Original Article | Gastrointestinal Imaging

eISSN 2005-8330
https://doi.org/10.3348/kjr.2021.0683
Korean J Radiol 2022;23(4):402-412

Korean Journal of Radiology
KJR

Image Quality and Lesion Detectability of Lower-Dose Abdominopelvic CT Obtained Using Deep Learning Image Reconstruction

June Park[1], Jaeseung Shin[2], In Kyung Min[3], Heejin Bae[2], Yeo-Eun Kim[1], Yong Eun Chung[2]

[1]Department of Radiology, Seoul Medical Center, Seoul, Korea; [2]Department of Radiology, Yonsei University College of Medicine, Seoul, Korea; [3]Biostatistics Collaboration Unit, Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Korea

Objective: To evaluate the image quality and lesion detectability of lower-dose CT (LDCT) of the abdomen and pelvis obtained using a deep learning image reconstruction (DLIR) algorithm compared with those of standard-dose CT (SDCT) images.
Materials and Methods: This retrospective study included 123 patients (mean age ± standard deviation, 63 ± 11 years; male:female, 70:53) who underwent contrast-enhanced abdominopelvic LDCT between May and August 2020 and had prior SDCT obtained using the same CT scanner within a year. LDCT images were reconstructed with hybrid iterative reconstruction (h-IR) and DLIR at medium and high strengths (DLIR-M and DLIR-H), while SDCT images were reconstructed with h-IR. For quantitative image quality analysis, image noise, signal-to-noise ratio, and contrast-to-noise ratio were measured in the liver, muscle, and aorta. Among the three different LDCT reconstruction algorithms, the one showing the smallest difference in quantitative parameters from those of SDCT images was selected for qualitative image quality analysis and lesion detectability evaluation. For qualitative analysis, overall image quality, image noise, image sharpness, image texture, and lesion conspicuity were graded using a 5-point scale by two radiologists. Observer performance in focal liver lesion detection was evaluated by comparing the jackknife free-response receiver operating characteristic figures-of-merit (FOM).
Results: LDCT (35.1% dose reduction compared with SDCT) images obtained using DLIR-M showed similar quantitative measures to those of SDCT with h-IR images. All qualitative parameters of LDCT with DLIR-M images but image texture were similar to or significantly better than those of SDCT with h-IR images. The lesion detectability on LDCT with DLIR-M images was not significantly different from that of SDCT with h-IR images (reader-averaged FOM, 0.887 vs. 0.874, respectively; $p = 0.581$).
Conclusion: Overall image quality and detectability of focal liver lesions is preserved in contrast-enhanced abdominopelvic LDCT obtained with DLIR-M relative to those in SDCT with h-IR.
Keywords: *Computed tomography; Image reconstruction; Deep learning; Radiation dose; Abdomen*

INTRODUCTION

Despite *in vivo* and *in vitro* biological evidence of radiation-induced genomic instability, such as DNA double-strand breaks and chromosomal aberrations [1], there is still a debate regarding the actual risk of low-dose radiation [2]. However, in terms of radiation protection, a conservative approach according to the "as low as reasonably achievable" principle is recommended [3-5]. CT examinations account for a large portion of medical radiation exposure, but continue to increase in use [6]. Therefore, several efforts have been made to reduce radiation exposure during CT studies, including optimization of the CT system through improvements in quantum detection and geometrical efficiency of the detector, appropriate usage of pre- and post-patient collimators and beam-shaping filters, and automatic exposure control [7]. In addition, iterative reconstruction (IR) algorithms have been investigated as a powerful method for lowering the radiation dose because of

Received: July 13, 2021 **Revised:** October 22, 2021
Accepted: October 31, 2021
Corresponding author: Yong Eun Chung, MD, PhD, Department of Radiology, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
• E-mail: yelv@yuhs.ac

Korean Journal of Radiology
KJR

their intrinsic ability to decrease image noise and artifacts [8-13].

Model-based IR, which models both system geometry and photon statistics, can reduce image noise and artifacts, thereby maximizing the potential for dose reduction [14]. However, long reconstruction times limit their clinical application. Currently, hybrid IR (h-IR), which reduces image noise in both image space and raw data space, which requires less demanding computations and shorter scan times, is widely utilized in routine clinical settings. However, h-IR algorithms also have certain limitations, such as unfamiliar image texture (plastic-like, blotchy, or pixelated) and nonlinearity in spatial resolution, which significantly vary depending on the local contrast and radiation dose [15]. Therefore, it is challenging to maintain sufficient diagnostic performance for detecting low-contrast lesions when IR methods are combined for aggressive radiation dose reduction [16,17].

A recently introduced deep learning image reconstruction (DLIR) algorithm (TrueFidelity, GE Healthcare) utilizes a deep neural network (DNN)-based model to differentiate noise from anatomical structures and emulate high-quality filtered back projection (FBP) images [18]. In the training process, the DLIR engine generates output images from a low-dose input sinogram, compares them with high-dose FBP images from the same objects, and repeatedly fine-tunes the parameters of the DNN to suppress image noise, retain the preferred noise texture, and improve spatial resolution. Some recent studies have suggested that DLIR algorithms can further improve diagnostic image quality when compared with that obtained using IR methods at the same radiation dose level [19-22]. Moreover, there is some evidence that it might be possible to maintain diagnostic image quality even at a reduced dose level using DLIR methods [23,24]. However, all these studies only evaluated the quantitative and qualitative measures of the reconstructed images. Hence, the purpose of this study was not only to evaluate the quantitative and qualitative image quality, but also to evaluate the detectability of focal liver lesions on LDCT images obtained using a DLIR algorithm compared with that on SDCT images.

MATERIALS AND METHODS

The Institutional Review Board of Yonsei University Severance Hospital (IRB No. 4-2020-1315) approved this single-center retrospective study, and the requirement for informed consent was waived. All patient data were anonymized prior to the analysis.

Study Population

In our institution, the LDCT protocol with DLIR was the routine abdominopelvic CT protocol for the GE Revolution CT machine in May 2020. Patients who underwent abdominopelvic CT with the LDCT protocol between May 2020 and August 2020 and who also underwent a prior abdominopelvic CT with the SDCT protocol within one year previously using the same CT scanner were included in this study (n = 123; mean age ± standard deviation [SD], 63 ± 11 years; male:female, 70:53). Quantitative and qualitative image analyses were performed in 111 patients (mean age ± SD, 63 ± 12 years; male:female, 62:49) after excluding 12 patients who underwent prior liver surgery (Fig. 1). Four patients who underwent radiofrequency ablation were included in the study population because we presumed that locoregional treatments, such as radiofrequency ablation, did not affect both quantitative and qualitative image analysis. Lesion conspicuity was evaluated in 45 of 111 patients who had at least one focal liver lesion.

