



Influence of deep learning image reconstruction and adaptive statistical iterative reconstruction-V on coronary artery calcium quantification

Yiran Wang¹, Hefeng Zhan¹, Jiameng Hou¹, Xueyan Ma¹, Wenjie Wu¹, Jie Liu¹, Jianbo Gao¹, Ying Guo², Yonggao Zhang¹

¹Department of Radiology, the First Affiliated Hospital of Zhengzhou University, Zhengzhou, China; ²GE Company, Beijing, China

Contributions: (I) Conception and design: Y Zhang, Y Wang; (II) Administrative support: J Gao; (III) Provision of study materials or patients: W Wu; (IV) Collection and assembly of data: Y Wang; H Zhan (V) Data analysis and interpretation: Y Wang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yonggao Zhang, PhD. Department of Radiology, the First Affiliated Hospital of Zhengzhou University, No. 1 East Jianshe Road, Zhengzhou 450052, China. Email: zyg01578@126.com.

Background: Deep learning image reconstruction (DLIR) and adaptive statistical iterative reconstruction-V (ASIR-V) has been used for cardiac computed tomography imaging. However, DLIR and ASIR-V may influence the quantification of coronary artery calcification (CAC).

Methods: CT images of 96 patients were reconstructed using filtered back projection (FBP), ASIR-V 50%, and three levels of DLIR [low (L), medium (M), and high (H)]. Image noise and the Agatston, volume, and mass scores were compared between the reconstructions. Patients were stratified into six Agatston score-based risk categories and five CAC percentile risk categories adjusted by Agatston score, age, sex, and race. The number of patients who were switched to another risk stratification group when ASIR-V and DLIR were used was compared. Bland-Altman plots were used to present the agreement of Agatston scores between FBP and the different reconstruction techniques.

Results: Compared to that with FBP, image noise was significantly decreased with ASIR-V 50%, and DLIR-L, -M, and -H (all $P < 0.001$). The Agatston, volume, and mass scores with ASIR-V 50% and DLIR-L, -M, and -H showed significant decreases in comparison to those calculated with FBP (all $P < 0.001$). Severity classification showed no significant differences between the five reconstruction techniques in any of the CAC score-based risk categories (all $P > 0.05$).

Conclusions: DLIR and ASIR-V show great potential for improving CT image quality, and appear to have no pronounced impact on CAC quantification or Agatston score-based risk stratification.

Keywords: Coronary artery disease; calcium score; deep learning image reconstruction (DLIR); adaptive statistical iterative reconstruction-V; risk stratification

Submitted Sep 24, 2021. Accepted for publication Dec 06, 2021.

doi: 10.21037/atm-21-5548

View this article at: <https://dx.doi.org/10.21037/atm-21-5548>

Introduction

Coronary artery calcium (CAC) quantification in non-contrast CT examination is a robust tool for stratifying cardiovascular risk. It has high predictive value for incident cardiovascular events in asymptomatic patients (1-3), especially those with intermediate cardiovascular risk. The

presence of CAC, a manifestation of plaque stabilization, increases the incidence of acute cardiovascular events independent of traditional risk factors (4,5). Predicting acute events based on the Agatston score produces higher accuracy than based on the Framingham risk factors (1). Coronary calcium scoring has become a common method

for monitoring the progression of coronary atherosclerotic plaques, and it is commonly used in patients who are ineligible for enhanced examinations.

Recent innovations in CT techniques have enabled reconstruction algorithms supplied by various vendors and by lowering the tube voltage/current to be extensively investigated in coronary CT angiography (CTA) (6,7). However, the reference protocol for CAC acquisition still requires 120 kVp presetting and filter back projection (FBP) reconstruction (8), and there are considerable concerns about medical radiation exposure and its resultant cancer risk (9,10).

Iterative reconstruction (IR) algorithms (11-13), which are available commercially from various vendors, can maintain adequate image quality while allowing for a low radiation dose. Adaptive statistical iterative reconstruction-V is partial model-based technique that overcomes the computational performance limitations of full-model iterative reconstruction (Veo; GE Healthcare) and reduces processing time, it has been suggested with modeling of objects and physics but pays less attention to the modelling of system optics, which is commonly used by GE company (13-15). However, these algorithms can affect the noise amplitude and texture features of images, particularly with high-strength IR (15,16).

In recent years, artificial intelligence (AI) has shown significant application promise in cardiovascular imaging to detect coronary calcium accurately and of great efficiency (17-19). Deep learning image reconstruction (20) (DLIR, TrueFidelity, GE Healthcare) based on a convolutional neural network (CNN), The features extracted by the CNN are learned by training the network rather than manual design, so that the CNN can extract more key features and achieve greater performance (19). DLIR has shown great potential in improving CT image reconstruction, and has been verified to reduce image noise by some phantom (21) and clinical (22,23) studies. Simultaneously, the image texture of DLIR is similar to that of high-dose FBP (21). However, the influence of such new reconstruction algorithms on CAC quantification is still unclear.

In the current study, we aimed to investigate the effects of DLIR and adaptive statistical IR-V (ASIR-V) on the quantitative evaluation of coronary calcification and subsequent cardiovascular risk stratification, and to compare image quality of DLIR and ASIR-V with those of FBP. We present the following article in accordance with the MDAR checklist (available at <https://dx.doi.org/10.21037/atm-21-5548>).

Methods

Study population

This retrospective study was approved by the Human Research Ethics Committee of the First Affiliated Hospital of Zhengzhou University (2021-KY-0043-002). All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). All patients included in the study signed a written informed consent form. A total of 108 patients who underwent coronary CTA due to suspected coronary artery disease were enrolled. Before CTA examination, all patients underwent electrocardiogram (ECG)-gated non-contrast CT CAC scan. Of the enrollees, 12 patients were excluded for the following reasons: coronary stent implantation (n=7), previous coronary artery bypass grafting (n=3), and prosthetic valve replacement (n=2). Ultimately, 96 patients were included in the study.

