Radiomics in the evaluation of lung nodules: Intrapatient concordance between full-dose and ultra-low-dose chest computed tomography

Pierre-Alexis Autrusseau, Aïssam Labani, Pierre de Marini, Pierre Leyendecker, Cédric Hintzpeter, Anne-Claire Ortlieb, Michael Calhoun, Ilya Goldberg, Catherine Roy, Mickaël Ohana

Version of Record: https://www.sciencedirect.com/science/article/pii/S2211568421000310
Manuscript_ebbc5ddda6a6b01da0a1773db88545b3

Radiomics in the evaluation of lung nodules: intra-patient concordance between full-dose and ultra-low-dose chest computed tomography

Short title:

Radiomics of lung nodules with ultra-low-dose CT acquisition

Pierre-Alexis AUTRUSSEAU [a], Aïssam LABANI [a] , Pierre DE MARINI [b] , Pierre LEYENDECKER [a] , Cédric HINTZPETER [a], Anne-Claire ORTLIEB [c], Michael CALHOUN[d], Ilya GOLDBERG [d], Catherine ROY [a], Mickael OHANA [a]

[a] Department of Diagnostic Imaging (Radio B), Hôpitaux Universitaires de Strasbourg, 67000 Strasbourg, France

[b] Department of Diagnostic Imaging (Radio A), Hôpitaux Universitaires de Strasbourg, 67000 Strasbourg, France

[c] Department of Radiology, Institut Paoli Calmettes, 13009 Marseille, France

[d] Mindshare Medical, 500 Yale Avenue North, Seattle, WA 98109, USA

Corresponding author: pierrealexis.autrusseau@chru-strasbourg.fr

Abstract

Purpose. The purpose of this study was to retrospectively evaluate the quantitative and qualitative intra-patient concordance of pulmonary nodule risk assessment by commercially available radiomics software between full-dose (DF) chest-CT and ultra-low-dose (ULD) chest CT.

Materials and Methods. Between July 2013 and September 2015, 68 patients (52 men and16 women; mean age, 65.5 ± 10.6 [SD] years; range: 35 – 87 years) with lung nodules ≥ 5 mm and < 30 mm who underwent the same day FD chest CT (helical acquisition; 120 kV; automated tube current modulation) and ULD chest CT (helical acquisition; 135 kV; 10 mA fixed) were retrospectively included. Each nodule on each acquisition was assessed by a commercial radiomics software providing a similarity malignancy index (mSI), classifying it as "benign-like" (mSI < 0.1); "malignant-like" (mSI > 0.9) or "undetermined" (0.1 ≤ mSI ≤ 0.9). Intra-patient qualitative agreement was evaluated with weighted Cohen Kappa test and quantitative agreement with intraclass correlation coefficient (ICC).

Results. Ninety-nine lung nodules with a mean size of 9.14 ± 4.3 (SD) mm (range: 5 - 25 mm) in 68 patients (mean 1.46 nodule per patient; range: 1–5) were assessed; mean mSI was 0.429 ± 0.331 (SD) (range: 0.001–1) with FD chest CT (22/99 [22%] "benign-like", 67/99 [68%] "undetermined" and 10/99 [10%] "malignant-like") and mean mSI was 0.487 ± 0.344 (SD) (range: 0.002–1) with ULD chest CT (20/99 [20%] "benign-like", 59/99 [60%] "undetermined" and 20/99 [20%] "malignant-like"). Qualitative and quantitative agreement of FD chest CT with ULD chest CT were "good" with Kappa value of 0.60 (95% CI: 0.46–0.74) and ICC of 0.82 (95% CI: 0.73–0.87), respectively.

Conclusion. A good agreement in malignancy similarity index can be obtained between ULD chest CT and FD chest CT using radiomics software. However, further studies must be done with more case material to confirm our results and elucidate the diagnostic capabilities of radiomics software using ULD chest CT for lung nodule characterization by comparison by FD chest CT.

Keywords: Radiomics; Pulmonary nodules; Nodule risk assessment; Ultra-low dose chest CT; Tomography, X-ray computed.

Abbreviations

CI: Confidence interval; CT: Computed tomography; DLP: Dose length product; FD: Full-dose; HU: Housfield unit; ICC: Intraclass correlation coefficient; IQR: Interquartile range; mSI: Malignancy similarity index; NLST: National lung screening trial; SD: Standard deviation; ULD: Ultra-low dose

1. Introduction

Incidental lung nodules are a common entity in everyday practice, all the more considering the continuously growing number of chest computed tomography (CT) examinations performed [1]. In addition, lung cancer screening programs with low-dose chest-CT are increasingly proposed to patients, with an average detection rate of 20% (range: 3–51%) but a rate of benign nodule of 90% [2]. In both situations (incidental and screening) the challenge is to discriminate between benign nodules that do not require follow-up and malignant nodules that must be managed promptly to provide a survival benefit.

Several scientific societies have published guidelines to manage pulmonary nodules. For incidental pulmonary nodules, The Fleischner Society has published revised guidelines in 2017 with the aim of being less aggressive in the management of small nodules, thus increasing the threshold of size for nodules requiring follow-up and increasing the follow-up intervals [3]. For lung cancer screening, several national trials, including the National Lung Screening Trial (NLST, USA) [4], Nederlans-Leuvens Longkanker Screenings Onderzoek (Nelson, Netherlands) [5], International Early Lung Cancer Action Program (iELCAP, USA) [6], Pan-Canadian Early Detection of Lung Cancer Study (PanCan, Canada) [7] or scientific societies, including the American College of Radiology (Lung-RADS) [8], give guidelines to manage pulmonary nodules in lung cancer screening, and follow the same policy of being less aggressive on small nodules, to avoid false-positive results. Nonetheless, follow-up CT examinations are needed in all these guidelines, as visual discriminators (essentially size, location and shape) have limited specificity.

Artificial intelligence and especially radiomics are promising techniques to overcome these challenges. Radiomics extracts a high number of imaging features to achieve better diagnostic performance [9,10]. Machine-learning algorithms are used to compare radiomics extracted

features with a large database of known malignant or benign pulmonary nodules, mostly from screening trials [11].

