University of Groningen

Deep Learning Reconstruction Shows Better Lung Nodule Detection for Ultra-Low-Dose Chest CT

Jiang, Beibei; Li, Nianyun; Shi, Xiaomeng; Zhang, Shuai; Li, Jianying; de Bock, Geertruida H; Vliegenthart, Rozemarijn; Xie, Xueqian

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

Deep Learning Reconstruction Shows Better Lung Nodule Detection for Ultra–Low-Dose Chest CT

Beibei Jiang, MD • Nianyun Li, BSc • Xiaomeng Shi, PhD • Shuai Zhang, PhD • Jianying Li, PhD • Geertruida H. de Bock, PhD • Rozemarijn Vliegenthart, MD, PhD • Xueqian Xie, MD, PhD

From the Department of Radiology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, 100 Haining Rd, Shanghai 200080, China (B.J., N.L., X.X.); CT Imaging Research Center, GE Healthcare China, Shanghai, China (X.S., S.Z., J.L.); and Departments of Epidemiology (G.H.d.B.) and Radiology (R.V.), University of Groningen, University Medical Center Groningen, Groningen, the Netherlands. Received February 28, 2021; revision requested April 15; revision received November 6; accepted November 15. **Address correspondence to** X.X. (e-mail: *xiexueqian@hotmail.com*).

Background: Ultra–low-dose (ULD) CT could facilitate the clinical implementation of large-scale lung cancer screening while minimizing the radiation dose. However, traditional image reconstruction methods are associated with image noise in low-dose acquisitions.

Purpose: To compare the image quality and lung nodule detectability of deep learning image reconstruction (DLIR) and adaptive statistical iterative reconstruction-V (ASIR-V) in ULD CT.

Materials and Methods: Patients who underwent noncontrast ULD CT (performed at 0.07 or 0.14 mSv, similar to a single chest radiograph) and contrast-enhanced chest CT (CECT) from April to June 2020 were included in this prospective study. ULD CT images were reconstructed with filtered back projection (FBP), ASIR-V, and DLIR. Three-dimensional segmentation of lung tissue was performed to evaluate image noise. Radiologists detected and measured nodules with use of a deep learning–based nodule assessment system and recognized malignancy-related imaging features. Bland-Altman analysis and repeated-measures analysis of variance were used to evaluate the differences between ULD CT images and CECT images.

Results: A total of 203 participants (mean age ± standard deviation, 61 years ± 12; 129 men) with 1066 nodules were included, with 100 scans at 0.07 mSv and 103 scans at 0.14 mSv. The mean lung tissue noise ± standard deviation was 46 HU ± 4 for CECT and 59 HU ± 4, 56 HU ± 4, 53 HU ± 4, 54 HU ± 4, and 51 HU ± 4 in FBP, ASIR-V level 40%, ASIR-V level 80% (ASIR-V-80%), medium-strength DLIR, and high-strength DLIR (DLIR-H), respectively, of ULD CT scans ($P < .001$). The nodule detection rates of FBP reconstruction, ASIR-V-80%, and DLIR-H were 62.5% (666 of 1066 nodules), 73.3% (781 of 1066 nodules), and 75.8% (808 of 1066 nodules), respectively ($P < .001$). Bland-Altman analysis showed the percentage difference in long diameter from that of CECT was 9.3% (95% CI of the mean: 8.0, 10.6), 9.2% (95% CI of the mean: 8.0, 10.4), and 6.2% (95% CI of the mean: 5.0, 7.4) in FBP reconstruction, ASIR-V-80%, and DLIR-H, respectively ($P < .001$).

Conclusion: Compared with adaptive statistical iterative reconstruction-V, deep learning image reconstruction reduced image noise, increased nodule detection rate, and improved measurement accuracy on ultra–low-dose chest CT images.

©RSNA, 2022

Online supplemental material is available for this article.

Lung cancer is the leading cause of cancer-related death at 18.4% (1). Volume CT screening can reduce the 10-year mortality of lung cancer by 24% (2). Despite the rapidly increasing demand for lung cancer screening, the cumulative burden of radiation exposure limits its wide implementation (3). On the premise of maintaining diagnostic performance, it is essential to reduce the radiation dose as much as possible. However, traditional image reconstruction methods, such as filtered back projection (FBP), do not reduce image noise with low-dose acquisitions (4). Iterative reconstruction, such as adaptive statistical iterative reconstruction-V (ASIR-V) and advanced modeled iterative reconstruction, or ADMIRE, was developed to reconstruct images and reduce noise (4,5). In low-dose CT settings, with absorbed doses from 1 to 2 mSv, iterative reconstruction can significantly reduce image noise and improve image quality (6). Further radiation dose reduction for low-dose CT screening may help encourage more use of this method. Ultra–low-dose (ULD) CT reduces the dose level to 0.13–0.49 mSv (7–9), still higher than that of chest radiography (0.03–0.1 mSv) (10). When using ULD CT at doses similar to those used in chest radiography, iterative reconstruction algorithms affect the display of subtle imaging features, limiting their applicability to this task (11).

Recently, deep learning image reconstruction (DLIR) technology based on convolutional neural networks has been introduced to improve image quality in low-dose settings (12–14). With these methods, studies have shown significant image quality improvement at coronary CT angiography (15) and thoracic (12), abdominal (16), and cerebral CT (17). The purpose of this study was to compare

Abbreviations

ASIR-V = adaptive statistical iterative reconstruction-V, ASIR-V-40% = ASIR-V level 40%, ASIR-V-80% = ASIR-V level 80%, BMI = body mass index, CECT = contrast-enhanced chest CT, DLIR = deep learning image reconstruction, DLIR-H = high-strength DLIR, DLIR-M = medium-strength DLIR, FBP = filtered back projection, ULD = ultra–low-dose

Summary

Compared with the traditional iterative reconstruction method, deep learning image reconstruction improved image quality and lung nodule detection at ultra–low-dose chest CT (0.07–0.14 mSv, similar to a single chest radiograph).

Key Results

- In a prospective ultra–low-dose CT study of 203 participants, compared with adaptive statistical iterative reconstruction-V (ASIR-V), deep learning image reconstruction (DLIR) reduced the background image noise by 21% (mean air background noise ± standard deviation, 23 HU ± 4 vs 29 HU ± 4; $P < .001$).
- The nodule detection rates for DLIR and ASIR-V were 75.8% (808 of 1066 nodules) and 73.3% (781 of 1066 nodules), respectively.
- The percentage difference of long diameter at DLIR and ASIR-V from that at contrast-enhanced chest CT was 6.2% and 9.2%, respectively ($P < .001$).

the image quality and lung nodule detectability of DLIR and the widely used ASIR-V in ULD CT.

Materials and Methods

Study Participants

Consecutive patients at Shanghai General Hospital from April to June 2020 were prospectively included and randomly divided into two subgroups (those scanned at 0.07 mSv and those at 0.14 mSv). The local institutional review board approved this study, and written informed consent was obtained. X.S., S.Z., and J.L. are employees of GE Healthcare, the CT manufacturer of this study, and had no control over the data or analysis. The inclusion criteria were as follows: *(a)* adult participants with indications for contrast-enhanced chest CT (CECT) and *(b)* adults in whom CECT and noncontrast ULD CT were performed. The exclusion criteria were *(a)* poor image quality caused by motion or respiratory artifacts, *(b)* an incomplete image reconstruction sequence, *(c)* unsuccessful lung parenchyma segmentation, and *(d)* body mass index (BMI) of 30 kg/m^2 or more. The included participants were prespecified to stratify by BMI according to the World Health Organization classification as follows: underweight (BMI <18.5 kg/m^2), normal weight (18.5–24.9 kg/m^2), and overweight (≥25 kg/m^2) (18). Assuming that our study is similar to the previous studies evaluating lung nodules at ULD CT based on approximately 200 patients (7,19), a sample size of 200 would provide sufficient statistical power.

