BMC Medical Imaging

RESEARCH

Open Access

Deep learning reconstruction improves computer-aided pulmonary nodule detection and measurement accuracy for ultra-low-dose chest CT

Jinhua Wang[1], Zhenchen Zhu[1], Zhengsong Pan[1,2], Weixiong Tan[3], Wei Han[4], Zhen Zhou[3], Ge Hu[5], Zhuangfei Ma[6], Yinghao Xu[6], Zhoumeng Ying[1,2], Xin Sui[1], Zhengyu Jin[1], Lan Song[1*†] and Wei Song[1*†]

Abstract

Purpose To compare the image quality and pulmonary nodule detectability and measurement accuracy between deep learning reconstruction (DLR) and hybrid iterative reconstruction (HIR) of chest ultra-low-dose CT (ULDCT).

Materials and Methods Participants who underwent chest standard-dose CT (SDCT) followed by ULDCT from October 2020 to January 2022 were prospectively included. ULDCT images reconstructed with HIR and DLR were compared with SDCT images to evaluate image quality, nodule detection rate, and measurement accuracy using a commercially available deep learning–based nodule evaluation system. Wilcoxon signed-rank test was used to evaluate the percentage errors of nodule size and nodule volume between HIR and DLR images.

Results Eighty-four participants (54 ± 13 years; 26 men) were finally enrolled. The effective radiation doses of ULDCT and SDCT were 0.16 ± 0.02 mSv and 1.77 ± 0.67 mSv, respectively ($P < 0.001$). The mean ± standard deviation of the lung tissue noises was 61.4 ± 3.0 HU for SDCT, 61.5 ± 2.8 HU and 55.1 ± 3.4 HU for ULDCT reconstructed with HIR-Strong setting (HIR-Str) and DLR-Strong setting (DLR-Str), respectively ($P < 0.001$). A total of 535 nodules were detected. The nodule detection rates of ULDCT HIR-Str and ULDCT DLR-Str were 74.0% and 83.4%, respectively ($P < 0.001$). The absolute percentage error in nodule volume from that of SDCT was 19.5% in ULDCT HIR-Str versus 17.9% in ULDCT DLR-Str ($P < 0.001$).

Conclusion Compared with HIR, DLR reduced image noise, increased nodule detection rate, and improved measurement accuracy of nodule volume at chest ULDCT.

Clinical trial number Not applicable.

Keywords Lung nodules, Tomography, X-ray computed, Deep learning reconstruction, Image quality

†Lan Song and Wei Song contributed equally to this work.

*Correspondence:
Lan Song
songl@pumch.cn
Wei Song
cjr.songwei@vip.163.com

Full list of author information is available at the end of the article

Introduction

Chest CT offers good visualization of the lung parenchyma and high diagnostic accuracy, so it is widely used for evaluating patients with suspected disease in the chest, accounting for a high proportion of the total radiation dose. Lung cancer has received widespread attention because it is the most common malignant tumor and the leading cause of cancer-related deaths worldwide [1]. In clinical practice, diagnostic-grade standard-dose CT (SDCT) examinations are common. During such examinations, incidental indeterminate nodules are often identified, requiring follow-up CT evaluations to monitor for changes [2], which could result in an undesirable cumulative burden of radiation exposure to these patient populations. Therefore, the performance of low-dose CT (LDCT) and ultra-low-dose CT (ULDCT) to detect pulmonary nodules are gradually being explored [3, 4].

It is generally thought that reducing radiation doses by applying the "as low as reasonably achievable" principle [5] is necessary. The traditional image reconstruction method, filtered back projection, markedly increases image noise and degrades image quality in reduced-dose acquisitions [6]. Therefore, iterative reconstruction (IR) algorithms, such as hybrid IR (HIR) and advanced model-based IR, have been developed to denoise images and maintain image quality, especially in thoracic LDCT [3, 7–9]. However, IR algorithms may affect the display of subtle imaging features and diagnostic accuracy [10, 11], thus limiting the clinical application potential of ULDCT if the radiation dose is further reduced.

With the rapid development of artificial intelligence (AI) and machine-learning techniques [12, 13], deep learning reconstruction (DLR) algorithms based on convolutional neural networks have been utilized to improve CT image quality [14]. Studies on the DLR algorithm have demonstrated better lung parenchyma assessments by thoracic CT [15–17], especially for pulmonary nodules. Quantitative measurements, such as nodule size and volume, are essential to evaluate a pulmonary nodule in an initial or follow-up CT examination. Currently, commercially available AI-aided diagnostic systems are implemented in clinical practice for nodule detection and measurement [18], increasing sensitivity and reducing the workload of radiologists. Recent studies have shown that AI assistance enhances nodule assessment performance, underscoring the great potential of AI-assisted systems in lung nodule evaluation [19, 20]. However, the incremental diagnostic value of the DLR algorithm combined with an AI-assisted system in chest ULDCT has not been fully explored. Therefore, the objective of this study was to evaluate the image quality, nodule detection rate and measurement accuracy of DLR in comparison to the widely used HIR in ULDCT, using non-contrast diagnostic chest SDCT as a reference.

Materials and methods

Study population

The Institutional Review Board of our hospital approved this prospective study (HS-2540). Consecutive participants at our hospital from October 2020 to January 2022 were prospectively included, and written informed consent was obtained. The inclusion criteria were as follows: *(a)* age > 18 years, *(b)* incidental detected pulmonary nodule detected previously for follow-up or suspected nodules on X-ray for further examination, and *(c)* agreed to receive SDCT and ULDCT at the same encounter. The exclusion criteria were as follows: *(a)* prior thoracic surgery, *(b)* diffuse pulmonary disease, *(c)* multiple lung metastases, and *(d)* severe motion or respiratory artifacts. In accordance with the World Health Organization's classification [21] of body mass index (BMI), the included participants were categorized into two groups: underweight/normal weight group (< 25 kg/m^2), and overweight/obese group (≥ 25 kg/m^2). A sample size of 75 was determined to provide sufficient statistical power for this experimental study [22].

CT examination, image reconstruction and radiation dose

Figure 1 shows the research pipeline steps. All participants underwent ULDCT after SDCT scanning with a 320-row detector CT scanner (Aquilion ONE GENESIS Edition; Canon Medical Systems). SDCT was performed at 120 kV and automatic mA settings, and ULDCT was performed at 120 kV and 20 mA. Detailed acquisition parameters are presented in Supplementary Note 1 and Table S1.

Five groups of CT images were reconstructed, including four groups of ULDCT images and one group of SDCT images. The DLR algorithm (Advanced Intelligent Clear-IQ Engine, Canon Medical Systems), detailed in Supplementary Note 2 and Fig. S1, and the HIR algorithm (adaptive iterative dose reduction 3D, Canon Medical Systems) were applied. The HIR and DLR algorithms provide several selectable reconstruction strength presets. The ULDCT image groups were reconstructed with HIR-Standard setting (HIR-Std), HIR-Strong setting (HIR-Str), DLR-Standard setting (DLR-Std), and DLR-Strong setting (DLR-Str). The SDCT image group reconstructed with HIR-Std was used as a reference for image quality and nodule detection/quantification assessment because it is commonly performed in our clinical practice. All images were reconstructed with a lung kernel, and the reconstruction thickness/interval was 1.0 mm/0.8 mm. The radiation dose parameters were recorded (Supplementary Note 3).