Given the increasing number of chest CT examinations, limitation of the radiation dose is of paramount importance. Low-dose chest CT and ultra-low dose (ULD) chest CT, which do not have a clear definition in the literature, are associated with a marked decrease in dose-length product (DLP) [12–14], play an important role in reducing the global burden of radiation. However, the diagnostic performance of radiomics algorithm with low dose CT, as compared to the reference standard, is unknown, and one might fear that the degraded image quality of low-dose CT has an impact on the performance of the radiomics algorithm [15].

The purpose of this study was to retrospectively evaluate the intra-patient concordance of pulmonary nodule characterization by commercially available radiomics software between full-dose (FD) chest CT and ULD chest CT.

2. Materials and methods

2.1 Study population

One hundred seventy patients from three prior institutional review board approved prospective monocentric studies were all reviewed for retrospective inclusion in this work.

The first study (July 2013 to May 2014) included 55 patients with an occupational exposure to asbestos for at least 15 years for screening of asbestos-related pleuropulmonary lesions[16]; the second study (April 2014 to September 2014) [17] and the third study (April 2015 to September 2015) included 51 and 64 patients who were referred for a clinically indicated unenhanced chest CT examination.

Each patient included underwent unenhanced FD and ULD chest CT examinations with two successive acquisitions. All 170 patients, with a total of 340 CT acquisitions (170 FD CT and 170 ULD CT), were retrospectively reviewed by the same operator (P.-A. A., radiologist with 5 years of experience in CT) to detect all lung nodules (solid, ground glass and mixed) with a maximum diameter > 5mm and < 30 mm. All patients with at least one lung nodule were retrospectively included in this study. The size, location, type and margins of all nodules were recorded by the same operator. Electronic health records were searched in December 2019 to look for a final diagnosis for the lung nodule. Nodules were categorized as malignant when a definitive histopathological confirmation (biopsy and/or surgery) was available. Nodules were

categorized as benign if stable or decreasing/disappearing from a subsequent CT examination with an at least 2-year interval.

This retrospective study was approved by the institutional review board (Hôpitaux Universitaires de Strasbourg, France) with permission to perform chart review and a waiver of written informed consent.

2.2 CT acquisition

For each patient, FD chest CT (helical acquisition; 120 kV; automated tube current modulation) and ULD chest CT (helical acquisition; 135 kV; 10 mA fixed) were obtained with a second-generation 320-row scanner (Aquilion® ONE Vision Edition, Canon Medical Systems). CT examinations were performed in inspiration with the arms raised above the head. CT examinations were acquired in prone or supine position depending of the clinical indication. FD and ULD CT were reconstructed in lung window (width = 1500 Housfield unit [HU]; center = - 700 HU) with a hard kernel and a slice thickness of 1 mm, using an iterative reconstruction algorithm (AIDR-3D; Canon Medical Systems).

2.3 Dosimetry and noise

Radiation-dose was retrieved from the CT scan dose report and expressed in dose-length-product (DLP). The effective dose was calculated by multiplying the DLP by the chest-specific conversion factor of 0.014 mSv / mGy.cm^{-1} [19].

Noise was defined as the standard deviation (SD) of the attenuation of air inside the tracheal lumen in the parenchymal reconstruction. It was obtained from the averaging of 3 consecutive measurements made by the same operator (P.-A. A.) on each acquisition.

2.4 Radiomics of lung nodules

Nodule analysis was performed with a research version of the RevealAI-Lung software (Mindshare Medical). RevealAI-Lung is a software analytics platform based on machine learning, trained with a population of lung nodules from low-dose CT and their known outcomes from the NLST database, to aid risk assessment of lung nodules. When the radiologist indicates a lung nodule, the software calculates image features from the indicated location and from the full image. The software calculates many features (> 1 k each) from multiple three-dimensional views of the scan and then uses ensemble machine learning

training techniques to relate these features to known clinical outcomes. In the end, it provides a relative malignancy risk as a "malignancy similarity index" (mSI), a variable ranging from 0 to 1. In accordance with the manufacturer guidelines and the design of the software, thresholds can be used that relate to sensitivity and specificity of existing nodule guidelines. An mSI < 0.1 may be considered likely benign (risk similar or below Lung-RADS Score 2), and a mSI > 0.9 is considered a high probability of malignancy (close to 50%). Between these two values, the manufacturer does not provide recommendation.

Results were therefore classified in three categories depending of the mSI value: "benign-like" – B (mSI < 0.1), "undetermined" – U (0.1–0.9), and "malignant like" – M (mSI > 0.9).

2.5 Data collection

Several data data were collected in a dedicated spreadsheet. They include patient sex, age and height (cm); dose length product (DLP) (mGy.cm); standard deviation (SD) of the air inside the trachea; number of pulmonary nodules; nodule size; nodule location; nodule type (solid, part-solid, ground glass); margins (regular, spiculated); nodule calcifications (presence or absence); follow-up CT data; results of histopathological analysis and mSI.

2.6 Statistical analysis

Statistical analysis was performed using R statistical software (v 3.4.5) (R Core Team 2019). Descriptive statistics were used to present results. The qualitative concordance (*i.e.* classification in B, U or M categories) between FD-CT and ULD-CT mSI results was evaluated using Cohen weighted kappa test [20]. The quantitative concordance (*i.e.* the number itself) between FD and ULD CT. mSI was evaluated using intraclass correlation coefficient (ICC). A graphical representation of the quantitative concordance was performed using Bland and Altman method [21]. Correlation between noise and mSI difference was evaluated using Spearman correlation test. Differences in radiation dose and noise between FD and ULD CT were evaluated using Student *t* test. A $P < 0.05$ was considered to indicate statistically significant differences.

