Original Article

Radiation dose reduction using deep learning-based image reconstruction for a low-dose chest computed tomography protocol: a phantom study

Yunsub Jung[1], Jin Hur[2], Kyunghwa Han[2], Yasuhiro Imai[3], Yoo Jin Hong[2], Dong Jin Im[2], Kye Ho Lee[2,4], Melissa Desnoyers[5], Brian Thomsen[5], Risa Shigemasa[5], Kyounga Um[1], Kyungeun Jang[1]

[1]Research Team, GE Healthcare Korea, Seoul, Republic of Korea; [2]Department of Radiology and Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea; [3]CT System Group, GE Healthcare Japan, Hino, Japan; [4]Department of Radiology, Dankook University Hospital, Cheonan, Republic of Korea; [5]Global Research Team, GE Healthcare US, Milwaukee, WI, USA

Contributions: (I) Conception and design: J Hur, Y Jung; (II) Administrative support: J Hur, Y Jung; (III) Provision of study materials or patients: J Hur, Y Jung; (IV) Collection and assembly of data: J Hur, Y Jung, K Han, Y Imai; (V) Data analysis and interpretation: J Hur, Y Jung, K Han, YJ Hong, DJ Im, KH Lee; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jin Hur, MD, PhD. Department of Radiology and Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea. Email: khuhz@yuhs.ac.

Background: The aim of this study was to compare the dose reduction potential and image quality of deep learning-based image reconstruction (DLIR) with those of filtered back-projection (FBP) and iterative reconstruction (IR) and to determine the clinically usable dose of DLIR for low-dose chest computed tomography (LDCT) scans.

Methods: Multi-slice computed tomography (CT) scans of a chest phantom were performed with various tube voltages and tube currents, and the images were reconstructed using seven methods to control the amount of noise reduction: FBP, three stages of IR, and three stages of DLIR. For subjective image analysis, four radiologists compared 48 image data sets with reference images and rated on a 5-point scale. For quantitative image analysis, the signal to noise ratio (SNR), contrast to noise ratio (CNR), nodule volume, and nodule diameter were measured.

Results: In the subjective analysis, DLIR-Low (0.46 mGy), DLIR-Medium (0.31 mGy), and DLIR-High (0.18 mGy) images showed similar quality to the FBP (2.47 mGy) image. Under the same dose conditions, the SNR and CNR were higher with DLIR-High than with FBP and all the IR methods (all $P<0.05$). The nodule volume and size with DLIR-High were significantly closer to the real volume than with FBP and all the IR methods (all $P<0.001$).

Conclusions: DLIR can improve the image quality of LDCT compared to FBP and IR. In addition, the appropriate effective dose for LDCT would be 0.24 mGy with DLIR-High.

Keywords: Low-dose chest computed tomography (LDCT); deep learning-based image reconstruction (DLIR); chest phantom

Submitted Jun 16, 2022. Accepted for publication Dec 20, 2022. Published online Feb 01, 2023.
doi: 10.21037/qims-22-618
View this article at: https://dx.doi.org/10.21037/qims-22-618

Introduction

The use of low-dose chest computed tomography (LDCT) screening has increased, and the implementation of lung cancer computed tomography (CT) has become important for diagnosing and managing small pulmonary nodules detected during screening. Therefore, radiation exposure from lung CT scans is of increasing concern to both the medical community and the general public (1-3). For effective and responsible lung cancer screening, LDCT should be performed using high-quality low-radiation-exposure CT.

Iterative reconstruction (IR) algorithms have become the mainstay reconstruction technique for LDCT (4,5). IR methods have been developed to optimize the quality of reconstructed images using iterative computations that include information about system statistics, X-ray physics, optics, and patient characteristics in the modeling (6-8). The IR technique allows patients to undergo scanning with radiation doses lower than those required for the traditional filtered back-projection (FBP) reconstruction algorithms while maintaining diagnostic quality.

With the development of artificial intelligence technology and graphics processing unit performance, medical image reconstruction methods using deep learning techniques are emerging (9,10). Deep learning-based image reconstruction (DLIR) can process a vast number of parameters, far more than human engineers can handle without user involvement. Therefore, it is possible to optimize various parameters and ensure stable image quality. Recently, artificial intelligence technologies have been proposed to enable dose reduction while maintaining the diagnostic performance of CT (11,12). The methods use training with high-quality FBP images (GE Healthcare) (13), training with IR images (Canon Medical Systems) (14), and training with noise-free IR or FBP images (non-CT device vendors) (13,14). Although each method pursues a slightly different purpose, these all methods have a common goal of replacing traditional IR methods with low-dose scanning. It is known that the DLIR technology using low-dose scanning guarantees the same quality as the conventional IR method.

Recent studies reported that the use of DLIR techniques can facilitate dose reduction while maintain the image quality in chest CT scan (15-17). However, the dose reduction potential using DLIR techniques for low dose chest CT scan was not fully assessed yet.

The purposes of this study were to investigate the dose reduction potential of DLIR techniques, compare the resulting image quality with FBP and IR images, and determine the clinically usable dose required by DLIR for an LDCT protocol.

Methods

Chest phantom and data acquisition

For this experimental phantom study, ethics review and institutional review board approval were waived.

To evaluate image quality according to image reconstruction technique, a multipurpose chest phantom (LUNGMAN; KYOTO KAGAKU, Japan) was used. Various structures (bone, heart, trachea, pulmonary vessel, and diaphragm) were included in the main body, along with 15 artificial, spherical nodules of three types (pure, urethane foam, −800 HU; sub-solid, urethane foam, −630 HU; solid, polyurethane, 100 HU) and five diameters (3, 5, 8, 10, 12 mm). The artificial nodules can be inserted anywhere in the phantom. All CT examinations were performed with a 256-slice CT scanner (Revolution CT, GE Healthcare) and the following parameters: gantry rotation time, 500 ms; display of field of view (DFOV), 350 mm; coverage, 160 mm; and slice thickness, 0.625 mm. Four fixed tube voltages (70, 80, 100, 120 kVp) and four fixed tube current-time products (15, 20, 25, 50 mAs) were chosen to vary the radiation dose of the chest CT images. The highest radiation exposure was selected from the LDCT protocol commonly used in clinical practice.