Participants with 1–50 nodules at CECT were further analyzed for nodule detection and measurement. Participants with 51 or more nodules were excluded because of the difficulty of manual counting. Finally, the imaging features of participants with malignant nodules confirmed with use of hematoxylin and eosin staining and immunohistochemical staining were studied. Figure 1 shows the research pipeline steps.

CT Scanning and Image Reconstruction

All participants underwent noncontrast ULD CT and CECT scanning with a 256-row volume CT scanner (Revolution CT, GE Healthcare) (Appendix E1 [online], Table E1 [online]). For the ULD CT scan, each participant's scanning protocol was randomly set to 70 kVp and 20 mA or 70 kVp and 40 mA, resulting in two different absorbed dose levels (0.07 mSv or 0.14 mSv, respectively). For CECT, each participant was injected with 40–60 mL of contrast medium (iopamidol [Iopamiro 300, Bracco]) at 3.0 mL/sec through the antecubital vein and scanned at 120 kVp and automatic mA, with a noise index of 20.

Six groups of CT images were reconstructed, including five groups of ULD CT images and one group of CECT images. The DLIR algorithm (TrueFidelity, GE Healthcare) was applied. The five ULD CT image groups were reconstructed with FBP, ASIR-V level 40% (ASIR-V-40%), ASIR-V level 80% (ASIR-V-80%), medium-strength DLIR (DLIR-M), and high-strength DLIR (DLIR-H). The CECT image group (ASIR-V level 50%) was used as a reference for image quality and nodule assessment. All images were reconstructed with a standard kernel, and the section thickness/interval was 1.25 mm/1.25 mm.

Image Quality Assessment

We three-dimensionally segmented the whole lung tissue and upper air background in vitro to evaluate the image noise (Appendix E1 [online], Fig E1 [online]), expressed as the standard deviation of CT values. Two radiologists (B.J. and a second radiologist) with 5 years of experience in chest imaging independently assessed subjective image quality with use of a five-point Likert scale (1, unacceptable; 2, poor; 3, moderate; 4, good; and 5, excellent) based on the European Guidelines on Quality Criteria for Computed Tomography (20).

Nodule Detection Rate and Measurement Accuracy

Based on the image quality evaluation results, one group with lower image noise and higher subjective image quality was selected from the ASIR-V-40% and ASIR-V-80% groups to represent ASIR-V, and one group from DLIR-M and DLIR-H to represent DLIR. A radiologist (B.J.) read all CT images to detect lung nodules with the aid of a deep learning–based nodule evaluation system (InferRead CT Lung 8.6, Infervision) and measured the long diameters and volumes of the nodules (Fig E2 [online]) on CECT, FBP, ASIR-V, and DLIR images. The detected nodules included three types: subsolid, solid, and calcified nodules (21). Nodules smaller than 30 mm were analyzed and divided into three categories: 3–6 mm, 6–10 mm, and 10 mm or larger. A general radiologist (X.X.) with 20 years of experience independently confirmed the detected nodules by using the results of CECT as a reference and measured the nodules at CECT as the reference for long diameter and volume.

Imaging Features

To investigate the influence of ULD CT on the diagnostic evidence of malignant nodules, we analyzed the malignancy-related
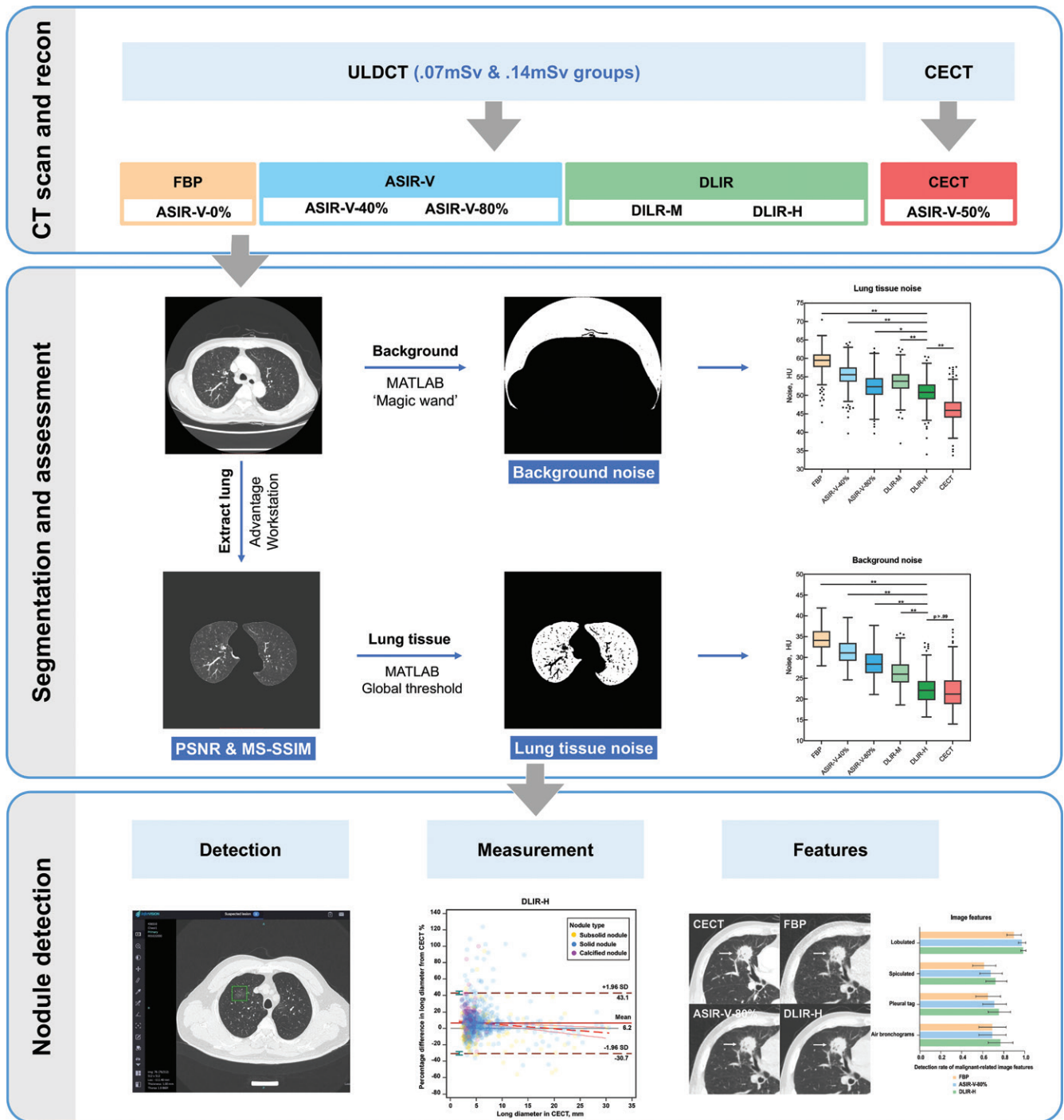
Figure 1: Overview of the research pipeline. ASIR-V = adaptive statistical iterative reconstruction-V, ASIR-V-0% = ASIR-V level 0%, ASIR-V-40% = ASIR-V level 40%, ASIR-V-50% = ASIR-V level 50%, ASIR-V-80% = ASIR-V level 80%, CECT = contrast-enhanced chest CT, DLIR = deep learning image reconstruction, DLIR-H = high-strength DLIR, DLIR-M = medium-strength DLIR, FBP = filtered back projection, MS-SSIM = multiscale structural similarity, PSNR = peak signal-to-noise ratio, recon = reconstruction, ULDCT = ultra–low-dose CT.

imaging features of histologically confirmed malignant nodules. Malignancy-related features included lobulated shapes, spiculated margins, pleural tags, and air bronchograms (22). These imaging features were independently identified by four general radiologists (X.X. and three others) with more than 10 years of experience.

