am J Roentgenol 2002; 179:833-836.

31. McWilliams A, Mayo J, MacDonald S, et al. Lung cancer screening: a different paradigm. Am J Respir Crit Care Med 2003; 168:1167-1173.

32. Kim HY, Shim YM, Lee KS, Han J, Yi CA, Kim YK. Persistent pulmonary

nodular ground-glass opacity at thin-section CT: histopathologic comparisons. Radiology 2007; 245:267-275.

33. Hasegawa M, Sone S, Takashima S, et al. Growth rate of small lung cancers detected on mass CT screening. Br J Radiol 2000; 73:1252-1259.

34. Aoki T, Nakata H, Watanabe H, et al. Evolution of peripheral lung adenocarcinomas: CT findings correlated with histology and tumor doubling time. AJR Am J Roentgenol 2000; 174:763-768.

35. Revel MP, Bissery A, Bienvenu M, Aycard L, Lefort C, Frija G. Are two-dimensional CT measurements of small noncalcified pulmonary nodules reliable? Radiology 2004; 231:453-458.

36. Yankelevitz DF, Reeves AP, Kostis WJ, Zhao B, Henschke CI. Small pulmonary nodules: volumetrically determined growth rates based on CT evaluation. Radiology 2000; 217:251-256.

37. Das M, Ley-Zaporozhan J, Gietema HA, et al. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. Eur Radiol 2007.

38. Wormanns D, Kohl G, Klotz E, et al. Volumetric measurements of pulmonary nodules at multi-row detector CT: in vivo reproducibility. Eur Radiol 2004; 14:86-92.

39. Kostis WJ, Reeves AP, Yankelevitz DF, Henschke CI. Three-dimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images. IEEE Trans Med Imaging 2003; 22:1259-1274.

40. Ko JP, Rusinek H, Jacobs EL, et al. Small pulmonary nodules: volume measurement at chest CT--phantom study. Radiology 2003; 228:864-870.

41. Usuda K, Saito Y, Sagawa M, et al. Tumor doubling time and prognostic assessment of patients with primary lung cancer. Cancer 1994; 74:2239-2244.

42. Schwartz M. A biomathematical approach to clinical tumor growth. Cancer 1961; 14:1272-1294.

43. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst 2000; 92:205-216.

44. Therasse P, Eisenhauer EA, Verweij J. RECIST revisited: a review of validation studies on tumour assessment. Eur J Cancer 2006; 42:1031-1039.

45. Hayes TG, Falchook GF, Varadhachary GR, et al. Phase I trial of oral talactoferrin alfa in refractory solid tumors. Invest New Drugs 2006; 24:233-240.

46. International Atomic Energy Agency. International basic safety standards for

protection against ionizing radiation and for the safety of radiation sources. IAEA safety series no. 115. Vienna, Austria: International Atomic Energy Agency, 1996.

47. Metz CE, Herman BA, Shen JH. Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Stat Med 1998; 17:1033-1053.

48. Shiraishi J, Katsuragawa S, Ikezoe J, et al. Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. AJR Am J Roentgenol 2000; 174:71-74.

49. Ruhl R, Wozniak MM, Werk M, et al. CsI-detector-based dual-exposure dual energy in chest radiography for lung nodule detection: results of an international multicenter trial. Eur Radiol 2008; 18:1831-1839.

50. Oda S, Awai K, Liu D, et al. Ground-glass opacities on thin-section helical CT: differentiation between bronchioloalveolar carcinoma and atypical adenomatous hyperplasia. AJR Am J Roentgenol 2008; 190:1363-1368.

51. Floyd CE, Jr., Warp RJ, Dobbins JT, 3rd, et al. Imaging characteristics of an amorphous silicon flat-panel detector for digital chest radiography. Radiology

2001; 218:683-688.

52. Hamer OW, Volk M, Zorger Z, Feuerbach S, Strotzer M. Amorphous silicon, flat-panel, x-ray detector versus storage phosphor-based computed radiography: contrast-detail phantom study at different tube voltages and detector entrance doses. Invest Radiol 2003; 38:212-220.