For lesion detectability assessment, among the 123 patients, 98 patients who had two additional imaging studies composed of routine-dose CT or MRI (one within 1 year before SDCT and the other within 1 year after LDCT) were initially screened. After excluding 14 patients who had more than 10 focal liver lesions, lesion detectability was evaluated in 84 patients. Details about determining the reference standard are provided in Supplement.

CT Image Acquisition

The SDCT and LDCT protocols were performed with the same parameters except for the noise index: rotation time, 0.5 seconds; spiral pitch factor, 0.992:1; table feed per rotation, 79.375 mm/rotation; total collimation width, 80 mm; detector configuration, 128 x 0.625 mm; tube voltage, 100 kVp; and automatic exposure control, Smart mA. The portal venous phase was acquired 55 seconds after the attenuation value of the aorta reached 100 Hounsfield unit (HU) or 30 s after the end of the late arterial phase. The noise index is a parameter that sets the SD of CT numbers within a region of interest (ROI) in a water phantom of a specific size. A proprietary algorithm in GE Healthcare modulates tube current values to maintain the same image noise level during the scanning process in consideration of the patient-specific attenuation characteristics on the

 https://doi.org/10.3348/kjr.2021.0683

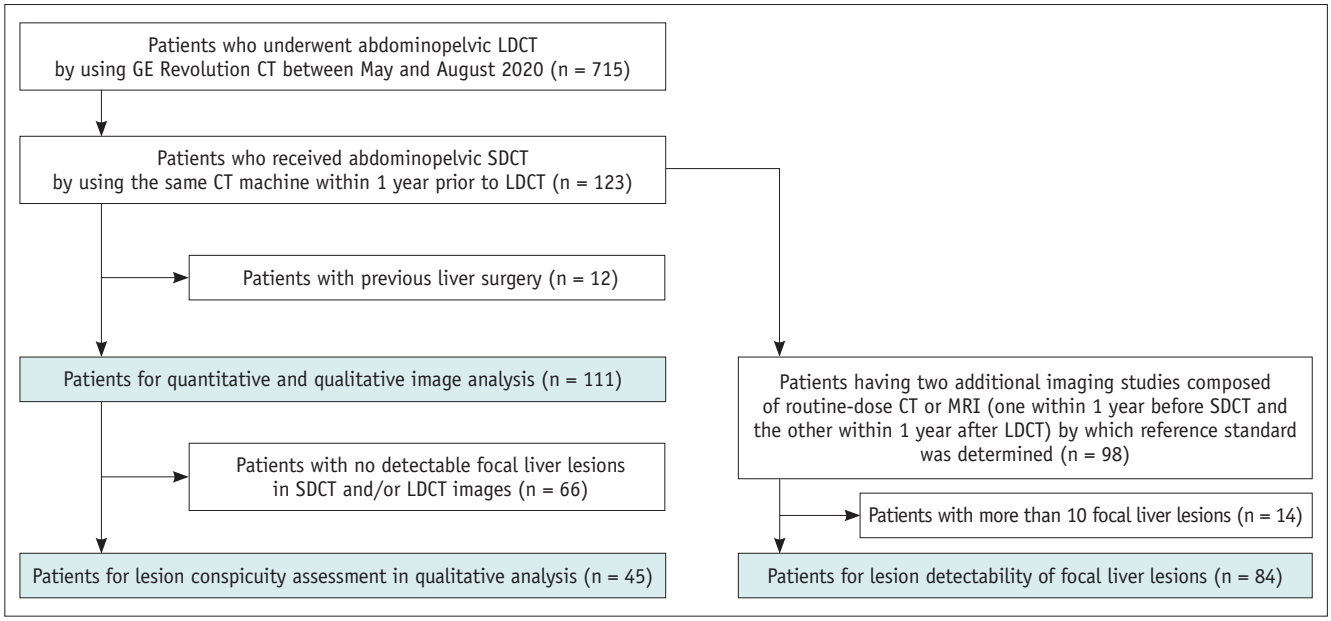

Fig. 1. Flow chart of patient selection. LDCT = lower-dose CT, SDCT = standard-dose CT

scout image [25]. The noise index of the LDCT protocol was 17, which was higher than that of the SDCT protocol at 14, and this difference was intentional to achieve an estimated dose reduction of 30%. LDCT images were reconstructed with h-IR (adaptive statistical iterative reconstruction-Veo [ASiR-V] 40%), DLIR (TrueFidelity) medium-strength (DLIR-M), and DLIR high-strength (DLIR-H), while SDCT images were reconstructed with h-IR. All CT images were reformatted with a slice thickness of 3.0 mm and slice interval of 3.0 mm.

Radiation Exposure

CT dose index volume (CTDIvol) and dose length product were extracted from the dose reports of all CT scans. Each patient's anteroposterior and lateral dimensions were measured at the level of the proximal right renal artery, and size-specific dose estimation (SSDE) was calculated by multiplying CTDIvol values and conversion factors based on the AAPM report 204 [26].

Quantitative Image Quality Analysis

A researcher measured the attenuation value and SD in the liver, paraspinal muscle, abdominal aorta, and subcutaneous fat [21,27]. The mean attenuation and SD values of the liver were calculated by averaging four measurements for each of the four liver sections: the left lateral, left medial, right anterior, and right posterior sections. A single ROI was measured on other anatomic structures: the paraspinal muscle, abdominal aorta, and subcutaneous fat. All ovoid ROIs (range 100–200 mm^2) were carefully placed in areas of homogeneous attenuation to avoid confounding structures such as large vessels, intramuscular fat, or vessel wall calcifications. All measurements were performed at the level of the umbilical portion of the left portal vein and were kept as constant as possible in their size and position between the SDCT and LDCT images (Fig. 2).

The signal-to-noise ratio (SNR) for the anatomy of interest was calculated using the following equation:

$$SNR_i = ROI_i/SD_i$$

where ROI_i is the attenuation value of the anatomical structure of interest, and SD_i is the SD value of the same structure. The contrast-to-noise ratio (CNR) relative to fat for the anatomical structure of interest was determined using the following equation:

$$CNR_i = (ROI_i - ROI_{fat})/SD_{fat}$$

The CNR between the focal liver lesions and background liver was not calculated in this study because it is difficult to draw ROIs accurately in small (< 5 mm) focal liver lesions.

