Radiology

Reduced-Dose Deep Learning Reconstruction for Abdominal CT of Liver Metastases

Corey T. Jensen, MD • Shiva Gupta, MD • Mohammed M. Saleh, MD • Xinming Liu, PhD • Vincenzo K. Wong, MD • Usama Salem, MD • Wei Qiao, PhD • Ehsan Samei, PhD • Nicolaus A. Wagner-Bartak, MD

From the Departments of Abdominal Imaging (C.T.J., S.G., M.M.S., V.K.W., U.S., N.A.W.B.), Physics (X.L.), and Biostatistics (W.Q.), the University of Texas MD Anderson Cancer Center, 1400 Pressler St, Unit 1473, Houston, TX 77030-4009; and Center for Virtual Imaging Trials, Carl E. Ravin Advanced Imaging Laboratories, Clinical Imaging Physics Group, Medical Physics Graduate Program, Departments of Radiology, Physics, Biomedical Engineering, and Electrical and Computer Engineering, Duke University Medical Center, Durham, NC (E.S.). Received July 19, 2021; revision requested September 7; revision received October 19; accepted October 28. **Address correspondence to** C.T.J. (e-mail: *cjensen@mdanderson.org*).

Supported by an institutional cancer center support grant from the National Institutes of Health/National Cancer Institute under award number P30CA016672 (NCI-2018-01272/NCT03151564).

C.T.J., X.L., and N.W.B. each received an in-kind grant from General Electric.

Conflicts of interest are listed at the end of this article.

Radiology 2022; 303:90–98 • https://doi.org/10.1148/radiol.211838 • Content codes: GI AI CT

Background: Assessment of liver lesions is constrained as CT radiation doses are lowered; evidence suggests deep learning reconstructions mitigate such effects.

Purpose: To evaluate liver metastases and image quality between reduced-dose deep learning image reconstruction (DLIR) and standard-dose filtered back projection (FBP) contrast-enhanced abdominal CT.

Materials and Methods: In this prospective Health Insurance Portability and Accountability Act–compliant study (September 2019 through April 2021), participants with biopsy-proven colorectal cancer and liver metastases at baseline CT underwent standard-dose and reduced-dose portal venous abdominal CT in the same breath hold. Three radiologists detected and characterized lesions at standard-dose FBP and reduced-dose DLIR, reported confidence, and scored image quality. Contrast-to-noise ratios for liver metastases were recorded. Summary statistics were reported, and a generalized linear mixed model was used.

Results: Fifty-one participants (mean age ± standard deviation, 57 years ± 13; 31 men) were evaluated. The mean volume CT dose index was 65.1% lower with reduced-dose CT (12.2 mGy) than with standard-dose CT (34.9 mGy). A total of 161 lesions (127 metastases, 34 benign lesions) with a mean size of 0.7 cm ± 0.3 were identified. Subjective image quality of reduced-dose DLIR was superior to that of standard-dose FBP ($P < .001$). The mean contrast-to-noise ratio for liver metastases of reduced-dose DLIR (3.9 ± 1.7) was higher than that of standard-dose FBP (3.5 ± 1.4) ($P < .001$). Differences in detection were identified only for lesions 0.5 cm or smaller: 63 of 65 lesions detected with standard-dose FBP (96.9%; 95% CI: 89.3, 99.6) and 47 lesions with reduced-dose DLIR (72.3%; 95% CI: 59.8, 82.7). Lesion accuracy with standard-dose FBP and reduced-dose DLIR was 80.1% (95% CI: 73.1, 86.0; 129 of 161 lesions) and 67.1% (95% CI: 59.3, 74.3; 108 of 161 lesions), respectively ($P = .01$). Lower lesion confidence was reported with a reduced dose ($P < .001$).

Conclusion: Deep learning image reconstruction (DLIR) improved CT image quality at 65% radiation dose reduction while preserving detection of liver lesions larger than 0.5 cm. Reduced-dose DLIR demonstrated overall inferior characterization of liver lesions and reader confidence.

Clinical trial registration no. NCT03151564

Online supplemental material is available for this article.

CT assessment of low-contrast liver lesions is one of the more challenging tasks in medical imaging. The detection and characterization of these lesions is more difficult in the setting of reduced CT radiation dose (1–5). For many years, there has been an attempt to maintain image quality using iterative reconstruction methods, such as adaptive statistical iterative reconstruction (ASIR, GE Healthcare), to reduce image noise relative to filtered back projection (FBP) while reducing radiation dose. Although reduced-dose imaging is sufficient in many clinical settings (6,7), the accuracy of CT in the assessment of small low-contrast liver lesions quickly becomes inadequate as radiation doses are lowered. Observer performance for low-contrast liver lesions, even with use of the more advanced hybrid reconstruction method, ASIR-V (GE Healthcare), has been shown to be inferior in examinations performed with a mean volume CT dose index ($CTDI_{vol}$) of 11.8 mGy, approximating the American College of Radiology Dose Index Registry levels, when compared directly to examinations using a mean $CTDI_{vol}$ of 25.8 mGy in a tertiary oncologic practice (1).

Although increasing radiation dose levels is a reasonable response to recently reported data (depending on baseline radiation dose levels in each practice) for patients in whom evaluation of small liver lesions is crucial (eg, in those with colorectal cancer in the presurgical setting), in the past few

Abbreviations

AV60 = ASIR-V 60%, $CTDI_{vol}$ = volume CT dose index, DLIR = deep learning image reconstruction, FBP = filtered back projection

Summary

At 65% reduced radiation dose, deep learning image reconstruction improved CT image quality with robust denoising but had inferior lesion characterization and inferior lesion detection for small liver lesions (≤0.5 cm) compared with standard-dose reconstruction.