All examinations were carried out as axial scans with a 160-mm wide detector because the nodules in the phantom can move as the table moves during a helical scan. Subtle changes in the position of the nodules can affect the accuracy of quantitative analyses, which require pixel-by-pixel calculations.

Image reconstruction

All the images were reconstructed using data from a single scanner. The images were reconstructed using adaptive statistical IR (ASiR-V, GE Healthcare) and DLIR (TrueFidelity, GE Healthcare). ASiR-V is a hybrid IR method that combines adaptive statistical IR with the physics model used for model-based IR, and the numbers that follow the V represent the blending ratio between the ASiR-V and FBP (i.e., ASiR-V30 = ASiR-V ×0.3 + FBP ×0.7). The DLIR was based on a deep neural network

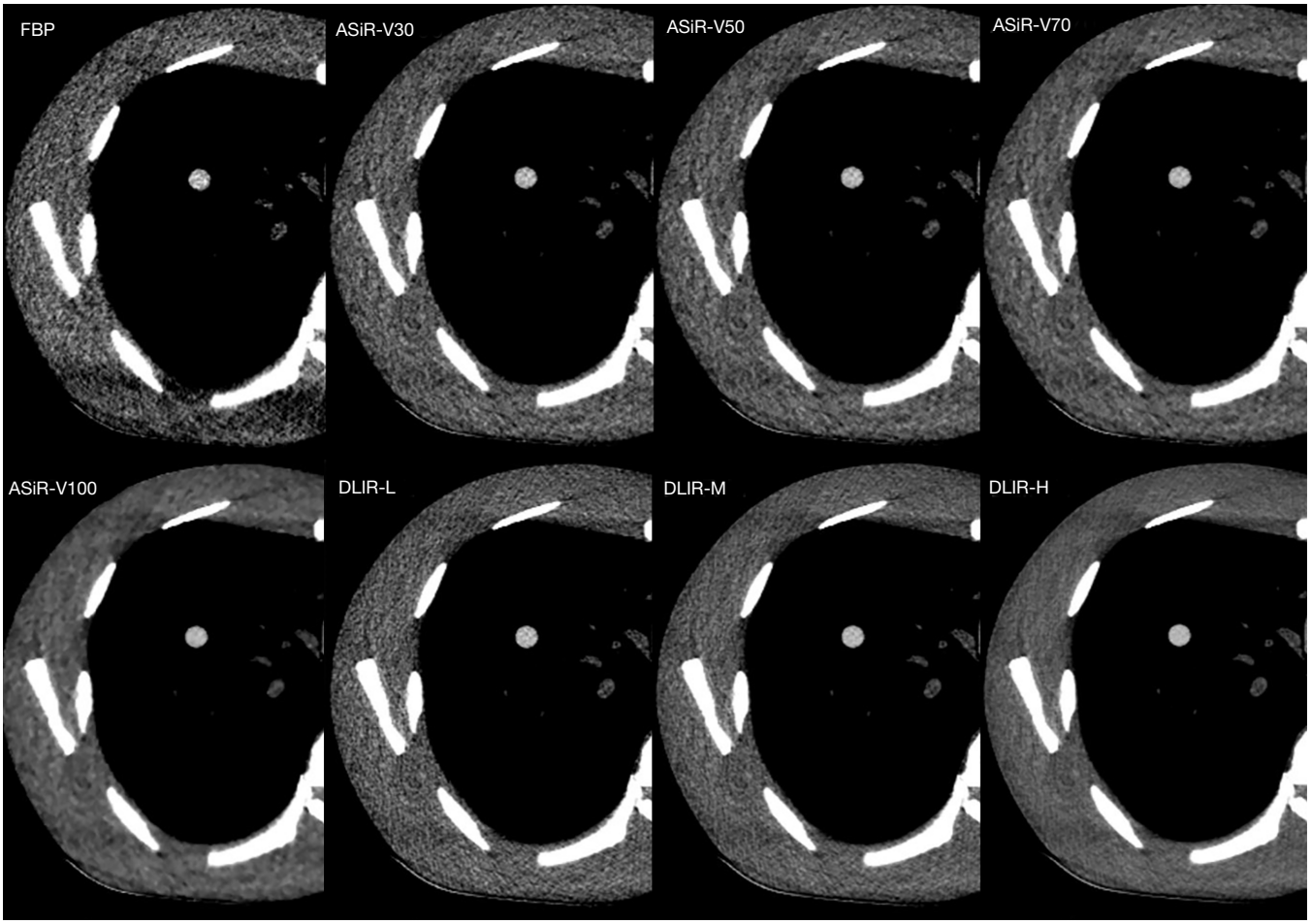

Figure 1 Nodule images according to reconstruction method. The images were scanned at a tube current of 25 mAs and tube potential of 120 kVp. ASiR-V, adaptive statistical iterative reconstruction, DLIR, deep learning-based image reconstruction; L, low; M, medium; H, high; FBP, filtered back-projection.

(DNN) that provides three selectable reconstruction strengths (low, medium, high) to control the amount of noise reduction. This DNN-based DLIR engine was formulated to apply the pre-computed and then fixed coefficients in an inference engine responsible for producing reconstructed images from the input sinogram data. The manufacturer currently provides only a standard kernel in the DLIR, so all reconstructions used that standard kernel. Each raw data set was reconstructed using eight reconstruction methods: FBP, ASiR-V (30, 50, 70, 100) and DLIR (low, medium, high) (*Figure 1*).

