European Journal of Radiology 141 (2021) 109808

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Image texture, low contrast liver lesion detectability and impact on dose: Deep learning algorithm compared to partial model-based iterative reconstruction

D. Racine [a,*,1], H.G. Brat [b,1], B. Dufour [b], J.M. Steity [c], M. Hussenot [d], B. Rizk [e], D. Fournier [b], F. Zanca [f]

[a] Institute of Radiation Physics (IRA), Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Rue du Grand-Pré 1, 1007 Lausanne, Switzerland
[b] Institut de Radiologie de Sion, Groupe 3R, Rue du scex, 2, 1950 Sion, Switzerland
[c] Centre d'imagerie de la Riviera, Groupe 3R, Rue des Moulins 5B, 1800 Vevey, Switzerland
[d] GE Medical Systems (Schweiz) AG, Europa-Strasse 31, 8152 Glattbrugg, Switzerland
[e] Centre d'Imagerie de Fribourg, Groupe 3R, Rue du Centre 10, 1752 Fribourg, Switzerland
[f] Palindromo Consulting, Willem de Croylaan 51, 3000 Leuven, Belgium

ARTICLE INFO

Keywords:
Computed tomography
Deep learning
Image quality
Model observer
Radiation dose

ABSTRACT

Objectives: To compare deep learning (True Fidelity, TF) and partial model based Iterative Reconstruction (ASiR-V) algorithm for image texture, low contrast lesion detectability and potential dose reduction.
Methods: Anthropomorphic phantoms (mimicking non-overweight and overweight patient), containing lesions of 6 mm in diameter with 20HU contrast, were scanned at five different dose levels (2,6,10,15,20 mGy) on a CT system, using clinical routine protocols for liver lesion detection. Images were reconstructed using ASiR-V 0% (surrogate for FBP), 60 % and TF at low, medium and high strength. Noise texture was characterized by computing a normalized Noise Power Spectrum filtered by an eye filter. The similarity against FBP texture was evaluated using peak frequency difference (PFD) and root mean square deviation (RMSD). Low contrast detectability was assessed using a channelized Hotelling observer and the area under the ROC curve (AUC) was used as figure of merit. Potential dose reduction was calculated to obtain the same AUC for TF and ASiR-V.
Results: FBP-like noise texture was more preserved with TF (PFD from -0.043mm-1 to -0.09mm-1, RMSD from 0.12mm-1 to 0.21mm-1) than with ASiR-V (PFD equal to 0.12 mm-1, RMSD equal to 0.53mm-1), resulting in a sharper image. AUC was always higher with TF than ASIR-V. In average, TF compared to ASiR-V, enabled a radiation dose reduction potential of 7%, 25 % and 33 % for low, medium and high strength respectively.
Conclusion: Compared to ASIR-V, TF at high strength does not impact noise texture and maintains low contrast liver lesions detectability at significant lower dose.

1. Introduction

Technological developments in computed tomography (CT) have transformed patient care for many diseases and clinical scenarios. The importance of CT is reflected in the hundreds of millions of scans performed each year [1]. However, its benefits in disease management come with the price of a very radiant nature – especially in younger patients or those who undergo repeated CT scans. Applying the ALARA (As Low As Reasonably Achievable) principle to appropriate CT examinations remains mandatory, as absence of risk using ionizing radiation has not been demonstrated, even at low dose [2]. CT contributes to most of the population radiation exposure from medical X-ray imaging; for

Abbreviations: ALARA, As Low As Reasonably Achievable; ASiR-V, adaptive statistical iterative reconstruction-V; AUC, area under the ROC curve; BMI, Body Mass Index; CHO, Channelized Hotelling Observer; CI, confidence interval; CNR, contrast-to-noise ratio; CT, computed tomography; CTDIvol, CT dose index; DDoG, difference of Gaussian; DLIR, deep learning iterative reconstruction; IR, iterative reconstruction; nNPSe, normalized noise power spectrum (NPS) with an eye filter; NI, Noise Index; PFD, peak frequency difference; RMSD, root mean square deviation; TF, True Fidelity.