Objective and subjective image quality assessment

The objective image noise in the whole lung tissue and upper air background in vitro, expressed as the standard

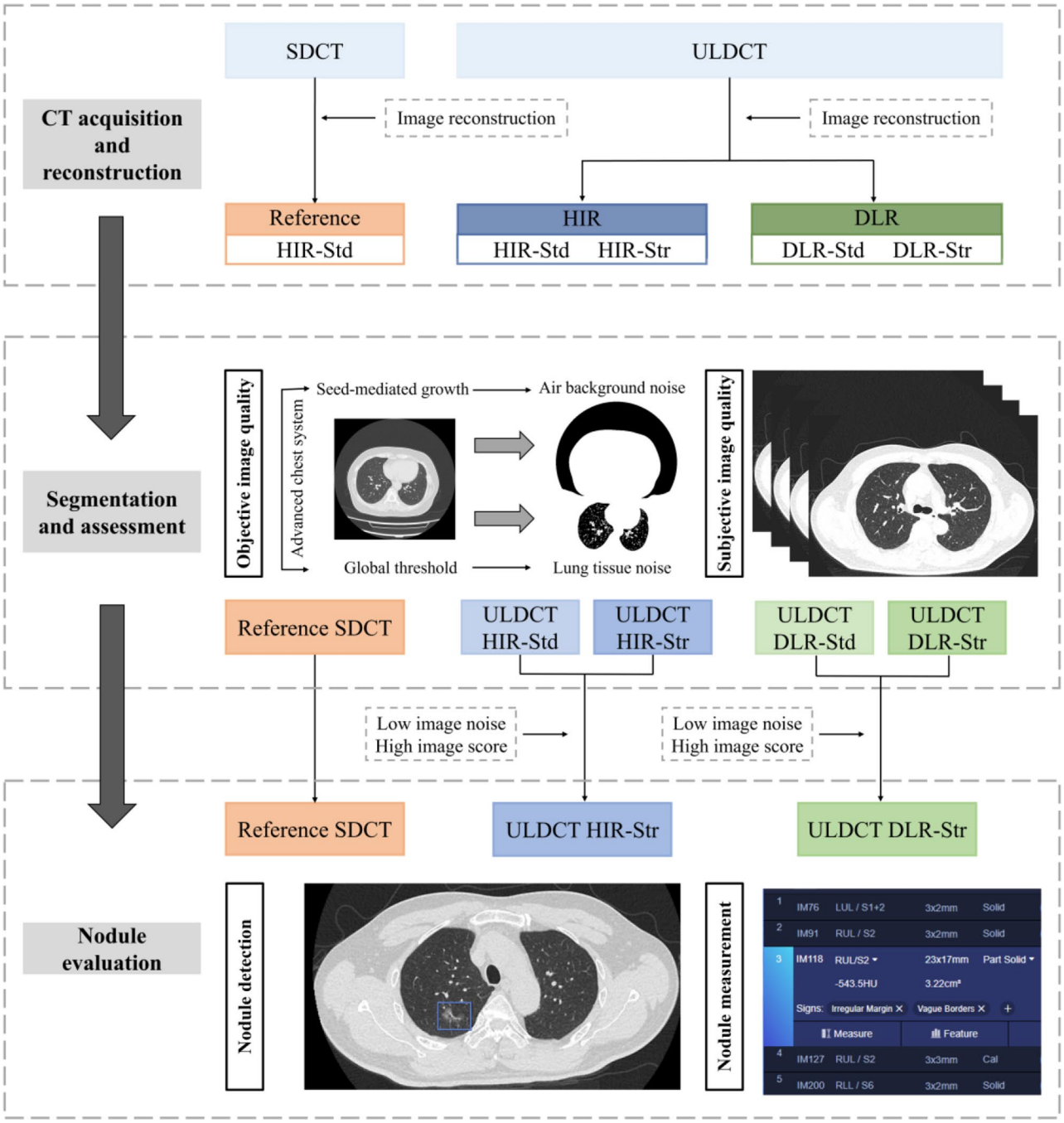

Fig. 1 Overview of the research pipeline. Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; Std, standard setting; Str, strong setting

deviation of CT values, was determined through three-dimensional segmentation (Supplementary Note 4, Fig. S2). For the subjective assessment of chest image quality, the de-identified images were presented in a randomized order to two chest radiologists (L.S. and J.W.) with 19 years and 4 years of experience using a 5-point Likert scale [16] (1, unacceptable; 2, poor; 3, moderate; 4, good; and 5, excellent).

Nodule detection rate and measurement accuracy analysis

Based on image quality results of ULDCT images, one group with lower image noise and higher subjective image quality score was selected from HIR-Std and HIR-Str images to represent HIR, and one group from DLR-Std and DLR-Str images to represent DLR following the same principle described by Jiang et al. [15]. After shielding participant information, CT images were transmitted to a commercially available deep learning–based

nodule evaluation AI system (Yitu AICare CT Chest system, Deepwise Inc., Supplementary Note 5, Fig. 2). This system has garnered official certification from the National Medical Products Administration (NMPA) in China (Registration No. 20213211094) for clinical use. The performance of the AI system has been published [23]. Automatic detection and segmentation of lung nodules were performed at the lung window setting (window width, 1200 HU; window level, – 600 HU). Then a senior chest radiologist (L.S.) independently read and confirmed all the detected nodules by using the results of SDCT as a reference with the aid of an AI system (Fig. 2), simulating the clinical scenario when we used an AI system. The long diameter and volume of the detected nodules measured by the same AI system at SDCT were used as references. Nodules with long diameter ≥ 3 mm but < 30 mm were included for further analysis, and nodules were categorized into two subgroups (3–6 mm; ≥ 6 mm) and three types (subsolid, solid, and calcified). According to previous study [15], a detection error of approximately 20% may be deemed acceptable in this study, with SDCT as the reference standard.

To analyze differences between the measured data and the reference data, the systematic percentage error (SPE) [15, 24] and absolute percentage error (APE) [25, 26] of long diameter and volume were determined as follows:

$$\mathrm{SPE} = (\mathrm{Vm} - \mathrm{Vr})/\mathrm{Vr} \times 100 \quad (1)$$

$$\mathrm{APE} = |(\mathrm{Vm} - \mathrm{Vr})/\mathrm{Vr}| \times 100 \quad (2)$$

where V_m represents the nodule long diameter/volume value measured by ULDCT, and V_r represents the reference nodule long diameter/volume value measured by SDCT.

Statistical analysis

Statistical tests were performed with SPSS 20.0 (IBM Corp.) and MedCalc 18.2 (MedCalc Software). Normality of data distribution was tested with the Shapiro–Wilk test. Continuous variables were expressed as mean ± SD or median and IQR according to the normality or non-normality of the distribution. Radiation dose between ULDCT and SDCT was compared using the Student's *t*-test for paired samples. The Friedman test with Bonferroni correction was used for image quality assessment among groups with different dose levels and reconstruction techniques. Cohen's kappa test (κ) was used to indicate the interobserver agreement of the two radiologists. The Fisher exact test was used to compare the detection rates of nodules. ULDCT and SDCT images were compared using Passing-Bablok regression and Bland–Altman analysis. The Wilcoxon signed-rank test was used to evaluate the APEs in long diameter and volume between HIR and DLR images. Univariate analyses were performed for factors affecting nodule detection, including age, sex, BMI, reconstruction algorithms, nodule size,