53. Kido S, Ikezoe J, Naito H, et al. Single-exposure dual-energy chest images with computed radiography. Evaluation with simulated pulmonary nodules. Invest Radiol 1993; 28:482-487.

54. Kido S, Ikezoe J, Naito H, et al. Clinical evaluation of pulmonary nodules with single-exposure dual-energy subtraction chest radiography with an iterative noise-reduction algorithm. Radiology 1995; 194:407-412.

55. Oda S, Awai K, Suzuki K, et al. Performance of radiologists in detection of small pulmonary nodules on chest radiographs: effect of rib suppression with a massive-training artificial neural network. AJR Am J Roentgenol 2009; 193:W397-402.

56. Fontana RS, Sanderson DR, Taylor WF, et al. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Mayo Clinic study. Am Rev Respir Dis 1984; 130:561-565.

57. Hakansson M, Bath M, Borjesson S, et al. Nodule detection in digital chest radiography: summary of the RADIUS chest trial. Radiat Prot Dosimetry 2005; 114:114-120.

58. Naidich D, Zerhouni E, Siegelman S. Computed tomography of the thorax. New York: Raven, 1984.

59. Ishigaki T, Sakuma S, Ikeda M. One-shot dual-energy subtraction chest imaging with computed radiography: clinical evaluation of film images. Radiology 1988; 168:67-72.

60. Kakeda S, Nakamura K, Kamada K, et al. Improved detection of lung nodules by using a temporal subtraction technique. Radiology 2002; 224:145-151.

61. Suzuki K, Shiraishi J, Abe H, MacMahon H, Doi K. False-positive reduction in computer-aided diagnostic scheme for detecting nodules in chest radiographs by means of massive training artificial neural network. Acad Radiol 2005; 12:191-201.

62. Suzuki K, Li F, Sone S, Doi K. Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network. IEEE Trans Med Imaging 2005; 24:1138-1150.

63. Suzuki K, Yoshida H, Nappi J, Dachman AH. Massive-training artificial neural network (MTANN) for reduction of false positives in computer-aided detection of polyps: Suppression of rectal tubes. Med Phys 2006; 33:3814-3824.

64. Suzuki K, Yoshida H, Nappi J, Armato SG, 3rd, Dachman AH. Mixture of expert 3D massive-training ANNs for reduction of multiple types of false positives in CAD for detection of polyps in CT colonography. Med Phys 2008; 35:694-703.

65. Li F, Engelmann R, Metz CE, Doi K, MacMahon H. Lung cancers missed on chest radiographs: results obtained with a commercial computer-aided detection program. Radiology 2008; 246:273-280.

66. Kuhlman JE, Collins J, Brooks GN, Yandow DR, Broderick LS. Dual-energy subtraction chest radiography: what to look for beyond calcified nodules. Radiographics 2006; 26:79-92.

67. Li F, Engelmann R, Doi K, MacMahon H. Improved detection of small lung cancers with dual-energy subtraction chest radiography. AJR Am J Roentgenol 2008; 190:886-891.

68. Tsubamoto M, Johkoh T, Kozuka T, et al. Temporal subtraction for the detection of hazy pulmonary opacities on chest radiography. AJR Am J Roentgenol 2002; 179:467-471.

69. Kakeda S, Moriya J, Sato H, et al. Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system. AJR Am J Roentgenol 2004; 182:505-510.

70. Aoki T, Tomoda Y, Watanabe H, et al. Peripheral lung adenocarcinoma: correlation of thin-section CT findings with histologic prognostic factors and survival. Radiology 2001; 220:803-809.

71. Kakinuma R, Ohmatsu H, Kaneko M, et al. Progression of focal pure ground-glass opacity detected by low-dose helical computed tomography screening for lung cancer. J Comput Assist Tomogr 2004; 28:17-23.

72. Awai K, Murao K, Ozawa A, et al. Pulmonary nodules: estimation of malignancy at thin-section helical CT--effect of computer-aided diagnosis on performance of radiologists. Radiology 2006; 239:276-284.

