BJR https://doi.org/10.1259/bjr.20201357

Received: 19 November 2020 | Revised: 19 April 2021 | Accepted: 06 May 2021

Cite this article as:
Tamura A, Mukaida E, Ota Y, Kamata M, Abe S, Yoshioka K. Superior objective and subjective image quality of deep learning reconstruction for low-dose abdominal CT imaging in comparison with model-based iterative reconstruction and filtered back projection. *Br J Radiol* 2021; **94**: 20201357.

FULL PAPER

Superior objective and subjective image quality of deep learning reconstruction for low-dose abdominal CT imaging in comparison with model-based iterative reconstruction and filtered back projection

[1]AKIO TAMURA, MD, [1]EISUKE MUKAIDA, MD, [2]YOSHITAKA OTA, RT, [2]MASAYOSHI KAMATA, RT, [2]SHUN ABE, RT and [1]KUNIHIRO YOSHIOKA, MD

[1]Department of Radiology, Iwate Medical University School of Medicine, Iwate, Japan
[2]Division of Central Radiology, Iwate Medical University Hospital, Iwate, Japan

Address correspondence to: Dr Akio Tamura
E-mail: *a.akahane@gmail.com*

Objective: This study aimed to conduct objective and subjective comparisons of image quality among abdominal computed tomography (CT) reconstructions with deep learning reconstruction (DLR) algorithms, model-based iterative reconstruction (MBIR), and filtered back projection (FBP).
Methods: Datasets from consecutive patients who underwent low-dose liver CT were retrospectively identified. Images were reconstructed using DLR, MBIR, and FBP. Mean image noise and contrast-to-noise ratio (CNR) were calculated, and noise, artifacts, sharpness, and overall image quality were subjectively assessed. Dunnett's test was used for statistical comparisons.
Results: Ninety patients (67 ± 12.7 years; 63 males; mean body mass index [BMI], 25.5 kg/m^2) were included. The mean noise in the abdominal aorta and hepatic parenchyma of DLR was lower than that in FBP and MBIR (p < .001). For FBP and MBIR, image noise was significantly higher for obese patients than for those with normal BMI. The CNR for the abdominal aorta and hepatic parenchyma was higher for DLR than for FBP and MBIR (p < .001). MBIR images were subjectively rated as superior to FBP images in terms of noise, artifacts, sharpness, and overall quality (p < .001). DLR images were rated as superior to MBIR images in terms of noise (p < .001) and overall quality (p = .03).
Conclusions: Based on objective and subjective comparisons, the image quality of DLR was found to be superior to that of MBIR and FBP on low-dose abdominal CT. DLR was the only method for which image noise was not higher for obese patients than for those with a normal BMI.
Advances in knowledge: This study provides previously unavailable information on the properties of DLR systems and their clinical utility.

INTRODUCTION

CT imaging has considerable diagnostic utility; however, due to the widespread use of CT and the development of advanced CT systems that enable scanning over wide areas and multiphase imaging, concerns remain regarding the increasing exposure of patients to radiation.[1,2] According to the Biologic Effects of Ionizing Radiation VII report, CT scanning can increase the lifetime risk of cancer by 0.34 to 1.30%, depending on age, sex, and site of examination.[3] In response to concerns about radiation exposure during CT scanning, the radiological community has been working to develop low-dose imaging techniques. This has been challenging because with filtered back projection (FBP), which is a widely used method for CT image reconstruction, image noise increases as the radiation dose decreases, making it difficult to assess structures in detail.

To solve this problem, CT vendors are developing technologies to improve image quality at a lower radiation dose.[4–14] Iterative reconstruction techniques are classified as hybrid and model based. While the former features more widespread use, its performance can still be improved—especially when compared against the superior image noise, image texture, and spatial resolution yielded by model-based iterative reconstruction (MBIR). Although time-consuming when reconstructing images, MBIR is not an

analytical-based, simple solution like the filtered back projection, which has a very high spatial resolution and a high capacity for noise reduction.[8,9,15] Importantly, this technique enables images of diagnostic quality to be generated at a low radiation dose; MBIR supports a reduction in radiation exposure of 22–70% in abdominal CT imaging.[9,13] However, a disadvantage of MBIR is that image reconstruction is a time-consuming process.

In recent years, artificial intelligence, especially deep learning, has attracted considerable attention, and multiple applications have been developed across various fields. One of the most advanced applications of deep learning is the improvement of image quality through noise reduction.[16–18] In the field of CT image reconstruction, new deep learning reconstruction (DLR) algorithms have recently been introduced. The Advanced Intelligent Clear-IQ Engine (AiCE), the first commercially available DLR tool developed for CT, uses deep convolutional neural networks to distinguish the true signal from the noise in the image.[19] For this deep learning approach, a high-dose MBIR image and image data corresponding to 12.5–75% of the dose of the target image are provided as training pairs, and statistical features that distinguish the signal from noise and artifacts are learned in the training process.[20] While DLR has great potential, there are still insufficient data available on the properties of DLR systems and their clinical utility. The purpose of this study was to perform objective and subjective comparisons of the use of FBP, vendor-specific MBIR (Forward Projected Model-Based Iterative Reconstruction Solution [FIRST]), and DLR (AiCE) systems for the reconstruction of low-dose abdominal CT scans.

METHODS

Patients and study design

Patients were included in the study if they were aged 18 years or older. Exclusion criteria were contraindications to iodinated contrast media, emergency cases, and hemodialysis or renal failure. The study design was approved by the Institutional Review Board, and the requirement for informed consent was waived due to the retrospective nature of the study.