Subjective image quality evaluation

Four radiologists with 6–12 years of experience in chest CT imaging independently evaluated all the CT data sets on the same display monitor (RadiForce MX193, EIZO). The readers were blinded to the number, location, and size of the simulated nodules, as well as to the reconstruction methods. The subjective image quality evaluation used a relative visual grading analysis with a reference image (18). Information (scan protocol and reconstruction method) about the reference image and the comparative image was deleted from the image viewer (Advanced Workstation, GE Healthcare), so only the position of the reference image was shown to the readers. The readers were asked to rate the overall image quality (defined as the reader's confidence in using a certain image to detect the presence or absence of lung lesions) on a 5-point scale compared to the reference image. The 5-point scale was rated as follows:

grade 1 (clearly inferior to the reference image), grade 2 (somewhat inferior to the reference image), grade 3 (equal to the reference image), grade 4 (somewhat superior to the reference image), and grade 5 (clearly superior to the reference image).

Subjective image quality and reconstruction methods

In this subjective analysis, we intended to evaluate the performance of the reconstruction methods and provide options for various radiation doses, so the FBP image at the corresponding dose was used as the reference image. This experiment contained 168 data sets made by combining the three nodule types (solid, sub-solid, and pure), four levels of tube potential (70, 80, 100, 120 kVp), and two levels of tube current (20 and 50 mAs). Each of those data sets was reconstructed using seven methods (FBP, ASiR-V30, 50, 70, DLIR-Low, -Medium, -High).

Subjective image quality and radiation dose

This subjective analysis was performed to find the dose level required to produce DLIR images usable for diagnosis. Therefore, we derived visual grades by comparing the reference image with DLIR images at various dose levels. The volume computerized tomography dose index ($CTDI_{vol}$) of the reference image used in this experiment was 2.47 mGy with FBP (120 kVp, 50 mAs), which was sufficient to diagnose an actual LDCT. To ensure that we used images with the same level as the reference image from those captured at various tube currents, we chose images (selected tube current) with an average of grade 3 or more in the visual grading analysis by our four evaluators. This experiment contained 48 data sets made by combining four levels of tube potential (70, 80, 100, 120 kVp) and four tube currents (15, 20, 25, 50 mAs). Each data set was reconstructed using the three DLIR methods (DLIR-Low, -Medium, -High).

Objective image quality evaluation

A total of 42 data sets were created by combining two tube potentials (70, 80 kVp) and three tube currents (15, 20, 25 mAs), and each data set was reconstructed using seven methods (FBP, ASiR-V30, 50, 70, DLIR-Low, -Medium, and -High). First, to measure the basic noise characteristics of each image, we measured the signal to noise ratio (SNR) and contrast to noise ratio (CNR) in the heart region of the phantom with a homogeneous medium using the pixel Hounsfield unit (HU) values. SNR and CNR were calculated using the following respective formulas: $SNR = |Mean_{heart}/\text{standard deviation (SD)}_{background}|$ and $CNR = |Mean_{heart} - Mean_{background}|/SD_{heart}$. Next, the volumes and diameters of 12-mm sized solid nodules were calculated in a semi-automatic manner using commercial software (AVIEW Research; CORELINE, Korea). The software requires the user to manually draw an initial seed line traversing the nodule, and then it automatically measures the volume and diameter of the nodule based on that line. Each measurement used the same algorithm (without changing parameters) and was performed twice by each of two physicians. The measurement error ratio (MER) was calculated to evaluate the measured size versus the real size of the nodule. The $MER_{volume/diameter}$ was calculated as follows: |measured volume or diameter/actual volume or diameter| ×100%, where the real diameter was provided by the phantom manufacturer, and the volume was calculated based on the provided diameter [volume = $(4/3) \times \pi \times \text{radius}^3$].

Statistical analysis

All statistical analyses were performed using statistical software (SPSS, version 22.0, IBM Corp., Armonk, NY, USA), and P values less than 0.05 were considered statistically significant. For continuous data, the results were showed as mean (min, max). The Kolmogorov-Smirnov test was used to check the normality (19), and showed that neither the subjective evaluation nor the objective evaluation followed the normal distribution. Because both subjective and objective evaluation data were nonparametric, Kruskal-Wallis testing was used to compare the image quality results among the reconstruction methods. In the subjective evaluation (visual grade analysis), we analyzed the differences in the image reconstruction methods according to dose and whether the pattern differed by dose using the average scores given by the four radiologists, which provided a reliable statistical evaluation. In the objective evaluation (volume, diameter measurement, SNR, CNR), differences among the image reconstruction methods were compared using Bonferroni analyses. Intraclass correlation coefficient (ICC) analysis was used to analyze intra-observer

Table 1 Average visual grading analysis results from four radiologists

Tube input*	ASiR-V			DLIR			P
	30	50	70	Low	Medium	High	
70 kVp, 20 mAs	3.33 [2, 4]	3.50 [2, 4]	3.67 [3, 5]	3.75 [2, 5]	4.08 [3, 5]	4.92 [4, 5]†	<0.001
70 kVp, 50 mAs	3.42 [3, 4]	3.75 [3, 4]	4.08 [4, 5]	3.75 [3, 5]	4.67 [4, 5]	5.00 [5, 5]†	<0.001
80 kVp, 20 mAs	3.42 [3, 4]	3.83 [3, 5]	4.33 [4, 5]	3.17 [3, 4]	4.17 [3, 5]	4.75 [4, 5]†	<0.001
80 kVp, 50 mAs	3.42 [3, 4]	3.83 [3, 5]	4.25 [4, 5]	3.92 [3, 5]	4.58 [3, 5]	4.83 [4, 5]†	<0.001
100 kVp, 20 mAs	3.67 [3, 4]	3.50 [3, 4]	4.42 [4, 5]	3.92 [3, 5]	4.58 [4, 5]	4.83 [4, 5]†	<0.001
100 kVp, 50 mAs	3.67 [3, 4]	4.08 [4, 5]	4.17 [4, 5]	4.08 [3, 5]	4.58 [4, 5]	4.83 [4, 5]†	<0.001
120 kVp, 20 mAs	3.50 [3, 5]	3.83 [3, 4]	4.17 [3, 5]	3.92 [3, 5]	4.58 [4, 5]	4.83 [4, 5]†	<0.001
120 kVp, 50 mAs	3.75 [3, 4]	4.00 [3, 5]	4.33 [4, 5]	4.33 [4, 5]	4.75 [4, 5]	4.75 [4, 5]†	<0.001
Total	3.63 [2, 4]	3.93 [2, 5]	4.24 [3, 5]	3.93 [2, 5]	4.54 [3, 5]	4.90 [4, 5]†	<0.001