73. Dougherty E. Digital image processing method. New York: Dekker, 1994.

74. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307-310.

75. de Hoop B, Gietema H, van Ginneken B, Zanen P, Groenewegen G, Prokop M. A comparison of six software packages for evaluation of solid lung nodules using semi-automated volumetry: What is the minimum increase in size to detect

growth in repeated CT examinations. Eur Radiol 2008.

76. Goodman LR, Gulsun M, Washington L, Nagy PG, Piacsek KL. Inherent variability of CT lung nodule measurements in vivo using semiautomated volumetric measurements. AJR Am J Roentgenol 2006; 186:989-994.

77. Winer-Muram HT, Jennings SG, Meyer CA, et al. Effect of varying CT section width on volumetric measurement of lung tumors and application of compensatory equations. Radiology 2003; 229:184-194.

78. Kostis WJ, Yankelevitz DF, Reeves AP, Fluture SC, Henschke CI. Small pulmonary nodules: reproducibility of three-dimensional volumetric measurement and estimation of time to follow-up CT. Radiology 2004; 231:446-452.

79. Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C. Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. Eur Radiol 2006; 16:781-790.

80. Das M, Ley-Zaporozhan J, Gietema HA, et al. Accuracy of automated volumetry of pulmonary nodules across different multislice CT scanners. Eur Radiol 2007; 17:1979-1984.

81. Ikeda K, Awai K, Mori T, Kawanaka K, Yamashita Y, Nomori H. Differential diagnosis of ground-glass opacity nodules: CT number analysis by three-dimensional computerized quantification. Chest 2007; 132:984-990.

82. Das M, Muhlenbruch G, Katoh M, et al. Automated volumetry of solid pulmonary nodules in a phantom: accuracy across different CT scanner technologies. Invest Radiol 2007; 42:297-302.

83. Goo JM, Tongdee T, Tongdee R, Yeo K, Hildebolt CF, Bae KT. Volumetric measurement of synthetic lung nodules with multi-detector row CT: effect of various image reconstruction parameters and segmentation thresholds on measurement accuracy. Radiology 2005; 235:850-856.

84. Gietema HA, Schaefer-Prokop CM, Mali WP, Groenewegen G, Prokop M. Pulmonary nodules: Interscan variability of semiautomated volume measurements with multisection CT-- influence of inspiration level, nodule size, and segmentation performance. Radiology 2007; 245:888-894.

85. Bolte H, Riedel C, Jahnke T, et al. Reproducibility of computer-aided volumetry of artificial small pulmonary nodules in ex vivo porcine lungs. Invest Radiol 2006; 41:28-35.

86. Oda S, Awai K, Murao K, et al. Computer-Aided Volumetry of Pulmonary

Nodules Showing Ground Glass Opacity at Multidetector Computed Tomography. *AJR* (in press)

87. Lillington GA. Management of solitary pulmonary nodules. Dis Mon 1991; 37:271-318.

88. Hayabuchi N, Russell WJ, Murakami J. Slow-growing lung cancer in a fixed population sample. Radiologic assessments. Cancer 1983; 52:1098-1104.

89. Lindell RM, Hartman TE, Swensen SJ, et al. Five-year lung cancer screening experience: CT appearance, growth rate, location, and histologic features of 61 lung cancers. Radiology 2007; 242:555-562.

90. Yankelevitz DF, Kostis WJ, Henschke CI, et al. Overdiagnosis in chest radiographic screening for lung carcinoma: frequency. Cancer 2003; 97:1271-1275.

91. Muhm JR, Miller WE, Fontana RS, Sanderson DR, Uhlenhopp MA. Lung cancer detected during a screening program using four-month chest radiographs. Radiology 1983; 148:609-615.

92. Heelan RT, Flehinger BJ, Melamed MR, et al. Non-small-cell lung cancer: results of the New York screening program. Radiology 1984; 151:289-293.

93. Nakata M, Sawada S, Saeki H, et al. Prospective study of thoracoscopic limited

resection for ground-glass opacity selected by computed tomography. Ann Thorac Surg 2003; 75:1601-1605; discussion 1605-1606.