Key Results

- In a prospective study of 51 participants (161 lesions [127 metastases, 34 benign lesions]) with colorectal liver metastases, 65% reduced-dose deep learning image reconstruction (DLIR) abdominal CT reconstructions were qualitatively superior to standard-dose filtered back projection (FBP) (odds ratio, 1.6; $P = .02$).
- Reduced-dose DLIR lesion characterization accuracy (67.1%; 108 of 161 lesions) was inferior to that of standard-dose FBP (80.1%; 129 of 161 lesions) ($P = .01$).
- Reduced-dose DLIR had inferior detection of small lesions (≤0.5 cm) (72.3%; 47 of 65) compared with standard-dose FBP (96.9%; 63 of 65) ($P < .001$).

years promising artificial intelligence–based reconstruction algorithms such as deep learning image reconstruction (DLIR) (True Fidelity, GE Healthcare) and Advanced intelligent Clear-IQ Engine (Canon Medical Solutions) have become commercially available. These reconstruction algorithms may mitigate the level of higher needed dose. They use neural network–based models intended to provide denoising with an image texture similar to that of FBP (DLIR) or model-based iterative reconstruction (Advanced intelligent Clear-IQ Engine) (8,9). Several studies have shown improved perceptual image quality and standard metrics (eg, contrast-to-noise ratio) using deep learning reconstructions when compared with FBP and iterative reconstruction (10–17); however, nonlinear image reconstruction algorithms such as these require thorough clinical and task-based assessments for an adequate comprehensive evaluation (18). Observer performance studies, such as assessment of low-contrast liver lesions, must be carried out to determine the degree to which these new methods will mitigate known issues and limitations of iterative reconstruction.

We hypothesized that DLIR would improve perceptual image quality, but that detection and characterization of small lesions would be inferior with reduced radiation dose DLIR compared with standard-dose FBP. The purpose of this study was to evaluate the detection and characterization of colorectal cancer liver metastases between reduced-dose DLIR and standard-dose FBP contrast-enhanced abdominal CT and to qualitatively compare image quality.

Materials and Methods

Our institutional review board approved this prospective Health Insurance Portability and Accountability Act–compliant study (NCI-2018–01272, ClinicalTrials.gov identifier: NCT03151564), and written informed consent was obtained. Our study was supported by an in-kind financial grant from General Electric. The authors maintained control of data collection, analysis, and submission.

Study Participants

On the basis of expected CT accuracy with an assumed discordant rate of 10% or higher, our power analysis indicated that 52 participants would provide a power of at least 80% with a one-sided type I error rate of 10%. We scanned participants during the accrual period of September 2019 to April 2021 using a convenience sample. Inclusion and exclusion criteria are listed in Figure 1. Each participant's age, sex, height, weight, and body mass index were recorded. Follow-up of the participants occurred through June 2021.

Radiation Dose Estimation

The $CTDI_{vol}$ for a 32-cm phantom was recorded from the scanner for each scan, and size-specific dose estimates were calculated (19).

Imaging Technique and Postprocessing

All participants underwent CT of the abdomen in the portal venous phase (Table 1). The reduced-dose parameters were set to approximate a 65% radiation dose reduction from the standard dose. The target reduction was chosen to approximate the 25th–75th percentile $CTDI_{vol}$ reported in the American College of Radiology Dose Index Registry ($CTDI_{vol}$, 9–19 mGy) (20). The standard- and reduced-dose scans were obtained in the craniocaudal direction and then, without interval delay, in the reverse direction during the same breath hold. The order of these same breath-hold scans was alternated between participants. Weight-based intravenous injection of contrast material was used with low-osmolar iohexol, 350 mg of iodine per milliliter (Omnipaque 350, GE Healthcare), using a setting of 0.6 g of iodine per kilogram and an injection duration setting of 40 seconds (injection

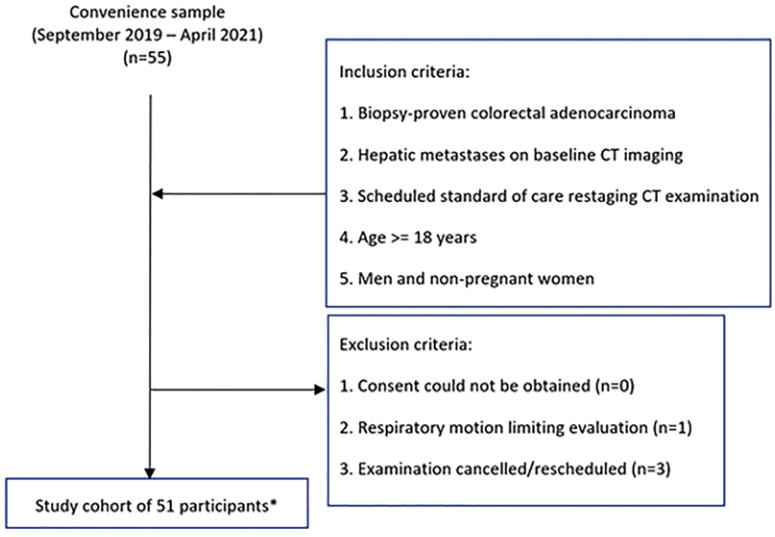

Figure 1: Study flowchart. * = Eight participants were found to have no lesions that fulfilled the study criteria, and two participants were excluded from lesion detection data because they had more than 20 liver lesions. Images from those participants still underwent qualitative and quantitative assessments.

Table 1: Imaging Settings for the Abdominal Protocol

Parameter	Datum
Scanner model	Revolution (GE Healthcare)
Scan mode	Helical, single source
Detector configuration (mm)	128 × 0.625
Beam collimation (mm)	80
Pitch	0.508:1
Rotation time (sec)	0.5–0.7
Table speed (mm/rotation)	40.64
Tube current modulation	
SD scan	Scan-by-scan settings to approximate prior standard scan
RD scan	Scan-by-scan settings to approximate 65% dose reduction
Tube potential (kV)	120
Acquisition section thickness (mm)	5, full mode
Reconstruction thickness and increment (mm)	2.5/2.5, plus mode
Reconstruction algorithms	
FBP	Standard
AV60	Iterative (hybrid)
Medium-strength DLIR	Deep learning image reconstruction
Reconstruction kernel	Standard plus

Note.—AV60 = ASIR-V (GE Healthcare) 60%, DLIR = deep learning image reconstruction, FBP = filtered back projection, RD = reduced dose, SD = standard dose.

rate, 3–6 mL/sec; intravenous contrast material volume, 110–170 mL) (21). Bolus tracking was used, with a 100-HU trigger value in the abdominal aorta at the level of the celiac artery and a scan delay of 46 seconds.