Statistical Analysis
Normally distributed continuous variables are expressed as means ± standard deviations and were evaluated with use of the Student t test. The Pearson χ^2 test was used to compare participant characteristics of nominal or categorical variables between the 0.07-mSv ULD CT and 0.14-mSv ULD CT groups. The Fisher exact test was used to compare the detection rates of nodules. Kruskal-Wallis one-way analysis of variance was used to compare qualitative image evaluations. Bland-Altman analysis and repeated-measures analysis of variance were used to evaluate the differences in long diameter and volume between ULD

Participant and Nodule Characteristics

Variable	All Participants	Participants Scanned at 0.07 mSv	Participants Scanned at 0.14 mSv	P Value*
No. of participants	203	100	103	
Age (y)[†]	61 ± 12	62 ± 11	60 ± 14	.20
Sex				.80
M	129 (63.5)	65 (65.0)	64 (62.1)	
F	74 (36.5)	35 (35.0)	39 (37.9)	
Body mass index (kg/m^2)[†]	23.1 ± 3.1	23.1 ± 3.4	23.1 ± 2.8	>.99
<18.5 kg/m^2	12 (5.9)	7 (7.0)	5 (4.9)	
18.5–24.9 kg/m^2	144 (70.9)	69 (69.0)	75 (72.8)	
≥25 kg/m^2	47 (23.2)	24 (24.0)	23 (22.3)	
No. of nodules	1066	596	470	
Type[‡]				.39
Subsolid	313 (29.4)	170 (28.5)	143 (30.4)	
Less than −400 HU	154 (14.4)	82 (13.8)	72 (15.3)	
−400 HU or greater	159 (14.9)	88 (14.8)	71 (15.1)	
Solid	661 (62.0)	367 (61.6)	294 (62.6)	
Calcified	92 (8.6)	59 (9.9)	33 (7.0)	
Size[‡]				.91
3–6 mm	761 (71.4)	425 (71.3)	336 (71.5)	
6–10 mm	193 (18.1)	110 (18.5)	83 (17.6)	
10–30 mm	112 (10.5)	61 (10.2)	51 (10.9)	
No. of nodules with histologic results	41	22	19	
Histologic subtype[‡]				.69
Benign and preinvasive lesions	14 (34)	10 (45)	4 (21)	
Malignant lesions	27 (66)	12 (55)	15 (79)	

Note.—Unless otherwise specified, data are numbers of participants, with percentages in parentheses.

* P values represent the comparison between participants scanned at 0.07 mSv and those scanned at 0.14 mSv.

[†] Data are means ± standard deviations.

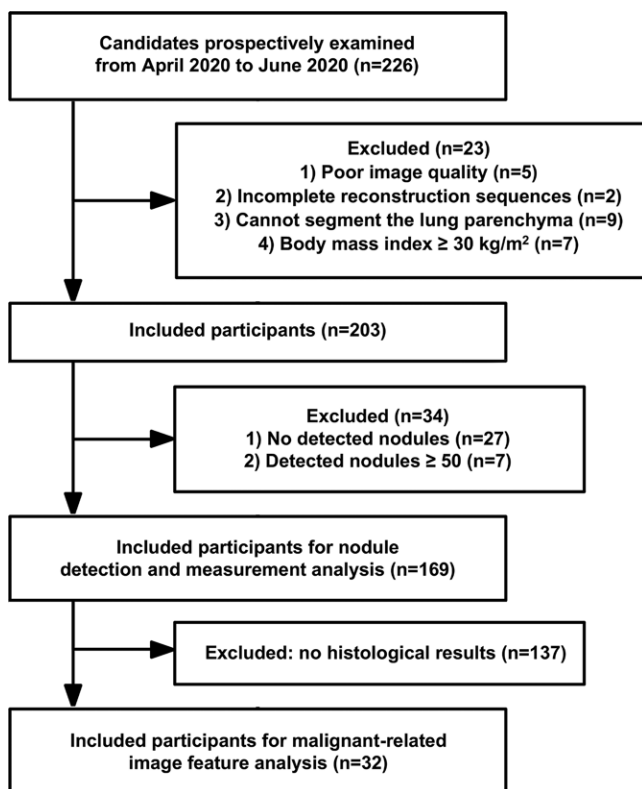[‡] Data are numbers of nodules, with percentages in parentheses.

Figure 2: Inclusion and exclusion flowchart.

CT images and CECT images. Multivariable logistic regression and a generalized linear mixed model were used to analyze the predictive value of reconstruction methods for nodule detection after adjusting for confounding factors on a per-nodule and per-participant basis, respectively. The intraclass correlation coefficient was used to indicate the interobserver consistency of subjective image quality assessment between the two radiologists and the observed malignancy-related features among the four radiologists. Two-tailed $P < .05$ was considered indicative of statistically significant difference. Two statistical packages (SPSS 25.0, IBM; MedCalc 19.1, MedCalc Software) were used for data analysis. B.J., X.S., and G.H.d.B. performed statistical analysis. Data generated or analyzed during the study are available from the corresponding author by request.