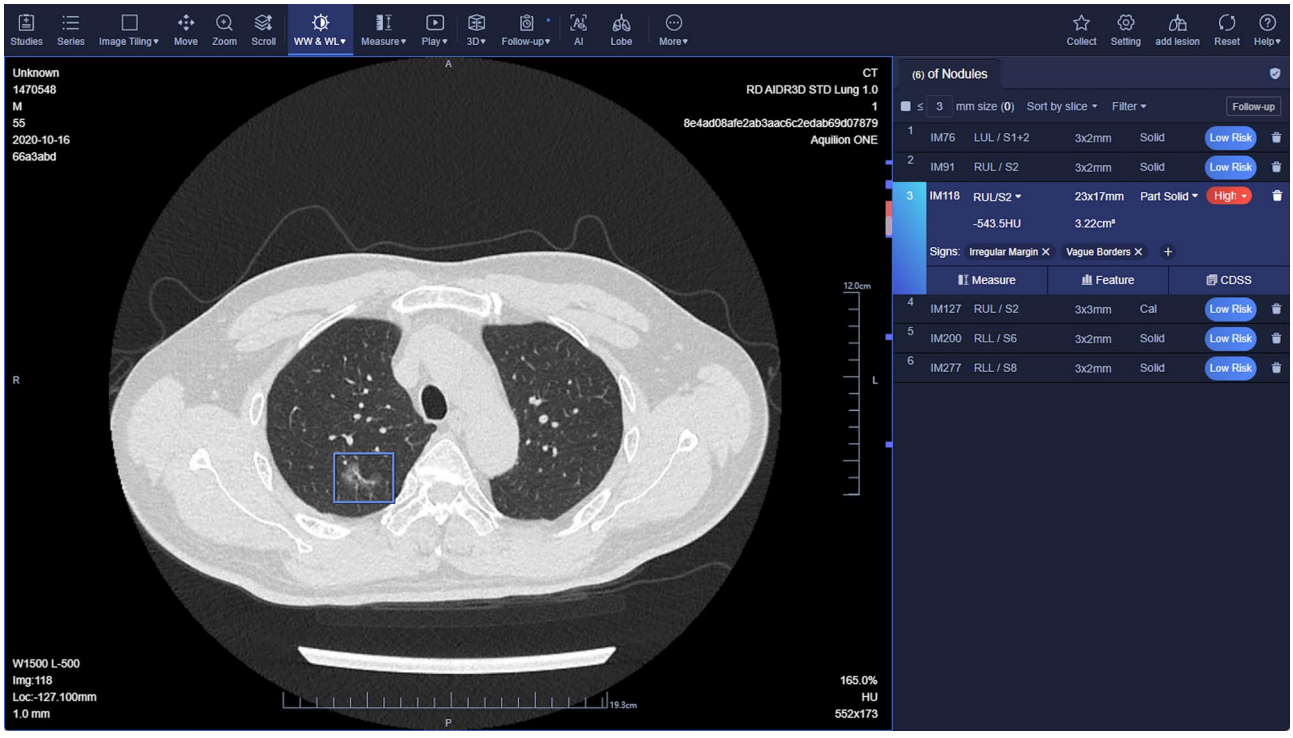

Fig. 2 Software interface of the deep learning–based nodule evaluation system. The nodule evaluation system automatically lists six lung nodules in this patient and displays the location, diameter, volume, and density type of these nodules after clicking the nodule list in the right panel

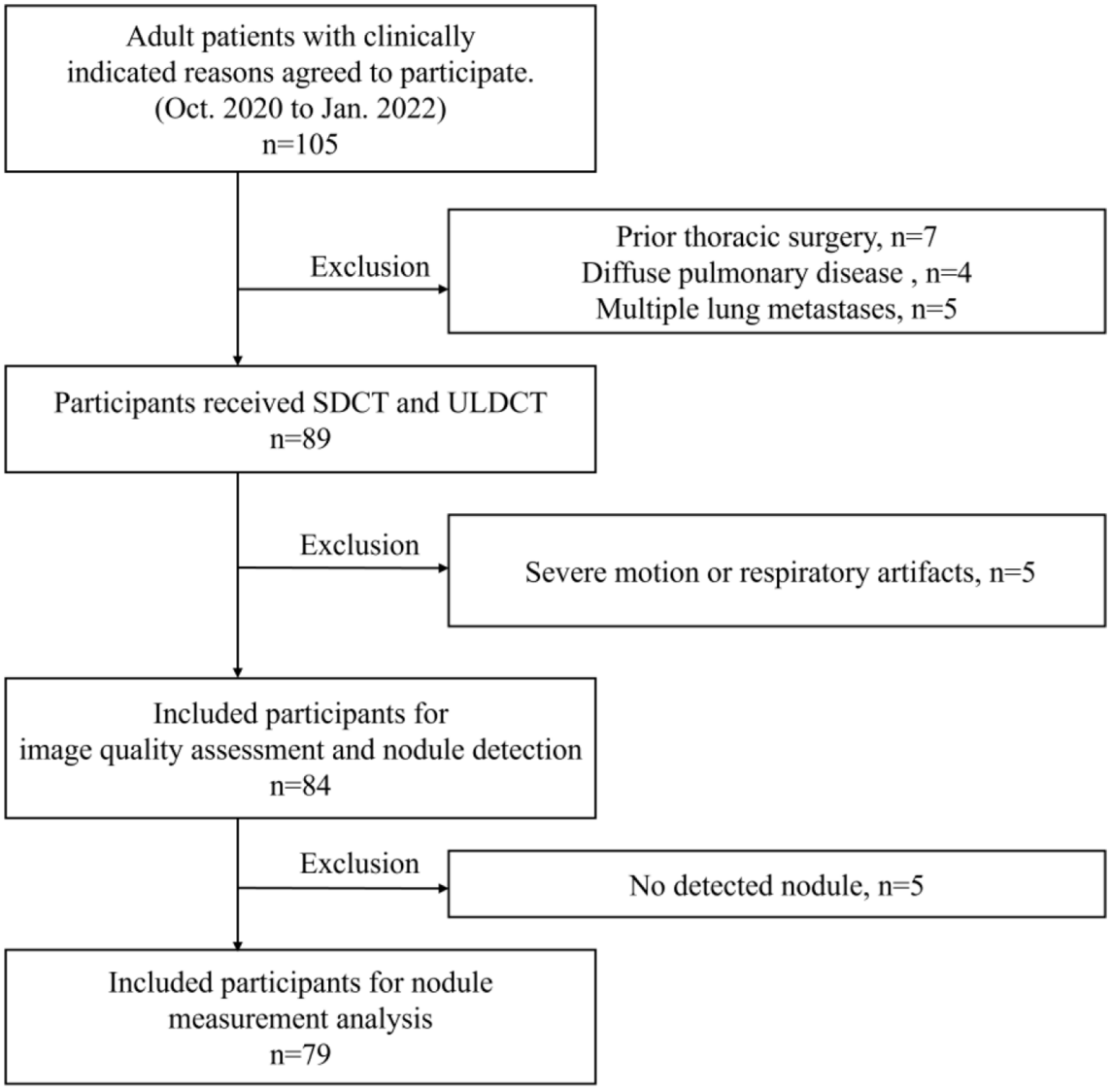

Fig. 3 Inclusion and exclusion flowchart of study participants. Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; n, number of patients

Table 1 Participants' characteristics and radiation dose

Variable	
Age, years (range)	54 ± 13 (26–85)
Sex, Men/Women	26/58
Height, m (range)	1.7 ± 0.8
Body weight, kg (range)	65.3 ± 11.6
BMI, kg/m^2 (range)	24.0 ± 3.0 (16.4–30.4)
$CTDI_{vol}$, mGy [SDCT]	4.2 ± 1.6
$CTDI_{vol}$, mGy [ULDCT]	0.4
DLP, mGy×cm [SDCT]	126.2 ± 47.6
DLP, mGy×cm [ULDCT]	11.3 ± 1.1
ED, mSv [SDCT]	1.77 ± 0.67
ED, mSv [ULDCT]	0.16 ± 0.02

Note: Unless otherwise specified, data are presented as the mean ± SD

Abbreviations: BMI, body mass index; $CTDI_{vol}$, volume CT dose index; DLP, dose-length product; ED, effective dose; SDCT, standard-dose CT; ULDCT, ultra-low-dose CT

and nodule type. Then, multivariable logistic regression was used to analyze the predictive value of reconstruction algorithms for nodule detection after adjusting for confounding factors for variables with $P < 0.10$ in the univariate analyses. Two-tailed $P < 0.05$ was considered to indicate statistical significance.