* Corresponding author.
E-mail address: damien.racine@chuv.ch (D. Racine).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ejrad.2021.109808
Received 19 January 2021; Received in revised form 12 April 2021; Accepted 30 May 2021
Available online 3 June 2021

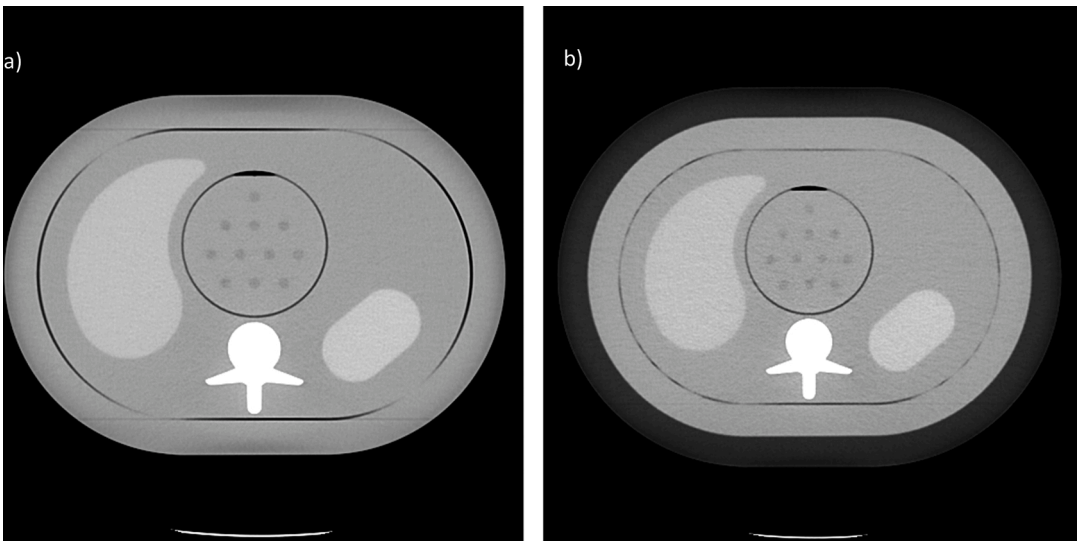

Fig. 1. CT image of the anthropomorphic QRM medium (a) and large (b) abdomen phantom acquired at 20 mGy and reconstructed with True Fidelity at high level. The image shown is at the slice-level of the 8 mm lesions, obtained by averaging 10 series to increase signal to noise ratio for visualization purpose.

example, in Switzerland, CT examinations represent only 10 % of studies, but 74 % of annual exposure to medical radiation [3].

CT manufacturers have implemented automatic tube current modulation, bow tie filters, lower tube tension (kVp), adaptive collimation and iterative reconstruction algorithms to reduce dose and optimize image quality [4–6]. Currently, iterative reconstruction (IR) is one of the most used, enabling substantial noise reduction compared to filtered back projection for high contrast diagnostic tasks, and allowing substantial dose reduction [7–11]. However, for low contrast tasks the advantage of using IR is limited, because of induced noise texture changes resulting in blurred images when compared to its FBP reconstruction [7,12–14]. In addition, some studies have reported that image quality and the evaluation of specific anatomical regions could be compromised by the use of IR algorithms [15,16]. Indeed, IR algorithms change noise texture but also noise amplitude, resulting in a false high-quality visual impression with potentially massive dose reduction, especially if detectability of small low contrast objects are not taken into account [12].

Recently, deep learning iterative reconstruction (DLIR) algorithms have been developed, with a proven potential of considerable noise reduction without changing the typical noise texture of FBP images, therefore allowing dose reduction obtained by IR algorithms whilst preserving FBP image quality [17–19].

Human observer studies are often used to evaluate the dose reduction potential especially when a new algorithm is used like DLIR. However, these studies are time consuming and require large samples to obtain precise results (inter-observer variability and intra-observer variability is large). To overcome these limitations and to objectively evaluate the image quality, mathematical model observers can be used as a surrogate for human observers for low contrast detectability. Channelized Hotelling Observer (CHO) model observer correlate well to human performances, especially for simple tasks such as a detection of a signal in a small region of interest for different anatomical area. Many studies have shown that model observer can correlate well with human performance include the detection of discs in white noise and clustered lumpy backgrounds [20], microcalcification in mammography [21], breast tomosynthesis [22] nodules in computed tomography (CT) ramp-noise spectrum [23], nodules in breast cone beam CT [24], lung region [25,11] and an abdominal CT acquisition [26–31].