Results

Participant Characteristics

Of 226 patients, 203 (mean age ± standard deviation, 61 years ± 12) were eligible, including 129 men (63.5%) and 74 women (36.5%) (Table). Twenty-three patients were excluded: five for poor image quality, two for incomplete CT reconstruction sequences, nine for unsuccessful lung parenchyma segmentation, and seven for BMI of 30 kg/m^2 or higher. Figure 2 shows the study flowchart. The 0.07-mSv ULD CT and 0.14-mSv ULD
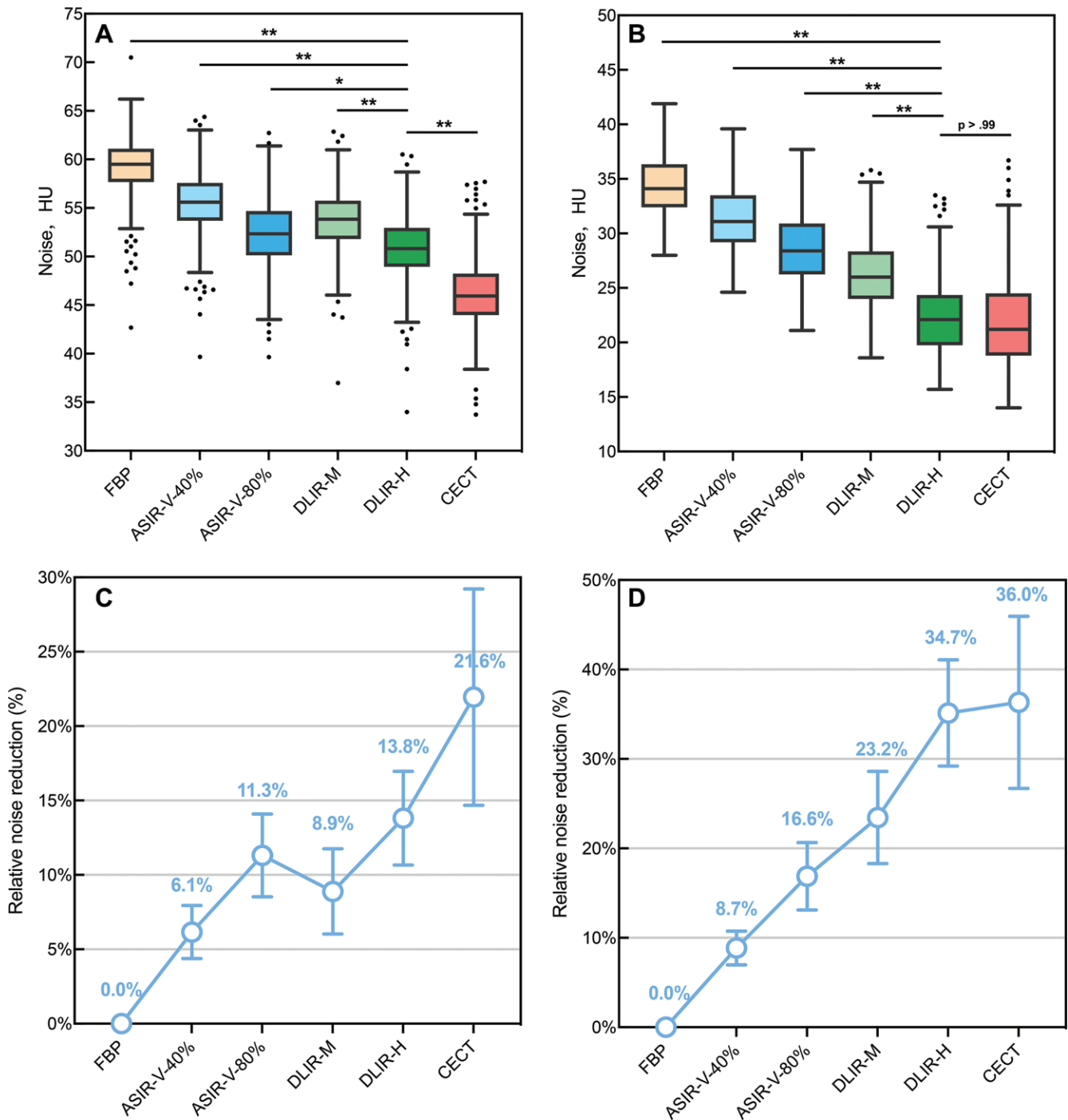
Figure 3: Image quality assessments according to image noise. **(A, B)** Box and whisker plots show lung tissue noise and background noise at contrast-enhanced chest CT (CECT) (reconstructed with use of adaptive statistical iterative reconstruction-V [ASIR-V] level 50%) as the reference and of filtered back projection (FBP), ASIR-V level 80% [ASIR-V-80%], and high-strength deep learning image reconstruction (DLIR-H) images at ultra–low-dose CT. The central line represents the median, box boundaries indicate the first and third quartiles, and the whisker boundaries extend 1.5 quartiles. Other points outside the whisker boundaries are plotted as outliers. **(C, D)** Dot plot shows relative noise reduction with the baseline FBP. Error bars represent the standard deviation. * = $P < .05$ and ** = $P < .001$. ASIR-V-40% = ASIR-V level 40%, DLIR-M = medium-strength DLIR.

CT groups comprised 100 and 103 participants, respectively. We found no differences in age, sex, or BMI between these two groups ($P > .05$). The mean BMI was 23.1 kg/m^2 ± 3.1. Of all 203 participants, 12 (5.9%), 144 (70.9%), and 47 (23.2%) were underweight, at normal weight, and overweight, respectively. The volume CT dose index values of CECT, 0.07-mSv ULD CT, and 0.14-mSv ULD CT were 4.9 mGy ± 0.7 (mean ± standard deviation), 0.13 mGy, and 0.27 mGy, respectively. The effective absorbed doses were 2.38 mSv ± 0.37 (mean ± standard deviation), 0.07 mSv, and 0.14 mSv, respectively. Compared with CECT, the absorbed doses of 0.07-mSv ULD CT and 0.14-mSv ULD CT were reduced by 97.0% and 94.0%, respectively (both $P < .001$).

Image Quality

The mean lung tissue noise ± standard deviation was 46 HU ± 4 for CECT and 59 HU ± 4, 56 HU ± 4, 53 HU ± 4, 54 HU ± 4, and 51 HU ± 4 for FBP, ASIR-V-40%, ASIR-V-80%, DLIR-M, and DLIR-H of ULD CT, respectively ($P <$ .001) (Fig 3, Table E2 [online]). Pairwise comparisons of FBP and ASIR-V-80% showed that the lung tissue noise of DLIR-H was the lowest, with a relative reduction of 13.8% from the baseline FBP level, which was better than the reduction of ASIR-V-80% (11.3% [$P =$ .03]). The relative noise reduction of DLIR-M was higher than that of ASIR-V-40% (8.9% vs 6.1% [$P <$ .001]).

The mean air background noise ± standard deviation of CECT images was 22 HU ± 5, while that of ULD CT FBP, ASIR-V-40%, ASIR-V-80%, DLIR-M, and DLIR-H was 35 HU ± 3, 32 HU ± 3, 29 HU ± 4, 27 HU ± 4, and 23 HU ± 4, respectively ($P <$ .001). Among all ULD CT reconstructions,

DLIR-H had the lowest background noise, with a relative reduction of 34.7% from the baseline FBP level, which was higher than that of ASIR-V-80% (16.6% [$P <$ .001]) and reached the CECT level (36.0% [$P >$ .99]). Thereafter, ASIR-V-80% and DLIR-H images were selected for nodule detection and measurement accuracy analysis because they showed better image quality than ASIR-V-40% and DLIR-M (Fig 4). Additionally, the image noise and background noise of lung tissues in the 0.14-mSv ULD CT group were lower than those in the 0.07-mSv ULD CT group (all $P <$ .05).

The subjective image quality scores of DLIR-M (4.1 for both radiologists) and DLIR-H (4.2 and 4.4) were higher than those of FBP (2.7 and 2.9), ASIR-V-40% (3.0 and 3.4), and ASIR-V-80% (3.8 and 4.0) (all $P <$ .001) (Table E2 [online]). The interobserver consistency was high (intraclass correlation coefficient = 0.85 [95% CI: 0.83, 0.86]).
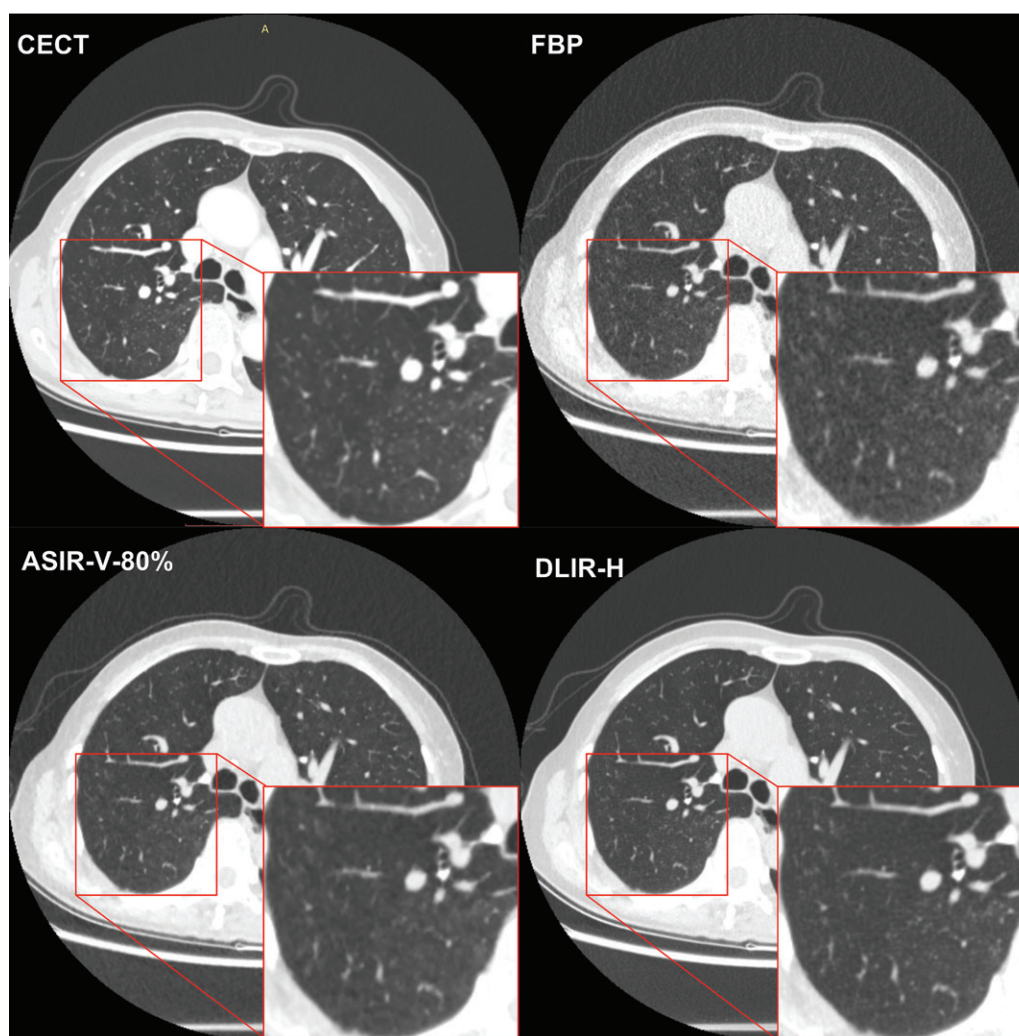