Six axial reconstructions of 2.5-mm section thickness were performed for each participant: three at our standard radiation doses (FBP, ASIR-V 60% [AV60], medium-strength DLIR) and three at reduced radiation doses (FBP, AV60, medium-strength DLIR). Previous studies showed that AV60 was subjectively preferred to FBP, adaptive statistical iterative reconstruction, and model-based iterative reconstruction (1,22); AV60 was thus chosen to be compared with the reference standard of FBP and with the new DLIR. An initial report suggested medium-strength DLIR to be the favored strength in the abdomen relative to the low- and high-strength DLIR (11). All images were reviewed under standard clinical conditions with a 3-megapixel monitor (MDNC-3421, Barco) using a picture archiving and communications system.

Image Analysis

Lesion detection and reference standard.—Three board-certified abdominal radiologists independently performed lesion detection in two primary series while blinded to all information except cancer diagnosis. The readers were fellowship trained in abdominal imaging and had 10 years (N.A.W.B.), 8 years (S.G.), and 6 years (V.K.W.) of additional experience reading abdominal CT scans. Randomized images from 51 participants with biopsy-proven colorectal cancer and liver metastases were assessed over four sessions, for a total review of 102 scans: standard-dose FBP versus reduced-dose DLIR. Only hypoattenuating liver lesions measuring 0.2–1.5 cm were to be marked. Herein, we variably refer to lesions as low contrast, meaning that their CT number is not markedly different from that of the adjacent liver parenchyma. Likert-type scores were given for each lesion with respect to characterization (ie, a score of 1 indicated definitely benign; 2, likely benign; 3, malignancy not excluded; 4, likely malignant; 5, definitely malignant) and related confidence in the diagnosis (ie, 1 for low confidence and up to 5 for high confidence). Reconstructions were presented in a randomized fashion, and review of the standard-dose FBP versus reduced-dose DLIR images in the same participant was separated by a delay of 2–3 weeks to minimize recall. There was no time limit for review, but the radiologists were instructed to read in a manner similar to clinical practice.

Two nonblinded consensus reviewers (fellowship-trained abdominal radiologists [C.T.J. and U.S., with 10 years and 4 years of additional experience beyond fellowship, respectively]) subsequently established the reference standard using the saved reader marks and all available clinical data. Any disagreements between the consensus reviewers, although none occurred, would have been decided by a third radiologist (M.M.S., an abdominal imaging fellow at the time of the study). Comparison was made to all available cross-sectional imaging examinations (CT, MRI, and PET/CT). All lesions identified were measured and classified by the consensus reviewers as metastatic or benign. For participants with 20 or more metastatic lesions, lesion performance was not assessed. For performance metrics, benign lesions scored as 3 or higher on the malignancy scale were considered false-positive results against the reference standard. Malignant lesions according to the reference standard that were either not identified by reviewers or scored 2 or lower on the malignancy scale were considered false-negative results (1,4).

Qualitative analysis.—After each primary series lesion detection and characterization, readers scored the series for overall image quality with respect to artifacts, image texture, and qualitative resolution (where a score of 5 indicated excellent image quality without related issues of concern; 4, minor issues not interfering with diagnostic decision making; 3, minor issues possibly interfering with diagnostic decision making; 2, major issues affecting visualization of major structures but diagnosis still possible; and 1, issues affecting diagnostic information).

For each participant, after the second lesion evaluation session was completed and the primary image quality score was given, all six reconstructions for that participant were displayed side-by-side on a single monitor and ranked (Appendix E1 [online]).

Quantitative analysis.—Three-dimensional spherical regions of interest were drawn in the liver and within the single largest liver metastasis per participant on each reconstruction by using GE Advantage Workstation software (version 3.2, GE Healthcare) (Appendix E2 [online]).

Statistical Analysis

Summary statistics for CT number of the liver, liver noise, signal-to-noise ratio for liver lesions, signal-to-noise ratio for liver metastases, and contrast-to-noise ratios for liver metastases were provided. Categorical data for lesion detection and accuracy were summarized. The McNemar test was used to test the marginal homogeneity in terms of diagnostic accuracy and lesion detection. The κ statistics were estimated between readers and consensus results for lesion accuracy. If any reader detected a lesion for a dose, then that lesion was considered detected at that dose. For lesion characterization, if two readers (ie, majority rule) classified a lesion as true malignant or true benign, then that lesion was considered accurate with the reference standard. Image quality ranking, signal-to-noise ratio, contrast-to-noise ratio, and noise were estimated and compared using the linear mixed model, where participant and reader were included as random effects. Tukey-Kramer adjustment was used to control the overall type I error rate at 5%. For the categorical outcome (eg, high vs low confidence level, score of 4–5 vs 1–3), the generalized estimating equation method was used to account for random effect of reader. All tests were two-sided and $P \leq .05$ was considered to indicate a statistically significant difference. Statistical analysis was carried out using software (SAS, version 9.4; SAS Institute).

Results

Participant and Lesion Characteristics

Among 55 enrolled participants, one was excluded because of respiratory motion that limited image evaluation and three were excluded because of cancelled or rescheduled examinations. The final study group consisted of 51 participants (31 men, 20 women) (Fig 1) with a mean age ± standard deviation of 57 years ± 13 (range, 25–90 years), a mean weight of 85 kg ± 17 (range, 49–122 kg), and a mean body mass index of 29 kg/m^2 ± 6 (range, 18–43 kg/m^2). Reference standard assessment showed 161 liver lesions (127 metastases, 34 benign lesions) with a mean size of 0.7 cm ± 0.3 (range, 0.2–1.5 cm) (Table 2). Eight participants had no lesions that fulfilled the study criteria, and two participants were excluded from lesion detection data because they had more than 20 liver lesions.