Results

Participants' characteristics and radiation dose

Eighty-four participants were finally enrolled according to the inclusion and exclusion flowchart shown in Fig. 3. Table 1 shows the participants' characteristics and radiation dose. Compared with SDCT, the effective dose of ULDCT was reduced by approximately 91.0% ($P < 0.001$). Of all 84 participants, 53 (63.1%) were divided into the underweight/normal group, and 31 (36.9%) were divided into the overweight/obese group.

Table 2 Objective and subjective image quality assessment

Variable	SDCT	ULDCT				*P* value*
		HIR-Std	HIR-Str	DLR-Std	DLR-Str	
Lung tissue noise, *HU*	61.4 ± 3.0	67.7 ± 2.5	61.5 ± 2.8	61.0 ± 2.4	55.1 ± 3.4	< 0.001
Background noise, *HU*	45.7 ± 5.1	58.1 ± 6.1	50.6 ± 6.2	46.1 ± 5.9	41.2 ± 6.2	< 0.001
Image quality radiologist 1	4.9 ± 0.3	3.3 ± 0.5	3.6 ± 0.5	4.3 ± 0.5	4.6 ± 0.5	< 0.001
Image quality radiologist 2	5.0 ± 0.1	3.4 ± 0.5	3.7 ± 0.5	4.4 ± 0.6	4.6 ± 0.5	< 0.001

Note: Data are presented as mean ± SD

Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; Std, standard setting; Str, strong setting

* *P* values reported are for the overall comparison of groups with different dose levels and reconstruction techniques

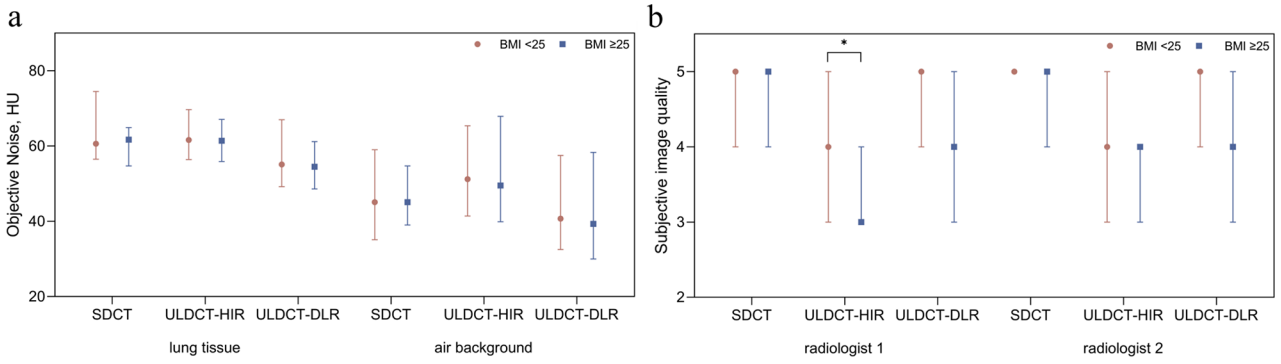

Fig. 4 Objective image quality assessments, including lung tissue noise, air background noise, and subjective image quality scores, performed by two radiologists on different BMI subgroups. Data are presented as median with range. * $P < 0.05$. Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; BMI, body mass index

Image quality

The CT values of the lung parenchyma were − 863 ± 19 HU for SDCT, -864 ± 20 HU for ULDCT DLR-Std, -862 ± 21 for ULDCT DLR-Str, -860 ± 18 HU for ULDCT HIR-Std and − 858 ± 20 HU for ULDCT HIR-Str, with no significant differences across different reconstruction methods ($P = 0.10$). The lung tissue noise and air background noise in SDCT images and four groups of ULDCT images are shown in Table 2. Pairwise comparisons showed that DLR-Str had the lowest lung tissue noise and background noise among all ULDCT reconstructions, even lower than the reference SDCT (all $P < 0.001$). The lung tissue noises of ULDCT DLR-Std and ULDCT HIR-Str were similar to the lung tissue noise of SDCT (all $P > 0.05$). The background noise of ULDCT DLR-Std was lower than ULDCT HIR-Str ($P < 0.001$) and reached the SDCT level ($P > 0.99$).

The subjective image quality scores of DLR-Std (4.3 and 4.4 rated by the two radiologists, respectively) and DLR-Str (4.6 for both radiologists) were higher than those of HIR-Std (3.3 and 3.4) and HIR-Str (3.6 and 3.7) (all $P < 0.001$) (Table 2). The subjective image quality scores of ULDCT DLR-Str images were comparable to SDCT images in both radiologists' evaluations ($P = 0.16$, $P = 0.054$). The interobserver agreement was excellent ($\kappa = 0.91$; 95% CI [0.88, 0.94]).

We found no difference between objective and subjective image quality in both the under/normal weight and overweight/obese groups (all $P > 0.05$), except that radiologist 1 scored the image quality of ULDCT HIR-Str in the overweight/obese group lower than that in the underweight/normal weight group ($P = 0.036$) (Fig. 4).

Nodule detection rate

Since ULDCT HIR-Str and ULDCT DLR-Str showed better objective and subjective image quality than ULDCT HIR-Std and ULDCT DLR-Std, they were selected to represent HIR and DLR for nodule detection and measurement accuracy analyses (Fig. 5). No nodules were detected in 5 participants during this examination. Therefore, 79 participants (94.0%) with 535 nodules (mean of seven and median of six nodules per participant, ranging from one to 30) detected by SDCT were eligible. For ULDCT, 396 of 535 nodules (74.0%) were detected with HIR-Str, and 446 of 535 nodules (83.4%) were detected with DLR-Str. Therefore, in the ULDCT group, the nodule detection rate of DLR-Str was greater than 80%, which was significantly higher than that of HIR-Str ($P < 0.001$) (Fig. 6; Table 3). Univariate and multivariable logistic regression showed that reconstruction algorithm, nodule size, and nodule type were the independent predictors for nodule detection (all $P < 0.05$) (Table 4). Compared to HIR, the detection rate of DLR