Qualitative Image Quality Analysis

Among the three different LDCT reconstruction algorithms, the one showing the smallest difference in the

https://doi.org/10.3348/kjr.2021.0683

Korean Journal of Radiology
KJR

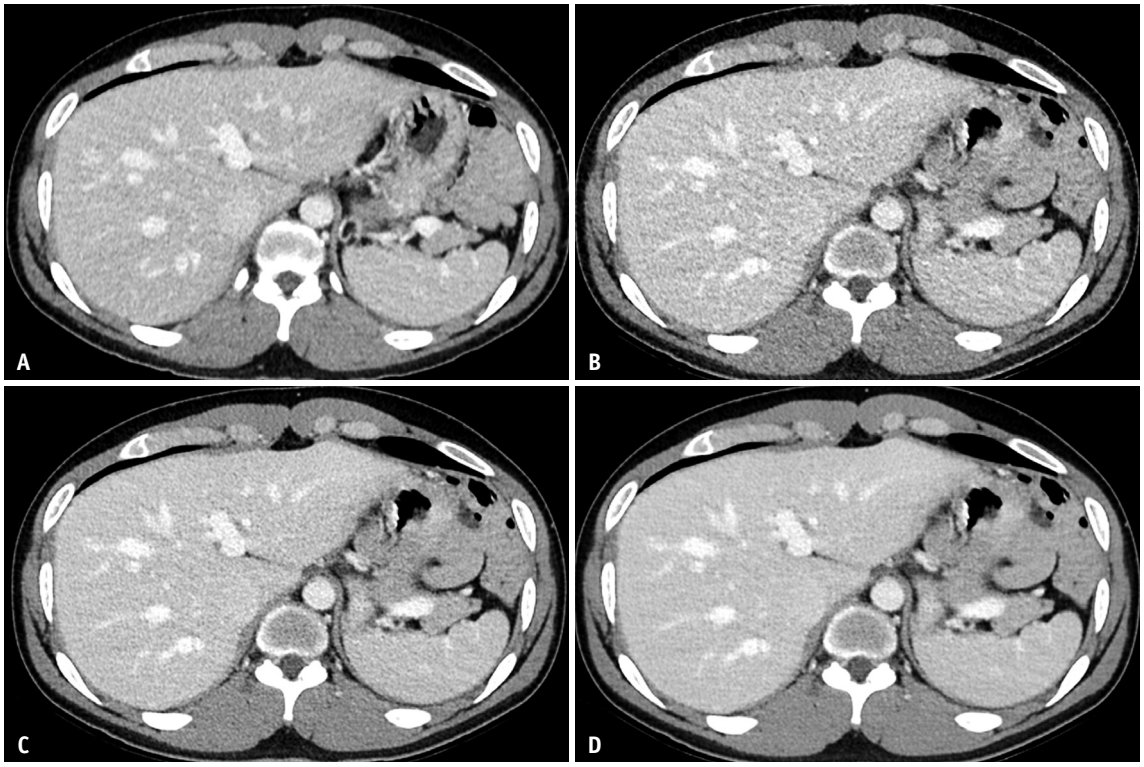

Fig. 2. Follow-up SDCT and LDCT protocol contrast-enhanced abdominopelvic CT images of a 61-year-old male with colon cancer; SDCT and LDCT studies were performed within one year of each other.
A-D. Axial CT images were taken at the same anatomic level to compare image quality between the SDCT protocol **(A)** and LDCT protocol **(B-D)**. SDCT images were reconstructed with h-IR **(A)**, while LDCT images were reconstructed with h-IR **(B)**, DLIR-M **(C)**, and DLIR-H **(D)**. Image noise in the liver in CT images **A, B, C,** and **D** is 10.8, 16.1, 12.8, and 8.4, respectively. Quantitative measures of image noise, signal-to-noise ratio and contrast-to-noise ratio, show equivalence between SDCT with h-IR images **(A)** and LDCT with DLIR-M images **(C)**. DLIR-H = deep learning image reconstruction high-strength, DLIR-M = deep learning image reconstruction medium-strength, h-IR = hybrid iterative reconstruction, LDCT = lower-dose CT, SDCT = standard-dose CT

quantitative parameters from those of SDCT images was selected for qualitative analysis. This was done to select images with the most similar image texture to the SDCT images. The LDCT obtained with DLIR-M was chosen, and it was displayed side by side with the SDCT with h-IR images. The images were independently reviewed by two board-certified radiologists (reader 1 and reader 2 with 19 and 6 years of experience in abdominal radiology, respectively) (Fig. 3). The readers were instructed to score the relative image quality of the LDCT with DLIR-M images compared with SDCT with h-IR images in terms of overall image quality, image noise, image sharpness, image texture, and lesion conspicuity. A score of 5 was given if the LDCT with DLIR-M images was markedly superior to the SDCT with h-IR images, 4 if marginally superior, 3 if equivalent; 2 if marginally inferior; and 1 if markedly inferior. Higher scores corresponded to better overall image quality, less image noise, better image sharpness, more preferred image texture, and better lesion conspicuity.

Lesion Detectability

Lesion detectability was also compared between LDCT with DLIR-M and SDCT with h-IR images. Two board-certified radiologists were asked to locate all liver lesions and measure the size of each lesion. For each lesion localization, radiologists rated the confidence of lesion detection using a 5-point Likert scale: 5, most likely present; 4, probably present, 3, indeterminate; 2, probably absent; and 1, most likely absent. A researcher temporarily stored the CT images for lesion detectability analysis in a mini-PACS system (Aquarius workstation, TeraRecon) to blind the radiation dose, reconstruction algorithm, study date, and all patient information from the readers. Cases were reviewed in a random order, and readers were permitted to scroll, zoom, and change window settings.