Radiation Dose

For standard-dose examinations, the mean $CTDI_{vol}$ was 34.9 mGy ± 10.9 (range, 10.1–54.9 mGy; 25th percentile, 26.1 mGy; 75th percentile, 45.0 mGy) and the mean size-specific dose estimate was 39.8 mGy ± 8.9 (range, 15.1–53.1 mGy). The mean dose reduction of the reduced-dose scan was 65.1% ± 2.8 (range, 53.7%–68.8%), with a mean $CTDI_{vol}$ of 12.2 mGy ± 3.8 (range, 3.4–18.9 mGy; 25th percentile, 9.3 mGy; 75th percentile, 15.9 mGy) and mean size-specific dose estimate of 13.8 mGy ± 3.2 (range, 5.1–19.7 mGy).

Intravenous Contrast Material Dose

The mean intravenous contrast material volume administered was 146 mL ± 18 (range, 123–170 mL), with a mean rate of 4.4 mL/sec ± 0.7 (range, 3.1–5.8 mL/sec).

Lesion Detection and Accuracy

Readers 1, 2, and 3 correctly detected 159, 150, and 150 lesions, respectively, during the standard-dose FBP evaluation

Table 2: Participant and Lesion Characteristics

Parameter	Value
No. of participants	51
Age (y)*	57 ± 13
Sex	
Male	31
Female	20
Height (cm)*	171 ± 8
Weight (kg)*	85 ± 17
Body mass index (kg/m^2)*	29 ± 6
Liver lesions	
No. of liver lesions	161
No. of malignant lesions	127
No. of benign lesions	34
Size (cm)*	0.7 ± 0.3

Note.—Except where indicated, data are numbers of participants or lesions.
* Data are means ± standard deviations.

Table 3: Performance of Standard-Dose CT using FBP versus Reduced-Dose CT using Medium-Strength DLIR

	Standard-Dose FBP			Reduced-Dose DLIR		
Performance Measure	Lesion Size ≤0.5 cm	Lesion Size 0.6–1 cm	Lesion Size >1 cm	Lesion Size ≤0.5 cm	Lesion Size 0.6–1 cm	Lesion Size >1 cm
Detection (%)	96.9 (89.3, 99.6) [63/65]	100 (95.0, 100) [72/72]	100 (85.8, 100) [24/24]	72.3 (59.8, 82.7) [47/65]	90.3 (81.0, 96.0) [65/72]	100 (85.8, 100) [24/24]
Sensitivity (%)	78.9 (62.7, 90.4) [30/38]	98.5 (91.8, 100.0) [65/66]	100.0 (85.2, 100.0) [23/23]	47.4 (31.0, 64.2) [18/38]	80.3 (68.7, 89.1) [53/66]	95.7 (78.1, 99.9) [22/23]
Specificity (%)	33.3 (16.5, 54.0) [9/27]	33.3 (4.3, 77.7) [2/6]	0.0 (0.0, 97.5) [0/1]	55.6 (35.3, 74.5) [15/27]	0.0 (0.0, 45.9) [0/6]	0.0 (0.0, 97.5) [0/1]
Accuracy (%)	60.0 (47.1, 72.0) [39/65]	93.1 (84.5, 97.7) [67/72]	95.8 (78.9, 99.9) [23/24]	50.8 (38.1, 63.4) [33/65]	73.6 (61.9, 83.3) [53/72]	91.7 (73.0, 99.0) [22/24]

Note.—Performance data are per lesion. Numbers in parentheses are 95% CIs, and numbers in brackets are numbers of lesions.
DLIR = deep learning image reconstruction, FBP = filtered back projection.

Table 4: False-Negative and False-Positive Lesion Data by Three Readers during Standard-Dose FBP and Reduced-Dose DLIR Evaluations

	Standard-Dose FBP				Reduced-Dose DLIR (Medium Strength)			
Lesion Size and Reader No.	No. of Lesions Diagnosed as Malignant	No. of False-Negative Lesions	No. of False-Positive Lesions	False-Positive Rate (%)	No. of Lesions Diagnosed as Malignant	No. of False-Negative Lesions	No. of False-Positive Lesions	False-Positive Rate (%)
≤0.5 cm								
1	45	11	18	40.0 (25.7, 55.7)	24	23	9	37.5 (18.8, 59.4)
2	37	10	9	24.3 (11.8, 41.2)	25	22	9	36.0 (18.0, 57.5)
3	52	7	21	40.4 (27.0, 54.9)	41	18	22	53.7 (37.4, 69.3)
0.6–1 cm								
1	69	2	5	7.3 (2.4, 16.1)	54	18	6	11.1 (4.2, 22.6)
2	67	1	2	3.0 (0.4, 10.4)	56	12	2	3.6 (0.4, 12.3)
3	67	3	4	6.0 (1.7, 14.6)	58	14	6	10.3 (3.9, 21.2)
>1 cm								
1	23	0	1	4.4 (0.1, 21.9)	24	0	1	4.2 (0.1, 21.1)
2	21	2	1	4.8 (0.1, 23.8)	21	2	0	0.0 (0.0, 16.1)
3	22	1	1	4.6 (0.1, 22.8)	22	2	1	4.6 (0.1, 22.8)

Note.—Numbers in parentheses are 95% CIs. DLIR = deep learning image reconstruction, FBP = filtered back projection.