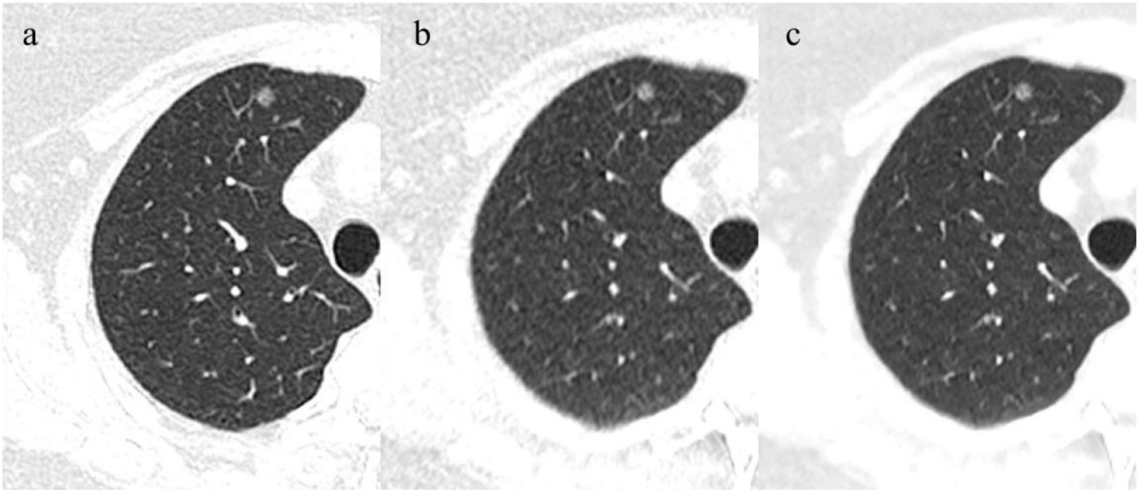

Fig. 5 Representative chest images for nodule detection and measurement by ULDCT DLR and ULDCT HIR compared with SDCT. A 6.4 mm subsolid nodule in a participant (BMI 26.1 kg/m^2) was detected by (**a**) SDCT, (**b**) ULDCT HIR-Str, and (**c**) ULDCT DLR-Str. Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; Std, standard setting; Str, strong setting

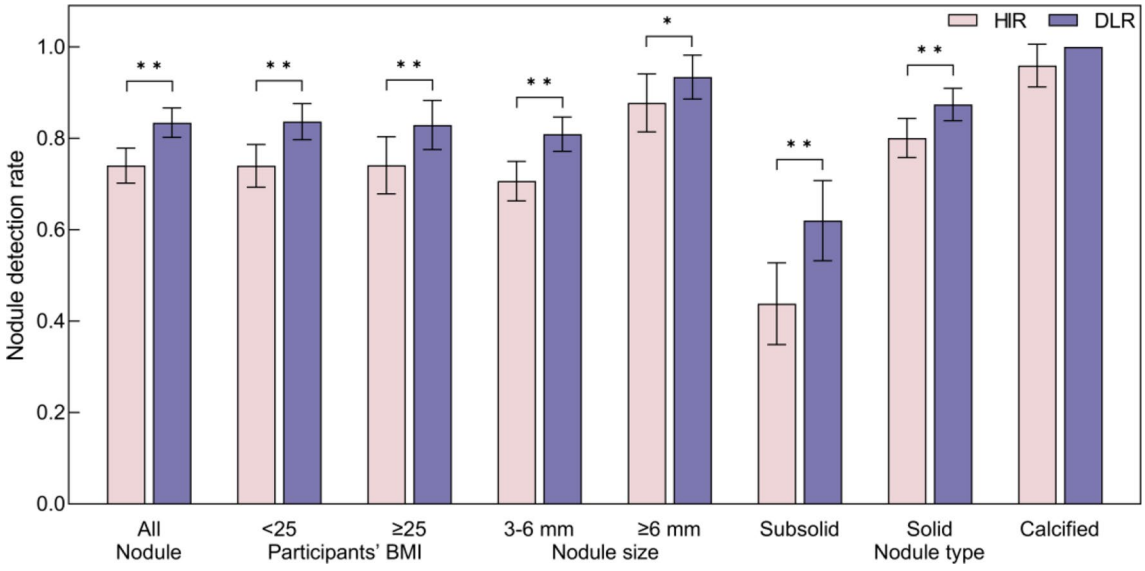

Fig. 6 Bar graphs show nodule detection rates of ULDCT DLR and ULDCT HIR compared with reference SDCT in relation to participant BMI, nodule size, and nodule type. Error bars represent the 95% CI of the mean. * $P < 0.05$, ** $P < 0.001$. Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; BMI, body mass index

was higher (odds ratio = 2.02, $P < 0.001$). For ULDCT, the false-positive rate was 2.0% (9 of 455 nodules) in DLR-Str and 1.5% (6 of 402 nodules) in HIR-Str ($P = 0.61$) (Table S2).

The detection rates of DLR-Str for 3–6 mm and ≥ 6 mm nodules were superior to that of HIR-Str ($P < 0.001$, $P = 0.03$, respectively). The detection rates of DLR-Str for subsolid and solid nodules were higher than those of HIR-Str (all $P < 0.001$). The nodule detection rates of ULDCT DLR-Str in the underweight/normal group and overweight/obese group were both higher than those of ULDCT HIR-Str (83.6% vs. 74.0%, 82.9% vs. 74.1%, both $P < 0.001$). The nodule detection rate of ULDCT in the overweight/obese group was comparable to that in the underweight/normal weight group. In univariate analysis, the nodule detection rate was not affected by participants' BMI ($P = 0.22$) (Table 4).

Nodule measurement accuracy

Figure 7 shows that ULDCT was correlated with SDCT in terms of nodule size and volume for different reconstruction algorithms (all $P < 0.001$). Bland–Altman analysis showed that the SPE of long diameter with reference to the long diameter determined by SDCT was 2.3% (95% CI of the mean: [0.8, 3.8]) for ULDCT HIR-Str and 1.1% (95% CI of the mean: [− 0.4, 2.6]) for ULDCT DLR-Str.

Table 3 Nodule detection rate of Ultra-Low-Dose CT compared with Standard-Dose CT

Variable	SDCT	ULDCT		*P* value*
		HIR-Str	DLR-Str	
Total	535	396 (74.0)	446 (83.4)	< 0.001
BMI (kg/m^2)				
< 25	342	253 (74.0)	286(83.6)	< 0.001
≥ 25	193	143(74.1)	160 (82.9)	< 0.001
Nodule size				
3–6 mm	429	303 (70.6)	347(80.9)	< 0.001
≥ 6 mm	106	93 (87.7)	99 (93.4)	0.03
Nodule type				
Subsolid	121	53 (43.8)	75(62.0)	< 0.001
Solid	341	273 (80.1)	298 (87.4)	< 0.001
Calcified	73	70 (95.9)	73 (100)	0.25

Note: Unless otherwise specified, data are numbers of nodules, with percentages in parentheses

Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; BMI, body mass index; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; Str, strong setting

* *P* value represents the significance between ULDCT HIR-Str and ULDCT DLR-Str scans

The SPE of the volume with reference to that of SDCT was 17.1% (95% CI of the mean: [13.7, 20.6]) for ULDCT HIR-Str and 8.5% (95% CI: [5.4, 11.5]) for ULDCT DLR-Str (Fig. 7). Analyses of SPE in different subgroups of BMI, nodule size, and nodule type are shown in Fig. S3.

Table 5 summarizes the median and IQR values for the APEs of nodule long diameter and volume. No differences of measurement error in nodule long diameter were found between ULDCT DLR-Str and ULDCT HIR-Str ($P = 0.70$). However, ULDCT DLR-Str showed lower volumetric measurement error from SDCT compared with ULDCT HIR-Str ($P < 0.001$).