CT technique

All scans were performed using a 320 row multidetector CT scanner (Aquilion ONE PRISM; Canon Medical Systems, Otawara, Japan) with the following parameters: 80 row × 0.5 mm, pitch of 0.813, 0.6 s rotation time (fixed), 120 kVp tube voltage, and automatic exposure control with a noise index of 20. The standard care of abdominal CT in our hospital is the MBIR protocol (FIRST BODY Mild), which is designed for low exposure, with a mean volume CT dose index (CTDIvol) of 7.9 mGy. This value was obtained by retrospectively evaluating the doses of 125 consecutive patients prior to the start of this study. Iohexol (Omnipaque 300; Daiichi-Sankyo, Tokyo, Japan), iomeprol (Iomeron 350; Eisai, Tokyo, Japan), or iopamidol (Iopamiron 370; Nihon Schering, Osaka, Japan) was administered via an antecubital vein using a 20-gauge catheter at a dose of 600 mgI/kg; the injection procedure was standard across all patients. The scan delays for arterial and portal venous phase imaging were determined using an automatic bolus-tracking program (Canon Medical Systems). The region of interest (ROI) cursor was placed in the aorta at the level of the diaphragmatic dome, and scanning for the arterial and portal venous phases began automatically at 20 and 60 s, respectively, after the trigger threshold of 100 Hounsfield Units was reached.

CT image reconstruction

Images were reconstructed from raw data using FBP (FC13), MBIR (FIRST BODY Mild; reconstruction time, 7 min), and DLR (AiCE BODY Mild; reconstruction time, 1 min) algorithms. Reconstructed images were created with soft tissue window settings, with a 3-mm slice thickness and interval. FBP, FIRST, and AiCE images were obtained for each of the 90 patients.

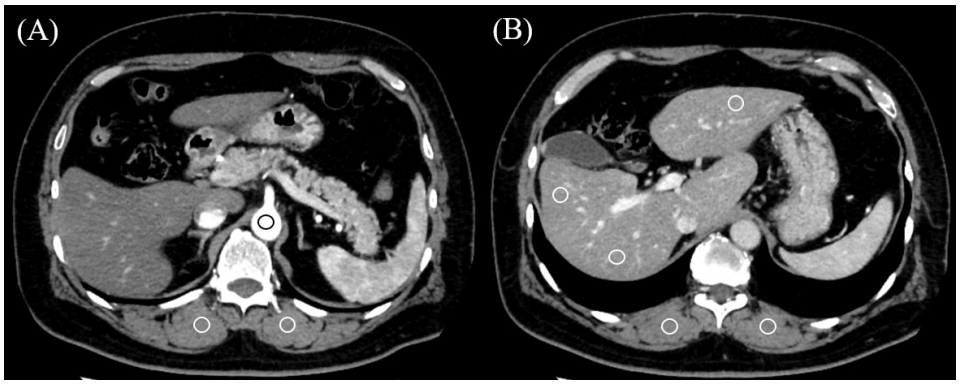
Analysis of image quality

Objective assessment: One CT technologist (M.K., with 30 years' experience in abdominal CT) measured the mean attenuation of the abdominal aorta, hepatic parenchyma, and bilateral erector spinae muscles within circular ROIs for each patient. The area of each ROI was set at approximately 150 ± 25 mm^2. We standardized the size of the ROIs using the copy-and-paste function of the workstation (Ziostation2; Ziosoft, Tokyo, Japan). Aortic attenuation in the arterial phase was measured immediately proximal to the level of the celiac trunk. Hepatic attenuation was measured in three separate areas (the left lobe and the anterior and posterior segments of the right lobe) at the level of the main portal vein in the portal venous phase. Areas that included focal changes in parenchymal density, large vessels, and prominent artifacts were avoided. Attenuation of the two erector spinae muscles was measured in the arterial and portal venous phases, and areas of macroscopic fat infiltration were avoided. Image noise was defined as the standard deviation of the attenuation value of the abdominal aorta and hepatic parenchyma. The contrast-to-noise ratio (CNR) of the abdominal aorta was calculated by subtracting the mean attenuation of the ROI of the erector spinae muscle from that of the ROI of the aorta and dividing this difference by the image noise of the erector spinae muscle. The CNR of the hepatic parenchyma was calculated similarly, using the ROI of the liver rather than the aorta (Figure 1).

To evaluate the modulation transfer function (MTF), image data were acquired from the CTP404 module of the Catphan 504 phantom (The Phantom Laboratory, Inc., Salem, NY) using a sensitometry target of Delrin®, a setting of 80 × 0.5 mm collimation, pitch of 0.813, 0.6 s rotation time, and 120 kVp/60 mA. Images reconstructed with FBP (FC13), FIRST BODY Mild, and AiCE BODY Mild, were acquired with a 3-mm slice thickness and interval for measuring the MTF. The MTF was calculated for each dataset via the radial edge method using a CT measure software (version 0.97b; Japanese Society of CT Technology, Hiroshima, Japan).

Subjective assessment: The three image sets were presented randomly to two blinded radiologists (M.N. and A.T.: 4 and 13 years' experience in abdominal CT, respectively). CT images were viewed with standard abdominal window settings (width, 300 Hounsfield Units; level, 50 Hounsfield Units). The radiologists independently graded the quality of portal phase axial CT images using a 5-point scale for image noise and artifacts,

Figure 1. Axial contrast-enhanced CT images obtained in a 57-year-old male during the arterial phase (a) show regions of interest (ROIs) manually drawn on aorta and bilateral erector spinae muscles; the portal phase (b) shows ROI drawn on liver and bilateral erector spinae muscles.



sharpness, and overall image quality, as follows: Image noise and artifacts: 1, unacceptable; 2, major; 3, minor; 4, very little; and 5, none. Sharpness, defined as the clarity of the liver contour: 1, blurry; 2, poorer than average; 3, average; 4, better than average; and 5, very sharp. Overall image quality: 1, unacceptable; 2, suboptimal; 3, average; 4, above average; and 5, excellent. Before image quality assessment, a training session including abdominal CT images from 10 patients was provided to familiarize the radiologists with the scoring system.