Evaluation results are expressed as mean [min, max]. *, $CTDI_{vol}$ values were 0.24, 0.63, 0.38, 0.96, 0.61, 1.53, 0.99 and 2.47 mGy for (70 kVp, 20 mAs), (70 kVp, 50 mAs), (80 kVp, 20 mAs), (80 kVp, 50 mAs), (100 kVp, 20 mAs), (100 kVp, 50 mAs), (120 kVp, 2 mAs), and (120 kVp, 50 mAs), respectively; †, highest value. ASiR-V, adaptive statistical iterative reconstruction; $CTDI_{vol}$, volume computerized tomography dose index; DLIR, deep learning-based image reconstruction.

and inter-observer differences.

Results

Subjective image quality

The subjective image quality scores at various dose levels are summarized in *Table 1* and Table S1. At all dose levels, DLIR-High and DLIR-Medium scored significantly higher than ASiR-V30, 50, and 70 (all P<0.05). DLIR-High had a higher score than DLIR-Low (P<0.01) and a slightly higher average score than DLIR-Medium, but the difference between DLIR-High and DLIR-Medium was not statistically significant (P=0.186). ASiR-V30 and DLIR-High showed the lowest and highest image-quality scores, respectively. As the dose increased, most reconstruction methods tended to have higher scores than FBP, but the differences between those values were not large for DLIR-Medium and DLIR-High. In all the experimental results, the interaction effect of dose and image reconstruction method was 0.004. The results of subjective image analysis to find the optimal dose for the DLIR images are listed in *Table 2* and Table S2. The average dose reduction ratio [(1 – $CTDI_{vol}$ of selected image/ $CTDI_{vol}$ of reference image) ×100%] calculated based on the $CTDI_{vol}$ values of the reference and selected images was 76.62% (70.04–81.38%), 80.87% (70.04–87.45%), and 82.19% (70.04–92.71%) for DLIR-Low, DLIR-Medium, and DLIR-High, respectively. The doses required to produce DLIR images at a level similar to that of the reference image (FBP, 2.47 mGy) were 0.46–0.74 mGy (DLIR-Low), 0.31–0.74 mGy (DLIR-Medium), and 0.18–0.74 mGy (DLIR-High).

Objective image quality

The measurement results for SNR and CNR are shown in *Table 3* and *Figure 2*. Under the same-dose conditions, SNR was higher with DLIR-High and DLIR-Medium than with FBP and all ASiR-V methods (all P<0.05); however, DLIR-Low did not show any significant difference from FBP or any ASiR-V method (all P>0.05). For CNR, DLIR-High was significantly higher than FBP and all ASiR-V methods (all P<0.05), but DLIR-Low and DLIR-Medium did not differ from the ASiR-V methods (all P>0.05).

The results for the volume and diameter measured for the 12 mm solid nodule are shown in *Table 3*. The nodule volume automatically measured by the fixed algorithm was significantly closer to the real volume in the DLIR-High image than those with FBP or any ASiR-V method (all P<0.001). DLIR-Low and DLIR-Medium were also closer to the real volume than the other methods (all ASiR-V *vs.* DLIR-Low, DLIR-Medium, P<0.001 and FBP *vs.* DLIR-Low, DLIR-Medium, P<0.01). In measuring the nodule diameter using the automatic measurement method, all DLIR images were significantly closer to the real diameter

Table 2 Visual grading analysis to find optimal dose for DLIR images

Test	Reconstruction mode	Tube potential (kVp)	Tested tube current (mAs)	Selected tube current* (mAs)	$CTDI_{vol}$ of selected image (mGy)	Dose reduction ratio† (%)
1	DLIR-L	70	15, 20, 25, 50	50	0.63	74.49
2	DLIR-L	80	15, 20, 25, 50	25	0.48	80.57
3	DLIR-L	100	15, 20, 25, 50	15	0.46	81.38
4	DLIR-L	120	15, 20, 25, 50	15	0.74	70.04
5	DLIR-M	70	15, 20, 25, 50	25	0.31	87.45
6	DLIR-M	80	15, 20, 25, 50	20	0.38	84.62
7	DLIR-M	100	15, 20, 25, 50	15	0.46	81.38
8	DLIR-M	120	15, 20, 25, 50	15	0.74	70.04
9	DLIR-H	70	15, 20, 25, 50	15	0.18	92.71
10	DLIR-H	80	15, 20, 25, 50	20	0.38	84.62
11	DLIR-H	100	15, 20, 25, 50	15	0.46	81.38
12	DLIR-H	120	15, 20, 25, 50	15	0.74	70.04

The $CTDI_{vol}$ of the reference image used in this experiment was 2.47 mGy with FBP at a tube current of 50 mAs and tube potential of 120 kVp. *, refers to the mAs value at which the average value of the four evaluators' grading scores is greater than three; †, dose reduction ratio is defined as (1 – $CTDI_{vol}$ of selected image/$CTDI_{vol}$ of reference image) ×100%. DLIR, deep learning-based image reconstruction; $CTDI_{vol}$, computerized tomography dose index volume; L, low; M, medium; H, high.