Discussion

In this clinical study, SDCT, ULDCT HIR-Str, and ULDCT DLR-Str datasets were compared in terms of image quality, nodule detection, and measurement accuracy to evaluate the dose reduction potential of DLR. With DLR, we performed chest ULDCT scans at 0.16 mSv, slightly higher than the range of 0.03–0.1 mSv used for chest radiography [27]. DLR demonstrated superior noise reduction compared to HIR in ULDCT while achieving image quality comparable to regular-dose conditions. Furthermore, when combined with an AI-assisted system, DLR enabled a higher nodule detection rate and more accurate nodule measurement than HIR, the commonly used reconstruction algorithm in current clinical practice. These findings suggest that DLR could provide potential clinical benefits for radiologists, enhancing visual reading and enabling more precise, individualized nodule management.

Prior studies have shown that the diagnostic accuracy of pulmonary nodules varies with reconstruction algorithms [4, 28, 29]. Advances in reconstruction algorithms have shown promise for the evaluation of lung nodules based on reduced radiation dose scans. The current study suggested that DLR was superior to HIR in detecting lung nodules ≥ 3 mm with ULDCT, irrespective of nodule characteristics (such as nodule type or nodule size), which was consistent with a previous study by Jiang et al. [15]. However, our results showed that the detection rates of nodules by ULDCT were not obviously affected by participants' BMI. The reasons may be that the objective image noise of the lung parenchyma and the image quality did not worsen between different BMI groups. Another two related studies by Ye et al. [4] and Miller et

Table 4 Analysis for the predictors of nodule detection in Ultra-Low-Dose CT

Variable	Detectable	Undetectable	Univariate logistic regression			Multivariable logistic regression		
			Coefficient	OR	*P* value	Coefficient	OR	*P* value
Age (years)	56 ± 12	55 ± 12	0.01	1.01	0.30			
Sex								
Male	81.2%	18.8%						
Female	77.5%	22.5%	-0.23	0.80	0.16			
BMI (kg/m^2)	23.9 ± 3.4	24.2 ± 2.8	-0.03	0.97	0.22			
Algorithm								
HIR- Str	74.0%	26.0%						
DLR-Str	83.4%	16.6%	0.57	1.76	< 0.001	0.71	2.02	< 0.001
Nodule size								
3–6 mm	75.8%	24.2%						
≥ 6 mm	90.6%	9.4%	1.12	3.07	< 0.001	1.847	6.341	< 0.001
Nodule type								
Subsolid	52.9%	47.1%						
Solid	83.7%	16.3%	1.52	4.58	< 0.001	1.96	7.09	< 0.001
Calcified	97.9%	2.1%	3.75	42.45	< 0.001	4.20	66.33	< 0.001

Note: The odds ratio of reconstruction algorithm, nodule size, and nodule type were calculated based on the respective first category listed in the table as reference category

Abbreviations: OR, odds ratio; BMI, body mass index; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; Str, strong setting

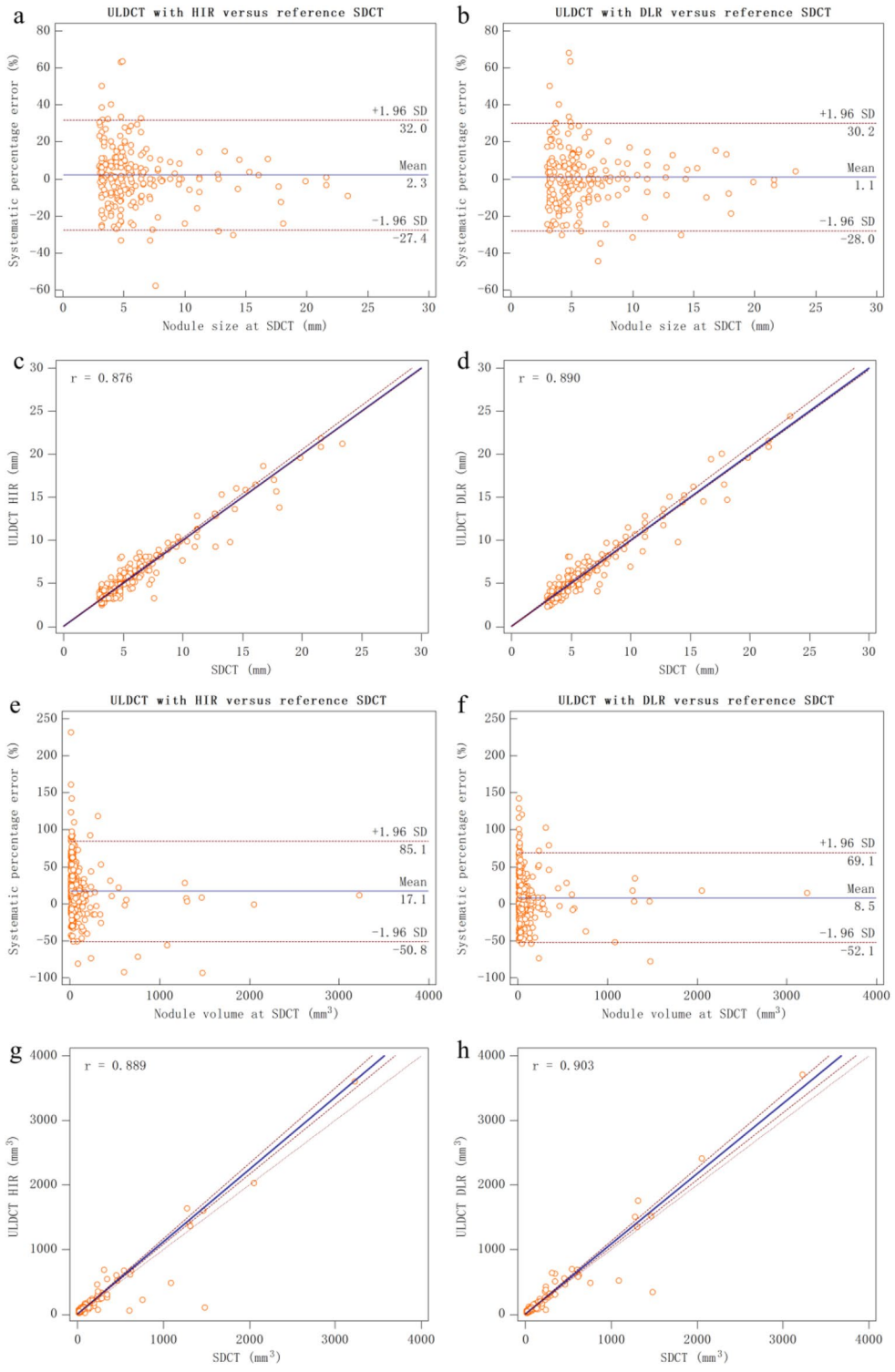

Fig. 7 Bland–Altman plots of SPE for (**a, b**) nodule size and (**e, f**) nodule volume measured by ULDCT and SDCT. Passing and Bablok regression analysis for (**c, d**) nodule size and (**g, h**) nodule volume between the ULDCT and SDCT. Abbreviations: SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; SPE, systematic percentage error

Table 5 APEs of nodule long diameter and volume between deep learning reconstruction and hybrid iterative reconstruction