The aim of our study was to evaluate the impact of a DLIR algorithm (True Fidelity, TF) on image texture and low contrast lesion detectability in the liver compared to partial model based iterative reconstruction (ASIR-V). Its potential for dose reduction without impairing lesion detectability was also evaluated.

Table 1
Detailed settings for image acquisition and reconstruction for the five investigated radiation dose levels.

Exposure settings	QRM Medium	QRM Large
$CTDI_{vol}$ (mGy)	2 / 6 / 10 / 15 / 20	2 / 6 / 10 / 15 / 20
Tube potential (kVp)	100	100
Noise Index (mA min – mA max)	70 (10–75) / 55 (100–480) / 42.5 (100–480) / 35 (100–480) / 30 (100–480)	70 (10–75) / 70 (10–220) / 65 (100–480) / 53 (100–480) / 46 (100–480)
Gantry revolution time (s)	0.5 (1 s for $CTDI_{vol}$ at 20 mGy)	0.5 (1 s for $CTDI_{vol}$ at 20 mGy)
Beam collimation (mm)	64 × 0.625	64 × 0.625
Pitch	0.984	0.984
Scan field of view (cm)	50 × 50	50 × 50
Display field of view (cm)	37 × 37	42 × 42
Section thickness (mm)	2.5 and 0.625	2.5 and 0.625
Section overlap (mm)	1.25/0.625	1.25/0.625
Kernel	Standard – option Plus*	Standard – option Plus*
Algorithm	ASiR-V 0%+, 60%, TF low, medium and high	ASiR-V 0%+, 60%, TF low, medium and high

* Plus mode gives you a thicker slice thickness than prescribed. +ASiR-V 0% is used as a surrogate for FBP algorithm. ASiR-V = adaptive statistical iterative reconstruction, $CTDI_{vol}$ = volume CT dose index, TF = True Fidelity.

2. Materials and methods

2.1. Experimental design, image acquisition and reconstruction protocol

To evaluate the low contrast detectability and noise texture, we used a semi-anthropomorphic abdomen phantom with uniform background (QRM, GMBH, Moehrendorf, Germany) mimicking attenuations produced by an adult patient. The central core dimension of this phantom were 20 and 30 cm in anteroposterior and lateral dimension, respectively. Two additional rings of 2.5 cm (size M) or 5 cm (size L) in thickness were added to vary patient's morphology and simulate non-overweight (Body Mass Index, BMI < 25 kg.m-2) or overweight (BMI > 25 kg.m-2) patients (Fig. 1). A cylindrical module containing spherical lesions of 8 and 6 mm in diameter with a contrast of 20 HU relative to the background was inserted in the centre of the phantom (eleven spheres per lesion diameter). These spherical lesions were used as surrogate of focal liver lesions to evaluate low contrast detectability.

The phantom was scanned on a CT Revolution Evo system (Revolution CT®, GE Healthcare) using standard institutional clinical CT protocols for liver tumour diagnostic tasks. Noise Index (NI) was set to obtain a displayed volume CT dose index (CTD_{Ivol}) equal to 2, 6, 10, 15 and 20 mGy. The different dose levels were chosen to frame the P50 local clinical diagnostic reference dose level (5.1 mGy for BMI < 25 kg.m^{-2} and 8.4 mGy for BMI > 25 kg.m^{-2} patients) [32]. The phantom was positioned at the isocenter of the CT unit and scanned ten times for each dose level, without changing its position between acquisitions.

CTDIvol were calculated according to recommendations from the International Electrotechnical Commission (IEC 60601-2-44) and displayed $CTDI_{vol}$ was used as a figure of merit for the patient dose exposure [33]. Images were reconstructed using a new DL reconstruction TrueFidelity (TF) algorithm at low, medium and high strength, and a partial model-based iterative reconstruction (adaptive statistical iterative reconstruction-V, ASiR-V) at 60 % strength [19]. According to previous studies, ASiR-V 60 % was a superior algorithm compared to FBP, ASIR and other ASiR-V levels [34–36]. Based on this results and after a dedicated institutional optimization procedure, it was decided that 60 % ASiR-V was used in clinical routine for abdominal acquisitions. The detailed parameters for image acquisition and reconstruction are reported in Table 1.