Table 3 Summary of SNR, CNR, and nodule measurements with different reconstruction methods and dose levels

Reconstruction	Tube input (kVp/mAs)	SNR, mean/SD	CNR, mean/SD	Volume, mean (95% CI)	Diameter, mean (95% CI)
FBP	70/15	0.84/0.03	21.56/0.21	887.9 (887.658, 888.142)	11.92 (11.919, 11.921)
	70/20	0.76/0.02	22.89/0.19	888.2 (887.958, 888.442)	11.93 (11.929, 11.931)
	70/25	0.82/0.03	22.41/0.22	893.4 (893.156, 893.644)	11.95 (11.949, 11.951)
	80/15	1.00/0.01	21.33/0.15	887.6 (887.358, 887.842)	11.92 (11.919, 11.921)
	80/20	1.14/0.03	20.29/0.21	892.3 (892.057, 892.544)	11.94 (11.939, 11.941)
	80/25	1.15/0.04	21.46/0.22	892.0 (891.757, 892.243)	11.94 (11.939, 11.941)
ASiR-V30	70/15	0.94/0.02	31.28/0.24	886.7 (886.458, 886.942)	11.92 (11.919, 11.921)
	70/20	0.91/0.02	32.73/0.11	887.3 (887.058, 887.542)	11.92 (11.919, 11.921)
	70/25	0.96/0.03	33.17/0.21	892.5 (892.256, 892.744)	11.95 (11.949, 11.951)
	80/15	1.24/0.01	33.03/0.31	888.5 (888.258, 888.742)	11.93 (11.929, 11.931)
	80/20	1.26/0.02	34.33/0.07	891.1 (890.857, 891.343)	11.94 (11.939, 11.941)
	80/25	1.29/0.03	34.77/0.25	891.1 (890.857, 891.343)	11.94 (11.939, 11.941)

Table 3 (*continued*)

Table 3 (*continued*)

Reconstruction	Tube input (kVp/mAs)	SNR, mean/SD	CNR, mean/SD	Volume, mean (95% CI)	Diameter, mean (95% CI)
ASiR-V50	70/15	1.04/0.03	35.28/0.31	887.0 (886.758, 887.242)	11.92 (11.919, 11.921)
	70/20	0.99/0.01	36.39/0.15	886.7 (886.458, 886.942)	11.92 (11.919, 11.921)
	70/25	1.06/0.05	37.67/0.24	892.5 (892.256, 892.744)	11.95 (11.949, 11.951)
	80/15	1.36/0.02	36.94/0.33	887.6 (887.358, 887.842)	11.92 (11.919, 11.921)
	80/20	1.38/0.03	40.01/0.17	890.5 (890.257, 890.743)	11.94 (11.939, 11.941)
	80/25	1.42/0.03	41.54/0.22	890.5 (890.257, 890.743)	11.94 (11.939, 11.941)
ASiR-V70	70/15	1.12/0.01	40.01/0.33	887.0 (886.758, 887.242)	11.92 (11.919, 11.921)
	70/20	1.09/0.03	42.81/0.17	885.5 (885.258, 885.742)	11.91 (11.909, 11.911)
	70/25	1.12/0.02	42.59/0.26	891.4 (891.157, 891.643)	11.94 (11.939, 11.941)
	80/15	1.51/0.04	42.82/0.11	887.3 (887.058, 887.542)	11.92 (11.919, 11.921)
	80/20	1.52/0.01	47.45/0.24	891.1 (890.857, 891.343)	11.94 (11.939, 11.941)
	80/25	1.57/0.03	47.70/0.08	890.8 (890.557, 891.043)	11.94 (11.939, 11.941)
DLIR-L	70/15	1.03/0.02	30.85/0.17	893.4 (893.156, 893.644)	11.95 (11.949, 11.951)
	70/20	1.04/0.02	32.52/0.28	895.8 (895.556, 896.044)	11.96 (11.959, 11.961)
	70/25	1.12/0.01	33.74/0.25	899.0 (898.755, 899.245)	11.97 (11.969, 11.971)
	80/15	1.47/0.02	35.64/0.11	890.8 (890.557, 891.043)	11.94 (11.939, 11.941)
	80/20	1.55/0.01	37.18/0.21	894.6 (894.356, 894.844)	11.95 (11.949, 11.951)
	80/25	1.62/0.02	38.23/0.17	895.5 (895.256, 895.744)	11.96 (11.959, 11.961)
DLIR-M	70/15	1.31/0.05	41.11/0.27	893.4 (893.156, 893.644)	11.95 (11.949, 11.951)
	70/20*	1.36/0.03	43.17/0.31	897.1 (896.904, 897.394)	11.97 (11.964, 11.966)
	70/25	1.43/0.02	44.48/0.18	897.5 (897.255, 897.745)	11.97 (11.969, 11.971)
	80/15	1.93/0.01	48.11/0.15	892.3 (892.057, 892.544)	11.94 (11.939, 11.941)
	80/20	2.07/0.01	50.17/0.28	895.5 (895.256, 895.744)	11.96 (11.959, 11.961)
	80/25	2.22/0.01	51.78/0.21	895.2 (894.956, 895.444)	11.96 (11.959, 11.961)
DLIR-H	70/15	2.21/0.03	66.01/0.33	896.9 (896.655, 897.145)	11.97 (11.969, 11.971)
	70/20	2.26/0.04	70.15/0.28	897.8 (897.555, 898.045)	11.97 (11.969, 11.971)
	70/25	2.42/0.02	74.05/0.18	901.0 (900.754, 901.246)	11.98 (11.979, 11.981)
	80/15	3.31/0.01	77.22/0.15	893.3 (893.056, 893.544)	11.95 (11.949, 11.951)
	80/20	3.53/0.01	82.35/0.32	899.0 (898.755, 899.245)	11.97 (11.969, 11.971)
	80/25	3.74/0.02	85.67/0.17	900.1 (899.854, 900.346)	11.98 (11.979, 11.981)

*, volume and diameter were measured twice by each of two physicians in a semi-automatic manner, and no inter-observer variability occurred except with DLIR-M at 70 kVp and 20 mAs. CI, confidence interval; FBP, filtered back-projection; ASiR-V, adaptive statistical iterative reconstruction; DLIR, deep learning-based image reconstruction; L, low; M, medium; H, high; SNR, signal to noise ratio; CNR, contrast to noise ratio; SD, standard deviation.