CT acquisition

All scans were acquired with a 256-slice CT scanner (Revolution CT, GE Healthcare, Waukesha, WI, USA). Patients with a resting heart rate >70 beats/min were given oral beta-blockers unless contraindications existed. Calcium scoring scans were performed using prospective ECG-gating, triggered at 75% of the R-R interval with the following acquisition parameters: a tube voltage of 120 kV, an automatic tube current modulation range of 100–720 mA, a noise index of 12, a collimation of 256 mm × 0.625 mm, a gantry rotation time of 280 ms, a scan coverage of 140 or 160 mm, and a display field of view of 25 mm.

The volume CT dose index ($CTDI_{vol}$) and dose length product (DLP) were recorded in the dose report. The effective radiation dose (ED) was estimated by multiplying the DLP by a body region-specific conversion coefficient ($k=0.014 \text{ mSv} \times \text{mGy}^{-1} \times \text{cm}^{-1}$).

Image reconstruction and analysis

Calcium score scanning images were reconstructed by standard FBP, ASIR-V at a level of 50%, and three strength levels of DLIR [low (L), medium (M), and high (H), respectively], with a section thickness of 2.5 mm and without padding.

Analysis of objective image quality parameters was performed on an offline workstation (Advantage Workstation Version 4.7, GE Healthcare). The signal and

Table 1 Baseline characteristics and radiological parameters of the study participants

Characteristics	Values
No. of patients	96
Age (y), mean \pm SD	65.3 \pm 10.2
Male, n (%)	58 (60.0)
BMI (kg/m ²), mean \pm SD	23.9 \pm 3.3
Hypertension, n (%)	37 (39.0)
Smoker, n (%)	19 (20.0)
Dyslipidemia, n (%)	53 (55.0)
Diabetes, n (%)	26 (27.0)
Mean heart rate (bpm), mean \pm SD	64.3 \pm 9.5
CTDI _{vol} (mGy), mean \pm SD	11.2 \pm 3.1
DLP (mGy·cm ⁻¹), mean \pm SD	175.8 \pm 48.8
ED (mSv), mean \pm SD	2.4 \pm 0.6

y, years; SD, standard deviation; BMI, body mass index; bpm, beats per minute; CTDI_{vol}, the volume CT dose index; DLP, dose-length product, ED, effective radiation dose.

image noise were measured as the mean attenuation value in Hounsfield units (HU) and standard deviation (noise) in the aortic root and the left ventricle, using a 0.5-cm² region of interest (ROI). Compare Viewer (GE Healthcare) was used to ensure that all images were at the same level. The signal-to-noise ratio (SNR) was calculated as the ratio of signal to noise.

CAC scoring was initially performed by a diagnostic cardiovascular physician with 3 years' experience using commercially available semiautomatic software (SmartScore 4.0, GE Healthcare). Calcified lesions were detected using this application. All pixels with a density equal to or greater than a fixed threshold (130 HU) and with an area of >1 mm² were marked automatically. Then, the calcified lesions were identified manually on the basis of their anatomical locations. Finally, the Agatston score, lesion volume (mm³), and mass (mg) were calculated automatically. All CAC scores were reconfirmed by a senior radiologist with 10 years' experience in cardiovascular imaging.

Patients were classified into the following five CAC score risk categories according to their Agatston score, as determined by FBP: very low risk (0), low risk [1–10], moderate risk [11–100], high risk [101–400], and very high risk (>400). According to the Multi-Ethnic Study of Atherosclerosis (MESA) (24), patients were grouped by

age, sex, and race into five Agatston score percentile-based risk categories (0–25, 26–50, 51–75, 76–90, and >90). The number of patients who were switched to another risk stratification group when ASIR-V and DLIR were used was recorded.

The Agatston, volume, and mass scores obtained by ASIR-V and different levels of DLIR (low, medium and high) algorithms and subsequent risk reclassification rates will be compared and the differences with the results obtained by the standard FBP algorithm will be determined.

Statistical analysis

Statistical analyses were performed using the SPSS 22.0 statistical software (SPSS, Chicago, IL, USA). Data were tested for normal distribution using the Shapiro-Wilk test. Continuous variables were expressed as mean \pm standard deviation or as median and interquartile range (IQR), whereas categorical variables were expressed as percentage and count. McNemar's test was used to assess the Agatston score classes and the Agatston score percentile-based risk categories. Noise measurements, the SNR, and CAC variables were compared using the Wilcoxon signed-rank test as well as the Bonferroni method for multiple comparisons. Bland-Altman plots were used to present the agreement of Agatston scores between FBP and the different reconstruction techniques. A P value of ≤ 0.05 indicated a significant difference.

Results

Patient characteristics

Baseline characteristics and radiological parameters of the included patients are listed in *Table 1*. The study population comprised 96 patients (60% males) with a mean age of 65.3 \pm 10.2 years, a mean body mass index (BMI) of 23.9 \pm 3.3 kg/m².

Image noise

In comparison to that with FBP, the mean image noise in the aortic root was decreased by 26.7% \pm 3.4% with ASIR-V 50% (used widely in clinical), and by 27.6% \pm 8.3%, 37.3% \pm 7.2%, and 46.4% \pm 5.4% with the three levels of DLIR (DLIR-L, -M, and -H, respectively), while the mean image noise in the left ventricular cavity decreased by 25.9% \pm 3.2% with ASIR-V 50%, and by 23.7% \pm 9.2%,