3. Results

3.1 Study population

Out of the 170 patients screened, 68 had at least one lung nodule responding to the inclusion criteria and were therefore retrospectively included. The final study population included 52 men and 16 women (mean age, 65.5 ± 10.6 [SD] years; age range: 35–87 years) with a mean number of 1.46 ± 1.12 (SD) nodule per patient (range: 1–7). A flowchart of the study is provided in **Figure 1.**

3.2 Pulmonary nodules

A total of 99 lung nodules > 5 mm and < 30 mm were detected in these 68 patients. All nodules were visible both on FD and ULD CT (**Figure 2**). Mean nodule size was 9.14 ± 4.3 mm (SD) (range: 5–25 mm), 16 were located in the right upper lobe, 5 in the middle lobe, 30 in the right lower lobe, 27 in the left upper lobe, and 21 in the left lower lobe. On CT, 90 nodules were solid nodules, 6 were part-solid nodules and 3 were ground glass nodules; 65 nodules had regular margins and 34 had irregular/spiculated margins. Eleven nodules were 11 partly calcified. (**Table 1**).

CT follow-up was available for 48/68 patients (70.6%), with a median follow-up for the entire cohort of 295 days (range: 0–1999 days; IQR: 0–992). Significant growth was noted in 8 nodules and histopathological confirmation was obtained by percutaneous biopsy, endoscopic biopsy or surgical lobectomy; 7 nodules were malignant (5 adenocarcinomas, 1 small-cell carcinoma and 1 carcinoid tumor) and 1 was a histopathologically proven, benign fibrous nodule. Median time between initial CT examination in our cohort and final histopathological confirmation was 196 days (range: 22–704 days; IQR: 72–404). Thirty-two nodules showed decrease in size or no changes at 2-years follow-up and were considered benign. Consequently, 33/99 nodules (33%) were considered benign. Fifty-nine nodules (59/99; 59.6%) had no or less than 2-years follow up and were considered undetermined.

3.3 Dosimetry and noise

With FD chest CT, mean DLP and effective dose were 238 ± 36 (SD) mGy.cm^2 (range: 84–793 mGy.cm^2) and 3.3 ± 1.9 (SD) mSv (range: 1.2–11.1 mSv) respectively. With ULD chest CT mean DLP and effective dose were 16 ± 1.6 (SD) mGy.cm^2 (range: 12.5–20 mGy.cm^2) and 0.23 ± 0.02 (SD) mSv (range: 0.18–0.28 mSv), respectively. Mean noise was 38 ± 10

(SD) (range: 22–62) with FD chest CT and 51 ± 13 (SD) (range: 18–74) with ULD chest CT. Significant differences between both groups were found for DLP, effective dose and noise ($P < 0.001$).

3.4 Radiomics of lung nodules

With FD chest CT, mean mSI was 0.429 ± 0.331 (SD) (range: 0.001 –1); 23/99 nodules (23.2%) were classified as "B", 12/99 (12.1%) as "M" and 64/99 (64.7%) as "U". With ULD chest CT, mean mSI was 0.487 ± 0.344 (SD) (range: 0.002–1); 20/99 nodules (20.2%) were classified as "B", 22/99 (22.2%) as "M" and 57/99 (57.6%) as "U".

Quantitative reproducibility of the mSI metric between FD and ULD chest CT was good, with an ICC of 0.82 (95% CI: 0.73–0.87). Mean bias at Bland-Altman analysis was 0.058 [95% CI: 0.018–0.097]. Bland and Altman plot is reported in **Figure 3**.

With ULD CT compared to FD CT, only one nodule classified as "B" shifted to "M". Twenty-three other nodules have changed by one rank ("B"-"U" or "U"-"M") **(Table 2)**. Qualitative reproducibility (*i.e.* classification of nodules in B, U and M categories based on the mSI) between FD and ULD chest CT was good, with a weighted Cohen's Kappa value of 0.60 (95% CI: 0.46–0.74).

No correlation between the noise difference and the mSI difference between FD and ULD chest CT was found ($r = -0.17$; $P = 0.03$).

In the histopathologically-proven malignant subgroup of 7/99 (7.1%) nodules, mean mSI was 0.621 ± 0.226 (SD) (range: 0.037–0.997) with FD chest CT and 0.598 ± 0.288 (SD) (0.056–0.999) with ULD chest CT. 6/7 (86%) nodules were classified "U" and 1/7 (14%) was classified "M" (this latter nodule being an adenocarcinoma). There was no group change between FD and ULD chest CT for all these 7 nodules.

In the benign subgroup of 32/99 (32.3%) lung nodules, mean mSI was 0.482 ± 0.333 (SD) (range: 0.004–0.993) with FD CT and 0.536 ± 0.366 (SD) (range: 0.039–1) with ULD-CT. With FD CT, 5/32 (15.6%) nodules were classified "B", 23/32 (71.9%) nodules were classified "U", and 4/32 (12.5%) nodules were classified "M". With ULD CT, 7/32 (21.9%) nodules were classified "B", 14/32 (43.8%) nodules were classified "U", and 11/32 (34.3%) nodules were classified "M". An example of mSI provided by the radiomics software on lung nodules is provided **Figure 4**.

4. Discussion

In the present retrospective study including 99 pulmonary nodules in 68 patients, we evaluated the intra-patient concordance of a radiomics-based software analysis for nodule characterization between FD and ULD chest-CT. This concordance was qualitatively and quantitatively good (Kappa =0.60 and ICC = 0.82). These findings suggest that the impact of image quality for quantitative radiomics-based analysis is most likely minor and does have a limited effect on the final analysis, at least in our cohort. In addition, we did not find any correlation between noise difference and mSI difference between FD and ULD chest CT, meaning that the degraded signal-to-noise ratio does not correlate with the slight absolute mSI difference we observed.

Artificial intelligence, deep learning and radiomics are relatively recent terms in modern radiology. Messerli et al. have already evaluated the applicability of dedicated software on degraded chest-CT acquisition, with computer-aided-detection of solid pulmonary nodule on ULD chest-CT, with similar sensitivity of 68% versus 70% with FD chest CT [22]. Ohno et al. demonstrated that iterative reconstruction algorithm significantly improves nodule detection rate with computer aided detection on ULD chest-CT, meaning that relatively “basic” software applications can accommodate a degraded image quality[23].