Statistical Analysis

The paired *t* test was used to compare continuous variables, such as patient characteristics and radiation

 https://doi.org/10.3348/kjr.2021.0683

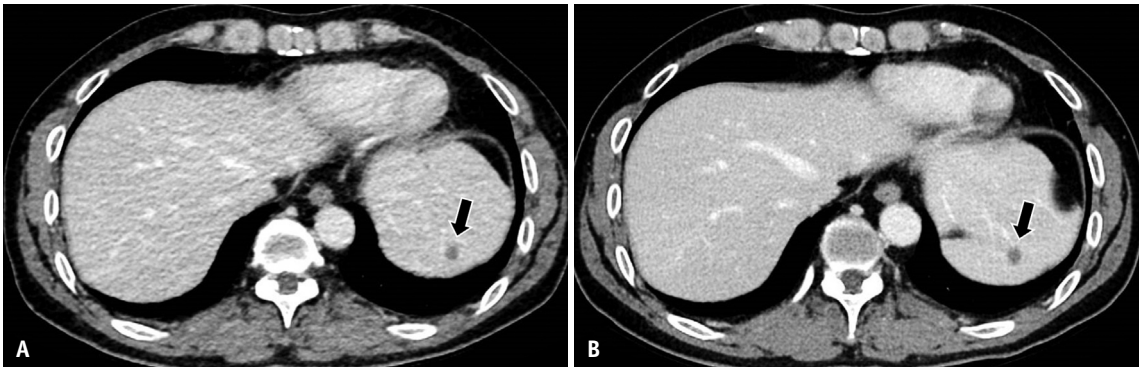

Fig. 3. A side-to-side comparison of qualitative measures between SDCT with h-IR images (A) and LDCT with DLIR-M images (B) of a 55-year-old male with colon cancer.
A, B. Black arrows show a focal liver lesion being assessed for lesion conspicuity. In terms of overall image quality and lesion conspicuity, all readers interpreted the LDCT with DLIR-M images **(B)** as being marginally superior (all score 4) to the SDCT with h-IR images **(A)**. Although the image texture of SDCT with h-IR images was preferred over that of LDCT with DLIR-M images, the qualitative parameters of LDCT with DLIR-M images including overall image quality, image noise, image sharpness, and lesion conspicuity were comparable or superior to those of SDCT with h-IR images. DLIR-M = deep learning image reconstruction medium-strength, h-IR = hybrid iterative reconstruction, LDCT = lower-dose CT, SDCT = standard-dose CT

exposure, between the LDCT and SDCT protocols. Repeated-measures analysis of variance with Dunnett's multiple comparison test was performed for image noise, SNR, and CNR. Furthermore, equivalence testing between SDCT with h-IR images and LDCT with DLIR-M images was performed. An equivalence between SDCT with h-IR images and LDCT with DLIR images was inferred when the two-sided 95% confidence interval (CI) of the mean difference (95% CI of difference) was within a pre-specified equivalence margin [28] (prespecified margins for the quantitative analysis are provided in Supplementary Table 1). The equivalence margin was set as the SD of measures in the SDCT with h-IR images because a mean difference smaller than this was considered clinically insignificant [29,30]. For the comparison of qualitative measures, a one-sample Wilcoxon signed-rank test was used. For lesion detectability analysis, jackknife free-response receiver operating characteristic (JAFROC) plots were used to calculate the figures-of-merit [31]. We used the Dorfman–Berbaum–Metz multireader-multicase method with a selection of fixed-readers random-case output in JAFROC software version 4.2.1 [32]. Interobserver agreement between the two readers was assessed using the bias-adjusted kappa (PABAK) statistic [33]. Statistical significance was set at $p < 0.05$. All statistical analyses except the JAFROC methods were performed using the R package (version 3.6.0; R Foundation for Statistical Computing) and SAS (version 9.4; SAS Institute Inc.).

RESULTS

Patient Characteristics

Patient characteristics, including age, sex, study interval, and underlying morbidities, are presented in Table 1. The mean study interval ± SD between SDCT and LDCT was 221 ± 91 days. Routine follow-up after treating primary colorectal tumors accounted for was the most common indication for abdominopelvic CT follow-up (104/111, 93.7%). Supplementary Tables 2 and 3 show the population characteristics for lesion conspicuity assessment (n = 45) and lesion detectability analysis (n = 84), respectively.

Radiation Exposure

Body weight, effective diameter, and scan length were not significantly different between the LDCT and SDCT protocols (all $p > 0.05$) (Table 1). The LDCT protocol resulted in a dose reduction of 35.1% in SSDE (6.6 ± 1.0 mGy) compared with that in the SDCT protocol (10.3 ± 1.6 mGy) (Table 2). Supplementary Tables 4 and 5 show the radiation exposure for lesion conspicuity assessment (n = 45) and lesion detectability analysis (n = 84), respectively.

Quantitative Image Quality Analysis

Image Noise

Image noise was higher on LDCT with h-IR images than SDCT with h-IR images (all $p < 0.001$), whereas it was lower on LDCT with DLIR-H images than SDCT with h-IR images (all $p < 0.001$) for all anatomical structures. Image noise

https://doi.org/10.3348/kjr.2021.0683

Korean Journal of Radiology
KJR

Table 1. Patient Characteristics

	Standard Dose	Lower Dose	*P*
Total number	111		N/A
Age, years*†	63 ± 12 (28–89)		N/A
Sex, male:female	62:49		N/A
Study interval, days*	221 ± 91 (45–365)		N/A
Body weight, kg*	63 ± 10 (40–89)	63 ± 11 (37–93)	0.981
Effective diameter, cm*‡	25.4 ± 2.4 (20.2–33.0)	25.5 ± 2.6 (19.2–33.4)	0.631
Scan length, cm*	58.9 ± 4.1 (51.7–68.2)	59.0 ± 4.2 (50.6–69.3)	0.853
Underlying oncologic disease			
Colorectal adenocarcinoma	104		N/A
Rectal NET or GIST	3		N/A
Small bowel NET	1		N/A
Colon polyps	1		N/A
Cutaneous malignant melanoma	2		N/A

Unless otherwise specified, the data are the number of patients. *Values are the mean ± standard deviation (range), †Age was recorded at the time of the lower-dose study, ‡Effective diameter = geometric mean of AP and LAT, where AP is the anteroposterior dimension and LAT is the lateral dimension of the body cross-section. GIST = gastrointestinal stromal tumor, NET = neuroendocrine tumor, N/A = non-applicable

Table 2. Radiation Exposure

	Standard Dose	Lower Dose	Dose Reduction		*P*
			Absolute	Relative, %	
SSDE, mGy	10.3 ± 1.6 (6.5–16.4)	6.6 ± 1.0 (4.5–10.8)	3.6 ± 0.9 (1.2–5.5)	35.1 ± 5.7 (12.9–47.6)	< 0.001
DLP, mGy·cm	423.4 ± 118.5 (202.8–917.9)	276.3 ± 83.8 (139.9–667.5)	147.1 ± 53.6 (40.8–296.6)	34.5 ± 7.7 (11.7–52.2)	< 0.001
CTDIvol, mGy	7.2 ± 1.7 (3.9–14.9)	4.6 ± 1.2 (2.7–9.9)	2.5 ± 0.8 (1.0–4.9)	34.9 ± 6.7 (12.9–50.1)	< 0.001

Data are presented as the mean ± standard deviation (range). CTDIvol = CT dose index volume, DLP = dose-length product, SSDE = size-specific dose estimate

was not different between the LDCT with DLIR-M images and SDCT with h-IR images for all anatomical structures (all $p > 0.05$) (Fig. 4A, Table 3). The equivalence test confirmed that image noise was similar between LDCT with DLIR-M images and SDCT with h-IR images (95% CI of difference: -0.449 to 0.200 HU, -0.544 to 0.210 HU, and -0.932 to -0.068 HU for the liver, muscle, and aorta, respectively) (Fig. 4B).