Statistical analysis

All numerical values are reported as mean ± standard deviation. We used Dunnett's test to compare quantitative and qualitative image analysis data. To evaluate the association between BMI and image noise or CNR in the portal phase, we categorized patients into four subgroups based on the World Health Organization classification[21]: BMI <18.5 kg/m^2, underweight (n = 1); 18.5 kg/m^2 ≤ BMI<25 kg/m^2, normal range (n = 45); 25 kg/m^2 ≤ BMI<30 kg/m^2, pre-obese (n = 33); and ≥30 kg/m^2, obese (n = 11). The weighted κ score was calculated to determine the interobserver agreement and was classified as follows: <0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost perfect or perfect agreement.[22] Statistical analysis was performed using the SPSS software (version 24; IBM Corp., Armonk, NY). A <i>p value < 0.05 was considered statistically significant.

RESULTS

Patients

Raw CT image data were obtained from 90 patients (63 males; age range, 19–89 years; mean age, 67 years; body weight range, 41–155 kg; mean body weight, 68.4 kg; height range, 145–185 cm; mean height, 163.5 cm; body mass index [BMI] range, 17.8–51.8 kg/m^2; mean BMI, 25.5 kg/m^2) who underwent liver CT (volume CT dose index range, 3.1–29.4 mGy; mean volume CT dose index, 8.2 mGy; dose-length product range, 78–1065 mGy-cm; mean dose-length product, 247.9 mGy-cm).

Objective assessment of image quality

The mean attenuation, image noise, and CNR for images reconstructed with each algorithm are presented in Table 1. There were no significant differences in the mean attenuation values measured in the aorta and hepatic parenchyma among the reconstruction algorithms (p = .90–.97). However, noise levels were significantly different among the three algorithms. Overall, the mean image noise measured in the abdominal aorta and hepatic parenchyma for AiCE images (11.1 ± 2.5 and 10.0 ± 0.9, respectively) was significantly lower than for FBP images (27.7 ± 5.0 and 23.1 ± 2.6, respectively) and FIRST images (16.9 ± 2.9 and 14.7 ± 1.6, respectively) (p < .001 for all comparisons). The CNR for the abdominal aorta and hepatic parenchyma was significantly higher for AiCE images (26.6 ± 6.7 and 4.0 ± 1.6, respectively) than for FBP (13.3 ± 3.0 and 2.1 ± 0.8, respectively) and FIRST images (18.9 ± 4.9 and 2.9 ± 1.1, respectively) (p < .001 for all comparisons). Table 2 shows the mean image noise and CNR for the three reconstruction algorithms when patients were classified according to the BMI group. Since there was only one underweight patient, we opted not to include this patient's data in the statistical analysis. For both FBP and FIRST, image noise was significantly higher in obese patients than in those with normal BMI (p = .001 and p = .008, respectively), but this effect was not observed for AiCE (p = .13). For all three reconstruction methods, the CNR was significantly lower in obese than in pre-obese patients, and in the pre-obese group than in the normal BMI group. MTF data for each algorithm are presented in Figure 2. The spatial resolution of images reconstructed with FIRST and AiCE was consistently higher than that of those reconstructed with FBP.

Subjective assessment of image quality

Table 3 shows the results of the subjective image quality assessments. On the 5-point scales, FIRST images received significantly higher scores than FBP images regarding noise, artifacts, sharpness, and overall image quality (p < .001). AiCE images received significantly higher scores than FIRST images with respect to noise (p < .001) and overall image quality (p = .03). Figure 3 shows an example of the superior noise reduction with

Table 1. Quantitative assessment of image quality for the sets of images reconstructed using FBP, FIRST, and AiCE

Parameter	FBP	FIRST	AiCE	FBP *vs* FIRST *p* value	FIRST *vs* AiCE *p* value
Mean CT number					
Aorta	340.1 ± 55.1	337.8 ± 55.5	336.1 ± 56.2	0.90	0.97
Hepatic parenchyma	105.1 ± 16.4	104.6 ± 16.4	103.8 ± 16.3	0.95	0.96
Image noise					
Aorta	27.7 ± 5.0	16.9 ± 2.9	11.1 ± 2.5	<0.001	<0.001
Hepatic parenchyma	23.1 ± 2.6	14.7 ± 1.6	10.0 ± 0.9	<0.001	<0.001
CNR					
Aorta	13.3 ± 3.0	18.9 ± 4.9	26.6 ± 6.7	<0.001	<0.001
Hepatic parenchyma	2.1 ± 0.8	2.9 ± 1.1	4.0 ± 1.6	<0.001	<0.001

AiCE, Advanced intelligent clear-IQ engine; CNR, Contrast-to-noise ratio; FBP, Filtered back projection; FIRST, Forward projected model-based iterative reconstruction solution.
Data are presented as mean values ± standard deviation. Dunnett's test was used for statistical comparisons. The mean volume CT dose index was 8.2 mGy.

AiCE reconstruction. In contrast to the noise and overall quality ratings, artifact scores were significantly higher for FIRST images than for AiCE images ($p < .001$). In one instance only, both readers scored the FBP image higher than the corresponding FIRST image in terms of artifacts (Figure 4). There was no significant difference between FIRST and AiCE scores in terms of the sharpness of the liver contour. Between the two readers, there was substantial to almost perfect interobserver agreement for all subjective image quality metrics with $\kappa = 0.86$, 0.90, 0.81, and 0.88 for noise, artifacts, sharpness, and overall image quality, respectively.