and 130, 125, and 124 lesions during reduced-dose DLIR evaluation. Overall, 159 of 161 lesions were detected at standard-dose FBP (98.8%; 95% CI: 95.6, 99.8) and 136 of 161 lesions were detected at reduced-dose DLIR (84.5%; 95% CI: 77.9, 89.7) ($P < .001$); this inferior detection with use of reduced-dose DLIR was due to inferior detection of small lesions (≤0.5 cm) ($P < .001$) (Table 3). The mean lesion confidence scores were 3.9 ± 1.3, 4.2 ± 1.1, and 4.5 ± 1.1 for standard-dose FBP and 3.1 ± 1.6, 3.5 ± 1.6, and 3.5 ± 1.6 for reduced-dose DLIR, respectively. Readers reported higher confidence scores with standard-dose FBP compared with reduced-dose DLIR (odds ratio, 0.32; 95% CI: 0.24, 0.43; $P < .001$). The overall lesion characterization accuracies were 80.1% (95% CI: 73.1, 86.0; 129 of 161 lesions) and 67.1% (95% CI: 59.3, 74.3; 108 of 161 lesions) for standard-dose FBP and reduced-dose DLIR ($P = .01$); the overall sensitivities were 92.9% (95% CI: 87.0, 96.7; 118 of 127 lesions) and 73.2% (95% CI: 64.7, 80.7; 93 of 127 lesions) for standard-dose FBP and reduced-dose DLIR; the overall specificities were 32.4% (95% CI: 17.4, 50.5; 11 of 34) and 44.1% (95% CI: 27.2, 62.1; 15 of 34) for standard-dose FBP and reduced-dose DLIR. False-

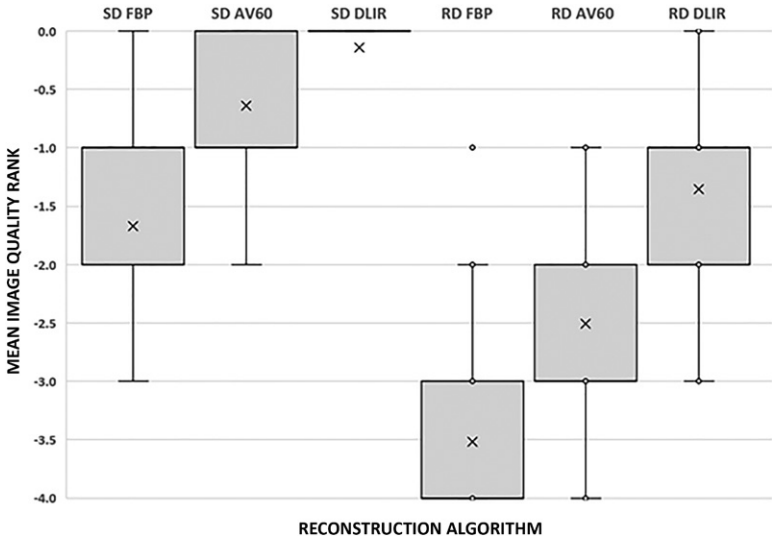

Figure 2: Box-and-whisker plot shows results of qualitative evaluation of overall image quality rank. A score of 0 was given for the best series, −1 for slightly inferior (no influence on diagnosis), −2 for mildly inferior (possible influence on diagnosis), −3 for moderately inferior (probable influence on diagnosis), and −4 for markedly inferior (impairing diagnosis). Mean image quality rank was significantly different between each reconstruction based on the pairwise comparison ($P < .001$). Box represents the interquartile range and the whiskers demonstrate the maximum and minimum data range excluding dots that represent outlier data points. The *X* is the data mean. AV60 = ASIR-V (GE Healthcare) 60%, DLIR = deep learning image reconstruction (with medium strength), FBP = filtered back projection, RD = reduced dose, SD = standard dose.

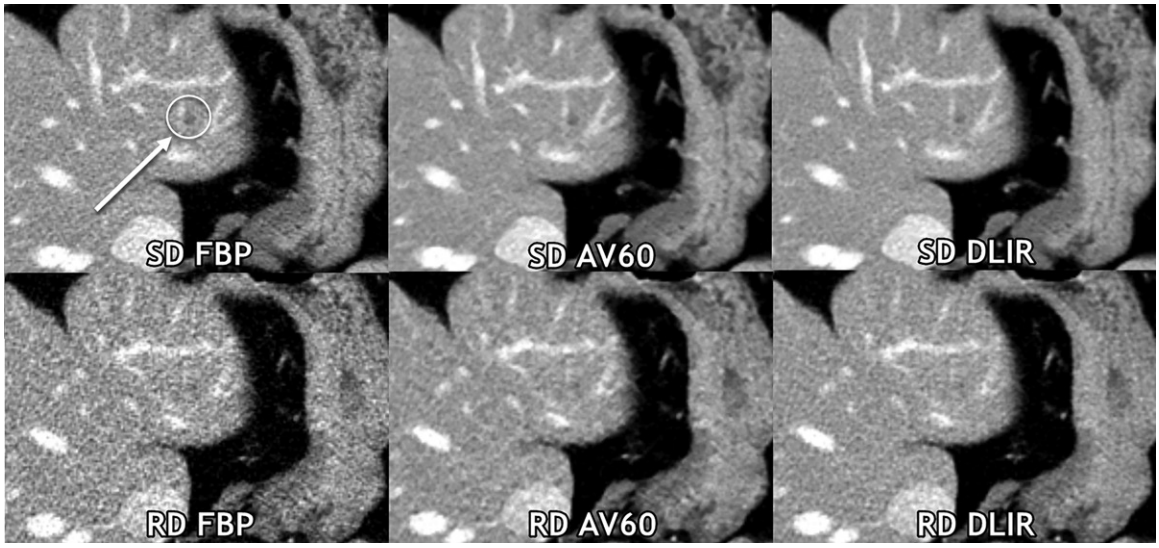

Figure 3: Axial contrast-enhanced CT images of the abdomen obtained with standard-dose (SD) filtered back projection (FBP) and reduced-dose (RD) deep learning image reconstruction (DLIR) with medium strength in the same breath hold. A 0.5-cm low-contrast left liver metastasis (arrow with circle) was missed by all three readers at reduced-dose deep DLIR and was detected by all readers at standard-dose FBP. Of note, all three readers qualitatively scored this reduced-dose DLIR scan as a 4 (superior to standard-dose FBP scores of 3 by each reader), even at the aggressive radiation dose reduction of 67% on this scan. Contrast-to-noise ratios for liver metastases in this participant for standard-dose FBP, standard-dose ASIR-V 60% (AV60), standard-dose DLIR, reduced-dose FBP, reduced-dose AV60, and reduced-dose DLIR were 3.6, 4.6, 4.7, 2.1, 3.3, and 3.4, respectively.