SNR

Unlike image noise, the SNR was lower on LDCT with h-IR images than on SDCT with h-IR images (all $p < 0.001$), while LDCT with DLIR-H images had higher SNR than that of SDCT with h-IR images (all $p < 0.001$), irrespective of the anatomical structure analyzed. The SNR did not significantly differ between LDCT with DLIR-M and SDCT with h-IR images for all anatomical structures (all $p > 0.05$) (Fig. 4C, Table 3). The equivalence test found similar SNR values for LDCT with DLIR-M and SDCT with h-IR images (95% CI of difference: -0.111 to 0.475, -0.101 to 0.240, and -0.236 to 0.789 for liver, muscle, and aorta, respectively) (Fig. 4D).

CNR

The CNR was lower on LDCT with h-IR images than on SDCT with h-IR images regardless of the anatomical structure analyzed (all $p < 0.001$). The CNR of LDCT images improved when DLIR-M was applied instead of h-IR and was significantly different from the CNR of SDCT with h-IR images for all anatomical areas (all $p < 0.001$) (Fig. 4E, Table 3). The equivalence test revealed similar CNR values for LDCT with DLIR-M and SDCT with h-IR images (95% CI of difference: 2.135–3.773, 1.452–2.649, and 2.319–4.571 for the liver, muscle, and aorta, respectively) (Fig. 4F). DLIR-H further increased CNR on LDCT images and showed the highest CNR among all reconstructions (all $p < 0.001$).

 https://doi.org/10.3348/kjr.2021.0683

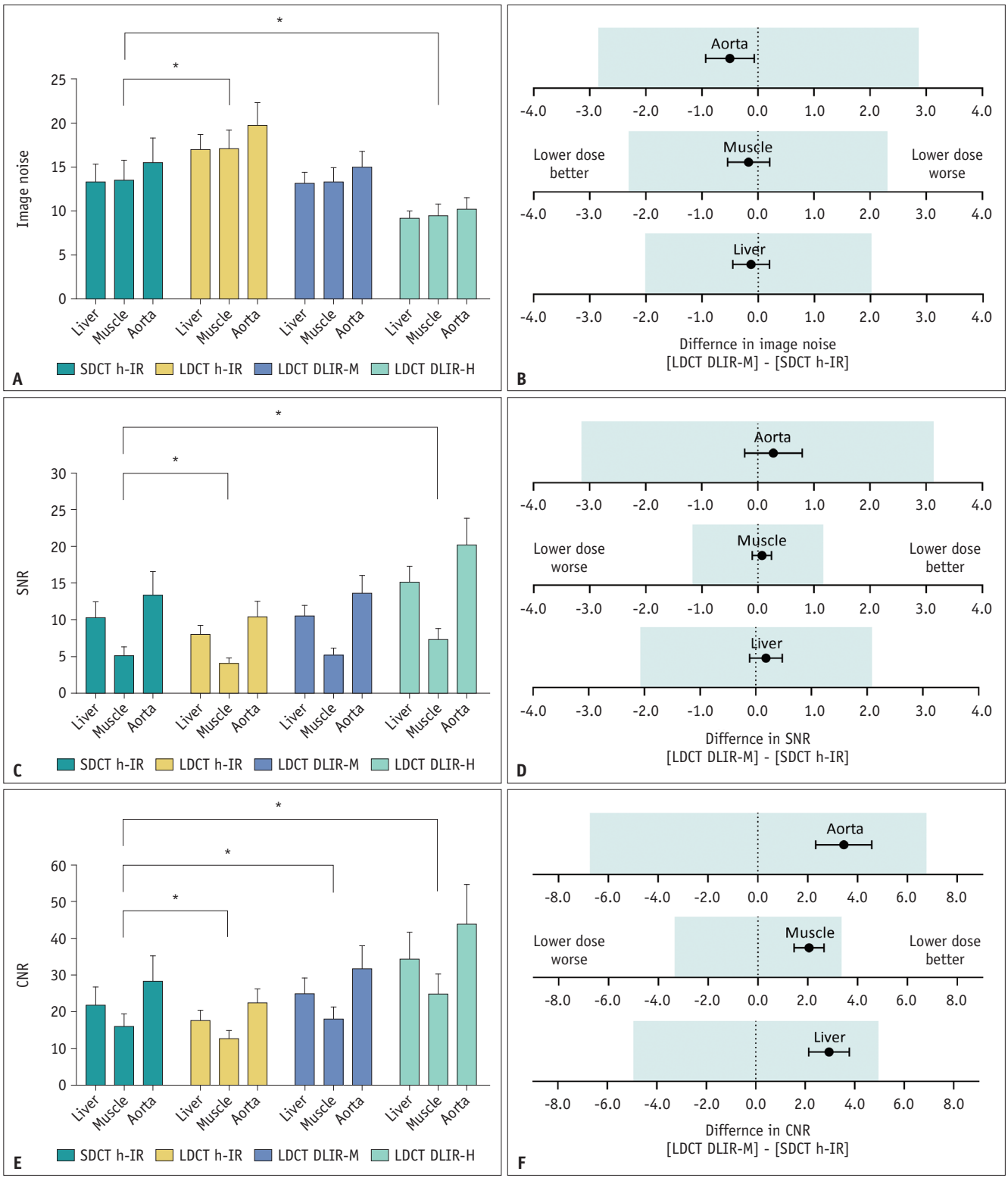

Fig. 4. Results of the quantitative image analysis. *Statistically significant difference ($p < 0.001$).
A, C. The bar graphs show no significant difference between SDCT with h-IR images and LDCT with DLIR-M images in terms of image noise **(A)** and SNR **(C)**. **B, D, F.** Equivalence tests show that LDCT with DLIR-M images are equivalent to SDCT with h-IR images with their prespecified margins (colored boxes in each figure part) in terms of image noise **(B)**, SNR **(D)**, and CNR **(F)**. **E.** Regarding CNR, all LDCT images are significantly different from SDCT with h-IR images regardless of reconstruction method, but the difference between the mean values in LDCT with DLIR-M images and SDCT with h-IR images is the smallest. CNR = contrast-to-noise ratio, DLIR-H = deep learning image reconstruction high-strength, DLIR-M = deep learning image reconstruction medium-strength, h-IR = hybrid iterative reconstruction, LDCT = lower-dose CT, SDCT = standard-dose CT, SNR = signal-to-noise ratio