However, radiomics software are working on a more complex and deep scale [24]. They generate complex high-dimensional data from CT or other imaging techniques, and extract quantitative descriptors from images which can be divided in first-order statistical outputs: distribution of values of individual voxels; second-order statistical outputs: textures feature for intra-tumoral and inter-tumoral heterogeneity; and higher-orders statistical method. Then, all these data are compared with machine-learning on large database of known histological tumors, and promise to increase the accuracy of diagnosis, prognosis and prediction of tumor response [24]. The promise is to have a far better reproducibility than radiologists, whose visual discriminators are limited and subjective [11]. While radiomics of pulmonary nodules is facilitated by the low attenuation of the surrounding lung, it has also been developed and evaluated in non-chest imaging, as for prostate cancer; hepatocellular carcinoma; glioblastoma and more [24].

Despite all the aforementioned advantages and promises, radiomics has also some limitations on the reproducibility of results. Most of radiomics software have been trained on the same dataset, specifically LIDC-IDRI [25], NLST [4] and from the National Cancer Institute

released in 2017, which limits the generalization of the results to other populations. Moreover, some studies have based their test-retest analysis on the same patients' images with the same CT acquisition parameters, which yields excellent results [26] but does not reflect current clinical practice. Berenguer et al. conducted an *in vitro* study on phantoms to evaluate the reproducibility of radiomics features by modifying data acquisition protocls like kV, mAs, pitch, kernel acquisition; and using different CT manufacturers; with a poor reproducibility except for the pitch variation [27]. These researchers concluded that out of 177 radiomics features tested, only 71 were reproducible [27]. Another study on radiomics features variability revealed that the quality and reproducibility of radiomics features depended on image acquisitions/reconstruction, and that standards protocols had to be established [28]. A third study studied the effect of simulated dose reduction on chest CT of patients with known lung cancer [29]. This study showed a significant deviation in all radiomics features [29]. This may suggest that the use of radiomics software with different acquisition protocols on different CT equipment may lead to inconsistent results.

Our study has some limitations due to the retrospective design, evaluated only on one CT model and manufacturer, and with only one type of radomics software. The fact that the software we evaluated was trained on the NLST database, meaning only low dose chest CT (below 1.5mSv, so still about 6 times the dose of ULD-CT in our cohort) were used, might have favored the similarity of results between FD and ULD chest CT examinations in our study. Further studies with more patients and CT acquisitions from different centers and manufacturers are needed. Another limitation related to relatively low number of nodules we analyzed. While the total number of lung nodules is enough to confirm a good concordance in radiomics analysis between FD CT and ULD CT examinations, the low number of confirmed lung nodules (27 benign and 7 malignant) and the retrospective nature of the study makes it impossible to draw any significant conclusions on the actual diagnostic performances of the radiomics software.

In conclusion, a good agreement in malignancy similarity index can be obtained between ULD chest CT and FD chest CT using radiomics software. However, further studies must be done with more case material to confirm our results and elucidate the diagnostic capabilities of radiomics software using ULD chest CT for lung nodule characterization by comparison by FD chest CT.

Human rights

The authors declare that the work described has been carried out in accordance with the Declaration of Helsinki of the World Medical Association revised in 2013 for experiments involving humans.

Informed consent and patient details

Institutional ethic committee approval was obtained. The authors declare that this report does not contain any personal information that could lead to the identification of the patients. Written informed consent from the patients was waived. The authors also confirm that the personal details of the patients have been removed.

Disclosure of interest

M. Calhoun is CEO of Mindshare Medical and I. Goldberg was CTO for Mindshare Medical from 2016 to 2019. The others authors have no conflict of interest to declare.

Funding

This work did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.Conflicts of interest

REFERENCES

[1] Brenner DJ. Computed tomography: an increasing source of radiation exposure. N Engl J Med 2007:8.

[2] Bach PB, Mirkin JN, Oliver TK, Azzoli CG, Berry DA, Brawley OW, et al. Benefits and harms of CT screening for lung cancer: a systematic review. JAMA 2012;307:2418.

[3] MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner society 2017. Radiology 2017;284:228–43.

[4] The national lung screening trial research team. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409.

[5] Ru Zhao Y, Xie X, de Koning HJ, Mali WP, Vliegenthart R, Oudkerk M. NELSON lung cancer screening study. Cancer Imaging 2011;11:S79–84.

[6] Chung M, Tam K, Wallace C, Yip R, Yankelevitz DF, Henschke CI; for the I-ELCAP Investigators. International early lung cancer action program: update on lung cancer screening and the management of CT screen-detected findings. AME Med J 2017;2:129.

[7] Tammemagi MC, Schmidt H, Martel S, McWilliams A, Goffin JR, Johnston MR, et al. Participant selection for lung cancer screening by risk modelling (the Pan-Canadian early detection of lung cancer [PanCan] study): a single-arm, prospective study. Lancet Oncol 2017;18:1523–31.

[8] Godoy MCB, Odisio EGLC, Truong MT, de Groot PM, Shroff GS, Erasmus JJ. Pulmonary nodule management in lung cancer screening. Radiol Clin North Am 2018;56:353–63.

[9] Wilson R, Devaraj A. Radiomics of pulmonary nodules and lung cancer. Transl Lung Cancer Res 2017;6:86–91.

[10] Nougaret S, Tardieu M, Vargas HA, Reinhold C, Vande Perre S, Bonanno N, et al. Ovarian cancer: an update on imaging in the era of radiomics. Diagn Interv Imaging 2019;100:647–655.

[11] Bartholmai BJ, Koo CW, Johnson GB, White DB, Raghunath SM, Rajagopalan S, et al. Pulmonary nodule characterization, including computer analysis and quantitative features. J Thorac Imaging 2015;30:139–156.

[12] Ludes C, Schaal M, Labani A, Jeung MY, Roy C, Ohana M. Ultra-low dose chest CT: the end of chest radiograph? Presse Med 2016;45:291–301.

[13] Martini K, Moon JW, Revel MP, Dangeard S, Ruan C, Chassagnon G. Optimization of acquisition parameters for reduced-dose thoracic CT: a phantom study. Diagn Interv Imaging 2020;101:269–279.

[14] Beregi JP, Greffier J. Low and ultra-low dose radiation in CT: opportunities and

limitations. Diagn Interv Imaging 2019;100:63–64.

[15] Kim H, Park CM, Lee M, Park SJ, Song YS, Lee JH, et al. Impact of reconstruction algorithms on CT radiomic features of pulmonary tumors: analysis of intra- and inter-reader variability and inter-reconstruction algorithm variability. PLoS One 2016;11:e0164924.