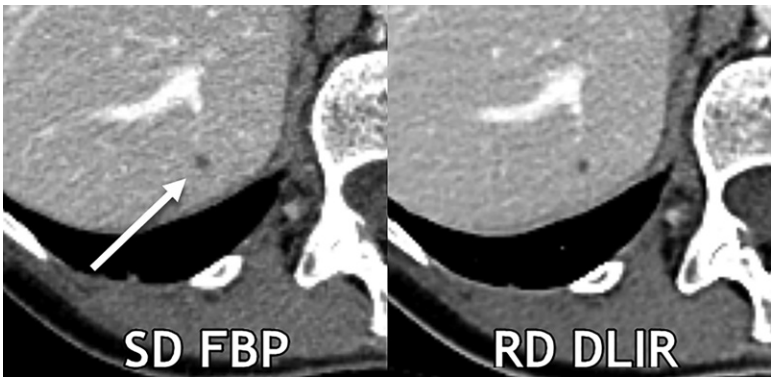

Figure 4: Axial contrast-enhanced CT images show a 0.3-cm liver cyst (arrow) that was detected by all readers on both scans. However, all three reader characterizations were false-positive at standard-dose (SD) filtered back projection (FBP), whereas the regular-dose (RD) deep learning image reconstruction (DLIR) with medium strength scan resulted in two true-negative characterizations and one false-positive characterization. At a radiation dose reduction of 66% on this scan, the cyst appears more conspicuous with DLIR, and each reader qualitatively scored the reduced-dose DLIR scan to be better than or equivalent to the standard-dose FBP scan. Contrast-to-noise ratios for liver metastases in this participant for standard-dose FBP and reduced-dose DLIR were 3.9 and 4.6, respectively.

negative and false-positive lesion data are listed according to reader and lesion size in Table 4. The estimated κ values for reader agreement of lesion accuracy for readers 1, 2, and 3 were moderate to good: 0.51, 0.72, and 0.51, respectively, for standard-dose FBP and 0.61, 0.73, and 0.54 for reduced-dose DLIR.

Qualitative Image Assessment

The mean image quality scores for standard-dose FBP for readers 1, 2, and 3 were 3.8 ± 0.4, 3.3 ± 0.5, and 3.1 ± 0.7, respectively, and the mean image quality scores for reduced-dose DLIR were 4.1 ± 0.5, 3.3 ± 0.6, and 3.4 ± 0.8. Reduced-dose DLIR received more high-quality scores than standard-dose FBP (odds ratio, 1.6; 95% CI: 1.1, 2.3; P = .02).

The side-by-side ranking data of series is shown in Figure 2 and Appendix E1 (online). Figure 3 demonstrates a reduced-dose DLIR scan that was scored as superior to the standard-dose FBP scan; however, a lesion was missed by all three readers only on the reduced-dose DLIR scan. Figure 4, conversely, shows images in a participant in whom the reduced-dose DLIR scan was scored as better than or equivalent to the standard-dose FBP scan; readers had fewer false-positive characterizations on the reduced-dose DLIR scan.

Quantitative Image Assessment

Hounsfield units and noise measurements with signal-to-noise ratio and contrast-to-noise ratio calculations are shown in Table 5 and Appendix E2 (online).

Discussion

Identification and assessment of small low-contrast liver lesions is an important task in oncologic staging that can be constrained as CT examination radiation doses are lowered. However, evidence suggests that new deep learning image reconstruction (DLIR) mitigates these effects. Our study demonstrates improvement of perceptual imaging quality, signal-to-noise ratio, and contrast-to-noise ratio with DLIR relative to ASIR-V 60% and filtered back projection. For liver lesions larger than 0.5 cm, reduced-dose DLIR maintained observer lesion detection with an aggressive dose reduction of 65% (mean reduced-dose vol-

Table 5: CT Number, Noise, Signal-to-Noise Ratio, and Contrast-to-Noise Ratio in the Abdomen according to Dose and Reconstruction Method

Variable	SD FBP	SD AV60	SD DLIR (Medium Strength)	RD FBP	RD AV60	RD DLIR (Medium Strength)
CT number (HU)*						
Liver	122 ± 18	122 ± 18	122 ± 18	121 ± 17	121 ± 18	121 ± 18
Metastases†	62 ± 16	63 ± 16	62 ± 17	63 ± 18	63 ± 17	63 ± 18
Noise‡	17 ± 3	11 ± 2	10 ± 2	25 ± 3	16 ± 3	15 ± 3
SNR§						
Liver	7.2 ± 1.6	11.6 ± 2.6	12.3 ± 2.9	5.0 ± 1.1	8.0 ± 1.9	8.4 ± 2.0
Metastases	3.6 ± 1.0	5.8 ± 1.8	5.9 ± 1.9	2.5 ± 0.7	4.0 ± 1.3	4.2 ± 1.4
CNR						
Metastases	3.5 ± 1.4	5.7 ± 2.2	6.0 ± 2.5	2.4 ± 1.0	3.7 ± 1.6	3.9 ± 1.7

Note.—Data are means ± standard deviations. AV60 = ASIR-V (GE Healthcare) 60%, CNR = contrast-to-noise ratio, DLIR = deep learning image reconstruction, FBP = filtered back projection, RD = reduced dose, SD = standard dose, SNR = signal-to-noise ratio.

* No significant difference in CT number was identified between reconstructions.

† The Hounsfield unit of the single largest representative lesion fulfilling the study criteria was measured in each applicable participant.

‡ Noise = Hounsfield unit standard deviation in the liver. Noise was significantly different between each reconstruction.

§ Signal-to-noise and contrast-to-noise ratios were significantly different between dose and between FBP when compared with AV60 and DLIR reconstructions ($P < .001$). Signal-to-noise and contrast-to-noise ratios ranked from highest to lowest are as follows: standard-dose medium-strength DLIR, standard-dose AV60, reduced-dose medium-strength DLIR, reduced-dose AV60, standard-dose FBP, and reduced-dose FBP.

ume CT dose index, 12.2 mGy). However, with reduced-dose DLIR, observer lesion characterization, confidence, and detection of small lesions (<0.5 cm) was inferior for that of low-contrast liver lesions.