DISCUSSION

The aim of this study was to perform both objective and subjective comparisons of image quality of low-dose abdominal CT images using DLR, MBIR, and FBP. Lower noise and a higher CNR were observed with deep learning than with the other two algorithms ($p < .001$). Furthermore, DLR was the only method for which image noise was not higher for obese patients than for those with a normal BMI.

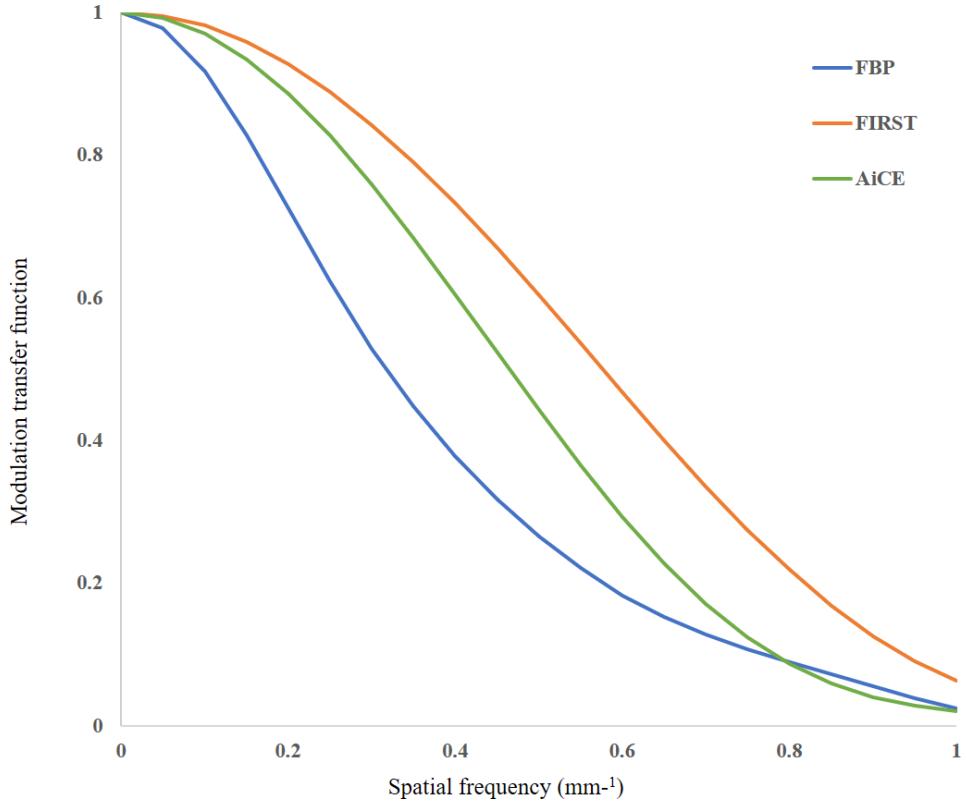
This study found that the noise reduction capability and low-contrast resolution of AiCE was superior to FIRST, and that the objective image noise measurements corresponded well with subjective image assessments. Akagi et al reported a comparison of DLR, hybrid-IR, and MBIR in ultra-high-definition CT of the abdomen and found significant noise reduction and CNR improvement in DLR.[22] Our results are similar to their report, but importantly, with AiCE, we further demonstrated that the image noise did not differ among the BMI groups. This is of particular clinical relevance because reducing noise in the CT images of patients with obesity can be challenging. Several techniques can improve the quality of the CT images of obese patients; BMI-adaptive scan protocols typically use higher kilovoltage acquisition and dual-source CT to reduce noise.[23,24] Unfortunately, both of these techniques involve a markedly increased radiation dose.

Table 2. Variation in image noise and CNR in the hepatic parenchyma with respect to BMI category

Parameter	Normal BMI ($n = 45$)	Pre-obese ($n = 33$)	Obese ($n = 11$)	Obese *vs* normal *p* value	Obese *vs* pre-obese *p* value
Image noise					
FBP	22.5 ± 1.9	23.2 ± 1.5	25.6 ± 5.5	0.001	0.01
FIRST	14.3 ± 1.5	14.9 ± 1.5	15.8 ± 1.5	0.008	0.16
AiCE	9.9 ± 0.8	10.0 ± 0.8	10.4 ± 1.2	0.13	0.18
CNR					
FBP	2.2 ± 0.7	2.2 ± 0.8	1.4 ± 0.8	0.002	0.002
FIRST	3.1 ± 0.9	3.0 ± 1.1	1.7 ± 1.3	<0.001	0.001
AiCE	4.3 ± 1.2	4.0 ± 1.6	2.4 ± 2.0	<0.001	0.006

AiCE, Advanced intelligent clear-IQ engine; BMI, body mass index; CNR, contrast-to-noise ratio; FBP, filtered back projection; FIRST, Forward projected model-based Iterative reconstruction solution.
Data are presented as mean values ± standard deviation. Dunnett's test was used for statistical comparisons. The mean volume CT dose index was 5.4 mGy, 9.4 mGy, and 16.5 mGy for the normal BMI, pre-obese, and obese groups, respectively.

Figure 2. Modulation transfer functions for CT images reconstructed with FBP, FIRST, and AiCE. Compared with FBP images, FIRST and AiCE images had a higher spatial resolution. AiCE, Advanced intelligent clear-IQ engine; FBP, Filtered back projection; FIRST, Forward projected model-based iterative reconstruction solution.