[16] Schaal M, Severac F, Labani A, Jeung M-Y, Roy C, Ohana M. Diagnostic performance of ultra-low-dose computed tomography for detecting asbestos-related pleuropulmonary diseases: prospective study in a screening setting. PLoS One 2016;11:e0168979.

[17] Ludes C, Labani A, Severac F, Jeung MY, Leyendecker P, Roy C, et al. Ultra-low-dose unenhanced chest CT: prospective comparison of high kV/low mA versus low kV/high mA protocols. Diagn Interv Imaging 2019;100:85–93.

[18] Meyer E, Labani A, Schaeffer M, Jeung M-Y, Ludes C, Meyer A, et al. Wide-volume versus helical acquisition in unenhanced chest CT: prospective intra-patient comparison of diagnostic accuracy and radiation dose in an ultra-low-dose setting. Eur Radiol 2019;29:6858–6866.

[19] Valentin J, editor. Managing patient dose in computed tomography. 4ed. 2000. Oxford: Pergamon Pr; 2001.

[20] Benchoufi M, Matzner-Lober E, Molinari N, Jannot A-S, Soyer P. Interobserver agreement issues in radiology. Diagn Interv Imaging 2020;101:639–641.

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

[22] Messerli M, Kluckert T, Knitel M, Rengier F, Warschkow R, Alkadhi H, et al. Computer-aided detection (CAD) of solid pulmonary nodules in chest x-ray equivalent ultralow dose chest CT - first in-vivo results at dose levels of 0.13 mSv. Eur Radiol 2016;85:2217–2224.

[23] Ohno Y, Aoyagi K, Chen Q, Sugihara N, Iwasawa T, Okada F, et al. Comparison of computer-aided detection (CADe) capability for pulmonary nodules among standard-, reduced- and ultra-low-dose CTs with and without hybrid type iterative reconstruction technique. Eur Radiol 2018;100:49–57.

[24] Nakaura T, Higaki T, Awai K, Ikeda O, Yamashita Y. A primer for understanding radiology articles about machine learning and deep learning. Diagn Interv Imaging 2020;101:765–770.

[25] Armato SG, McLennan G, Bidaut L, McNitt-Gray MF, Meyer CR, Reeves AP, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans: the LIDC/IDRI thoracic CT database of lung nodules. Med Phys 2011;38:915–931.

[26] Balagurunathan Y, Gu Y, Wang H, Kumar V, Grove O, Hawkins S, et al. Reproducibility and prognosis of quantitative features extracted from CT Images. Transl

Oncol 2014;7:72–87.

[27] Berenguer R, Pastor-Juan M del R, Canales-Vázquez J, Castro-García M, Villas MV, Mansilla Legorburo F, et al. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. Radiology 2018;288:407–415.

[28] Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features: Invest Radiol 2015;50:757–765.

[29] Hepp T, Othman A, Liebgott A, Kim JH, Pfannenberg C, Gatidis S. Effects of simulated dose variation on contrast-enhanced CT-based radiomic analysis for non-small cell lung cancer. Eur J Radiol 2020;124:108804.

FIGURE LEGENDS

Figure 1. Study flowchart. FD indicates full dose; ULD indicates ultra-low dose.

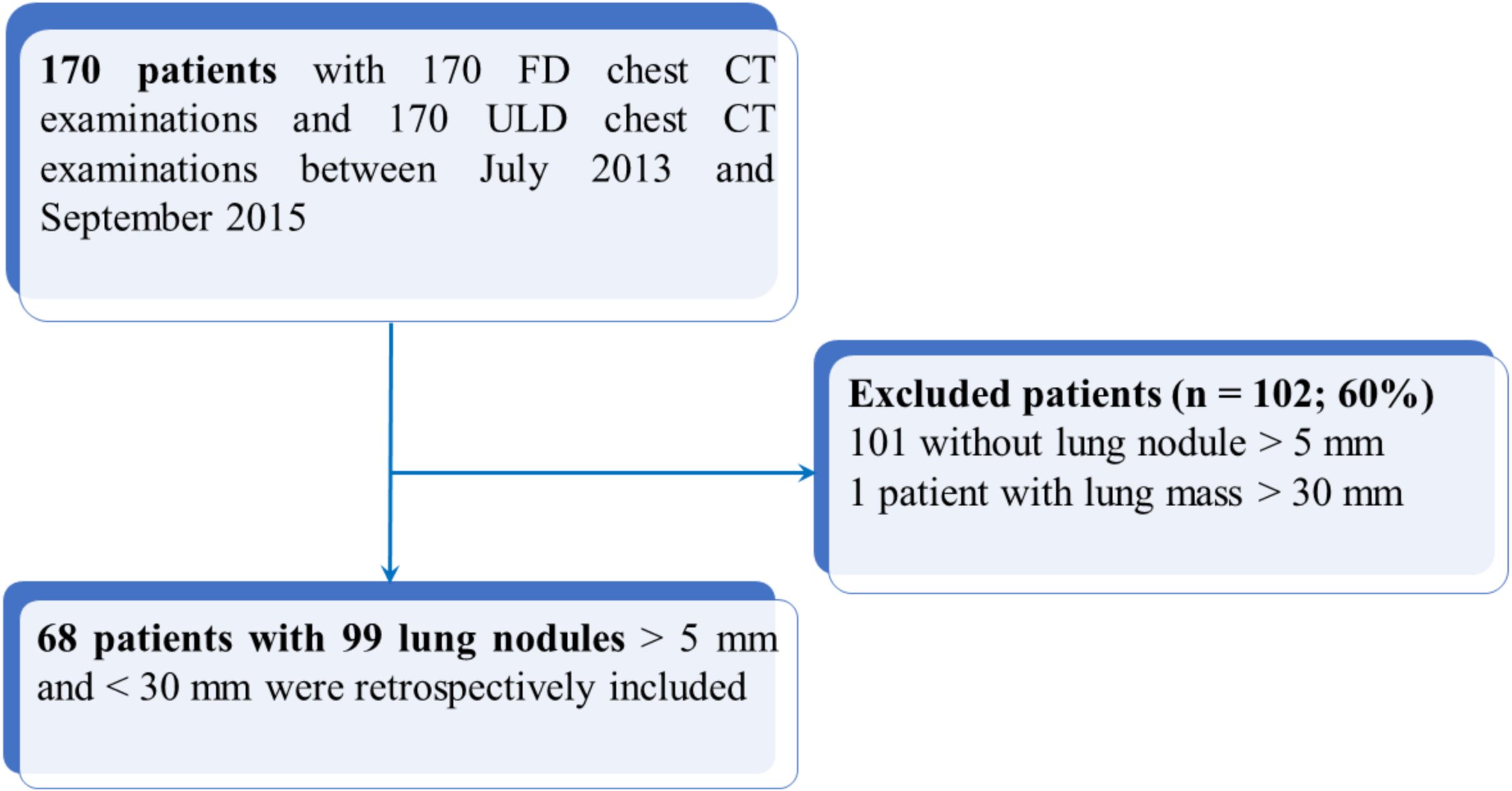
Figure 2. Visual discriminators between full dose and ultra-low dose CT. A, B, 58-year-old man who underwent the same day unenhanced full-dose chest CT **(A)** and unenhanced ultra-low-dose chest CT (**B**). The part-solid lung nodule (arrows) shows same characteristics on both CT acquisitions. **C, D**, 63-year-old woman with a smoking history who underwent the same day full-dose chest CT (**C**) and ultra-low-dose chest CT **(D)**. Both CT acquisitions shows partially calcified lesion (arrows) of the right lower lobe. **E, F**, A 47-year-old man with history of cancer who underwent the same day full-dose chest CT **(E)** and ultra-low-dose chest-CT **(F)**. A spiculated nodule with irregular margins is visible in the right lower lobe (arrows) displaying same characteristics on both CT acquisitions.