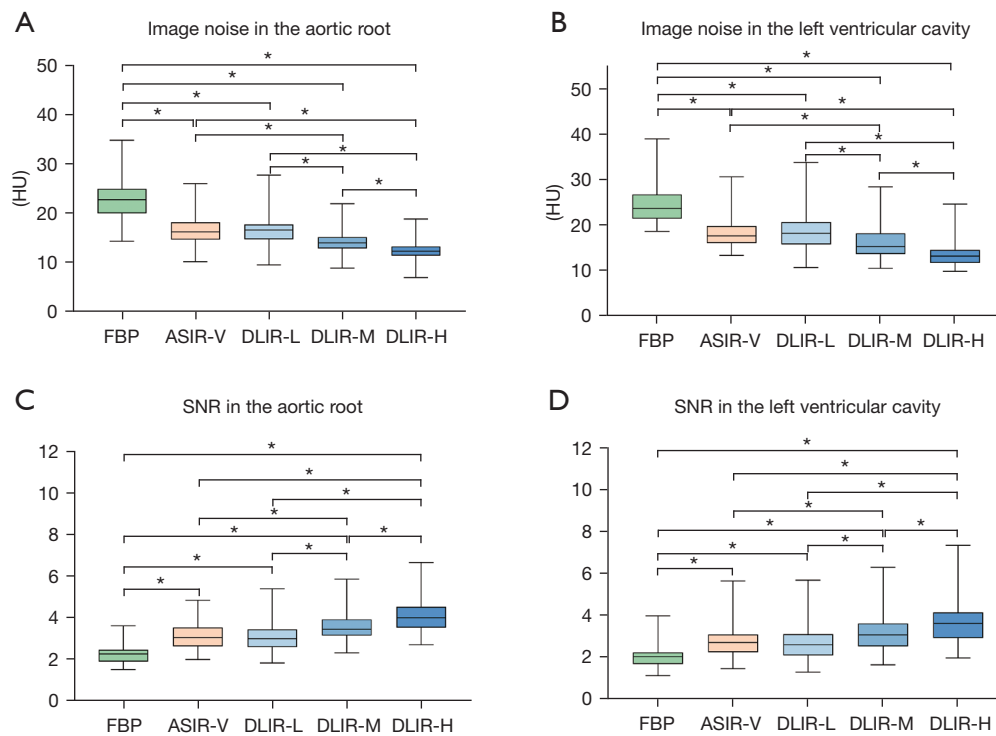


Figure 1 Box-and-whisker plots of quantitative comparisons between FBP, ASIR-V, and three levels of DLIR. (A) Image noise in the aortic root; (B) image noise in the left ventricular cavity; (C) SNR in the aortic root; (D) SNR in the left ventricular cavity. The figure shows data including the median, IQR between the 25th and 75th percentiles (box), and 1.5 IQR range (whisker). *, $P < 0.05$ between the reconstructions. FBP, filtered back projection; SNR, signal-to-noise ratio; IQR, interquartile range; ASIR-V, adaptive statistical iterative reconstruction-V; DLIR-L, DLIR-M, and DLIR-H: deep learning image reconstruction at low, medium, and high levels, respectively.

$34.6\% \pm 7.7\%$, and $44.9\% \pm 5.7\%$ with DLIR-L, M, and H, respectively (all $P < 0.05$).

The image noise was significantly decreased with DLIR-M and DLIR-H as compared to ASIR-V 50% ($P < 0.05$). However, no significant difference in image noise was observed between DLIR-L and ASIR-V 50% ($P > 0.05$). As shown in *Figures 1, 2*, there were significant differences in image noise between the three levels of DLIR (all $P < 0.05$), with DLIR-H showing the greatest noise reduction. The mean signal in the aortic root and the left ventricle remained constant across the various groups ($P > 0.05$). Consequently, with DLIR, the SNRs of the aortic root and left ventricular cavity showed significant rises compared to those measured using FBP; these rises increased with the strength of DLIR (aortic root: from 3.1 ± 0.7 to 4.3 ± 1.1 with DLIR-L, DLIR-M and DLIR-H, respectively; left ventricular cavity: from 2.8 ± 0.8 to 3.7 ± 1.2 with DLIR-L, DLIR-M and DLIR-H, respectively (all $P < 0.05$). The SNR also showed a significant rise with both DLIR-M and

DLIR-H compared to ASIR-V 50%, and with ASIR-V 50% compared to FBP ($P < 0.05$, in both the aortic root and left ventricular cavity). Data relating to the mean CT values and SNRs in the aortic root and the left ventricle can be found in *Table S1*. *Table 2* displays the data of relative differences (expressed as mean percentages \pm SD) between the targeted and referenced reconstruction groups.

CAC scoring

Patients with an Agatston score = 0 in FBP had the same score using ASIR-V and DLIR. Therefore, to compare the Agatston score, calcium volume, and CAC mass between the five reconstruction techniques, 89 patients with an Agatston score of > 1 were included, and the data are shown in *Table 3*.

The median Agatston calcium score was 199 (IQR, 32–451), 188 (IQR, 26–422), 196 (IQR, 30–426), 192.0 (IQR, 28–419), and 185 (IQR, 28–408) with FBP, ASIR-V