Figure 4: Representative chest CT images from deep learning image reconstruction (DLIR) and adaptive statistical iterative reconstruction-V (ASIR-V) compared with contrast-enhanced chest CT (CECT). A 72-year-old man had a body mass index of 19.0 kg/m², with a clinical history of lung cancer surgery in the right middle lobe 5 years prior. The four images include a CECT image (reconstructed with use of ASIR-V level 50%) as the reference and filtered back projection (FBP), ASIR-V level 80% (ASIR-V-80%), and high-strength DLIR (DLIR-H) images at 0.14-mSv ultra–low-dose CT. The red outlined areas show the details of the magnified images of lung parenchyma. The DLIR-H image shows clearer lung parenchyma and more details than the ASIR-V-80% image, similar to the CECT image.
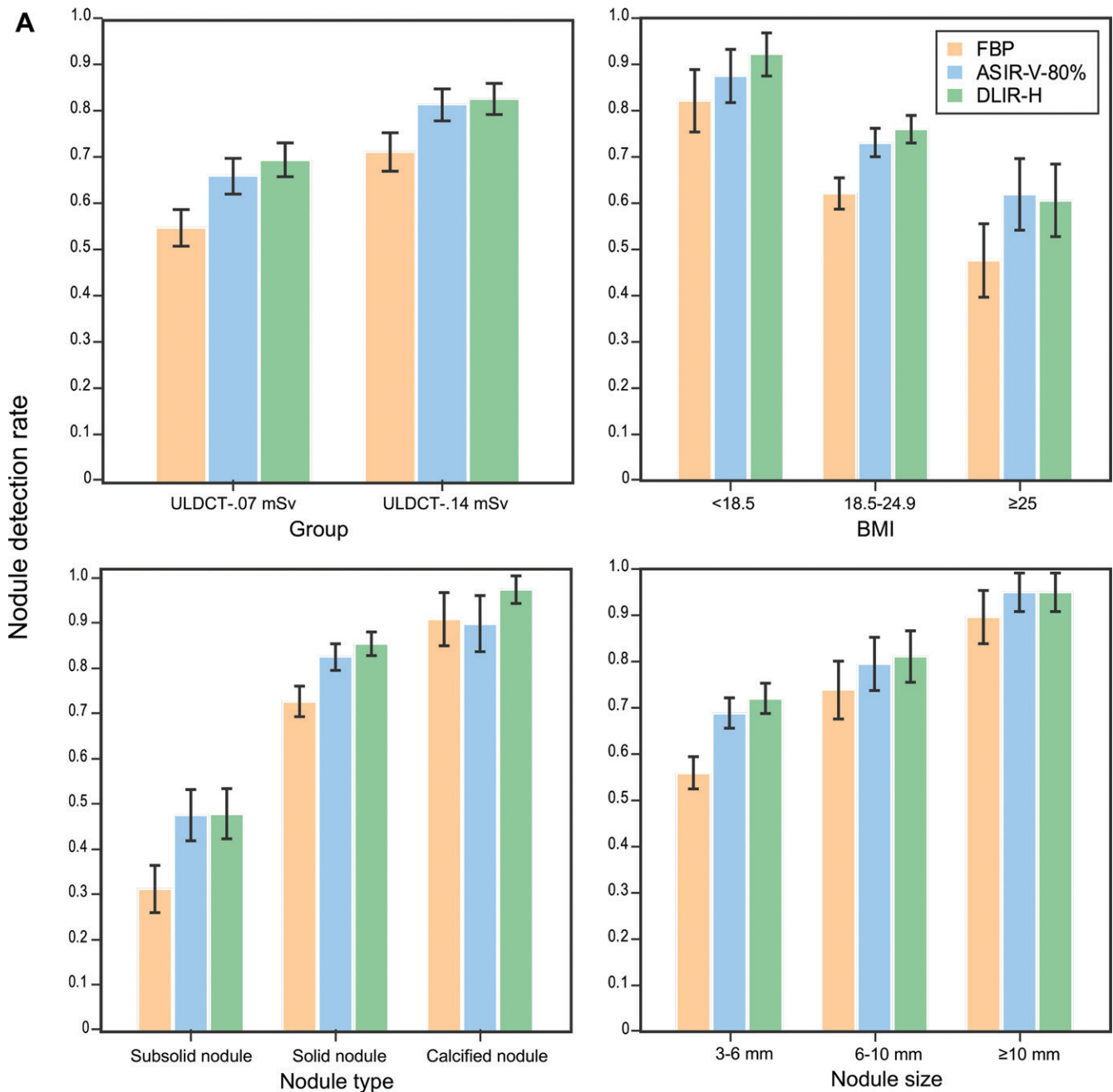
Figure 5: **(A)** Bar graphs show nodule detection rates of deep learning image reconstruction (DLIR) and adaptive statistical iterative reconstruction-V (ASIR-V) compared with filtered back projection (FBP) in relation to radiation dose, participant body mass index (BMI), nodule type, and nodule size. Error bars represent the 95% CI of the mean. ASIR-V-80% = ASIR-V level 80%, DLIR-H = high-strength DLIR, ULDCT = ultra–low-dose CT *(Fig 5 continues).*