2.2. Noise texture measurements

Before characterizing detectability, it is important to consider the visual aspect of the image. In clinical routine, the change in texture or image appearance can minimize the impact of new technologies on the optimization process. The normalized noise power spectrum (NPS) with an eye filter (nNPSe) was used to characterize the different noise textures with FBP as gold standard [37]. NPS was calculated using an homemade algorithm compute using the IGOR Pro 6 scripting language (Wavemetrics, Inc., Portland, OR, USA) according to recommendations from the International Commission on Radiation Units and Measurements reports 54 and 87, using region of interest in the uniform background of the phantom QRM M [38,39]. Since the image noise is not stationary, several ROIs (64 × 64 pixels) at different position in the x–y plane were used. To take into account variable perception of noise by a human observer, each NPS was filtered by a human visual response function, V (ρ), then it was normalized by its integral on all frequencies (nNPSe). The human visual response function parameters are those proposed by Solomon et al. [37].

We computed the mean dose level nNPSe for each reconstruction algorithm used in the study to obtain a general conclusion. Subsequently the shape of the mean nNPSe curves of each algorithm was compared for root mean square deviation (RMSD) to FBP according to the following formula:

$$RMSD = \sqrt{\frac{\sum_r^R \left(x_{(r)} - y_{(r)}\right)^2}{N}}$$

Where x_r represents the nNPSe value at the frequency "r" for the FBP algorithm and y_r represents the nNPSe value at the frequency "r" for TF at low, medium or high strength or ASiR-V at 60 %, N represents the total number of sample values.

For an equivalent noise texture, curves have similar shape and overlap, with an RMSD equal to 0.

The frequency at which the nNPSe (r) had its peak value ("peak frequency") was also computed, and the peak frequency difference (PFD) was calculated with respect to FBP. A positive PFD implies a sharper algorithm and a negative PFD implies a smoother algorithm than FBP. Low RMSD and PFD values imply similar texture against FBP.

2.3. Task-based image quality assessment

2.3.1. Channelized Hotelling Observer

Low contrast detectability was assessed by CHO model observer computed on an homemade program in python 3.3 and validated with an international comparison, with 10 dense difference of Gaussian (DDoG) channels [27,40]. However, as CHO model observer are more efficient than human observers for simple detection task in uniform background, it is necessary to adjust the detection outcomes of model observers by adding internal noise. In this study, the internal noise is diagonal matrix multiplied by a proportional factor (p) added on the

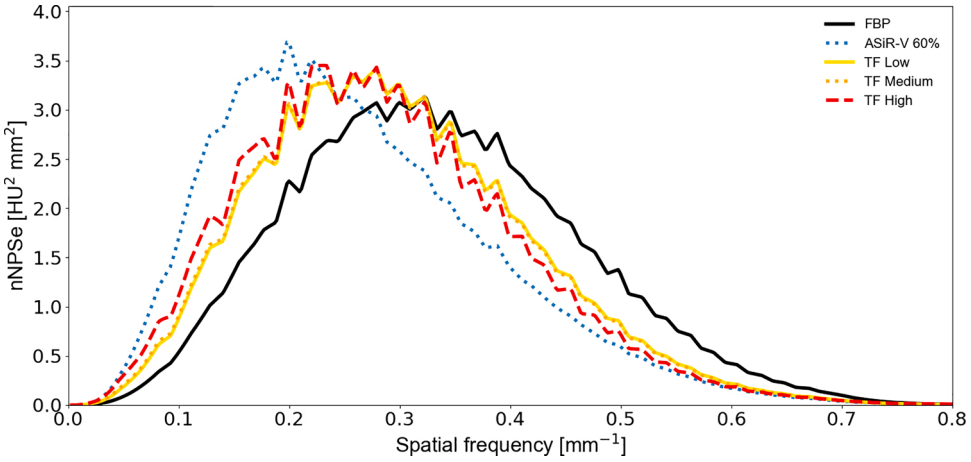

Fig. 2. Plots of radially averaged nNPSe(r) curve for FBP, ASiR-V 60 % and TF algorithms. These curves are normalized, to highlight differences in spatial frequency content, independently of differences in noise magnitude.

Table 2
RMSD and PFD values between the nNPSe(r) of FBP, ASiR-V and TF algorithms. The more RMSD and PFD are close to 0, the more the algorithm preserves FBP-like image noise texture. A negative PFD implies that the algorithm is smother than the FBP algorithm.