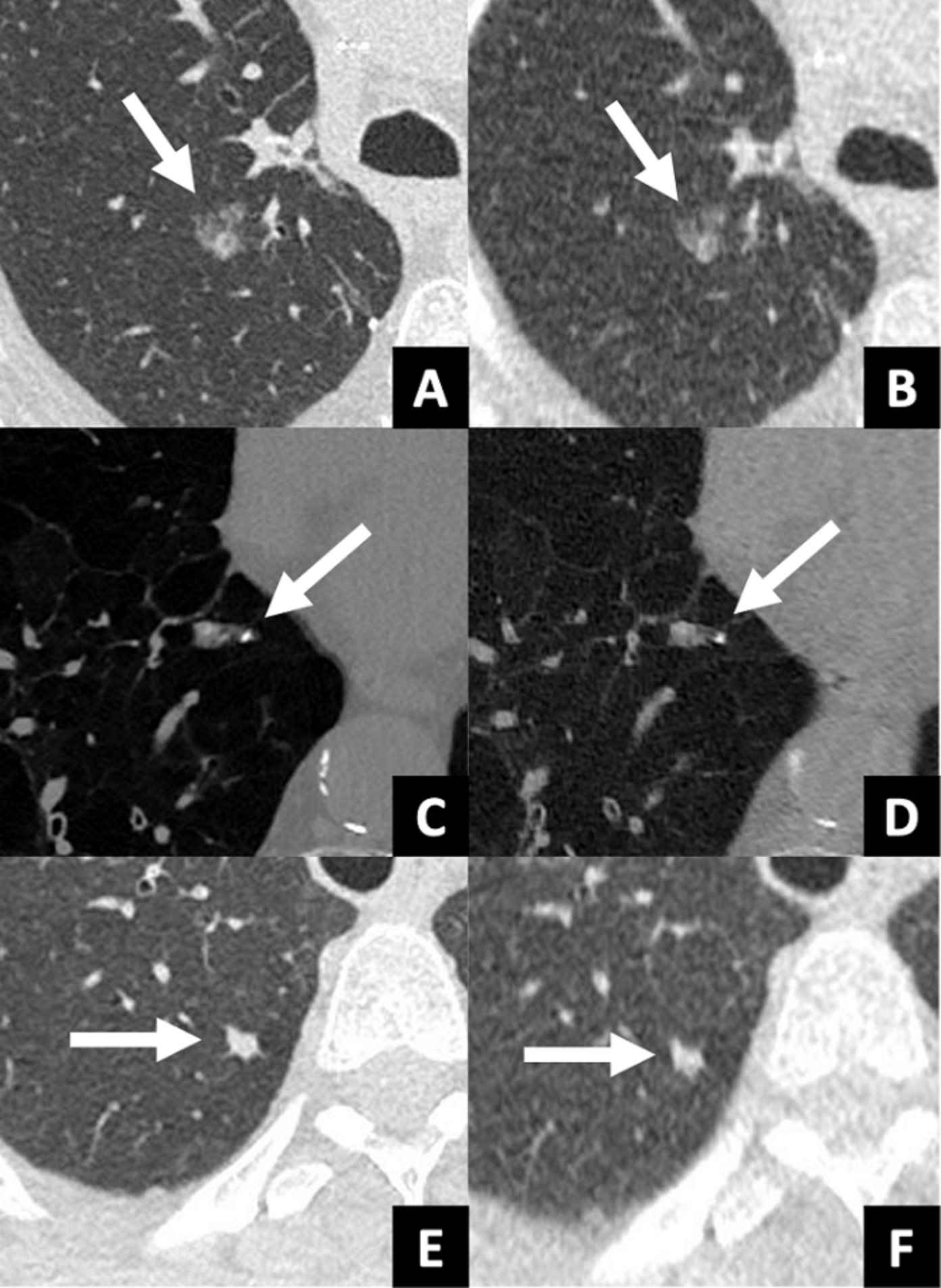
Figure 3. Bland and Altman plot. Blue points: values; continuous blue line: bias; spotted blue line: 95% bias confidence interval; spotted red line: 95% confidence interval. Note that the bias is close to zero indicating good quantitative concordance.