Nodule Detection Rate

For nodule detection analysis, 169 of 203 participants (83.3%) with 1066 nodules (mean of six and median of four nodules per participant, ranging from one to 39) detected at CECT were eligible. For ULD CT, 666 of 1066 (62.5%), 781 of 1066 (73.3%), and 808 of 1066 (75.8%) nodules were detected on the FBP, ASIR-V-80%, and DLIR-H images, respectively (Fig 5A, Table E3 [online]). Among the 100 participants who underwent 0.07-mSv ULD CT, 596 nodules were detected at the reference CECT and 329 (55.2%), 396 (66.4%), and 417 nodules (70.0%) were detected at FBP, ASIR-V-80%, and DLIR-H, respectively. Among the 103 participants who under-

went 0.14-mSv ULD CT, 470 nodules were detected at CECT and 337 (71.7%), 385 (81.9%), and 391 nodules (83.2%) were detected at FBP, ASIR-V-80%, and DLIR-H, respectively. The detection rate of DLIR-H in both ULD CT groups was higher than that of FBP and ASIR-V-80% ($P < .001$). After adjusting for age, BMI, radiation dose, nodule size, and nodule type, multivariable logistic regression showed that the odds ratio of DLIR-H in predicting nodule detection on a per-nodule basis was 2.46 ($P < .001$), slightly higher than the 2.04 ($P < .001$) for ASIR-V-80% (Table E4 [online]). On a per-participant basis, the odds ratio of DLIR in predicting nodule detection was 2.69 ($P < .001$), higher than the 2.19 ($P < .001$) of ASIR-V.
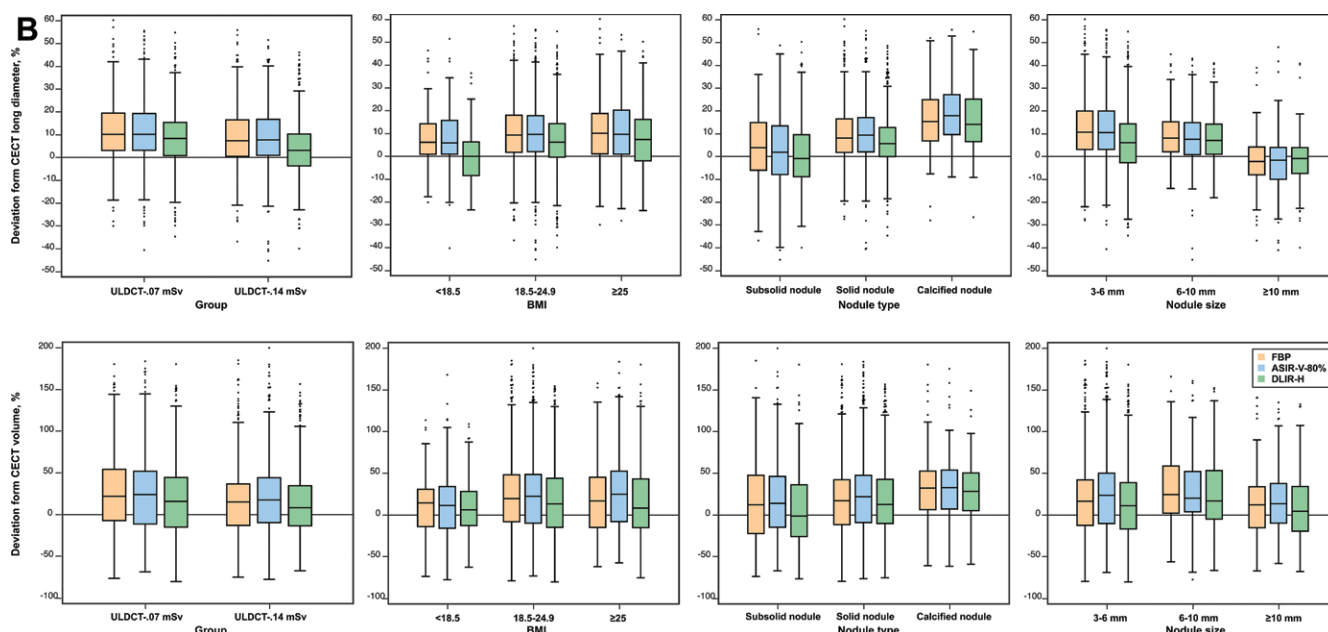
Figure 5 *(continued).* **(B)** Box and whisker plots show deviation of the measured diameter and volume at ultra–low-dose CT (ULDCT) from those at contrast-enhanced chest CT (CECT). The central line represents the median, the box boundaries indicate the first and third quartiles, and the whisker boundaries extend 1.5 quartiles. Other points outside the whisker boundaries are plotted as outliers. ASIR-V-80% = adaptive statistical iterative reconstruction-V level 80%, BMI = body mass index, DLIR-H = high-strength deep learning image reconstruction, FBP = filtered back projection.

For ULD CT, the false-positive rate of DLIR-H was 1.8% (15 of 823 nodules), which was lower than that of ASIR-V-80% (7.4% [62 of 843 nodules]; $P < .001$) and FBP (1.9% [13 of 679 nodules]; $P = .90$).

The detection rates of DLIR-H in the underweight, normal-weight, and overweight groups were 92.2% (119 of 129 nodules), 76.1% (595 of 782 nodules), and 60.6% (94 of 155 nodules), respectively. At 0.14 mSv, the detection rates of DLIR-H in the underweight and normal-weight groups were 99% (85 of 86 nodules) and 81.7% (272 of 333 nodules), respectively. At 0.07 mSv, the detection rate of DLIR-H in the underweight group was 79% (34 of 43 nodules). However, the detection rate in the overweight group was seriously affected by poor image quality and was only 57.7% (60 of 104 nodules) and 67% (34 of 51 nodules) at 0.07 mSv and 0.14 mSv, respectively.

The detection rates of DLIR-H for subsolid, solid, and calcified nodules were 48.2% (151 of 313 nodules), 85.8% (567 of 661 nodules), and 98% (90 of 92 nodules), respectively. Among subsolid nodules, the detection rate of DLIR-H for nodules with CT values of −400 HU or greater was 69.8% (111 of 159 nodules). At 0.14 mSv, the detection rate of DLIR-H for solid nodules was 94.9% (279 of 294 nodules). At 0.07 mSv, the detection rate of DLIR-H was reduced to 78.5% (288 of 367 nodules). The detection rates of DLIR-H for nodules 3–6 mm, 6–10 mm, and 10 mm or larger were 71.7% (546 of 761 nodules), 80.8% (156 of 193 nodules), and 94.6% (106 of 112 nodules), respectively.

Measurement Accuracy
Bland-Altman analysis showed that the percentage difference in the long diameter from that of CECT was 9.3% (95% CI of the mean: 8.0, 10.6), 9.2% (95% CI of the mean: 8.0, 10.4), and 6.2% (95% CI of the mean: 5.0, 7.4) in FBP reconstruction, ASIR-V-80%, and DLIR-H, respectively ($P < .001$) (Fig E3A [online]). Bland-Altman analysis showed that the percentage difference in volume from that of CECT was 19.0% (95% CI of the mean: 14.6, 23.5), 21.0% (95% CI: 16.9, 25.1), and 14.4% (95% CI: 10.6, 18.2) in the three reconstructions, respectively (Fig E3B [online]).

DLIR-H showed the least error in nodule size from CECT compared with that at FBP and ASIR-V-80% (Fig 5B). DLIR-H slightly underestimated the long diameter and volume of subsolid nodules by 1.1% and 0.9%, respectively, but overestimated those of solid (6.2% and 16.7%) and calcified nodules (14.1% and 28%). The overestimation of DLIR-H nodule measurements at 0.14 mSv was lower than at 0.07 mSv (long diameter, 3.5% vs 8.9%; volume, 11.0% vs 17.9%; both $P < .001$).

Malignancy-related Imaging Features
Postsurgical histologic results were available for 41 nodules in 32 of 203 participants (15.8%), including 14 benign and pre-invasive lesions and 27 malignant lesions (Table, Table E4 [online]). Four radiologists evaluated 27 malignant nodules and found 276 malignancy-related features at CECT, including 79 lobulated shapes, 79 spiculated margins, 66 pleural tags, and 52 air bronchograms. The overall intraclass correlation coefficient among the four radiologists in determining malignancy-related features was 0.77 (95% CI: 0.73, 0.80), and the intraclass correlation coefficients for lobulated shapes, spiculated margins, pleural tags, and air bronchograms were 0.75, 0.70, 0.74, and 0.79, respectively.