	RMSD (mm^{-1})	PFD (mm^{-1})
ASiR-V vs FBP	0.53	−0.122
TF Low vs FBP	0.12	−0.043
TF medium vs FBP	0.12	−0.043
TF High at 2.5 mm vs FBP	0.21	−0.090
TF High at 0.625 mm vs FBP	0.19	−0.064

covariance matrix. The p factor was calibrated on the data from the inter-comparison study of Ba et al. [41,40] and it was equal to 200.

Then, to establish reliable confidence interval (CI) estimators of CHO performance and to overcome the bias associated with point estimates, the CHO performance was calculated using the method developed by Wunderlich et al. [42]

For this study, after evaluating the minimal number of images needed [40], 110 signal-absent images and 110 signal-present images were respectively used to compute 110 decision variables for the signal-absent image category and for the signal-present image category. Finally, the SNR was converted in Area under the ROC curve using the following formula:

$$AUC = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{SNR_t}{2}\right)$$

where Φ is the normal cumulative distribution function.

The AUC was used as figure of merit and its value ranged from 0.5 to 1.0. The area under a ROC curve quantifies the ability to discriminate images with signal and images without signal. When it is impossible to distinguish the two groups of images, the AUC is equal to 0.5, when the detectability (distinction) is perfect the AUC is equal to one.

2.4. Statistical analysis

The potential for reducing the radiation dose of TF at different strength and for two slice thicknesses as reported in Table 1 were compared to ASiR-V applying three steps :