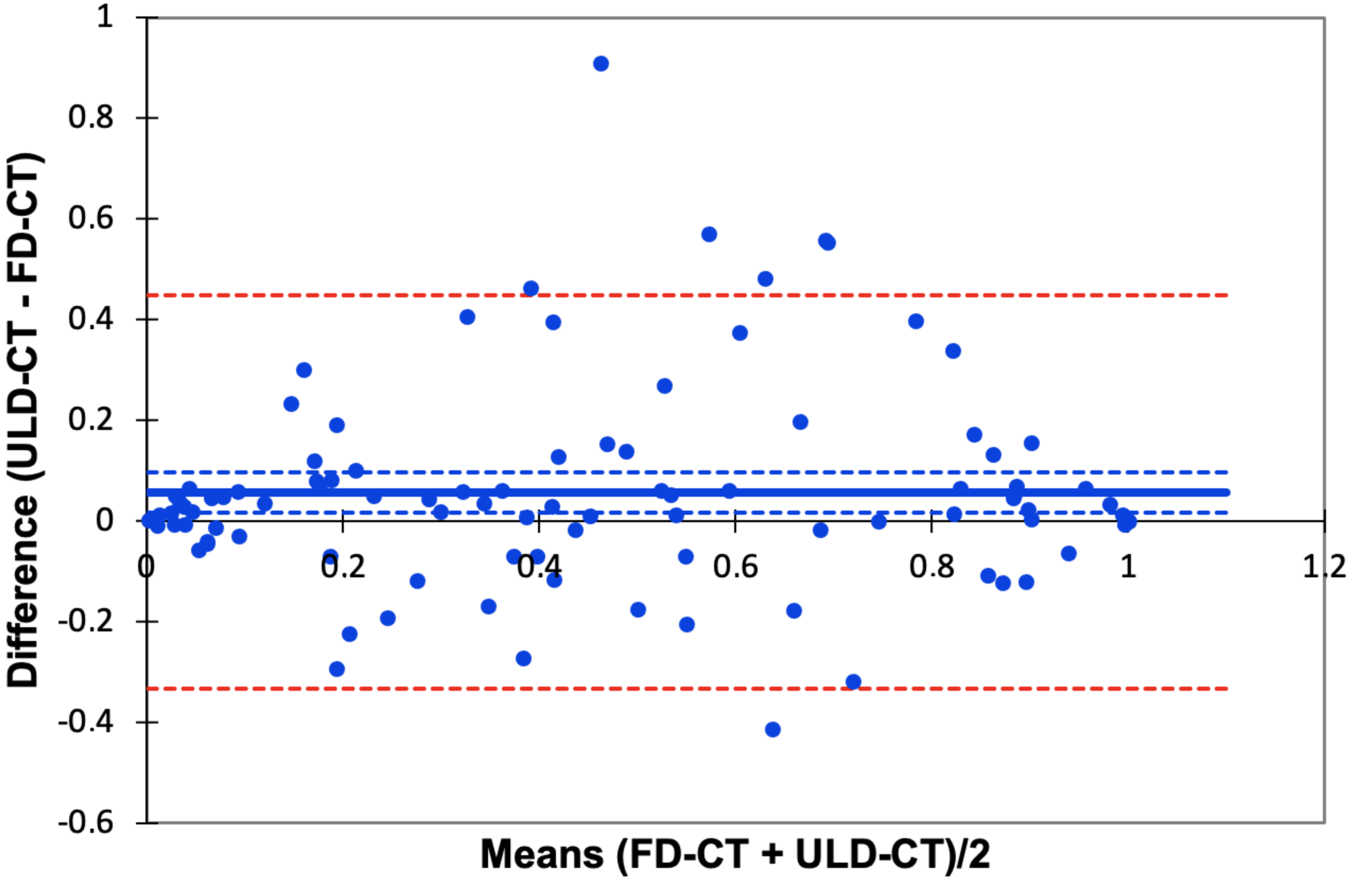
Figure 4. Examples of mSI provided by the radiomics software. A, B, C, 78-year-old man with smoking history who underwent the same day unenhanced full-dose chest CT and unenhanced ultra-low-dose chest CT. Full-dose chest CT image in the axial plane shows a spiculated lesion of the left pulmonary apex (arrow) (mSI = 0.997) **(A),** and ultra-low dose CT image show the same nodule (arrow) (mSI = 0.999) **(B).** On PET/CT examination, the nodule shows intense uptake of ^{18}F-FDG (arrow) **(C)**. Results of histopathological analysis were consistent with adenocarcinoma. **D, E, F,** 48-year-old woman with a single lung nodule who underwent the same day unenhanced full-dose chest CT and unenhanced ultra-low dose chest CT. Full-dose chest CT image in the axial plane shows 5-mm nodule (arrow) with regular margins of the right lower lobe (mSI = 0.022) **(D)**, and ultra-low-dose CT image (mSI = 0.052) shows the same nodule (arrow) with similar features **(E).** Follow-up CT image at 1999 days shows a stable lesion (arrow) in size with a new calcification **(F)**. **G, H, I**, 51-year-old man who underwent the same day unenhanced full-dose chest CT and unenhanced ultra-low dose chest CT. Full-dose chest CT image in the axial plane shows irregular lung lesion (arrow) in the right lower lobe (mSI = 0.923) **(G)**, and ultra-low-dose CT image shows the same nodule (arrow) with similar features (mSI = 0.989) **(H)**. Follow-up at 729 days shows a decrease in size of the lesion (arrow) consistent with benign lesion **(I)**.