Variable	SDCT†	APE of long diameter		*P* value*	APE of volume		*P* value*
		ULDCT HIR-Str	ULDCT DLR-Str		ULDCT HIR-Str	ULDCT DLR-Str	
Total	394	6.8% (0, 16.7)	6.7% (0, 16.7)	0.70	19.5% (10.0, 37.5)	17.9% (7.6, 33.4)	< 0.001
BMI (kg/m^2)							
< 25	251	6.4% (0, 15.8)	5.5% (0, 14.8)	0.27	18.4% (9.5, 35.4)	17.0% (7.3, 32.7)	< 0.001
≥ 25	143	6.9% (0.3, 19.4)	8.8% (0.7, 19.2)	0.40	22.8% (11.1, 43.2)	19.4% (8.1, 36.9)	0.001
Nodule Size							
3–6 mm	301	7.0% (0, 19.0)	7.9% (0, 18.2)	0.58	21.1% (10.7, 38.9)	20.3% (9.0, 33.8)	< 0.001
≥ 6 mm	93	5.4% (1.1, 12.5)	3.8% (0.0, 10.2)	0.69	16.9% (8.8, 33.8)	13.0% (5.6, 29.7)	< 0.001
Nodule Type							
Subsolid	53	5.4% (1.4, 14.3)	5.4% (1.5, 13.2)	0.32	14.9% (6.8, 31.7)	19.1% (7.3, 33.5)	0.86
Solid	271	6.9% (0, 16.7)	7.2% (0, 16.7)	0.76	20.0% (10.7, 37.5)	18.2% (7.3, 33.8)	< 0.001
Calcified	70	7.1% (0, 20.0)	7.2% (0, 20.0)	0.51	27.6% (11.2, 44.4)	16.0% (8.1, 30.4)	0.001

Note: Data are presented as median and IQR

Abbreviations: APE, absolute percentage measurement error; BMI, body mass index; SDCT, standard-dose CT; ULDCT, ultra-low-dose CT; HIR, hybrid iterative reconstruction; DLR, deep learning reconstruction; Str, strong setting

† Number of nodules detected on both SDCT and ULDCT scans

* *P* value represents the significance between ULDCT HIR-Str and ULDCT DLR-Str scans

al. [30] revealed that the BMI of patients had no noticeable effects on the detectability of nodules when BMI was less than 30 kg/m^2. Our study showed similar results. Lenfant et al. [31] demonstrated a slower decrease of signal-to-noise ratio and contrast-to-noise ratio per unit dose with DLR than with HIR. Zhao et al. [32] reported that DLR could maintain the image quality even in overweight interstitial lung disease patients at low radiation dose. These studies indicated that DLR may be valuable in populations with high BMI, which is one of the major concerns in chest ULDCT.

The SPE gives information about signed differences (i.e., systematic overestimation or underestimation), which were observed in some previous studies of ULDCT and LDCT with SDCT as a reference [4, 24, 33]. In this study, DLR showed lower SPEs in the measurement of nodule diameter and volume compared with HIR. Similar results were reported by Jiang et al. [15], who compared the systematic percentage differences between non-contrast ULDCT and contrast-enhanced SDCT. However, a retrospective study [34] found that the three-dimensional volumetric measurement of pulmonary nodules could be affected after administration of contrast medium, and the measured nodule volume increased on contrast-enhanced CT. Moreover, deviations with inverse signs would cancel each other out when using SPE [35]. Therefore, we performed non-contrast CT scans at both standard dose and ultra-low dose for each participant, and the APE between ULDCT and SDCT was calculated to evaluate the differences more accurately. The results indicated no statistical difference between ULDCT HIR and ULDCT DLR in the measurement of diameter, but the APEs of nodule volume were shown to be diminished by DLR. Volume has been reported to be more sensitive for evaluation of nodule size than diameter [36]. ULDCT DLR could measure nodule volume more accurately than ULDCT HIR, providing the potential to reflect the true growth rate of pulmonary nodules.

DLR has proven to be highly effective in noise reduction and smoothing, making it a valuable tool for enhancing CT image quality. However, the nodule detection rate of ULDCT, even with DLR, is 16% lower compared to SDCT. While DLR effectively reduces image noise and achieves image quality comparable to SDCT, it may also alter image texture or fine details, potentially impacting AI detection performance. This is particularly relevant for nodules with relatively low attenuation contrast against the surrounding lung parenchyma and those with ill-defined lesion margins [37]. Recent studies have shown that AI software can assist radiologists in diagnosing pulmonary nodules, with performance now reaching or even surpassing that of radiologists [38, 39]. This suggests that AI assistance holds great promise for lung nodule assessment, maintaining consistent performance to provide radiologists with decision support. With DLR providing images with lower noise and higher quality than HIR, AI-assisted systems have the potential to enhance nodule detection sensitivity, thereby reducing radiologists' workload. By pre-screening CT images and flagging suspicious nodules, AI systems help to ease the increasing burden associated with the growing number of CT examinations. This enables radiologists to focus on reviewing AI-detected findings, enhancing diagnostic efficiency. Nonetheless, optimizing AI-assisted detection systems specifically for ULDCT images remains essential to improve nodule detection performance.

Our study had several limitations. First, the body size of the participants in this study was generally small, and

only one participant had a BMI over 30 kg/m^2. Further studies with larger populations including more obese participants are required to verify these results and evaluate the potential of DLR. Second, the ULDCT results were generated by the AI-assisted system, as the primary objective was to compare the detectability of lung nodules using the same AI-assisted detection system across different reconstruction algorithms. In clinical practice, AI-aided diagnosis and visual confirmation could be used together as complementary reading methods to minimize false-negatives [37, 40]. Third, the reconstruction algorithms were from a single vendor, results may differ among different vendors.

In conclusion, DLR was superior to HIR in terms of image noise and image quality at ULDCT as low as 0.16 mSv. DLR also increased nodule detection rate and improved nodule measurement accuracy in comparison with HIR when using the AI system.

Abbreviations

BMI	Body mass index
$CTDI_{vol}$	Volume CT dose index
DLP	Dose-length product
ED	Effective dose
SDCT	Standard-dose CT
ULDCT	Ultra-low-dose CT

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12880-025-01746-6.

Supplementary Material 1

Author contributions

Conceptualization: Song L. Data curation: Wang J, Song L. Formal analysis: Zhu Z, Pan Z. Funding acquisition: Song L, Song W. Investigation: Hu G, Ying Z, Sui X. Methodology: Wang J. Software: Tan W, Zhou Z, Ma Z, Xu Y. Supervision: Jin Z. Validation: Han W. Writing-original draft: Wang J. Writing-review & editing: Song L, Song W. All authors read and approved the final manuscript.

Funding

This work has supported by the National Natural Science Foundation of China (NSFC No. 82171934) and the National High Level Hospital Clinical Research Funding (2022-PUMCH-B-069).

Data availability

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was performed in accordance with the Declaration of Helsinki, and approved by the ethics committee of Peking Union Medical College Hospital (HS-2540). Informed consent was obtained from all patients who participated in this study. Clinical trial number: not applicable. This study was performed in accordance with the Declaration of Helsinki, and approved by the ethics committee of Peking Union Medical College Hospital (HS-2540).

Consent for publication

Not applicable.

Informed consent

All patients who participated in this study. All the images, tables, and recordings are consented to be published.

Competing interests

The authors declare no competing interests.

Ethics approval

The Institutional Review Board of our hospital approved this prospective study (HS-2540).