https://doi.org/10.3348/kjr.2021.0683

Korean Journal of Radiology
KJR

Table 3. Quantitative Measurements of Image Quality

	SD h-IR	LD h-IR	LD DLIR-M	LD DLIR-H	Adjusted *P*		
					LD h-IR vs. SD h-IR	LD DLIR-M vs. SD h-IR	LD DLIR-H vs. SD h-IR
Image noise, HU							
Liver	13.29 ± 2.02	17.04 ± 1.66	13.17 ± 1.22	9.17 ± 0.83	< 0.001	0.854	< 0.001
Muscle	13.54 ± 2.26	17.08 ± 2.14	13.37 ± 1.55	9.47 ± 1.35	< 0.001	0.804	< 0.001
Aorta	15.54 ± 2.78	19.76 ± 2.58	15.04 ± 1.77	10.21 ± 1.31	< 0.001	0.143	< 0.001
SNR							
Liver	10.33 ± 2.12	8.09 ± 1.16	10.51 ± 1.53	15.12 ± 2.20	< 0.001	0.604	< 0.001
Muscle	5.17 ± 1.16	4.13 ± 0.69	5.24 ± 0.93	7.42 ± 1.44	< 0.001	0.838	< 0.001
Aorta	13.40 ± 3.21	10.45 ± 2.10	13.68 ± 2.40	20.28 ± 3.56	< 0.001	0.698	< 0.001
CNR							
Liver	22.13 ± 4.69	17.83 ± 2.66	25.09 ± 4.35	34.61 ± 7.33	< 0.001	< 0.001	< 0.001
Muscle	16.20 ± 3.41	13.03 ± 2.03	18.25 ± 3.23	25.11 ± 5.42	< 0.001	< 0.001	< 0.001
Aorta	28.50 ± 6.79	22.69 ± 3.76	31.94 ± 6.19	44.25 ± 10.63	< 0.001	< 0.001	< 0.001

Data are mean ± standard deviation. All *p* values are multiplicity-adjusted. Statistical significance was set at $p < 0.05$. CNR = contrast-to-noise ratio, DLIR-H = deep learning image reconstruction high-strength, DLIR-M = deep learning image reconstruction medium-strength, h-IR = hybrid iterative reconstruction, HU = Hounsfield unit, LD = lower dose, SD = standard dose, SNR = signal-to-noise ratio

Table 4. Qualitative Measurements of Image Quality

	Reader 1		Reader 2	
	Score	*P*	Score	*P*
Overall image quality	3.26 ± 0.48	< 0.001	3.46 ± 0.52	< 0.001
Image noise	3.54 ± 0.58	< 0.001	3.63 ± 0.57	< 0.001
Image sharpness	2.99 ± 0.29	> 0.999	3.04 ± 0.19	0.134
Image texture	2.82 ± 0.39	< 0.001	2.76 ± 0.43	< 0.001
Lesion conspicuity	3.09 ± 0.36	0.219	3.29 ± 0.51	0.001

Data are mean ± standard deviation. A score of 3 was assigned if the LDCT with DLIR-M images were equivalent to SDCT with h-IR images. Higher scores reflected better overall image quality, less image noise, better image sharpness, more preferred image texture, and better lesion conspicuity. A one-sample Wilcoxon signed-rank test was used to reject the null hypothesis: the median of qualitative scores for LDCT with DLIR-M images equals a hypothetical value of 3. DLIR-M = deep learning image reconstruction medium-strength, h-IR = hybrid iterative reconstruction, LDCT = lower-dose CT, SDCT = standard-dose CT

Table 5. FOM Values and DBM-MRMC Significance Test Obtained from JAFROC Analysis

	SDCT h-IR	LDCT DLIR-M	*P*
Reader 1	0.868 (0.814, 0.922)	0.889 (0.833, 0.945)	0.420
Reader 2	0.880 (0.824, 0.953)	0.886 (0.833, 0.938)	0.832
Reader average	0.874 (0.824, 0.923)	0.887 (0.838, 0.937)	0.581

Numbers in parentheses represent the 95% confidence interval for FOM or differences. DBM-MRMC = Dorfman–Berbaum–Metz multireader-multicase, DLIR-M = deep learning image reconstruction medium-strength, FOM = figure-of-merit, h-IR = hybrid iterative reconstruction, JAFROC = jackknife free-response receiver operating characteristic, LDCT = lower-dose CT, SDCT = standard-dose CT

Qualitative Image Quality Analysis

Compared to SDCT with h-IR images, LDCT with DLIR-M images showed better overall image quality, less image noise, and no significant difference in image sharpness for both readers (Table 4). The image texture of the LDCT with DLIR-M images was not preferred to that of SDCT with h-IR images for both readers. Lesion conspicuity on LDCT with DLIR-M images was not significantly different from that on SDCT with h-IR images for reader 1 and was better than that of SDCT with h-IR images for reader 2. The PABAK values of interobserver agreement were 0.59 for overall image quality, 0.89 for image sharpness, 0.72 for image texture, 0.70 for image noise, and 0.69 for lesion conspicuity, respectively.

Lesion Detectability

A total of 99 true-positive lesions with a mean size ± SD of 5.9 ± 4.6 mm (range, 3.0–31.7 mm) were found in 34 patients (Supplementary Table 3). Observer performance in the detection of focal liver lesions was not significantly different between LDCT with DLIR-M images and SDCT with h-IR images for both readers ($p > 0.05$) (Table 5).

DISCUSSION

This study showed that by using a DLIR-M, image noise, SNR, and CNR could be maintained with a 35.1% radiation

https://doi.org/10.3348/kjr.2021.0683

dose reduction relative to that of the SDCT protocol. In the qualitative analysis, LDCT with DLIR-M images showed better overall image quality, less image noise, and similar or better qualitative scores for image sharpness and lesion conspicuity, although the image texture was not preferred. The lesion detectability of focal liver lesions on LDCT with DLIR-M was not significantly different from that of SDCT with h-IR.