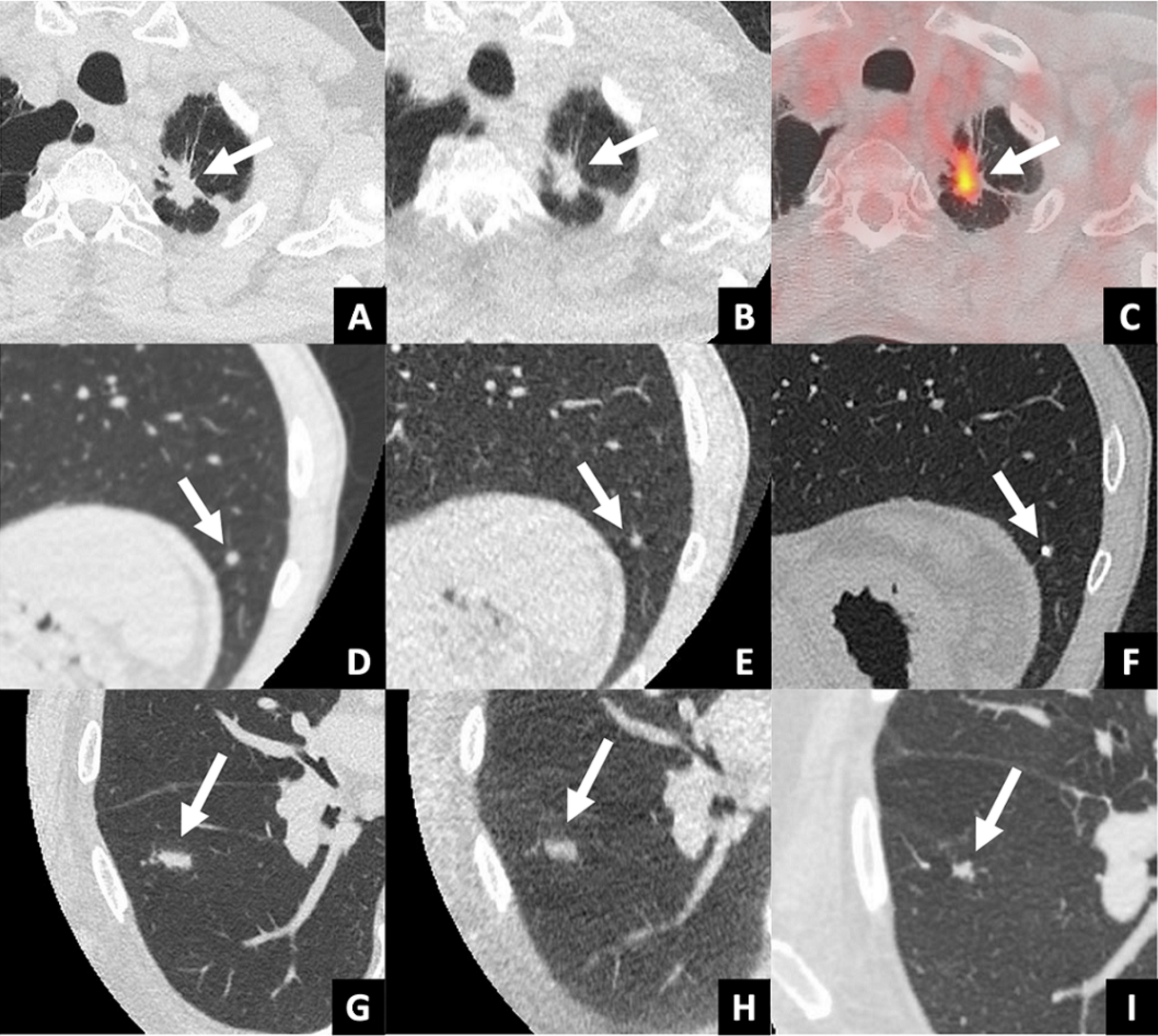
Table 1. Nodules characteristics of the cohort according to histopathological proof and follow-up.

Table 2. Correlation table of malignancy risk between full dose and ultra-low dose CT.


170 patients with 170 FD chest CT examinations and 170 ULD chest CT examinations between July 2013 and September 2015
Excluded patients (n = 102; 60%)
101 without lung nodule > 5 mm
1 patient with lung mass > 30 mm
68 patients with 99 lung nodules > 5 mm and < 30 mm were retrospectively included


A
B
C
D
E
F


1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.2
0.4
0.6
0.8
1
1.2
Difference (ULD-CT - FD-CT)
Means (FD-CT + ULD-CT)/2


A
B
C
D
E
F
G
H
I

	Malignant	Undetermined	Benign	Total
Mean size (mm)	13.7 ± 6.7 [6–25]	7 ± 3.8 [5–25]	10 ± 4 [5–20]	**9.14 ± 4.3 [5–25]**
Regular	3 (3/7; 42.8%)	41 (41/59; 69.5%)	21 (21/33; 63.6%)	**65 (65/99; 65.7%)**
Irregular	4 (4/7; 57.2%)	18 (18/59; 30.5%)	12 (12/33; 36.4%)	**34 (34/99; 34.3%)**
Calcifications	0 (0/7; 0%)	7 (7/59; 11.2%)	4 (4/33; 12.1%)	**11 (11/99; 11.1%)**
RUL	2 (2/7; 28.5%)	10 (10/59; 16.9%)	4 (4/33; 12.1%)	**16 (16/99; 16.1%)**
ML	1 (1/7; 14.3%)	2 (2/59; 3.4%)	2 (2/33; 6.1%)	**5 (5/99; 5.1%)**
RLL	1 (1/7; 14.3%)	15 (15/59; 25.4%)	14 (14/33; 42.4%)	**30 (30/99; 30.3%)**
LUL	2 (2/7; 28.5%)	14 (14/59; 23.7%)	11 (11/33; 33.3%)	**27 (27/99; 27.3%)**
LLL	1 (1/7; 14.3%)	18 (18/59; 30.6%)	2 (2/33; 6.1%)	**21 (21/99; 21.2%)**
Total	**7 (7/99; 7.1%)**	**59 (59/99; 59.6%)**	**33 (33/99; 33.3%)**	**99**

RUL: Right upper lobe; ML: Middle lobe; RLL: Right lower lobe; LUL: Left upper lobe; LLL: Left lower lobe. Size expressed in millimeters. Quantitative variables are expressed as means ± standard deviations; numbers in brackets are ranges. Qualitative variables are expressed as raw numbers; numbers in parentheses are proportions followed by percentages

		ULD CT		
		B	**U**	**M**
FD CT	**B**	17 (1)	5 (0.5)	1 (0)
	U	3 (0.5)	49 (1)	12 (0.5)
	M	0 (0)	3 (0.5)	9 (1)

ULD CT: ultra-low-dose CT; FD CT: full-dose CT; B: benign-like; U: undetermined; M: malignant-like. Qualitative variables are expressed as raw numbers; Numbers in parentheses are weights used for weighted kappa.