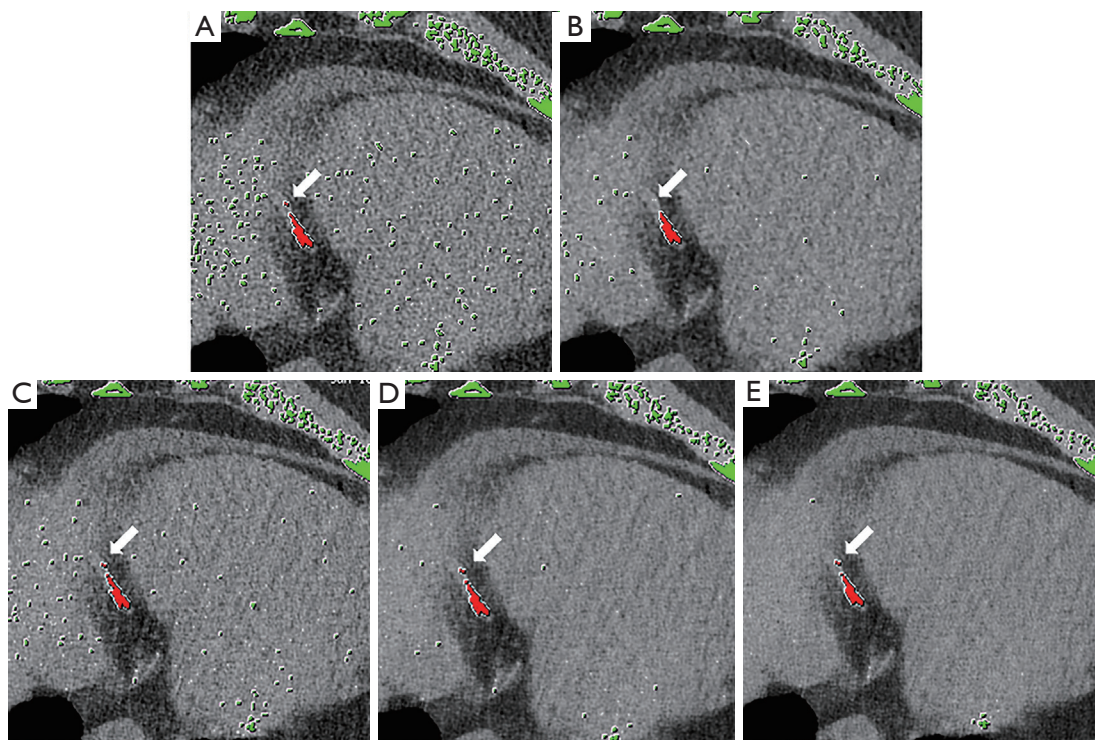


Figure 2 Axial cardiac CT images of a 75-year-old man (body mass index, 26.5 kg/m²) reconstructed using FBP, ASiR-V 50%, and three levels of DLIR (DLIR-L, M, and H, respectively). The section thickness was 2.5 mm. Image noise was 31.1, 23, 26.2, 20.9, and 18.1 HU in the left ventricle from (A) to (E). At the level of the distal segment of the RCA, calcified lesions were marked manually in red dots. Scattered false-positive dots were marked in green automatically on FBP (A), ASiR-V 50% (B), and DLIR (with incremental strength levels) (C-E) images. One calcified plaque could not be detected in the ASiR-V 50% group (white arrows). FBP, filtered back projection; ASiR-V, adaptive statistical iterative reconstruction-V; RCA, right coronary artery; DLIR-L, DLIR-M, and DLIR-H: deep learning image reconstruction at low, medium, and high levels, respectively.

Table 2 Quantitative analysis of image noise and SNR for the reconstruction techniques

Reconstruction	Quantitative image quality metrics (%)			
	Noise in AR (HU)	SNR in AR	Noise in LV (HU)	SNR in LV
ASiR-V 50% vs. FBP	-26.7±3.4*	36.5±5.8*	-25.9±3.2*	35±5.5*
DLIR-L vs. FBP	-27.6±8.3*	39.9±16.1*	-23.7±9.2*	32.1±18.3*
DLIR-M vs. FBP	-37.3±7.2*	61.3±17.7*	-34.6±7.7*	54.3±18.4*
DLIR-H vs. FBP	-46.4±5.4*	88.2±21.7*	-44.9±5.7*	82.5±19.9*
DLIR-L vs. ASiR-V 50%	-1.4±9.7	2.4±10.1	2.9±10.9	-2.3±11.6
DLIR-M vs. ASiR-V 50%	-14.6±8.7*	18.1±11.4*	-11.8±8.5*	14.1±11.7*
DLIR-H vs. ASiR-V 50%	-26.9±6.4*	37.8±13.5*	-25.7±6.4*	35.1±11.2*

Data given are mean relative differences ± standard deviation (SD) (expressed in percentages) between the targeted and the referenced reconstruction groups. *, P<0.05. SNR, signal-to-noise ratio; AR, aortic root, LV, left ventricular cavity; FBP, filtered back projection; ASiR-V 50%, adaptive statistical iterative reconstruction-V at 50% intensity; DLIR-L, deep learning image reconstruction at low level; DLIR-M, deep learning image reconstruction at medium level; DLIR-H, deep learning image reconstruction at high level; HU, Hounsfield units.

Table 3 Agatston, volumetric, and mass scores using the different reconstruction techniques

Reconstruction technique	FBP	ASIR-V	DLIR-L	DLIR-M	DLIR-H
Agatston score					
Mean	328.8±416.1	314.3±407.2	321.9±410.9	315.6±402.1	307.5±392.8
Median	199 (32, 451)	188 (26, 422)	196 (30, 426)	192 (28, 419)	185 (28, 408)
Volumetric score (mm ³)					
Mean	132.8±158.9	126.3±154.8	122.6±149.6	119.6±146.7	116.8±143.1
Median	99 (20.8, 179.5)	95 (15, 165)	95.5 (15.8, 173)	93 (15, 169.8)	91 (14, 164.8)
Mass score (mg)					
Mean	52.9±72.8	51.3±71.7	52.3±72.8	51.6±71.6	51.0±70.9
Median	32 (7, 63)	29 (6, 61)	31 (6.5, 63)	30 (6.5, 59)	30 (6.5, 59)

Data given is mean ± standard deviation (SD) or interquartile range (IQR). FBP, filtered back projection; ASIR-V 50%, adaptive statistical iterative reconstruction-V at 50% intensity; DLIR-L, deep learning image reconstruction at low level; DLIR-M, deep learning image reconstruction at medium level; DLIR-H, deep learning image reconstruction at high level.

50%, DLIR-L, DLIR-M, and DLIR-H, respectively. The relative differences for ASIR-V 50%, DLIR-L, DLIR-M, and DLIR-H to FBP were 5.5%, 1.5%, 3.5%, and 7.0%, respectively (all $P < 0.001$). The Agatston score was higher with DLIR-L and smaller with DLIR-H than with ASIR-V 50% (both $P < 0.001$). However, no difference was observed between the Agatston score of images reconstructed with DLIR-M and ASIR-V 50% ($p = 0.23$). However, the Agatston score showed significant differences between the three DLIR reconstruction levels (all $P < 0.05$).

The median CAC volumes on ASIR-V and DLIR images differed significantly from those on FBP images {99 [20.8, 179.5], 95 [15, 165], 95.5 [15.8, 173], 93 [15, 169.8] vs. 91 [14, 164.8] mm³}. Post-hoc analysis showed significant differences between the five techniques (all $P < 0.05$). The median mass scores were 32 [7, 63], 29 [6, 61], 31 [6.5, 63], 30 [6.5, 59], and 30 [6.5, 59] mg for FBP, ASIR-V 50%, DLIR-L, DLIR-M, and DLIR-H, respectively. Post-hoc analysis revealed that ASIR-V and three DLIR reconstruction levels differed significantly from FBP in mass score (all $P < 0.05$). However, DLIR-M and DLIR-H exhibited no significant difference compared to ASIR-V 50% in mass score ($P = 0.186$ and 0.858 , respectively). The mass score was higher with DLIR-L than with ASIR-V ($P < 0.05$), but no significant difference in mass score was observed between the three DLIR reconstruction levels (all $P > 0.05$).