In abdominal CT images, high image noise may obscure subtle low-contrast lesions in parenchymal organs,[25,26] increasing the risk of missing liver tumors.[27–30] In this regard, our observation that DLR, in contrast to MBIR and FBP, was able to maintain low noise levels even in patients with a higher BMI is encouraging. Further investigation of the ability of AiCE to support the diagnosis of liver lesions in obese patients is recommended.

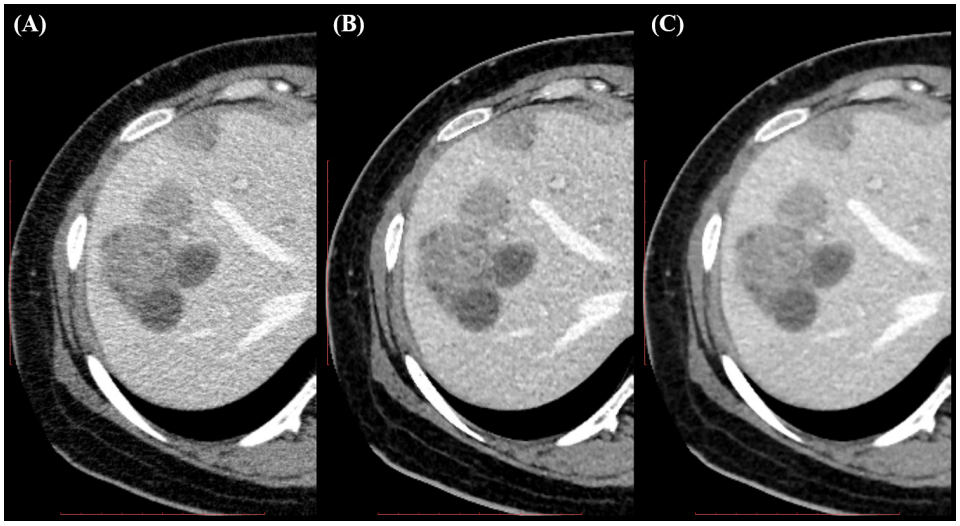
An analysis of the subjective results revealed a lower level of artifacts on FIRST and AiCE images than on FBP images. FIRST performed better than AiCE regarding artifacts in general. MBIR models the optics and geometry of the CT system, including the response of the X-ray tube and detector, and incorporates the physical properties of scattering and crosstalk.[31] This modeling reduces image noise and artifacts and improves spatial resolution. In contrast, it is important to note that the accuracy of DLR reproduction is highly dependent on the training data set that the algorithm uses to train the model. In the AiCE, high-quality MBIR images and low-dose images are used as teacher data, but at this point, it is thought that the network of deep learning is being constructed with priority given to noise removal rather than artifact improvement. Simulation-based studies have shown that such algorithms can lead to serious misdiagnosis and incorrect addition of anatomical features when tested against lesions or anatomical textures that are not present in the training dataset.[32] In the present clinical study, no such cases were found in AiCE,

Table 3. Subjective assessment of image quality for the sets of images reconstructed using FBP, FIRST, and AiCE

Parameter	FBP	FIRST	AiCE	FBP *vs* FIRST *p* value	FIRST *vs* AiCE *p* value
Noise	2.5 ± 0.8	3.5 ± 0.7	4.0 ± 0.6	<0.001	<0.001
Artifacts	2.4 ± 0.6	4.3 ± 0.6	3.8 ± 0.5	<0.001	<0.001
Sharpness	2.6 ± 0.7	3.9 ± 0.7	4.0 ± 0.6	<0.001	0.60
Overall image quality	2.5 ± 0.6	3.8 ± 0.6	3.9 ± 0.5	<0.001	0.03

AiCE, Advanced intelligent clear-IQ engine; FBP, Filtered back projection; FIRST, Forward projected model-based iterative reconstruction solution. Data are presented as mean values ± standard deviation. Dunnett's test was used for statistical comparisons. κ = 0.86, 0.90, 0.81, and 0.88 for noise, artifacts, sharpness, and overall image quality, respectively.

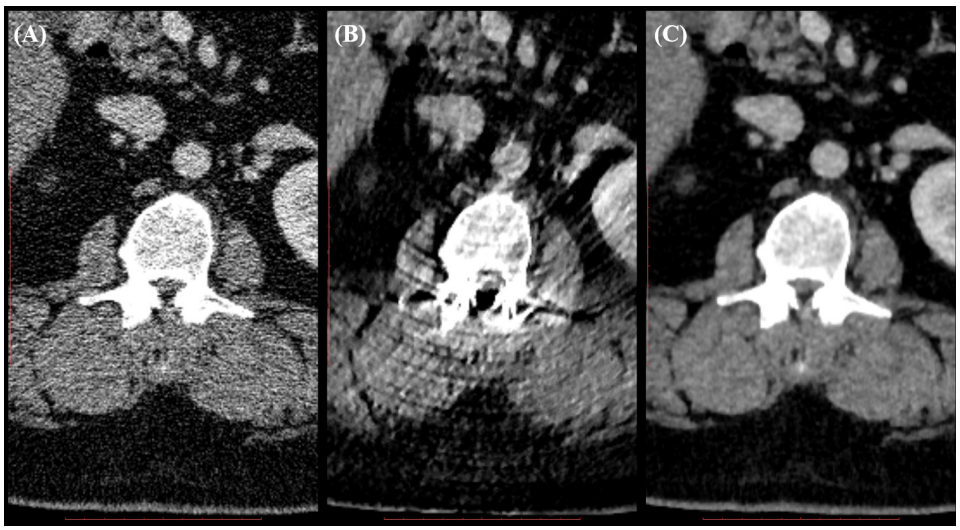
Figure 3. Axial contrast-enhanced abdominal CT images obtained in a 38-year-old female with liver metastases from gastric cancer. Images were obtained with a preset soft tissue window (width, 300 HU; level, 50 HU). Differences in image quality can be observed between (a) FBP, (b) FIRST, and (c) AiCE reconstructions of the same anatomical location. The noise level is appreciably lower in (c). AiCE, Advanced intelligent clear-IQ engine; FBP, Filtered back projection; FIRST, Forward projected model-based iterative reconstruction solution.