The detection rate of all malignancy-related features for DLIR-H was 81.5% (225 of 276 features), higher than that of FBP (72.5% [200 of 276 features]) and ASIR-V-80% (77.5% [214 of 276 features]) ($P = .04$) (Fig 6, Table E6 [online]). The
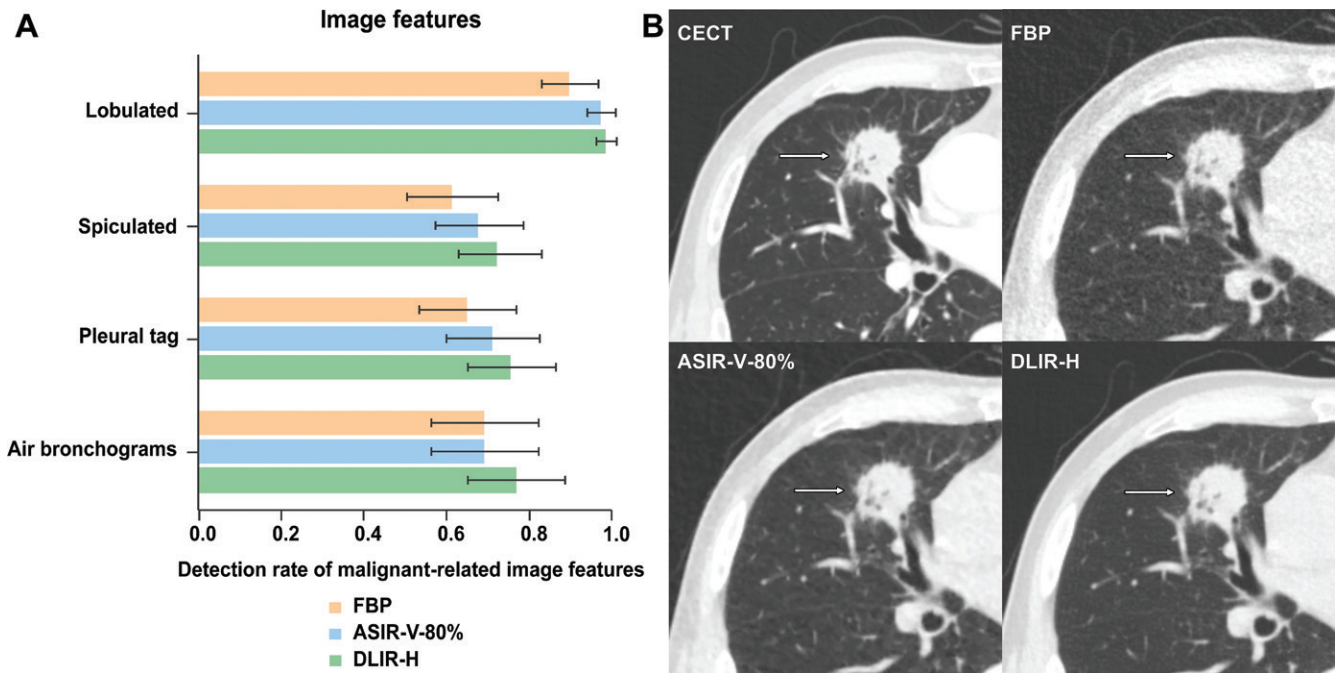
Figure 6: Malignancy-related imaging features. **(A)** Bar graph shows detection rate of malignancy-related imaging features among filtered back projection (FBP), adaptive statistical iterative reconstruction-V level 80% (ASIR-V-80%), and high-strength deep learning image reconstruction (DLIR-H). **(B)** A 64-year-old man with a body mass index of 23.9 kg/m² was hospitalized for surgery because of the CT finding of a pulmonary nodule (28.2 mm in long diameter in the right middle lobe), which was histologically proven as invasive adenocarcinoma. The array shows a contrast-enhanced chest CT (CECT) image and three 0.14-mSv ultra–low-dose CT images of FBP, ASIR-V-80%, and DLIR-H. The DLIR-H image shows that this nodule (arrow) has three malignancy-related imaging features, including a lobulated shape, spiculated margin, and air bronchograms.

detection rates of DLIR-H for the four features were 99% (78 of 79 lobulated shapes), 72% (57 of 79 spiculated margins), 76% (50 of 66 pleural tags), and 77% (40 of 52 air bronchograms).

If at least one malignancy-related feature could be detected in each nodule, we considered the nodule to have a malignant tendency. Thus, the diagnosis rates of malignant tendency identified with use of CECT, FBP reconstruction, ASIR-V-80%, and DLIR-H were 92.6% (100 of 108 nodule evaluations by four radiologists), 79.6% (86 of 108), 86.1% (93 of 108), and 88.0% (95 of 108), respectively.

Discussion

We evaluated deep learning image reconstruction (DLIR) in chest radiography–equivalent ultra–low-dose CT in terms of image quality and nodule detection and compared this method to the traditional method, adaptive statistical iterative reconstruction-V (ASIR-V). DLIR showed better image noise reduction both in lung tissue noise (13.8% vs 11.3%; P = .03) and air background noise (34.7% vs 16.6%; $P < .001$) than ASIR-V from the baseline of the filtered back projection (FBP) level. DLIR also showed a higher detection rate of lung nodules compared with those of FBP (75.8% vs 62.5%; $P < .001$) and ASIR-V (73.3%; P = .18), provided higher nodule measurement accuracy compared with that of ASIR-V in long diameter (6.2% vs 9.2%; $P < .001$) and volume (14.4% vs 21.0%; $P < .001$), and displayed more malignancy-related imaging features of nodules than ASIR-V (81.5% [225 of 276 features] vs 77.5 [214 of 276 features]; P = .04).

With DLIR, we performed chest CT scans at 0.07 and 0.14 mSv, lower than the 0.13–0.49 mSv reported in previous ULD

CT studies (7–9) and equivalent to the 0.03–0.1 mSv in chest radiography (10). The difference in denoising performance between DLIR and iterative reconstruction was investigated. Several studies have reported that DLIR can reduce more noise than iterative reconstruction in conventional-dose CT (13,17,23) and low-dose CT (12,15). Benz et al (15) compared the image noise between DLIR-H and ASIR-V in low-dose coronary CT angiography and found that DLIR-H yielded the higher image quality.

We found that the long diameter and volume of nodules were over- or underestimated in ULD CT with CECT as a reference, which was also observed in some previous studies of ULD CT (24) and low-dose CT (25,26). In our study, DLIR-H slightly underestimated the long diameter and volume of subsolid nodules but overestimated the solid and calcified nodules. Due to the partial volume effect, there is a transition zone between high-attenuation (nodules) and low-attenuation (pulmonary parenchyma) objects on CT images, which is important for accurate volumetry (27). Since the measured values were determined by the segmentation algorithm, the transition zone around solid or calcified nodules would lead to overestimation. However, subsolid nodules (especially pure ground-glass nodules) have blurred margins and low attenuation and are therefore indistinguishable from lung parenchyma, which would reduce the accuracy of nodule segmentation, resulting in underestimation.

Subsequently, we observed malignancy-related imaging features at different ULD CT reconstruction sequences. Among them, DLIR images had the most abundant subtle features, and the overall malignant feature detection rate was 81.5%.

According to the World Health Organization, BMI of 30 kg/m^2 or greater is considered to indicate obesity and chronic disease. Because our ULD CT study was preferentially carried out in a normal population for screening purposes, participants with obesity were not included. Previous studies on ULD CT also excluded participants with obesity (28,29). Vardhanabhuti et al (30) found that the detectability of pulmonary nodules was reduced in individuals with higher BMIs due to higher image noise, and they concluded that ULD CT may not be suitable for overweight people. In our study, seven of 226 candidates had a BMI higher than 30 kg/m^2. The obesity incidence rate was not high and therefore did not affect the screening target of ULD CT.