Bland-Altman plots comparing the Agatston score with the mean and limits of agreement for images reconstructed with ASIR-V and DLIR with that for images reconstructed

with FBP are shown in *Figure 3*.

In the risk classification analysis, 5 (5.6%), 3 (3.4%), 5 (5.6%), and 5 (5.6%) patients with a CAC score of ≥ 1 were reassigned to another risk category when ASIR-V 50%, DLIR-L, DLIR-M, and DLIR-H were used, respectively. No patients moved by two or more risk categories. The detailed data are available in *Table 4*. Severity classification revealed no significant differences between the reconstruction techniques (all $P > 0.05$).

Furthermore, no significant differences were observed between FBP, ASIR-V 50%, and the different levels of DLIR when the CAC percentile risk categories were adjusted for Agatston score, age, sex, and race (all $P > 0.05$). Eighty-five patients aged between 45 and 85 years old were included according to the MESA study. When CAC scores were acquired with ASIR-V 50% and DLIR L, M, and H, a change in risk category was observed in 4 (4.7%), 3 (3.5%), 4 (4.7%), and 5 (5.9%) patients, respectively (*Table 5*). For one 57-year-old patient, the Agatston score decreased from 1 (low risk, 50–75th percentile) with FBP to 0 (very low risk, 0–25th percentile) with ASIR-V 50%, DLIR-M, and DLIR-H; however, the patient's Agatston score with DLIR-L stayed the same.

Discussion

Our study demonstrated that compared to the standard FBP algorithm, DLIR and ASIR-V reduced the Agatston score, calcium mass, and calcium volume. However, although the Agatston score decreased significantly with ASIR-V and

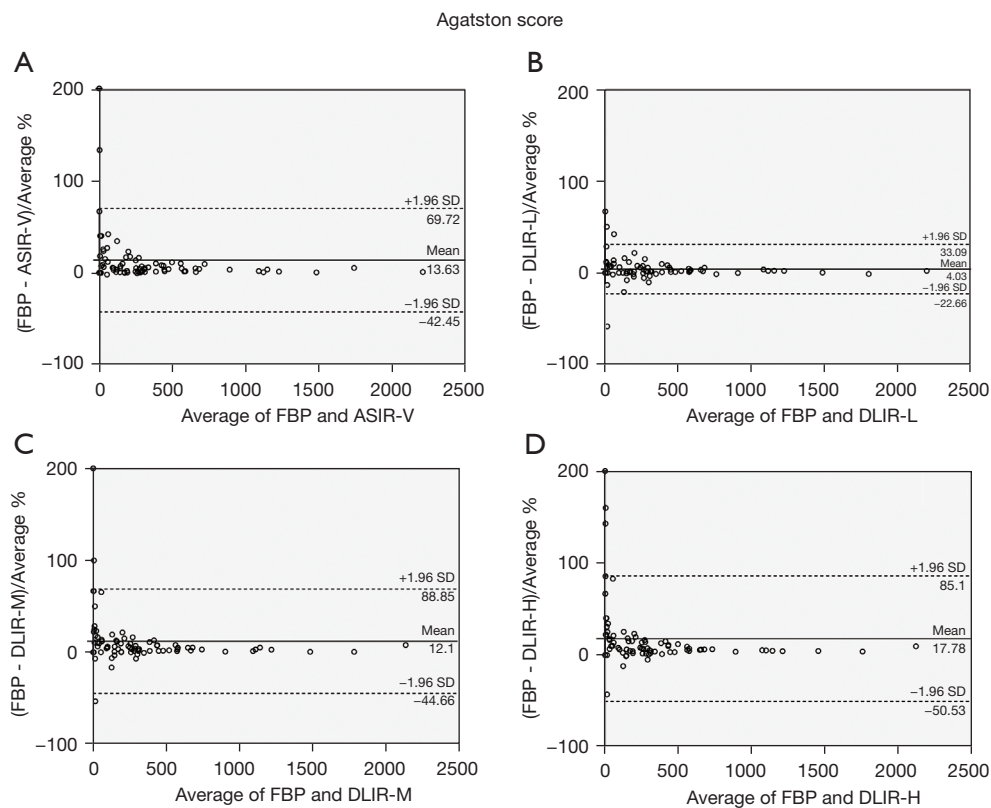


Figure 3 Bland-Altman plots comparing all Agatston scores measured using ASIR-V 50% and three levels of DLIR (DLIR-L, M, and H, respectively) and using FBP. The limit of agreement and mean difference for the Agatston scores are shown in (A-D). FBP, filtered back projection; ASIR-V, adaptive statistical iterative reconstruction-V; DLIR-L, DLIR-M, and DLIR-H: deep learning image reconstruction at low, medium, and high levels, respectively.

increasing DLIR strength levels, no significant differences were observed in cardiovascular risk stratification based on Agatston scores calculated using the different reconstruction techniques. Furthermore, DLIR and ASIR-V also resulted in significant noise reduction, but the signal remained constant, thus improving the SNR.

The improvement of image quality (for instance, by reducing image noise) has been reported to result in a decrease in CAC score (25-28). However, the widespread existence of image noise may hinder the dependable quantification of CAC. Several studies have confirmed the increasing value of CAC score to informing cardiovascular risk stratification and treatment decisions, including downstream detection (6). With the introduction of the CAC data and reporting system (CAC-DRS) (29), an expert consensus published by the Cardiovascular Computed Tomography Society, correct and reproducible CAC acquisition has become essential to make clinical follow-up

accurate (30).