but very strong augmentation ring artifacts were found in the FIRST images of one patient. This artifact was not observed on the AiCE images of the same patient. The cause of this artifact remains uncertain, but a possible explanation could be that the noise model in the FIRST algorithm misrecognized the artifact as a signal and, therefore, performed an inappropriate iteration. Gaddikeri et al reported strong artifacts due to thyroid shielding in head and neck CT images reconstructed with MBIR, but not with hybrid iterative reconstruction, and suggested that these artifacts are partly due to the MBIR noise model.[33]

Subjective scores regarding the clarity of the liver contour were high for both FIRST and AiCE images. This may reflect the MTF data obtained from the phantom experiment, which showed that both the FIRST and AiCE have a higher spatial resolution than FBP. The spatial resolution of FIRST images was slightly but consistently higher than that of AiCE images at all spatial frequency values. However, there was no appreciable difference between the two on visual assessment.

Figure 4. A 42-year-old male with chronic hepatitis. Axial contrast-enhanced abdominal CT images of the same anatomical location were reconstructed with (a) FBP, (b) FIRST, and (c) AiCE. A strong augmented ring artifact is visible in the image reconstructed using FIRST, but not in those reconstructed using FBP or AiCE. AiCE, Advanced intelligent clear-IQ engine; FBP, Filtered back projection; FIRST, Forward projected model-based iterative reconstruction solution.



The clinical benefits of CT are clear, and the number of scans conducted is increasing every year.[1] However, there are concerns regarding radiation-induced malignancies resulting from CT use. In the last decade, the development of hybrid iterative reconstruction and MBIR has made it possible to achieve a substantial dose reduction.[4–14,34] In this study, overall abdominal CT image quality was higher for AiCE than FIRST images; therefore, the use of AiCE may potentially allow for further radiation dose reductions while maintaining the same image quality. Several noise reduction algorithms have been developed with the aim of reducing radiation dose; these have typically been applied to CT images obtained using routine tube voltage settings (120 kV). In contrast-enhanced CT, the use of low tube voltage imaging in combination with various noise reduction methods can reduce exposure without compromising image quality or contrast.[6,7,35,36] However, owing to increased noise and artifacts, these techniques have a limited use in overweight patients.[6,7,37] While BMI-adaptive scan protocols have been developed to improve image quality in obese patients, these typically use higher kilovoltage acquisition and dual-source CT.[23,38] Unfortunately, both of these techniques involve a markedly high radiation dose. Since AiCE allows for noise reduction even in patients with a BMI ≥ 25 kg/m^2, the use of AiCE may potentially support adequate noise reduction in CT images of overweight patients, even when using low tube voltages. The reconstruction times with MBIR were originally between 15 min and 1 h, but due to recent technological advances, reconstruction times are now more appropriate for a clinical context.[8,9,15] Since AiCE reconstruction takes only approximately 1 min and as such takes only slightly more time than FBP, AiCE may be more appropriate than FIRST for clinical use, particularly given its superior noise reduction, CNR, and overall image quality compared with the latter.

There are several limitations of this study. The main limitation is the fact that because the AiCE is vendor specific, it might not be possible to generalize the findings to other vendors. Recently, GE has also developed the DRL algorithm "TrueFidelity" and has reported on their image evaluation. Jensen et al reported that TrueFidelity improves the image quality through noise reduction and decreased CNR without altering the texture on abdominal CT.[39] DLR seems to be a promising technology, but the methods still constitute a developing field of technology; hence, this study could be considered preliminary. Second, the mean BMI in our study was lower than the global average.[40] Third, we did not compare the reconstruction algorithms with respect to diagnostic accuracy because the number of patients with liver tumors was low. Future studies are needed to evaluate the use of this DLR algorithm in the diagnosis of liver tumors. In addition, we used CNR to evaluate the low-contrast resolution, but CNR may not always correlate with visual evaluation. A promising way to evaluate the dose reduction capability of IR algorithms is to use computational models of human observers. These models reproduce the response of the human visual system to the image system and, unlike CNR, calculate a detectability index (d′) that incorporates the attributes of image resolution and noise texture.[41]

In summary, our data suggest that images reconstructed with the AiCE, a novel deep learning CT reconstruction method, are of considerably higher quality than images reconstructed with either MBIR or traditional FBP. Considering the excellent performance of this novel algorithm observed during phantom testing as well as on abdominal CT in clinical patients regardless of BMI, DLR appears to be promising for low-dose CT applications. Further studies are required to develop reliable low-dose CT protocols for routine clinical use and to systematically analyze the performance of the AiCE in the diagnosis of various conditions.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Kazumasa Arakita and Mr. Shinsuke Tsukagoshi from Canon Medical Systems for their technical expertise. We would like to thank Editage for English language editing.

FUNDING

This work was supported by JSPS KAKENHI (Grant Number 19K08160) and the Regional medical research of Iwate.