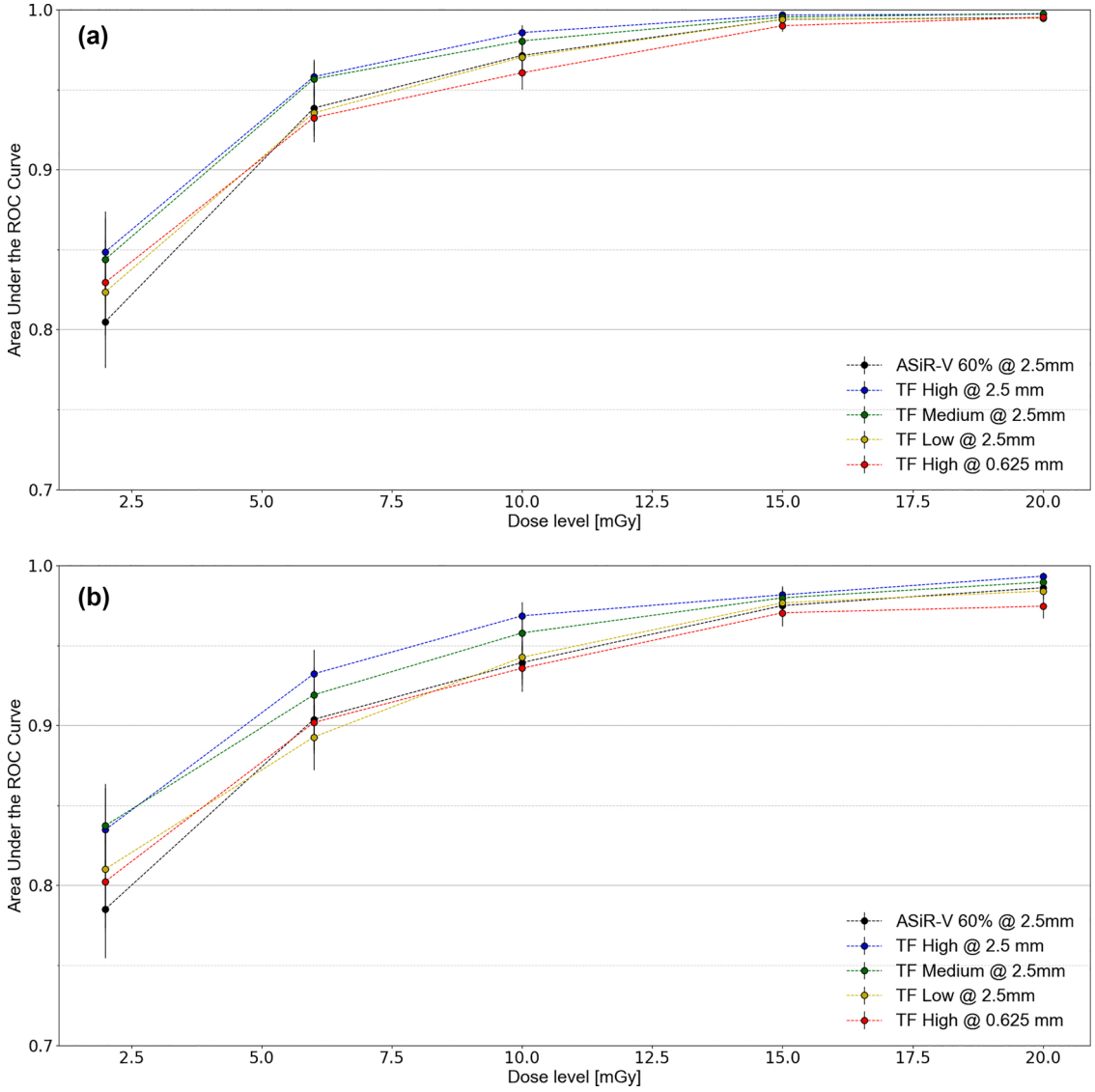

Fig. 3. Area under the ROC curve as a function of the CTDIvol and algorithms investigated for 8-mm-diameter (a) and 6-mm-diameter (b) lesions for QRM M. TF High significantly outperformed ASiR-V 60 %. Increasing dose substantially enhanced detectability for all algorithms. For visualization purpose the AUC scale started at 0.7.

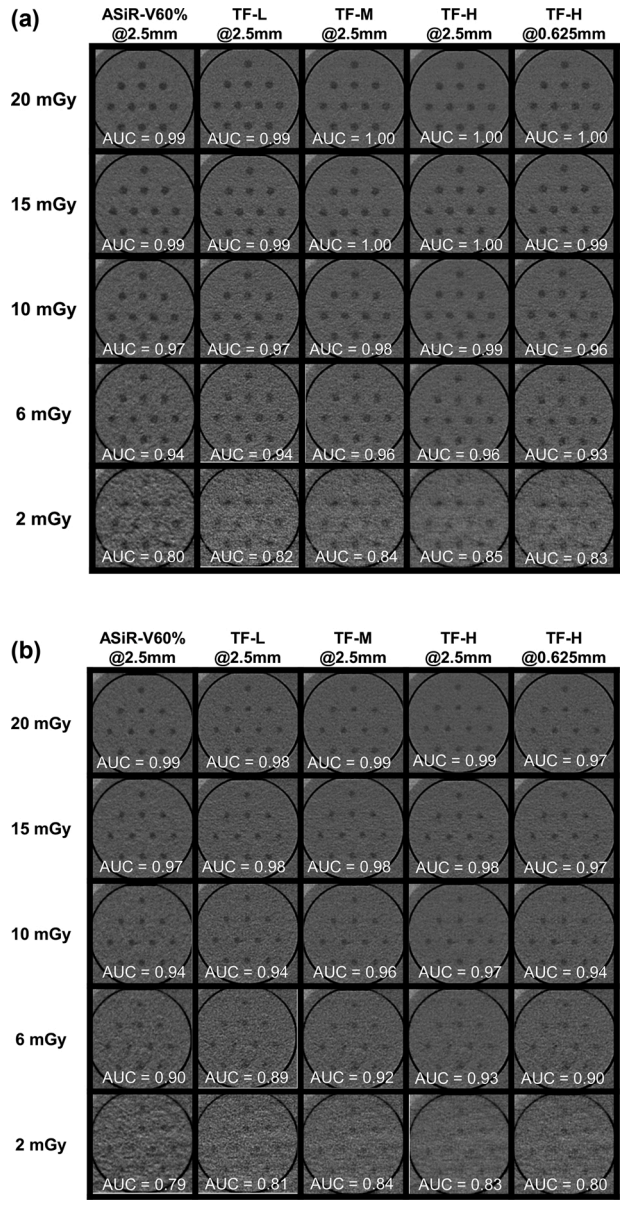

Fig. 4. On axial CT images, a 3.5 × 3.5 cm2 region of interest of the cylindrical module containing spherical lesions of 8 (a) and 6 mm (b) in diameter for QRM M is reconstructed with ASiR-V at 60 % and TF at the low, medium and high strengths as function of the dose level. Areas under the ROC curve result obtained with CHO model observer for each category are indicated in each square. Note that for visualization purpose images are obtained by averaging 10 series to increase signal to noise ratio.