Tube voltage, tube current, slice thickness and reconstruction technique are the main factors influencing CAC quantification (14). Several studies have shown that lower tube voltage led to lower CAC score (31,32), while appropriately lowering tube current helps to reduce radiation dose with little risk reclassifications (7). Thinner slice thickness improved the detection of small calcifications and increased noise-level, which may result in more false-positives (33). Also, multiple studies have reported varying impact of iterative reconstruction on CAC quantification and risk stratification (25-28).

The Agatston score is widely used in quantitative assessment of CAC under low tube voltage or tube currents. Considering the ultimate goal to implement advanced image reconstruction is to reduce radiation exposure, Gräni *et al.* attempted to reduce the tube voltage to lower levels (70 and 80 kVp) while applying tube-adapted thresholds to

Table 4 Effects of ASIR-V 50%, DLIR-L, DLIR-M, and DLIR-H on Agatston score-based risk categories

Severity class with FBP	Severity classification with ASIR-V and DLIR														
	ASIR-V 50%			DLIR-L			DLIR-M			DLIR-H					
	Very low	Low	Moderate	High	Very high	Very low	Low	Moderate	High	Very high	Very low	Low	Moderate	High	Very high
Very low (n=7)	7	0	0	0	0	7	0	0	0	0	7	0	0	0	0
Low (n=12)	1*	11	0	0	0	1*	11	0	0	0	1*	11	0	0	0
Moderate (n=15)	0	2*	13	0	0	2*	13	0	0	0	0	2*	13	0	0
High (n=32)	0	0	1*	31	0	0	0	1*	31	0	0	0	1*	31	0
Very high (n=30)	0	0	0	1*	29	0	0	0	30	0	0	0	1*	29	0
Total (n=96)	8	13	14	32	29	7	14	14	31	30	8	13	14	32	29

*, numbers indicate patients who changed severity classification with the use of ASIR-V and DLIR. mod., moderate; ASIR-V 50%, adaptive statistical iterative reconstruction-V at 50% intensity; DLIR-L, DLIR-M, and DLIR-H: deep learning image reconstruction at low, medium, and high levels, respectively.

Table 5 Effects of ASIR-V 50%, DLIR-L, DLIR-M, and DLIR-H on CAC percentile risk categories adjusted by Agatston score, age, sex, and race

Severity class with FBP	Severity classification with ASIR-V and DLIR														
	ASIR-V 50%			DLIR-L			DLIR-M			DLIR-H					
	<25	26-50	51-75	76-90	>90	<25	26-50	51-75	76-90	>90	<25	26-50	51-75	76-90	>90
<25 (n=9)	9	0	0	0	0	9	0	0	0	0	9	0	0	0	0
26-50 (n=7)	0	7	0	0	0	0	7	0	0	0	0	7	0	0	0
51-75 (n=17)	1*	0	16	0	0	0	0	17	0	0	1*	0	16	0	0
76-90 (n=27)	0	0	2*	25	0	0	0	2*	25	0	0	0	2*	25	0
>90 (n=25)	0	0	0	1*	24	0	0	0	1*	24	0	0	0	1*	24
Total (n=89)	10	7	18	26	24	9	7	19	26	24	10	7	18	26	24

*, numbers indicate patients who changed severity classification with use of ASIR-V and DLIR. ASIR-V 50%, adaptive statistical iterative reconstruction-V at 50% intensity; DLIR-L, DLIR-M, and DLIR-H: deep learning image reconstruction at low, medium, and high levels, respectively.

reduce the radiation dose by 80% with little effect on risk classification (31). Nada Sulaiman used ASIR-V combined with low tube current to reduce radiation dose by 26% but enables stable Agatston score (7), while Kamani *et al.* applied low KV combined with ASIR-V introduces marked underestimation of Agatston score and clinically unacceptably high reclassification rates as compared with standard-dose CAC scoring (34).

With the rapid development of multi-slice spiral CT technology, new reconstruction algorithms have continued to appear. Reconstruction algorithms supplied by various vendors can have different effects on the assessment of cardiovascular risk. In their study using the sinogram-affirmed IR algorithm (25), Kurata *et al.* reported that the Agatston score was decreased by 48% compared to that calculated using FBP. Messerli *et al.* (26) also noted significant reductions of 39%, 41%, and 40% in the Agatston, volume, and mass scores, respectively, with high-level ADMIRE. In the present study, the Agatston, volume, and mass scores were decreased by 7%, 8.1%, and 6.3%, respectively, with high-strength DLIR as compared to FBP. Szilveszter *et al.* (27) used data from cardiovascular screening in asymptomatic patients to assess the risk reclassification rate by virtue of an IR technique. They observed that the reclassification rate was moderate (2.4%) for risk categories derived from the CAC; however, 6 out of 63 (9.5%) patients in the symptomatic test cohort were assigned to another risk category. Gebhard *et al.* (28) reported that compared with FBP, reconstruction with increasing ASIR levels significantly reduced the Agatston and CAC volume scores, whereas the CAC mass score remained unchanged. They hypothesized that IR reconstructions made the borders of the lesion smoother and the center appear denser. Contrary to the results observed with ASIR, the decreases in CAC mass score with DLIR and ASIR-V were similar to those observed for Agatston score and CAC volume score in our study, with a small but significant reduction in mean mass score also being observed. The ambiguity of these results should prompt caution toward the adoption of new reconstruction algorithms in clinical practice. In future, a standardized process for different vendors or reconstruction-based CAC value adjustment is needed to minimize the influence of reconstruction algorithms on cardiovascular risk stratification.

At present, the disadvantage of mainstream IR is the resulting change in image texture. Multiple studies (15,16) have suggested that with the increase of IR strength, “blotchy”, “plastic-looking”, or “unnatural” noise texture,

caused by differences in the noise power spectrum of FBP and IR, can appear in images, thus hindering the image interpretation.

As a branch of machine learning and AI, DLIR overcomes the modeling limitations of IR by handling complex models and thousands of parameters during the training process (20). DLIR uses deep neural networks (DNNs) trained with high-quality FBP data to generate images that accurately match the ground truth data, and retains a noise texture similar to that of high-dose FBP images (35). With the increase of reconstruction strength, the noise decreases and the SNR increases significantly. Recently published comparison studies (36,37) of DLIR versus IR in chest plain scan and abdominal multi-phase enhanced scan have reported that DLIR may aid in reducing the radiation dose, with the noise level remaining the same or significantly decreasing.