REFERENCES

1. Smith-Bindman R, Miglioretti DL, Johnson E. Use of diagnostic imaging studies and associated radiation exposure for patients enrolled in large integrated healthcare systems, 1996–2010. *JAMA* 2002; **307**: 2400–9.
2. McCollough CH, Primak AN, Braun N, Kofler J, Yu L, Christner J. Strategies for reducing radiation dose in CT. *Radiol Clin North Am* 2009; **47**: 27–40. doi: https://doi.org/10.1016/j.rcl.2008.10.006
3. National Research Council.*Health risks from exposure to low levels of ionizing radiation: BEIR VII Phase 2*. Washington, DC: The National Academies Press; 2006.
4. Hara AK, Wellnitz CV, Paden RG, Pavlicek W, Sahani DV. Reducing body CT radiation dose: beyond just changing the numbers. *AJR Am J Roentgenol* 2013; **201**: 33–40. doi: https://doi.org/10.2214/AJR.13.10556
5. Schindera ST, Odedra D, Raza SA, Kim TK, Jang H-J, Szucs-Farkas Z, et al. Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? *Radiology* 2013; **269**: 511–8. doi: https://doi.org/10.1148/radiol.13122349
6. Nakaura T, Awai K, Maruyama N, Takata N, Yoshinaka I, Harada K, et al. Abdominal dynamic CT in patients with renal dysfunction: contrast agent dose reduction with low tube voltage and high tube current-time product settings at 256-detector row CT. *Radiology* 2011; **261**: 467–76. doi: https://doi.org/10.1148/radiol.11110021
7. Nakaura T, Nakamura S, Maruyama N, Funama Y, Awai K, Harada K, et al. Low contrast agent and radiation dose protocol for hepatic dynamic CT of thin adults at 256-detector row CT: effect of low tube voltage and hybrid iterative reconstruction algorithm on image quality. *Radiology* 2012; **264**: 445–54. doi: https://doi.org/10.1148/radiol.12111082