According to previous phantom and human studies, DLIR could reduce image noise effectively and improve overall image quality compared to those when using h-IR methods such as ASiR-V, and these effects seemed to increase as the strength of DLIR increases [19-21,34]. Some phantom studies have shown that low-contrast spatial resolution increases as the strength of DLIR decreases [19,34], and a human study reported that diagnostic confidence or lesion conspicuity did not significantly differ between DLIR-H and DLIR-M. Furthermore, slight blurring was observed for lesions smaller than 5 mm and tiny vessels on some DLIR-H images [21]. Along with the reasons mentioned above, LDCT with DLIR-M was chosen for qualitative and lesion detectability analyses because it showed the smallest differences in quantitative measures and might show the most similar image texture relative to that on SDCT with h-IR images. In addition to diagnostic performance, image texture itself is an important factor in determining whether to use a particular reconstruction method in daily practice because some radiologists tend not to prefer artificial noise texture [15,35]. In this study, we also found that the overall image quality of and lesion detectability on LDCT with DLIR-M images were not different or better than those of SDCT with h-IR images despite a reduction in radiation dose.

In terms of image texture, phantom studies have shown similar results of FBP-like noise texture being preserved in the dose reduction protocol with the application of DLIR methods [19,20,34]. According to these previous studies, the noise power spectrum (NPS), which visualizes the distribution of noise variance in terms of spatial frequencies for describing the noise amplitude and texture properties of CT images [36], only marginally shifted to lower spatial frequencies after DLIR was applied, whereas NPS with h-IR produced substantially lower frequency contents. Because low-contrast lesion detectability depends not only on spatial resolution but also on noise texture [15], it is worth noting that DLIR methods maintain FBP-like noise texture at clinically usable low-dose radiation levels. Similar to these previous studies, our study demonstrated that quantitative image noise was not significantly different between SDCT with h-IR images and LDCT with DLIR-M images, and lesion detectability for focal liver lesions was preserved even after a reduction in radiation dose. Unlike other quantitative parameters, the image texture of LDCT with DLIR-M images did not show a reader preference over SDCT with h-IR images in our study. This may be due to the shift in NPS, which can occur in DLIR-M. NPS was not evaluated in this study, and future studies may include quantitative comparison of NPS between DLIR and h-IR images.

Several attempts have been made to exploit the dose reduction potential of DLIR algorithms in abdominal CT scans. Nam et al. [23] reported that a chest CT protocol with DLIR-H showed image quality comparable to that of an abdominal CT protocol with an h-IR for upper abdominal scans with a radiation dose of less than 50%. Cao et al. [24] showed that DLIR-H methods reduced image noise and maintained overall image quality compared with images reconstructed using h-IR for abdominal CT at an extremely low dose (76% dose reduction). Both studies evaluated the most potent DLIR-H, hoping to show that DLIR algorithms were superior to standard ASiR methods when aiming for aggressive dose reduction. However, previous studies did not evaluate the diagnostic performance for focal liver lesions, which are the main hurdle for dose reduction due to the innate low contrast between the liver parenchyma and focal liver lesions. Our results suggest that the detection of focal liver lesions does not deteriorate with approximately 35.1% dose reduction using DLIR-M. Further studies to identify aggressive dose reduction protocols that do not compromise diagnostic accuracy are needed to validate our findings.

Our study had several limitations. First, this study was conducted retrospectively at a single institution. Second, there was a time difference between LDCT and SDCT, which might have led to biological differences in patients. However, body profiles, such as weight and effective diameter, which are known to significantly influence image noise, did not differ between the two CT scans in this study. Hence, any possible differences due to the time gap may have been minimized. Third, since the qualitative analysis was conducted in a side-to-side manner and not fully blinded, there was a concern that bias may have been introduced into the results. Fourth, only the medium strength of DLIR was evaluated comparatively in qualitative image analysis and lesion detectability assessment. The effects of other strengths of DLIR on qualitative measures and lesion detectability should be evaluated in further

https://doi.org/10.3348/kjr.2021.0683

Korean Journal of Radiology
KJR

studies. Fifth, we only evaluated the detectability of hypoattenuating liver lesions. This may not reflect the actual clinical situation of abdominal CT evaluations. Therefore, it is mandatory to conduct a future analysis that includes hypervascular and extrahepatic lesions.

In conclusion, the overall image quality and detectability of focal liver lesions could be preserved in contrast-enhanced abdominopelvic LDCT obtained with DLIR-M compared to those of SDCT with h-IR images.

Supplement

The Supplement is available with this article at https://doi.org/10.3348/kjr.2021.0683.

Availability of Data and Material

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

Conflicts of Interest

Yong Eun Chung who is on the editorial board of the *Korean Journal of Radiology* was not involved in the editorial evaluation or decision to publish this article. All remaining authors have declared no conflicts of interest.

Author Contributions

Conceptualization: June Park, Yong Eun Chung. Data curation: June Park, Yong Eun Chung. Formal analysis: June Park, Jaeseung Shin, Yong Eun Chung. Investigation: all authors. Methodology: all authors. Supervision: Yeo-Eun Kim, Yong Eun Chung. Validation: June Park, Jaeseung Shin, Yong Eun Chung. Visualization: June Park, Yong Eun Chung. Writing—original draft: all authors. Writing—review & editing: all authors.

ORCID iDs

June Park
https://orcid.org/0000-0003-1131-9730
Jaeseung Shin
https://orcid.org/0000-0002-6755-4732
In Kyung Min
https://orcid.org/0000-0002-8070-0116
Heejin Bae
https://orcid.org/0000-0002-1227-8646
Yeo-Eun Kim
https://orcid.org/0000-0002-6876-8871
Yong Eun Chung
https://orcid.org/0000-0003-0811-9578