To the best of our knowledge, the effects of DLIR on CAC score-based cardiovascular risk re-stratification have yet to be reported. In this study, only small differences in CAC score were observed between the reconstruction techniques. Bland–Altman analysis showed that the Agatston scores with ASIR-V and incremental DLIR levels showed good agreement with those calculated with FBP, and there was no systematic bias. We confirmed that compared to FBP, ASIR-V and DLIR can significantly reduce noise while having little effect on Agatston score-based risk assessment or CAC percentile risk categories adjusted for Agatston score, age, sex, and race.

Our study has several limitations. First, this is a single-center study, and we only investigated the effects of specific reconstruction algorithms from one CT vendor. Second, the BMI of the patients in our cohort was low (about 24 kg/m², with small standard deviations) compared to that of Western European cohorts (38). However, we used an automatic tube current technique adapted to BMI to obtain stable image quality parameters for comparison. Third, we did not prospectively use low tube voltage/tube current techniques combined with DLIR in diagnostic accuracy compared to FBP in standard doses. Studies to evaluate the potential of novel reconstruction technology in reducing the radiation dose without affecting calcium quantification are needed in future in order to address the continuously increasing concerns regarding radiation exposure. Finally, the symptomatic patients included in this study were predominantly moderate risk, with the median Agatston score of these patients being 199 and only 7 patients having a CAC score of 0. Therefore, our study has potentially

limited applicability. However, including a large number of patients without CAC would have only added limited information for CAC comparison.

In conclusion, compared with traditional FBP, DLIR and ASIR-V can reduce the CAC score, while having little effect on cardiac risk categorization. As advanced reconstruction methods, DLIR and ASIR-V show great potential in improving CT image quality. DLIR and ASIR-V appear to have no pronounced impact on the quantification of CAC and subsequent risk stratification based on the Agatston score.

Acknowledgments

Funding: The study was supported by grant from Scientific Research Project plan of Henan Province (No. 20A320079).

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at <https://dx.doi.org/10.21037/atm-21-5548>

Data Sharing Statement: Available at <https://dx.doi.org/10.21037/atm-21-5548>

Conflicts of Interest: The authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/atm-21-5548>). The authors report that the Deep Learning Imaging Reconstruction is the latest generation of reconstruction algorithms from GE Company and the study may have implications for the clinical application of this algorithm. YG is an employee of GE Company. The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The questionnaire and methodology for this study was approved by the Human Research Ethics Committee of the First Affiliated Hospital of Zhengzhou University (ethics approval number: 2021-KY-0043-002). All procedures performed in this study involving human participants were in accordance with the Declaration of Helsinki (as revised in 2013). All patients included in the study signed a written informed consent form.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. *N Engl J Med* 2008;358:1336-45
2. Hecht HS. Coronary artery calcium scanning: past, present, and future. *JACC Cardiovasc Imaging* 2015;8:579-96.
3. Greenland P, Blaha MJ, Budoff MJ, et al. Coronary Calcium Score and Cardiovascular Risk. *J Am Coll Cardiol* 2018;72:434-47.
4. Ahmadi A, Argulian E, Leipsic J, et al. From Subclinical Atherosclerosis to Plaque Progression and Acute Coronary Events: JACC State-of-the-Art Review. *J Am Coll Cardiol* 2019;74:1608-17.
5. Budoff MJ, Young R, Burke G, et al. Ten-year association of coronary artery calcium with atherosclerotic cardiovascular disease (ASCVD) events: the multi-ethnic study of atherosclerosis (MESA). *Eur Heart J* 2018;39:2401-8.
6. Tesche C, Duguay TM, Schoepf UJ, et al. Current and future applications of CT coronary calcium assessment. *Expert Rev Cardiovasc Ther* 2018;16:441-53.
7. Sulaiman N, Soon J, Park JK, et al. Comparison of low-dose coronary artery calcium scoring using low tube current technique and hybrid iterative reconstruction vs. filtered back projection. *Clin Imaging* 2017;43:19-23.
8. Hecht HS, Cronin P, Blaha MJ, et al. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. *J Thorac Imaging* 2017;32:W54-66.
9. Messenger B, Li D, Nasir K, et al. Coronary calcium scans and radiation exposure in the multi-ethnic study of atherosclerosis. *Int J Cardiovasc Imaging* 2016;32:525-9.
10. Tatsugami F, Higaki T, Fukumoto W, et al. Radiation dose reduction for coronary artery calcium scoring at