8. Deák Z, Grimm JM, Treitl M, Geyer LL, Linsenmaier U, Körner M, et al. Filtered back projection, adaptive statistical iterative reconstruction, and a model-based iterative reconstruction in abdominal CT: an experimental clinical study. *Radiology* 2013; **266**: 197–206. doi: https://doi.org/10.1148/radiol.12112707
9. Volders D, Bols A, Haspeslagh M, Coenegrachts K. Model-Based iterative reconstruction and adaptive statistical iterative reconstruction techniques in abdominal CT: comparison of image quality in the detection of colorectal liver metastases. *Radiology* 2013; **269**: 469–74. doi: https://doi.org/10.1148/radiology.13130002
10. Pickhardt PJ, Lubner MG, Kim DH, Tang J, Ruma JA, del Rio AM, et al. Abdominal CT with model-based iterative reconstruction (MBIR): initial results of a prospective trial comparing ultralow-dose with standard-dose imaging. *AJR Am J Roentgenol* 2012; **199**: 1266–74. doi: https://doi.org/10.2214/AJR.12.9382
11. Son JH, Kim SH, Yoon J-H, Lee Y, Lim Y-J, Kim S-J. Comparison of model-based iterative reconstruction, adaptive statistical iterative reconstruction, and filtered back projection for detecting hepatic metastases on submillisievert low-dose computed tomography. *J Comput Assist Tomogr* 2017; **41**: 644–50. doi: https://doi.org/10.1097/RCT.0000000000000577
12. Hassani C, Ronco A, Prosper AE, Dissanayake S, Cen SY, Lee C. 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**: 548–56. doi: https://doi.org/10.2214/AJR.17.19245
13. Kataria B, Althén JN, Smedby Örjan, Persson A, Sökjer H, Sandborg M. Assessment of image quality in abdominal CT: potential dose reduction with model-based iterative reconstruction. *Eur Radiol* 2018; **28**: 2464–73. doi: https://doi.org/10.1007/s00330-017-5113-4
14. Tamura A, Nakayama M, Ota Y, Kamata M, Hirota Y, Sone M, et al. Feasibility of thin-slice abdominal CT in overweight patients using a vendor neutral image-based denoising algorithm: assessment of image noise, contrast, and quality. *PLoS One* 2019; **14**: e0226521. doi: https://doi.org/10.1371/journal.pone.0226521
15. Katsura M, Matsuda I, Akahane M, Sato J, Akai H, Yasaka K, et al. Model-Based iterative reconstruction technique for radiation dose reduction in chest CT: comparison with the adaptive statistical iterative reconstruction technique. *Eur Radiol* 2012; **22**: 1613–23. doi: https://doi.org/10.1007/s00330-012-2452-z
16. Umehara K, Ota J, Ishimaru N, Ohno S, Okamoto K, Suzuki T. Super-Resolution convolutional neural network for the improvement of the image quality of magnified images in chest radiographs. *Proc SPIE* 2017; **10133**: 1–7.
17. Zeng K, Yu J, Wang R, Li C, Tao D. Coupled deep autoencoder for single image super-resolution. *IEEE Trans Cybern* 2017; **47**: 27–37. doi: https://doi.org/10.1109/TCYB.2015.2501373
18. Isogawa K, Ida T, Shiodera T, Takeguchi T. Deep shrinkage convolutional neural network for adaptive noise reduction. *IEEE Signal Process Lett* 2018; **25**: 224–8. doi: https://doi.org/10.1109/LSP.2017.2782270
19. Higaki T, Nakamura Y, Zhou J, Yu Z, Nemoto T, Tatsugami F, et al. Deep learning reconstruction at CT: phantom study of the image characteristics. *Acad Radiol* 2020; **27**: 82–7. doi: https://doi.org/10.1016/j.acra.2019.09.008
20. Lenfant M, Chevallier O, Comby P-O, Secco G, Haioun K, Ricolfi F, et al. Deep learning versus iterative reconstruction for CT pulmonary angiography in the emergency setting: improved image quality and reduced radiation dose. *Diagnostics* 2020; **10**: E558. doi: https://doi.org/10.3390/diagnostics10080558
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. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol* 2019; **11**: 6163–71.
23. Leschka S, Stinn B, Schmid F, Schultes B, Thurnheer M, Baumueller S, et al. Dual source CT coronary angiography in severely obese patients: trading off temporal resolution and image noise. *Invest Radiol* 2009; **44**: 720–7. doi: https://doi.org/10.1097/RLI.0b013e3181b46f1a
24. Chinnaiyan KM, McCullough PA, Flohr TG, Wegner JH, Raff GL. Improved noninvasive coronary angiography in morbidly obese patients with dual-source computed tomography. *J Cardiovasc Comput Tomogr* 2009; **3**: 35–42. doi: https://doi.org/10.1016/j.jcct.2008.11.003
25. Funama Y, Awai K, Miyazaki O, Nakayama Y, Goto T, Omi Y, et al. Improvement of low-contrast detectability in low-dose hepatic multidetector computed tomography using a novel adaptive filter: evaluation with a computer-simulated liver including tumors. *Invest Radiol* 2006; **41**: 1–7. doi: https://doi.org/10.1097/01.rli.0000188026.20172.5d
26. Christianson O, Winslow J, Frush DP, Samei E. Automated technique to measure noise in clinical CT examinations. *AJR Am J Roentgenol* 2015; **205**: W93–9. doi: https://doi.org/10.2214/AJR.14.13613
27. Schindera ST, Nelson RC, Toth TL, Nguyen GT, Toncheva GI, DeLong DM, et al. Effect of patient size on radiation dose for abdominal MDCT with automatic tube current modulation: phantom study. *AJR Am J Roentgenol* 2008; **190**: W100–5. doi: https://doi.org/10.2214/AJR.07.2891
28. Schindera ST, Torrente JC, Ruder TD, Hoppe H, Marin D, Nelson RC, et al. Decreased detection of hypovascular liver tumors with MDCT in obese patients: a phantom study. *AJR Am J Roentgenol* 2011; **196**: W772–6. doi: https://doi.org/10.2214/AJR.10.5351
29. Schindera ST, Diedrichsen L, Müller HC, Rusch O, Marin D, Schmidt B, et al. Iterative reconstruction algorithm for abdominal multidetector CT at different tube voltages: assessment of diagnostic accuracy, image quality, and radiation dose in a phantom study. *Radiology* 2011; **260**: 454–62. doi: https://doi.org/10.1148/radiol.11102217
30. Schindera ST, Odedra D, Mercer D, Thipphavong S, Chou P, Szucs-Farkas Z, et al. Hybrid iterative reconstruction technique for abdominal CT protocols in obese patients: assessment of image quality, radiation dose, and low-contrast detectability in a phantom. *AJR Am J Roentgenol* 2014; **202**: W146–52. doi: https://doi.org/10.2214/AJR.12.10513
31. Thibault J-B, 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. doi: https://doi.org/10.1118/1.2789499
32. Antun V, Renna F, Poon C, Adcock B, Hansen AC. On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci U S A* 2020; **117**: 30088–95. doi: https://doi.org/10.1073/pnas.1907377117
33. Gaddikeri S, Andre JB, Benjert J, Hippe DS, Anzai Y. Impact of model-based iterative reconstruction on image quality of contrast-enhanced neck CT. *AJNR Am J Neuroradiol* 2015; **36**: 391–6. doi: https://doi.org/10.3174/ajnr.A4123
34. Geyer LL, Schoepf UJ, Meinel FG, Nance JW, Bastarrika G, Leipsic JA, et al. State of the art: iterative CT reconstruction techniques. *Radiology* 2015; **276**: 339–57. doi: https://doi.org/10.1148/radiol.2015132766

35. Marin D, Nelson RC, Samei E, Paulson EK, Ho LM, Boll DT, et al. Hypervascular liver tumors: low tube voltage, high tube current multidetector CT during late hepatic arterial phase for detection--initial clinical experience. *Radiology* 2009; **251**: 771–9. doi: https://doi.org/10.1148/radiol.2513081330
36. Ehman EC, Guimarães LS, Fidler JL, Takahashi N, Ramírez-Giraldo JC, Yu L, et al. Noise reduction to decrease radiation dose and improve conspicuity of hepatic lesions at contrast-enhanced 80-kV hepatic CT using projection space denoising. *AJR Am J Roentgenol* 2012; **198**: 405–11. doi: https://doi.org/10.2214/AJR.11.6987
37. Fletcher JG. Adjusting kV to reduce dose and improve image quality. Technology Assessment Institute, Summit on CT dose, 2011. 2011. Available from: https://www.aapm.org/meetings/2011CTS/documents/Fletcher_AAPM_2011-AutokV.pdf.
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–74. doi: https://doi.org/10.2307/2529310
39. 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**: 50–7. doi: https://doi.org/10.2214/AJR.19.22332
40. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the global burden of disease study 2013. *Lancet* 2014; **384**: 766–81. doi: https://doi.org/10.1016/S0140-6736(14)60460-8
41. Christianson O, Chen JJS, Yang Z, Saiprasad G, Dima A, Filliben JJ, et al. An improved index of image quality for task-based performance of CT iterative reconstruction across three commercial implementations. *Radiology* 2015; **275**: 725–34. doi: https://doi.org/10.1148/radiol.15132091