The qualitative portion of our study corroborates results of previous studies that also indicated improved image quality with use of DLIR. For example, phantom analyses have demonstrated a preserved DLIR noise power spectrum similar in appearance to FBP and maintained high-contrast spatial resolution (10,23). However, progressively higher strengths of DLIR have been reported to result in minor blurring of tiny liver lesions and vessels (Appendix E3 [online]) (11). This parallels one phantom analysis indicating a reduction of low-contrast task transfer function for DLIR relative to FBP, albeit less so than the reduction by AV60 (23). In a study conducted with a mean $CTDI_{vol}$ of 10.5 mGy, DLIR was qualitatively favored over ASIR-V 40% for abdominal evaluation (15). Another study specifically assessing a pancreatic cancer sample (mean $CTDI_{vol}$, 12 mGy) showed improved observer performance with progressively higher strengths of DLIR compared with FBP and AV60 (16).

Although this mounting evidence indicates clinical improvements with DLIR across disciplines (24,25), appropriate radiation dose levels for other clinical tasks must be assessed. In a phantom analysis by Racine et al (26) using a detectability index, a potential dose reduction of 67% was suggested by using high-strength DLIR compared with FBP ($CTDI_{vol}$: 1, 3, and 7 mGy). A follow-up phantom study with model observer ($CTDI_{vol}$ range, 2–20 mGy) indicated a dose reduction potential of 25% for medium-strength DLIR compared with AV60 (27). A phantom study by Greffier et al (28) using dose levels up to a $CTDI_{vol}$ of 15 mGy proposed a potential reduction of 46%–56% with use of high-strength DLIR compared with ASIR-V 50%. Importantly, the degree to which any practice can reduce radiation doses will depend on their baseline and intended clinical task.

In oncologic imaging, liver evaluation of potential metastatic disease is crucial (29,30). In this setting, diagnostic image quality, rather than radiation dose reduction, is the main concern (31). Our study reconfirms that radiation dose levels in the American College of Radiology Dose Index Registry are too low for the clinical task of small low-contrast liver lesion evaluation (1,4). Our reduced-dose scans were obtained to approximate the 25th–75th percentile $CTDI_{vol}$ levels reported by sites in the College of Radiology Dose Index Registry for contrast-enhanced CT scans of the chest, abdomen, and pelvis ($CTDI_{vol}$, 9–19 mGy) (32). Our oncology practice currently uses reduced-dose levels only in patients with treated leukemia, lymphoma, and testicular cancer because these patient subgroups have a lower relative pretest probability of disease, particularly in the liver. On the basis of our results and the literature, further evaluation of dose levels that might preserve liver lesion evaluation are suggested around a mean $CTDI_{vol}$ of 20 mGy. In addition, it will be useful to consider further specification of data sent to the Dose Index Registry such that perhaps higher-dose oncologic CT examinations are listed separately from general examinations (eg, appendicitis evaluation) to allow for better guidance regarding the appropriate oncologic achievable doses (ie, 50th percentile from the Dose Index Registry) and diagnostic reference levels (ie, the 75th percentile from the Dose Index Registry) (32).

Our study had limitations. First, unlike in normal clinical practice, readers characterized lesions without reference to previous examinations; thus, reader accuracy is likely underestimated for our study in this regard. However, the opposite effect may be true because of a reference standard bias, which predominately used CT. Second, as with any clinical CT evaluation in which radiation exposure is a requisite concern for patient safety, our study assessed only two dose levels. Third, human observer studies such as this are inherently limited by human factors such as reader fatigue; thus, only a limited number of

key reconstructions were assessed. Fourth, the results directly apply to only one vendor scanner model and associated reconstructions. Fifth, although each participant had biopsy-proven colorectal adenocarcinoma, the characterization of liver lesions in the reference standard was based on imaging diagnosis.

In conclusion, deep learning image reconstruction (DLIR) improved subjective CT image quality, signal-to-noise ratio, and contrast-to-noise ratio relative to ASIR-V and filtered back projection (FBP) and maintained observer lesion detection at a 65% radiation dose reduction for lesions larger than 0.5 cm relative to standard-dose FBP. Although these improvements from DLIR likely potentiate radiation dose reduction in many clinical scenarios, our results demonstrated that DLIR did not maintain observer lesion detection for very small low-contrast lesions (≤0.5 cm), lesion characterization, or lesion confidence. Therefore, we advise using caution with regard to radiation reduction and the choosing of appropriate radiation dose levels for patients in whom small liver lesion evaluation is important. Future studies should be carried out to determine appropriate radiation dose levels in other clinical scenarios and to assess other dose levels for evaluation of low-contrast liver lesions.

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

Data sharing. All data generated or analyzed during the study are included in the published paper.

Disclosures of conflicts of interest: C.T.J. No relevant relationships. **S.G.** No relevant relationships. **M.M.S.** No relevant relationships. **X.L.** Data safety monitoring board or advisory board at MD Anderson Cancer Center. **V.K.W.** No relevant relationships. **U.S.** No relevant relationships. **W.Q.** No relevant relationships. **E.S.** Relationships with GE, Siemens, Imalogix, 12Sigma, SunNuclear, Nanox, Metis Health Analytics, Cambridge University press, and Wiley and Sons. **N.A.W.B.** No relevant relationships.