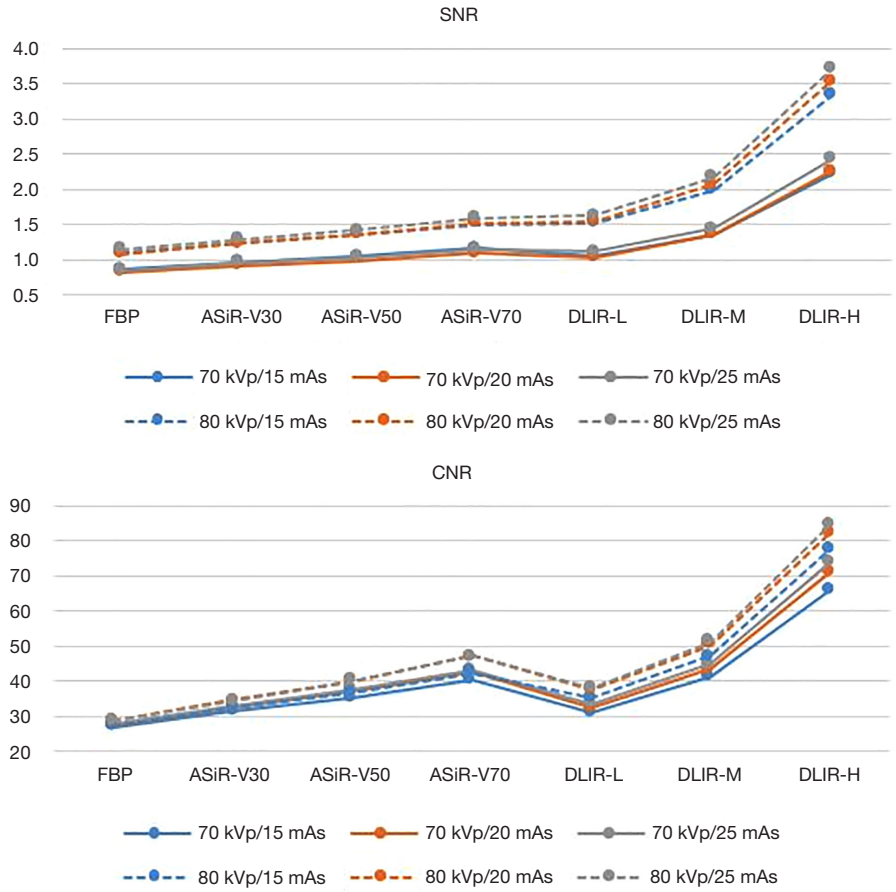

Figure 2 SNR and CNR measurements according to image reconstruction method at various doses. In the FBP reference image, scanned at 120 kVp, 50 mAs, the SNR =1.92 and CNR =42.55. ASiR-V, adaptive statistical iterative reconstruction, DLIR, deep learning-based image reconstruction; L, low; M, medium; H, high; SNR, signal to noise ratio; CNR, contrast to noise ratio; FBP, filtered back-projection.

than were FBP or any ASiR-V image (all P<0.001). In particular, in the semi-automatically measured volumes and diameters, we confirmed that the average MER of DLIR-Low, -Medium, and -High was higher than that of the reference image (120 kVp, 50 mAs) at a higher dose (*Table 4*). In the semi-automatic measurement of the nodule volume, we found no intra-observer variability (ICC =1), and all but two of the inter-observer measurements were consistent (ICC =0.9825 at 70 kVp, 20 mAs and ICC =0.9995 at 80 kVp, 15 mAs). In the semi-automatic diameter measurement, there was no intra-observer variability (ICC =1), and the inter-observer measurements were mostly consistent (ICC =0.9885 at 70 kVp, 20 mAs).

Discussion

Preserving image quality while reducing radiation dose is important because of the increased use of LDCT scans (20). Various attempts have been made to reduce the radiation dose from CT imaging by lowering the tube current and voltage, implementing automatic exposure control, reducing the number of radiation projections, imaging filtering, and development of efficient image reconstruction methods (21-29).

In this study, we confirmed that the DLIR technique quantitatively and qualitatively improves the image quality of low-dose chest CT images compared with the conventional FBP and IR techniques. The subjective evaluation confirmed that the DLIR images were consistently superior to the conventional FBP and IR images under various dose conditions (0.24 to 2.47 mGy). In the subjective analysis, DLIR images using a lower dose [DLIR-Low (0.46–0.74 mGy), DLIR-Medium (0.31–0.74 mGy), and DLIR-High (0.18–0.74 mGy)] had quality comparable to

Table 4 Measured value versus the real value of the nodule

Variables	FBP	ASiR-V			DLIR			Reference*
		30	50	70	Low	Medium	High	
MER_{volume}	96.73 [96.44, 97.08]	97.21 [97.54, 96.90]	97.94 [98.31, 97.67]	98.24 [98.52, 97.87]	98.90 [99.36, 98.45]	98.94 [99.20, 98.62]	99.25 [99.58, 98.73]	98.74 [98.74, 98.74]
$MER_{diameter}$	96.94 [97.08, 96.83]	98.19 [98.33, 98.08]	98.60 [98.75, 98.50]	99.40 [99.50, 99.25]	99.63 [99.75, 99.50]	99.65 [99.75, 99.50]	99.75 [99.83, 99.58]	99.58 [99.58, 99.58]

The MER results are expressed as mean [max, min]. *, 2.47 mGy with FBP at a tube current of 50 mAs and tube potential of 120 kVp. MER_{volume} = |measured volume/actual volume| ×100%; $MER_{diameter}$ = |measured diameter/actual diameter| ×100%. MER, measurement error ratio; FBP, filtered back-projection; ASiR-V, adaptive statistical iterative reconstruction; DLIR, deep learning-based image reconstruction.

that of the reference FBP (2.47 mGy) image. We found that a dose of 0.24 mGy (70 kV and 20 mAs) would be appropriate for LDCT when the DLIR-High technique was applied, which is 92.7% of radiation dose reduction compared to the reference radiation dose of 2.47 mGy. The image quality tendency obtained through subjective evaluation was objectively supported by the SNR, CNR, nodule diameter, and nodule volume values obtained using an automatic measurement algorithm.