- 320-detector CT with adaptive iterative dose reduction 3D. *Int J Cardiovasc Imaging* 2015;31:1045-52.
11. Halliburton SS, Tanabe Y, Partovi S, et al. The role of advanced reconstruction algorithms in cardiac CT. *Cardiovasc Diagn Ther* 2017;7:527-38.
 12. Pontone G, Muscogiuri G, Andreini D, et al. Impact of a New Adaptive Statistical Iterative Reconstruction (ASIR)-V Algorithm on Image Quality in Coronary Computed Tomography Angiography. *Acad Radiol* 2018;25:1305-13.
 13. Benz DC, Gräni C, Mikulicic F, et al. Adaptive Statistical Iterative Reconstruction-V: Impact on Image Quality in Ultralow-Dose Coronary Computed Tomography Angiography. *J Comput Assist Tomogr* 2016;40:958-63.
 14. Vonder M, van der Werf NR, Leiner T, et al. The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: A systematic review. *J Cardiovasc Comput Tomogr* 2018;12:352-63.
 15. Samei E, Richard S. Assessment of the dose reduction potential of a model-based iterative reconstruction algorithm using a task-based performance metrology. *Med Phys* 2015;42:314-23.
 16. Verdun FR, Racine D, Ott JG, et al. Image quality in CT: From physical measurements to model observers. *Phys Med* 2015;31:823-43.
 17. Lin A, Kolossváry M, Motwani M, et al. Artificial Intelligence in Cardiovascular Imaging for Risk Stratification in Coronary Artery Disease. *Radiol Cardiothorac Imaging* 2021;3:e200512.
 18. Liu P, Wang M, Wang Y, et al. Impact of Deep Learning-based Optimization Algorithm on Image Quality of Low-dose Coronary CT Angiography with Noise Reduction: A Prospective Study. *Acad Radiol* 2020;27:1241-8.
 19. Datong C, Minghui L, Cheng J, et al. Coronary Calcium Detection Based on Improved Deep Residual Network in Mimics. *J Med Syst* 2019;43:119.
 20. Hsieh J, Liu E, Nett B, et al. A New Era of Image Reconstruction: TrueFidelity. White Paper (JB68676XX). GE Healthcare, 2019.
 21. 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;30:3951-9.
 22. 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;14:444-51.
 23. Akagi M, Nakamura Y, Higaki T, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol* 2019;29:6163-71.
 24. McClelland RL, Chung H, Detrano R, et al. Distribution of coronary artery calcium by race, gender, and age: results from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation* 2006;113:30-7.
 25. Kurata A, Dharampal A, Dedic A, et al. Impact of iterative reconstruction on CT coronary calcium quantification. *Eur Radiol* 2013;23:3246-52.
 26. Messerli M, Rengier F, Desbiolles L, et al. Impact of Advanced Modeled Iterative Reconstruction on Coronary Artery Calcium Quantification. *Acad Radiol* 2016;23:1506-12.
 27. Szilveszter B, Elzomor H, Károlyi M, et al. The effect of iterative model reconstruction on coronary artery calcium quantification. *Int J Cardiovasc Imaging* 2016;32:153-60.
 28. Gebhard C, Fiechter M, Fuchs TA, et al. Coronary artery calcium scoring: Influence of adaptive statistical iterative reconstruction using 64-MDCT. *Int J Cardiol* 2013;167:2932-7.
 29. Hecht HS, Blaha MJ, Kazerooni EA, et al. CAC-DRS: Coronary Artery Calcium Data and Reporting System. An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT). *J Cardiovasc Comput Tomogr* 2018;12:185-91.
 30. Oudkerk M, Stillman AE, Halliburton SS, et al. Coronary artery calcium screening: current status and recommendations from the European Society of Cardiac Radiology and North American Society for Cardiovascular Imaging. *Int J Cardiovasc Imaging* 2008;24:645-71.
 31. Gräni C, Vontobel J, Benz DC, et al. Ultra-low-dose coronary artery calcium scoring using novel scoring thresholds for low tube voltage protocols-a pilot study. *Eur Heart J Cardiovasc Imaging* 2018;19:1362-71.
 32. Groen JM, Kofoed KF, Zacho M, et al. Calcium score of small coronary calcifications on multidetector computed tomography: results from a static phantom study. *Eur J Radiol* 2013;82:e58-e63.
 33. Mantini C, Maffei E, Toia P, et al. Influence of image reconstruction parameters on cardiovascular risk reclassification by Computed Tomography Coronary Artery Calcium Score. *Eur J Radiol* 2018;101:1-7.
 34. Kamani CH, Huang W, Lutz J, et al. Impact of Adaptive Statistical Iterative Reconstruction-V on Coronary Artery Calcium Scores Obtained From Low-Tube-Voltage Computed Tomography - A Patient Study. *Acad Radiol* 2020. [Epub ahead of print]. doi: 10.1016/

- j.acra.2020.10.025.
35. Chartrand G, Cheng PM, Vorontsov E, et al. Deep Learning: A Primer for Radiologists. *Radiographics* 2017;37:2113-31.
 36. Kim JH, Yoon HJ, Lee E, et al. Validation of Deep-Learning Image Reconstruction for Low-Dose Chest Computed Tomography Scan: Emphasis on Image Quality and Noise. *Korean J Radiol* 2021;22:131-8.
 37. Cao L, Liu X, Li J, et al. A study of using a deep learning image reconstruction to improve the image quality

Cite this article as: Wang Y, Zhan H, Hou J, Ma X, Wu W, Liu J, Gao J, Guo Y, Zhang Y. Influence of deep learning image reconstruction and adaptive statistical iterative reconstruction-V on coronary artery calcium quantification. *Ann Transl Med* 2021;9(23):1726. doi: 10.21037/atm-21-5548

- of extremely low-dose contrast-enhanced abdominal CT for patients with hepatic lesions. *Br J Radiol* 2021;94:20201086.
38. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative image reconstruction techniques for CT coronary artery calcium quantification: comparison with traditional filtered back projection in vitro and in vivo. *Radiology* 2014;270:387-93.

(English Language Editor: J. Reylonds)

Table S1 Quantitative image analysis of CT value, image noise, and SNR for five reconstruction groups

Reconstruction	Quantitative image quality metrics					
	Signal AR (HU)	Noise AR (HU)	Signal-to-noise ratio in AR	Signal LV (HU)	Noise LV (HU)	Signal-to-noise ratio in LV
FBP	50.8±8.7	22.9±4.6	2.3±0.5	48.4±8.3	24.7±4.8	2.0±0.6
ASIR-V 50%	50.7±8.5	16.8±3.6	3.1±0.7	48.3±8.3	18.3±3.9	2.8±0.8
DLIR-L	50.8±8.7	16.5±3.8	3.2±0.8	48.2±8.4	18.9±4.9	2.7±0.9
DLIR-M	50.7±8.6	14.2±2.7	3.7±0.8	48.2±8.4	16.1±3.7	3.2±1.0
DLIR-H	50.7±8.5	12.3±2.6	4.3±1.1	48.1±8.5	13.6±3.2	3.7±1.2

Data given is mean ± SD. AR, aorta root; LV, left ventricular cavity; FBP, filtered back projection; ASIR-V 50%, adaptive statistical iterative reconstruction-V at 50% intensity; DLIR-L, deep learning image reconstruction at low level; DLIR-M, deep learning image reconstruction at medium level; DLIR-H, deep learning image reconstruction at high level; HU, Hounsfield units.