References

1. Jensen CT, Wagner-Bartak NA, Vu LN, et al. Detection of colorectal hepatic metastases is superior at standard radiation dose CT versus reduced dose CT. Radiology 2019;290(2):400–409.
2. Smith TB, Solomon J, Samei E. Estimating detectability index *in vivo*: development and validation of an automated methodology. J Med Imaging (Bellingham) 2018;5(3):031403.
3. 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(3):777–787.
4. Pooler BD, Lubner MG, Kim DH, et al. Prospective evaluation of reduced dose computed tomography for the detection of low-contrast liver lesions: direct comparison with concurrent standard dose imaging. Eur Radiol 2017;27(5):2055–2066.
5. Fält T, Söderberg M, Hörberg L, et al. Simulated dose reduction for abdominal CT with filtered back projection technique: effect on liver lesion detection and characterization. AJR Am J Roentgenol 2019;212(1):84–93.
6. Yan C, Xu J, Liang C, et al. Radiation dose reduction by using CT with iterative model reconstruction in patients with pulmonary invasive fungal infection. Radiology 2018;288(1):285–292.
7. Niu YT, Mehta D, Zhang ZR, et al. Radiation dose reduction in temporal bone CT with iterative reconstruction technique. AJNR Am J Neuroradiol 2012;33(6):1020–1026.
8. Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. RadioGraphics 2017;37(7):2113–2131.
9. Mohammadinejad P, Mileto A, Yu L, et al. CT noise-reduction methods for lower-dose scanning: strengths and weaknesses of iterative reconstruction algorithms and new techniques. RadioGraphics 2021;41(5):1493–1508.
10. Szczykutowicz TP, Nett B, Cherkezyan L, et al. Protocol optimization considerations for implementing deep learning CT reconstruction. AJR Am J Roentgenol 2021;216(6):1668–1677.
11. Jensen CT, Liu X, Tamm EP, et al. Image quality assessment of abdominal CT by use of new deep learning image reconstruction: initial experience. AJR Am J Roentgenol 2020;215(1):50–57.
12. 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(8):5533–5543.
13. Brady SL, Trout AT, Somasundaram E, Anton CG, Li Y, Dillman JR. Improving image quality and reducing radiation dose for pediatric CT by using deep learning reconstruction. Radiology 2021;298(1):180–188.
14. Hsieh J, Liu E, Nett B, Tang J, Thibault J, Sahney S. A new era of image reconstruction: TrueFidelity—Technical white paper on deep learning image reconstruction. GE Healthcare Systems. https://www.gehealthcare.com/-/jssmedia/040dd213fa89463287155151fdb01922.pdf. Published July 2019. Accessed March 30, 2021.
15. Parakh A, Cao J, Pierce TT, Blake MA, Savage CA, Kambadakone AR. Sinogram-based deep learning image reconstruction technique in abdominal CT: image quality considerations. Eur Radiol 2021;31(11):8342–8353.
16. Lyu P, Neely B, Solomon J, et al. Effect of deep learning image reconstruction in the prediction of resectability of pancreatic cancer: diagnostic performance and reader confidence. Eur J Radiol 2021;141:109825.
17. Nakamura Y, Higaki T, Tatsugami F, et al. Deep learning-based CT image reconstruction: initial evaluation targeting hypovascular hepatic metastases. Radiol Artif Intell 2019;1(6):e180011.
18. Cheng Y, Smith TB, Jensen CT, Liu X, Samei E. Correlation of algorithmic and visual assessment of lesion detection in clinical images. Acad Radiol 2020;27(6):847–855.
19. McCollough C, Bakalyar DM, Bostani M, et al. Use of water equivalent diameter for calculating patient size and size-specific dose estimates (SSDE) in CT: the report of AAPM Task Group 220. AAPM Rep 2014;2014:6–23.
20. American College of Radiology. Executive Summary Report (Jul-Dec 2020). Dose Index Registry. Reston, Va: American College of Radiology, 2020.
21. Jensen CT, Blair KJ, Wagner-Bartak NA, et al. Comparison of abdominal computed tomographic enhancement and organ lesion depiction between weight-based scanner software contrast dosing and a fixed-dose protocol in a tertiary care oncologic center. J Comput Assist Tomogr 2019;43(1):155–162.
22. Goodenberger MH, Wagner-Bartak NA, Gupta S, et al. Computed tomography image quality evaluation of a new iterative reconstruction algorithm in the abdomen (adaptive statistical iterative reconstruction-V) a comparison with model-based iterative reconstruction, adaptive statistical iterative reconstruction, and filtered back projection reconstructions. J Comput Assist Tomogr 2018;42(2):184–190.
23. 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(9):3961–3971.
24. Benz DC, Benetos G, Rampidis G, et al. Validation of deep-learning image reconstruction for coronary computed tomography angiography: Impact on noise, image quality and diagnostic accuracy. J Cardiovasc Comput Tomogr 2020;14(5):444–451.
25. Kim I, Kang H, Yoon HJ, Chung BM, Shin NY. Deep learning-based image reconstruction for brain CT: improved image quality compared with adaptive statistical iterative reconstruction-Veo (ASIR-V). Neuroradiology 2021;63(6):905–912.
26. Racine D, Becce F, Viry A, et al. Task-based characterization of a deep learning image reconstruction and comparison with filtered back-projection and a partial model-based iterative reconstruction in abdominal CT: a phantom study. Phys Med 2020;76:28–37.
27. Racine D, Brat HG, Dufour B, et al. Image texture, low contrast liver lesion detectability and impact on dose: Deep learning algorithm compared to partial model-based iterative reconstruction. Eur J Radiol 2021;141:109808.

28. Greffier J, Hamard A, Pereira F, et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. Eur Radiol 2020;30(7):3951–3959.
29. Abdalla EK, Vauthey JN, Ellis LM, et al. Recurrence and outcomes following hepatic resection, radiofrequency ablation, and combined resection/ablation for colorectal liver metastases. Ann Surg 2004;239(6):818–825; discussion 825–827.
30. Choti MA, Sitzmann JV, Tiburi MF, et al. Trends in long-term survival following liver resection for hepatic colorectal metastases. Ann Surg 2002;235(6):759–766.
31. Brenner DJ, Shuryak I, Einstein AJ. Impact of reduced patient life expectancy on potential cancer risks from radiologic imaging. Radiology 2011;261(1):193–198.
32. Kanal KM, Butler PF, Sengupta D, Bhargavan-Chatfield M, Coombs LP, Morin RL. U.S. diagnostic reference levels and achievable doses for 10 adult CT examinations. Radiology 2017;284(1):120–133.