Because dose reduction alters image quality, it is crucial to find a proper balance between dose reduction and diagnostic performance. Our experimental results are crucial to understand the effects of acquisition and reconstruction parameters on required dose and image quality. DLIR reconstruction could improve image quality in terms of image noise, SNR, and CNR and allow clinicians to obtain values similar to those obtained with the reference, higher-dose CT protocol. The quantitative noise evaluation and image quality with the DLIR method found in our study have been reported in other studies (15-17,30-34). Those studies showed image quality evaluation results according to the image reconstruction method under various dose conditions. However, our study aim was to find the low-dose condition (level) for DLIR images usable in actual clinical practice and to evaluate various image reconstruction methods in low-dose conditions. Because our study used a phantom, the environment was not that of the actual clinical image. However, because CT scans of various doses cannot be performed multiple times on actual patients, our results have value.

A potential concern when screening for lung cancer is that changes in the apparent shape of a nodule caused by differences in the dose and image reconstruction method could have a net effect on the final Lung Image Reporting and Data System coding and subsequent patient management. In this study, the nodule measurement results using an automated algorithm showed that the low-dose DLIR images did not change the shapes of structures such as nodules compared with those of the sufficient-dose FBP and IR images. In other words, the DLIR approach can further reduce the patient dose to ultra-low levels, improving the safety profile of lung cancer screening examinations without sacrificing screening or diagnostic performance.

Our study has several limitations. First, only phantom experiments were performed, with no clinical verification. Because one of our main goals in this study was to compare images of a single object obtained using various doses, we used the phantoms. Clinical studies should be performed using the optimal dose level for DLIR imaging identified in this study. Second, our quantitative evaluation in this study used only a solid nodule with a diameter of 12 mm. We made that choice because nodules smaller than 12 mm and the sub-solid and pure nodules produced segmentation failure or unreliable segmentation results. Third, our study was conducted using a specific equipment company's DLIR method, and we did not compare its performance against methods or equipment from other companies. Currently, two CT manufacturers have commercialized DLIR reconstruction technology, and the DLIRs of the two manufacturers have different characteristics. The DLIR used in this study is trained using high-dose FBP images, and the other manufacturer's method uses model-based reconstruction images as training data. Therefore, the characteristics of the two image reconstruction methods are likely to differ.

In conclusion, the DLIR technique quantitatively and qualitatively improved the image quality from LDCT scans at a lower radiation dose than required by the conventional FBP and IR techniques. It was possible to reduce the radiation dose of 2.47 mGy by about 74.5% (DLIR-Low),

87.5% (DLIR-Medium), and 92.7% (DLIR-High) using the DLIR technique without compromising image quality. We found that a dose of 0.24 mGy (70 kV and 20 mAs) would be the appropriate effective radiation dose for LDCT when using the DLIR-High technique. This method can be used to reduce the radiation burden of periodic lung cancer screening examinations.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-618/coif). YJ, YI, MD, BT, RS, KU, KJ are employee of GE Healthcare, the manufacturer of the CT system used in this study. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The research object of this study is a phantom. Ethics committee approval and informed consent are not required.

References

1. Rampinelli C, De Marco P, Origgi D, Maisonneuve P, Casiraghi M, Veronesi G, Spaggiari L, Bellomi M. Exposure to low dose computed tomography for lung cancer screening and risk of cancer: secondary analysis of trial data and risk-benefit analysis. BMJ 2017;356:j347.
2. Nishizawa K, Iwai K, Matsumoto T, Sakashita K, Limima TA, Tateno Y, Miyamoto T, Shimura A, Takagi H. Estimation of the exposure and a risk–benefit analysis for a CT system designed for a lung cancer mass screening unit. Radiat Prot Dosimetry 1996;67:101-8.
3. Mascalchi M, Sali L. Lung cancer screening with low dose CT and radiation harm-from prediction models to cancer incidence data. Ann Transl Med 2017;5:360.
4. Greffier J, Macri F, Larbi A, Fernandez A, Khasanova E, Pereira F, Mekkaoui C, Beregi JP. Dose reduction with iterative reconstruction: Optimization of CT protocols in clinical practice. Diagn Interv Imaging 2015;96:477-86.
5. Tang H, Yu N, Jia Y, Yu Y, Duan H, Han D, Ma G, Ren C, He T. Assessment of noise reduction potential and image quality improvement of a new generation adaptive statistical iterative reconstruction (ASIR-V) in chest CT. Br J Radiol 2018;91:20170521.
6. Geyer LL, Schoepf UJ, Meinel FG, Nance JW Jr, Bastarrika G, Leipsic JA, Paul NS, Rengo M, Laghi A, De Cecco CN. State of the Art: Iterative CT Reconstruction Techniques. Radiology 2015;276:339-57.
7. Thibault JB, Sauer KD, Bouman CA, Hsieh J. A three-dimensional statistical approach to improved image quality for multislice helical CT. Med Phys 2007;34:4526-44.
8. Katsura M, Sato J, Akahane M, Matsuda I, Ishida M, Yasaka K, Kunimatsu A, Ohtomo K. Comparison of pure and hybrid iterative reconstruction techniques with conventional filtered back projection: image quality assessment in the cervicothoracic region. Eur J Radiol 2013;82:356-60.
9. Wu D, Kim K, Li Q. Computationally efficient deep neural network for computed tomography image reconstruction. Med Phys 2019;46:4763-76.
10. Zhang HM, Dong B. A Review on Deep Learning in Medical Image Reconstruction. Journal of the Operations Research Society of China 2020;8:311-40.
11. Kuanar S, Athitsos V, Mahapatra D, Rao KR, Akhtar Z, Dasgupta D. Low dose abdominal CT image reconstruction: an unsupervised learning based approach. 2019 IEEE International Conference on Image Processing (ICIP); 22-25 September 2019; Taipei, Taiwan. IEEE, 2019.
12. Shin YJ, Chang W, Ye JC, Kang E, Oh DY, Lee YJ, Park JH, Kim YH. Low-Dose Abdominal CT Using a Deep Learning-Based Denoising Algorithm: A Comparison with CT Reconstructed with Filtered Back Projection or Iterative Reconstruction Algorithm. Korean J Radiol 2020;21:356-64.
13. Lim WH, Choi YH, Park JE, Cho YJ, Lee S, Cheon JE, Kim WS, Kim IO, Kim JH. Application of Vendor-Neutral Iterative Reconstruction Technique to Pediatric Abdominal Computed Tomography. Korean J Radiol