Funding Statement

None

REFERENCES

1. Huang L, Snyder AR, Morgan WF. Radiation-induced genomic instability and its implications for radiation carcinogenesis. *Oncogene* 2003;22:5848-5854
2. Vaiserman A, Koliada A, Zabuga O, Socol Y. Health impacts of low-dose ionizing radiation: current scientific debates and regulatory issues. *Dose Response* 2018;16:1559325818796331
3. Sodickson A, Baeyens PF, Andriole KP, Prevedello LM, Nawfel RD, Hanson R, et al. Recurrent CT, cumulative radiation exposure, and associated radiation-induced cancer risks from CT of adults. *Radiology* 2009;251:175-184
4. Mullenders L, Atkinson M, Paretzke H, Sabatier L, Bouffler S. Assessing cancer risks of low-dose radiation. *Nat Rev Cancer* 2009;9:596-604
5. National Research Council. *Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2*. Washington DC: National Academies Press, 2006:406
6. Brenner DJ, Hall EJ. Computed tomography--an increasing source of radiation exposure. *N Engl J Med* 2007;357:2277-2284
7. Yu L, Liu X, Leng S, Kofler JM, Ramirez-Giraldo JC, Qu M, et al. Radiation dose reduction in computed tomography: techniques and future perspective. *Imaging Med* 2009;1:65-84
8. Solomon J, Marin D, Roy Choudhury K, Patel B, Samei E. Effect of radiation dose reduction and reconstruction algorithm on image noise, contrast, resolution, and detectability of subtle hypoattenuating liver lesions at multidetector CT: filtered back projection versus a commercial model–based iterative reconstruction algorithm. *Radiology* 2017;284:777-787
9. Patino M, Fuentes JM, Singh S, Hahn PF, Sahani DV. Iterative reconstruction techniques in abdominopelvic CT: technical concepts and clinical implementation. *AJR Am J Roentgenol* 2015;205:W19-W31
10. Singh S, Kalra MK, Hsieh J, Licato PE, Do S, Pien HH, et al. Abdominal CT: comparison of adaptive statistical iterative and filtered back projection reconstruction techniques. *Radiology* 2010;257:373-383
11. Sagara Y, Hara AK, Pavlicek W, Silva AC, Paden RG, Wu Q. Abdominal CT: comparison of low-dose CT with adaptive statistical iterative reconstruction and routine-dose CT with filtered back projection in 53 patients. *AJR Am J Roentgenol* 2010;195:713-719
12. Lim K, Kwon H, Cho J, Oh J, Yoon S, Kang M, et al. Initial

phantom study comparing image quality in computed tomography using adaptive statistical iterative reconstruction and new adaptive statistical iterative reconstruction V. *J Comput Assist Tomogr* 2015;39:443-448
13. Kwon H, Cho J, Oh J, Kim D, Cho J, Kim S, et al. The adaptive statistical iterative reconstruction-V technique for radiation dose reduction in abdominal CT: comparison with the adaptive statistical iterative reconstruction technique. *Br J Radiol* 2015;88:20150463
14. Ehman EC, Yu L, Manduca A, Hara AK, Shiung MM, Jondal D, et al. Methods for clinical evaluation of noise reduction techniques in abdominopelvic CT. *Radiographics* 2014;34:849-862
15. Mileto A, Guimaraes LS, McCollough CH, Fletcher JG, Yu L. State of the art in abdominal CT: the limits of iterative reconstruction algorithms. *Radiology* 2019;293:491-503
16. Schindera ST, Odedra D, Raza SA, Kim TK, Jang HJ, Szucs-Farkas Z, et al. Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? *Radiology* 2013;269:511-518
17. McCollough CH, Yu L, Kofler JM, Leng S, Zhang Y, Li Z, et al. Degradation of CT low-contrast spatial resolution due to the use of iterative reconstruction and reduced dose levels. *Radiology* 2015;276:499-506
18. Hsieh J, Liu E, Nett B, Tang J, Thibault JB, Sahney S. *A new era of image reconstruction: TrueFidelity™. Technical white paper on deep learning image reconstruction*. Milwaukee: GE Healthcare, 2019
19. Racine D, Becce F, Viry A, Monnin P, Thomsen B, Verdun FR, 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
20. Greffier J, Hamard A, Pereira F, Barrau C, Pasquier H, Beregi JP, et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. *Eur Radiol* 2020;30:3951-3959
21. Jensen CT, Liu X, Tamm EP, Chandler AG, Sun J, Morani AC, et al. Image quality assessment of abdominal CT by use of new deep learning image reconstruction: initial experience. *AJR Am J Roentgenol* 2020;215:50-57
22. Noda Y, Iritani Y, Kawai N, Miyoshi T, Ishihara T, Hyodo F, et al. Deep learning image reconstruction for pancreatic low-dose computed tomography: comparison with hybrid iterative reconstruction. *Abdom Radiol (NY)* 2021;46:4238-4244
23. 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;31:5533-5543
24. Cao L, Liu X, Li J, Qu T, Chen L, Cheng Y, et al. A study of using a deep learning image reconstruction to improve the image quality of extremely low-dose contrast-enhanced abdominal CT for patients with hepatic lesions. *Br J Radiol* 2021;94:20201086
25. McCollough CH, Bruesewitz MR, Kofler JM Jr. CT dose reduction and dose management tools: overview of available options. *Radiographics* 2006;26:503-512
26. American Association of Physicists in Medicine. Size-specific dose estimates (SSDE) in pediatric and adult body CT examinations. AAPM.org Web site. https://www.aapm.org/pubs/reports/RPT_204.pdf. Accessed June 2, 2021
27. Park C, Choo KS, Jung Y, Jeong HS, Hwang JY, Yun MS. CT iterative vs deep learning reconstruction: comparison of noise and sharpness. *Eur Radiol* 2021;31:3156-3164
28. Ahn S, Park SH, Lee KH. How to demonstrate similarity by using noninferiority and equivalence statistical testing in radiology research. *Radiology* 2013;267:328-338
29. Nakamura Y, Narita K, Higaki T, Akagi M, Honda Y, Awai K. Diagnostic value of deep learning reconstruction for radiation dose reduction at abdominal ultra-high-resolution CT. *Eur Radiol* 2021;31:4700-4709
30. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Awai K. Deep learning reconstruction of equilibrium phase CT images in obese patients. *Eur J Radiol* 2020;133:109349
31. Chakraborty DP, Berbaum KS. Observer studies involving detection and localization: modeling, analysis, and validation. *Med Phys* 2004;31:2313-2330
32. Dorfman DD, Berbaum KS, Metz CE. Receiver operating characteristic rating analysis. Generalization to the population of readers and patients with the jackknife method. *Invest Radiol* 1992;27:723-731
33. Brennan RL, Prediger DJ. Coefficient kappa: some uses, misuses, and alternatives. *Educ Psychol Meas* 1981;41:687-699
34. 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;47:3961-3971
35. Chen LH, Jin C, Li JY, Wang GL, Jia YJ, Duan HF, et al. Image quality comparison of two adaptive statistical iterative reconstruction (ASiR, ASiR-V) algorithms and filtered back projection in routine liver CT. *Br J Radiol* 2018;91:20170655
36. International Commission on Radiation Units and Measurements. ICRU report no. 87: radiation dose and image-quality assessment in computed tomography. *J ICRU* 2012;12:1-149

https://doi.org/10.3348/kjr.2021.0683