Our study has limitations. First, because of ethics concerns, noncontrast CT scans of high and low radiation were not repeated for each participant. Instead, we used conventional-dose CECT as a reference from a routine chest CT examination (noncontrast and contrast-enhanced scans). Second, a head-to-head comparison between conventional-dose and ULD noncontrast CT was not performed. Third, we used a deep learning–based system to detect and measure lung nodules. This system was trained based on conventional-dose data sets, not on ULD CT data.

Compared with adaptive statistical iterative reconstruction-V, deep learning image reconstruction (DLIR) improved the image quality by reducing image noise and increasing the lung nodule detection rate. DLIR also improved the nodule measurement accuracy and displayed more detailed malignancy-related imaging features of nodules. With the introduction of this specific DLIR algorithm, it is feasible to use chest radiography–equivalent chest CT acquisition as low as 0.14 mSv in underweight and normal-weight participants to detect and characterize lung nodules, which is crucial to facilitate the clinical implementation of large-scale lung cancer screening. Nevertheless, only a limited number of nodules had pathologic results in this study, and with larger sample size, the results may have been different. More studies are needed to explore the clinical application and image optimization of DLIR.

Author contributions: Guarantors of integrity of entire study, **B.J., N.L., X.X.**; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **B.J., N.L., X.S., R.V., X.X.**; clinical studies, **B.J., N.L., G.H.d.B.**; statistical analysis, **B.J., N.L.**; and manuscript editing, **B.J., N.L., X.S., J.L., R.V., X.X.**

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68(6):394–424 [Published correction appears in CA Cancer J Clin 2020;70(4):313.].

2. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020;382(6):503–513.

3. Bach PB, Mirkin JN, Oliver TK, et al. Benefits and harms of CT screening for lung cancer: a systematic review. JAMA 2012;307(22):2418–2429.

4. Thibault JB, Sauer KD, Bouman CA, Hsieh J. A three-dimensional statistical approach to improved image quality for multislice helical CT. Med Phys 2007;34(11):4526–4544.

5. Willemink MJ, Leiner T, de Jong PA, et al. Iterative reconstruction techniques for computed tomography part 2: initial results in dose reduction and image quality. Eur Radiol 2013;23(6):1632–1642.

6. Tang H, Liu Z, Hu Z, et al. Clinical value of a new generation adaptive statistical iterative reconstruction (ASIR-V) in the diagnosis of pulmonary nodule in low-dose chest CT. Br J Radiol 2019;92(1103):20180909.

7. Messerli M, Kluckert T, Knitel M, et al. Ultralow dose CT for pulmonary nodule detection with chest x-ray equivalent dose – a prospective intra-individual comparative study. Eur Radiol 2017;27(8):3290–3299.

8. Zanon M, Pacini GS, de Souza VVS, et al. Early detection of lung cancer using ultra-low-dose computed tomography in coronary CT angiography scans among patients with suspected coronary heart disease. Lung Cancer 2017;114:1–5.

9. Carey S, Kandel S, Farrell C, et al. Comparison of conventional chest x ray with a novel projection technique for ultra-low dose CT. Med Phys 2021;48(6):2809–2815.

10. Quaia E, Baratella E, Cernic S, et al. Analysis of the impact of digital tomosynthesis on the radiological investigation of patients with suspected pulmonary lesions on chest radiography. Eur Radiol 2012;22(9):1912–1922.

11. Geyer LL, Schoepf UJ, Meinel FG, et al. State of the art: iterative CT reconstruction techniques. Radiology 2015;276(2):339–357.

12. Kim JH, Yoon HJ, Lee E, Kim I, Cha YK, Bak SH. Validation of deep-learning image reconstruction for low-dose chest computed tomography scan: emphasis on image quality and noise. Korean J Radiol 2021;22(1):131–138.

13. Jensen CT, Liu X, Tamm EP, et al. Image quality assessment of abdominal CT by use of new deep learning image reconstruction: initial experience. AJR Am J Roentgenol 2020;215(1):50–57.

14. Liu P, Wang M, Wang Y, et al. Impact of deep learning-based optimization algorithm on image quality of low-dose coronary CT angiography with noise reduction: a prospective study. Acad Radiol 2020;27(9):1241–1248.

15. Benz DC, Benetos G, Rampidis G, et al. Validation of deep-learning image reconstruction for coronary computed tomography angiography: impact on noise, image quality and diagnostic accuracy. J Cardiovasc Comput Tomogr 2020;14(5):444–451.

16. Racine D, Becce F, Viry A, et al. Task-based characterization of a deep learning image reconstruction and comparison with filtered back-projection and a partial model-based iterative reconstruction in abdominal CT: a phantom study. Phys Med 2020;76:28–37.

17. Kim I, Kang H, Yoon HJ, Chung BM, Shin NY. Deep learning-based image reconstruction for brain CT: improved image quality compared with adaptive statistical iterative reconstruction-Veo (ASIR-V). Neuroradiology 2021;63(6):905–912.

18. World Health Organization. The SuRF Report 2: Surveillance of Chronic Disease Risk Factors: Country-Level Data and Comparable Estimates. https://apps.who.int/iris/bitstream/handle/10665/43190/9241593024_eng.pdf. Published 2005. Accessed September 30, 2021.

19. Ludwig M, Chipon E, Cohen J, et al. Detection of pulmonary nodules: a clinical study protocol to compare ultra-low dose chest CT and standard low-dose CT using ASIR-V. BMJ Open 2019;9(8):e025661.

20. Bongartz G. European guidelines on quality criteria for computed tomography. http://www.drs.dk/guidelines/ct/quality/htmlindex.htm. Accessed August 26, 2021.

21. Ettinger DS, Wood DE, Aggarwal C, et al. NCCN Guidelines Insights: Non-Small Cell Lung Cancer, Version 1.2020. J Natl Compr Canc Netw 2019;17(12):1464–1472.

22. Nair A, Bartlett EC, Walsh SLF, et al. Variable radiological lung nodule evaluation leads to divergent management recommendations. Eur Respir J 2018;52(6):1801359.

23. Tatsugami F, Higaki T, Nakamura Y, et al. Deep learning-based image restoration algorithm for coronary CT angiography. Eur Radiol 2019;29(10):5322–5329.

24. Jin S, Zhang B, Zhang L, Li S, Li S, Li P. Lung nodules assessment in ultra-low-dose CT with iterative reconstruction compared to conventional dose CT. Quant Imaging Med Surg 2018;8(5):480–490.

25. Chen B, Barnhart H, Richard S, Colsher J, Amurao M, Samei E. Quantitative CT: technique dependence of volume estimation on pulmonary nodules. Phys Med Biol 2012;57(5):1335–1348.
26. 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(8):1979–1984.
27. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: a comparative phantom study. Int J Cardiovasc Imaging 2008;24(5):547–556.
28. Kroft LJM, van der Velden L, Girón IH, Roelofs JJH, de Roos A, Geleijns J. Added value of ultra–low-dose computed tomography, dose equivalent to chest x-ray radiography, for diagnosing chest pathology. J Thorac Imaging 2019;34(3):179–186.
29. Ye K, Zhu Q, Li M, Lu Y, Yuan H. A feasibility study of pulmonary nodule detection by ultralow-dose CT with adaptive statistical iterative reconstruction-V technique. Eur J Radiol 2019;119:108652.
30. Vardhanabhuti V, Pang CL, Tenant S, Taylor J, Hyde C, Roobottom C. Prospective intra-individual comparison of standard dose versus reduced-dose thoracic CT using hybrid and pure iterative reconstruction in a follow-up cohort of pulmonary nodules—effect of detectability of pulmonary nodules with lowering dose based on nodule size, type and body mass index. Eur J Radiol 2017;91:130–141.