2019;20:1358-67.
14. Shan H, Padole A, Homayounieh F, Kruger U, Khera RD, Nitiwarangkul C, Kalra MK, Wang G. Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction. Nat Mach Intell 2019;1:269-76.
15. 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;22:131-8.
16. Wang H, Li LL, Shang J, Song J, Liu B. Application of deep learning image reconstruction in low-dose chest CT scan. Br J Radiol 2022;95:20210380.
17. Nam JG, Hong JH, Kim DS, Oh J, Goo JM. Deep learning reconstruction for contrast-enhanced CT of the upper abdomen: similar image quality with lower radiation dose in direct comparison with iterative reconstruction. Eur Radiol 2021;31:5533-43.
18. Ludewig E, Richter A, Frame M. Diagnostic imaging--evaluating image quality using visual grading characteristic (VGC) analysis. Vet Res Commun 2010;34:473-9.
19. Yadolah D. Kolmogorov-Smirnov Test. In: Dodge Y. The Concise Encyclopedia of Statistics. 1st edition. Springer, 2008:283-7.
20. Ravenel JG, Scalzetti EM, Huda W, Garrisi W. Radiation exposure and image quality in chest CT examinations. AJR Am J Roentgenol 2001;177:279-84.
21. Kubo T, Lin PJ, Stiller W, Takahashi M, Kauczor HU, Ohno Y, Hatabu H. Radiation dose reduction in chest CT: a review. AJR Am J Roentgenol 2008;190:335-43.
22. Hamberg LM, Rhea JT, Hunter GJ, Thrall JH. Multi-detector row CT: radiation dose characteristics. Radiology 2003;226:762-72.
23. Thomas CK, Mühlenbruch G, Wildberger JE, Hohl C, Das M, Günther RW, Mahnken AH. Coronary artery calcium scoring with multislice computed tomography: in vitro assessment of a low tube voltage protocol. Invest Radiol 2006;41:668-73.
24. Mulkens TH, Bellinck P, Baeyaert M, Ghysen D, Van Dijck X, Mussen E, Venstermans C, Termote JL. Use of an automatic exposure control mechanism for dose optimization in multi-detector row CT examinations: clinical evaluation. Radiology 2005;237:213-23.
25. Kalra MK, Rizzo S, Maher MM, Halpern EF, Toth TL, Shepard JA, Aquino SL. Chest CT performed with z-axis modulation: scanning protocol and radiation dose. Radiology 2005;237:303-8.
26. Gao Y, Bian Z, Huang J, Zhang Y, Niu S, Feng Q, Chen W, Liang Z, Ma J. Low-dose X-ray computed tomography image reconstruction with a combined low-mAs and sparse-view protocol. Opt Express 2014;22:15190-210.
27. Hu Z, Gao J, Zhang N, Yang Y, Liu X, Zheng H, Liang D. An improved statistical iterative algorithm for sparse-view and limited-angle CT image reconstruction. Sci Rep 2017;7:10747.
28. Kachelriess M, Watzke O, Kalender WA. Generalized multi-dimensional adaptive filtering for conventional and spiral single-slice, multi-slice, and cone-beam CT. Med Phys 2001;28:475-90.
29. Kubo T, Nishino M, Kino A, Yoshimura N, Lin PJ, Takahashi M, Raptopoulos V, Hatabu H. 3-dimensional adaptive raw-data filter: evaluation in low dose chest multidetector-row computed tomography. J Comput Assist Tomogr 2006;30:933-8.
30. Solomon J, Lyu P, Marin D, Samei E. Noise and spatial resolution properties of a commercially available deep learning-based CT reconstruction algorithm. Med Phys 2020;47:3961-71.
31. Racine D, Becce F, Viry A, Monnin P, Thomsen B, Verdun FR, Rotzinger DC. 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.
32. Greffier J, Hamard A, Pereira F, Barrau C, Pasquier H, Beregi JP, Frandon J. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. Eur Radiol 2020;30:3951-9.
33. Higaki T, Nakamura Y, Zhou J, Yu Z, Nemoto T, Tatsugami F, Awai K. Deep Learning Reconstruction at CT: Phantom Study of the Image Characteristics. Acad Radiol 2020;27:82-7.
34. Sun J, Li H, Li J, Yu T, Li M, Zhou Z, Peng Y. Improving the image quality of pediatric chest CT angiography with low radiation dose and contrast volume using deep learning image reconstruction. Quant Imaging Med Surg 2021;11:3051-8.

Cite this article as: Jung Y, Hur J, Han K, Imai Y, Hong YJ, Im DJ, Lee KH, Desnoyers M, Thomsen B, Shigemasa R, Um K, Jang K. Radiation dose reduction using deep learning-based image reconstruction for a low-dose chest computed tomography protocol: a phantom study. Quant Imaging Med Surg 2023;13(3):1937-1947. doi: 10.21037/qims-22-618