Author details

[1]Department of Radiology, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 1, Shuaifuyuan, Dongcheng District, Beijing 100730, China
[2]4+4 Medical Doctor Program, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[3]Deepwise AI Lab, Beijing Deepwise & League of PhD Technology Co.Ltd, Beijing, China
[4]Department of Epidemiology and Health Statistics, Institute of Basic Medicine Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[5]Theranostics and Translational Research Center, National Infrastructures for Translational Medicine, Institute of Clinical Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China
[6]Canon Medical System (China), No.3, Xinyuan South Road, Chaoyang District, Beijing, China

Received: 19 December 2024 / Accepted: 22 May 2025
Published online: 30 May 2025

References

1. Bray F, Laversanne M, Sung H et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024.
2. MacMahon H, Naidich DP, Goo JM, et al. Guidelines for management of incidental pulmonary nodules detected on CT images: from the Fleischner society 2017. Radiology. 2017;284(1):228–43.
3. 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.
4. Ye K, Zhu Q, Li M, et al. A feasibility study of pulmonary nodule detection by ultralow-dose CT with adaptive statistical iterative reconstruction-V technique. Eur J Radiol. 2019;119:108652.
5. Pedersen AE, Kusk MW, Knudsen GH, et al. Collimation border with U-Net segmentation on chest radiographs compared to radiologists. Radiography (Lond). 2023;29(3):647–52.
6. Kubo T, Ohno Y, Kauczor HU, et al. Radiation dose reduction in chest CT–review of available options. Eur J Radiol. 2014;83(10):1953–61.
7. Xu Y, Zhang TT, Hu ZH, et al. Effect of iterative reconstruction techniques on image quality in low radiation dose chest CT: a Phantom study. Diagn Interv Radiol. 2019;25(6):442–50.
8. Vardhanabhuti V, Pang CL, Tenant S, et al. 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;91:130–41.
9. Nam JG, Ahn C, Choi H, et al. Image quality of ultralow-dose chest CT using deep learning techniques: potential superiority of vendor-agnostic post-processing over vendor-specific techniques. Eur Radiol. 2021;31(7):5139–47.
10. Geyer LL, Schoepf UJ, Meinel FG, et al. State of the art: iterative CT reconstruction techniques. Radiology. 2015;276(2):339–57.
11. Mileto A, Guimaraes LS, McCollough CH, et al. State of the Art in abdominal CT: the limits of iterative reconstruction algorithms. Radiology. 2019;293(3):491–503.
12. Zhou HY, Yu Y, Wang C, et al. A transformer-based representation-learning model with unified processing of multimodal input for clinical diagnostics. Nat Biomed Eng. 2023;7(6):743–55.

13. Zhao G, Feng Q, Chen C, et al. Diagnose like a radiologist: hybrid Neuro-Probabilistic reasoning for Attribute-Based medical image diagnosis. IEEE Trans Pattern Anal Mach Intell. 2022;44(11):7400–16.
14. Koetzier LR, Mastrodicasa D, Szczykutowicz TP, et al. Deep learning image reconstruction for CT: technical principles and clinical prospects. Radiology. 2023;306(3):e221257.
15. Jiang B, Li N, Shi X, et al. Deep learning reconstruction shows better lung nodule detection for Ultra-Low-Dose chest CT. Radiology. 2022;303(1):202–12.
16. Kim JH, Yoon HJ, Lee E, et al. 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):131–8.
17. Wang J, Sui X, Zhao R, et al. Value of deep learning reconstruction of chest low-dose CT for image quality improvement and lung parenchyma assessment on lung window. Eur Radiol. 2024;34(2):1053–64.
18. Qi LL, Wang JW, Yang L, et al. Natural history of pathologically confirmed pulmonary subsolid nodules with deep learning-assisted nodule segmentation. Eur Radiol. 2021;31(6):3884–97.
19. Venkadesh KV, A A A Setio A, Schreuder, et al. Deep learning for malignancy risk Estimation of pulmonary nodules detected at Low-Dose screening CT. Radiology. 2021;300(2):438–47.
20. Li R, Xiao C, Huang Y et al. Deep learning applications in computed tomography images for pulmonary nodule detection and diagnosis: A review. Diagnostics (Basel) 2022;12(2).
21. World Health Organization. Obesity: preventing and managing the global epidemic. Report of a WHO consultation. Geneva, Switzerland: World Health Organization technical report series 2000.
22. Kang H. Sample size determination and power analysis using the G*Power software. J Educ Eval Health Prof. 2021;18:17.
23. Cui S, Ming S, Lin Y, et al. Development and clinical application of deep learning model for lung nodules screening on CT images. Sci Rep. 2020;10(1):13657.
24. Han D, Cai J, Heus A, et al. Detection and size quantification of pulmonary nodules in ultralow-dose versus regular-dose CT: a comparative study in COPD patients. Br J Radiol. 2023;96(1144):20220709.
25. Kim SK, Kim C, Lee KY, et al. Accuracy of Model-Based iterative reconstruction for CT volumetry of Part-Solid nodules and solid nodules in comparison with filtered back projection and hybrid iterative reconstruction at various dose settings: an anthropomorphic chest Phantom study. Korean J Radiol. 2019;20(7):1195–206.
26. Mikayama R, Shirasaka T, Kojima T, et al. Deep-learning reconstruction for ultra-low-dose lung CT: volumetric measurement accuracy and reproducibility of artificial ground-glass nodules in a Phantom study. Br J Radiol. 2022;95(1130):20210915.
27. 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):1912–22.
28. Hassani C, Ronco A, Prosper AE, et al. Forward-Projected Model-Based iterative reconstruction in screening Low-Dose chest CT: comparison with adaptive iterative dose reduction 3D. AJR Am J Roentgenol. 2018;211(3):548–56.
29. Zhang M, Qi W, Sun Y, et al. Screening for lung cancer using sub-millisievert chest CT with iterative reconstruction algorithm: image quality and nodule detectability. Br J Radiol. 2018;91(1090):20170658.
30. Miller AR, Jackson D, Hui C, et al. Lung nodules are reliably detectable on ultra-low-dose CT utilising model-based iterative reconstruction with radiation equivalent to plain radiography. Clin Radiol. 2019;74(5):409. e17-409.e22.
31. Lenfant M, Chevallier O, Comby PO et al. (2020) Deep learning versus iterative reconstruction for CT pulmonary angiography in the emergency setting: improved image quality and reduced radiation dose. Diagnostics (Basel) 10(8).
32. Zhao R, Sui X, Qin R, et al. Can deep learning improve image quality of low-dose CT: a prospective study in interstitial lung disease. Eur Radiol. 2022;32(12):8140–51.
33. Sui X, Meinel FG, Song W, et al. Detection and size measurements of pulmonary nodules in ultra-low-dose CT with iterative reconstruction compared to low dose CT. Eur J Radiol. 2016;85(3):564–70.
34. Honda O, Johkoh T, Sumikawa H, et al. Pulmonary nodules: 3D volumetric measurement with multidetector CT–effect of intravenous contrast medium. Radiology. 2007;245(3):881–7.
35. Kuhnigk JM, Dicken V, Bornemann L, et al. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. IEEE Trans Med Imaging. 2006;25(4):417–34.
36. Song YS, Park CM, Park SJ, et al. Volume and mass doubling times of persistent pulmonary subsolid nodules detected in patients without known malignancy. Radiology. 2014;273(1):276–84.
37. Silva M, Schaefer-Prokop CM, Jacobs C, et al. Detection of subsolid nodules in lung Cancer screening: complementary sensitivity of visual reading and Computer-Aided diagnosis. Invest Radiol. 2018;53(8):441–9.
38. Gao S, Xu Z, Kang W, et al. Artificial intelligence-driven computer aided diagnosis system provides similar diagnosis value compared with Doctors' evaluation in lung cancer screening. BMC Med Imaging. 2024;24(1):141.
39. de Margerie-Mellon C, Chassagnon G. Artificial intelligence: A critical review of applications for lung nodule and lung cancer. Diagn Interv Imaging. 2023;104(1):11–7.
40. Snoeckx A, Franck C, Silva M, et al. The radiologist's role in lung cancer screening. Transl Lung Cancer Res. 2021;10(5):2356–67.

Publisher's note