- First, for each algorithm, using a bootstrap method, a curve fit in terms of AUC as a function of radiation dose ($CTDI_{vol}$) with a shape function of a*x^c / (x^c + b) was carried out from data provided by the CHO model observer (a, b and c are the parameters of the function and x represented CTDIvol). The fit was performed with the scipy optimize package from python 3.3.
- Second, we calculated which $CTDI_{vol}$ obtained the same AUC levels with TF (different strengths and slice thicknesses) as with ASiR-V at 60 % at P50 local diagnostic reference level for BMI < 25 $kg.m^{-2}$ (5.1 mGy) for QRM medium and at local diagnostic reference level for BMI > 25 $kg.m^{-2}$ (8.4 mGy) for QRM Large [32].
- Third, from 10 000 bootstrap realizations, we constructed a distribution that represents the $CTDI_{vol}$ difference between TF (different strengths and slice thicknesses) and ASiR-V 60 %. The difference was considered statistically significant if the 95 % confidence interval [95 % CI] did not contain "0".

3. Results

3.1. Noise texture

Fig. 2 presents nNPSe curves for the different reconstruction modes. nNPSe root mean square deviation and peak frequency shifts for TF and ASiR-V in comparison with FBP are summarized in Table 2. In comparison to FBP, TF maintained noise texture with slight RMSD (≤0.21 mm^{-1}) and PFD (≤0.09 mm^{-1}).

In contrast, change in noise texture is more pronounced with ASiR-V (0.53 mm^{-1} for RMSD and 0.12 mm^{-1} for PFD).

3.2. Radiation dose optimization

Fig. 3 shows AUC as a function of $CTDI_{vol}$ values for ASiR-V and TF for (a) 8 mm and (b) 6 mm lesion sizes on M-sized and corresponding images (Fig. 4a,b). Fig. 5a,b show AUC as a function of $CTDI_{vol}$ and corresponding images (Fig. 6a,b) for L-sized phantom. As expected, the detectability (AUC) increased when dose and lesion size increased and was always higher with TF than ASiR-V.

For 6 mm lesions, considering the P50 local clinical DRL for BMI ≤ 25 patients (5.1 mGy $CTDI_{vol}$), AUC was higher for TF (low: 0.889 ± 0.026; high: 0.924 ± 0.023) than ASIR-V 60 % (0.889 ± 0.029). For BMI > 25 patients (8.4 mGy CTDIvol), results were similar for TF (low: 0.853 ± 0.033; high: 0.884 ± 0.030) compared to ASIR-V60 (0.829 ± 0.035).

For 8 mm lesions, at 5.1 mGy CTDIvol, AUC was higher for TF (low: 0.926 ± 0.023; high: 0.951 ± 0.018) than ASIR-V 60 % (0.926 ± 0.023). For BMI > 25 patients (8.4 mGy $CTDI_{vol}$), results were similar for TF (low: 0.909 ± 0.026; high: 0.939 ± 0.020) compared to ASIR-V60 (0.884 ± 0.030).

Table 3 shows the potential dose reduction obtained with TF at different strengths and slice thicknesses compared to ASiR-V. Mean $CTDI_{vol}$ difference for 8 and 6 mm lesion size was +2.5 % for TF low, −23.5 % for TF medium and −29 % for TF high strength for QRM medium. For QRM large, mean $CTDI_{vol}$ difference was −17 %, −26 %, −36.5 % for low, medium and high strength, respectively. A statistically significant difference was only observed for TF high strength, with greatest benefit at a slice thickness of 2.5 mm for a 6 mm in diameter sphere, with a dose reduction of −33 % [−66 %; −1%]. For TF high strength with a slice thickness of 0.625 mm, despite the fourfold thinner slices, a small but not significantly dose reduction can be reached (−2% [−51 % ; +51 %]).

4. Discussion

To our knowledge, this study represents the first anthropomorphic phantom characterization for low contrast lesion detectability of a new deep learning-based CT reconstruction algorithm using a task-based paradigm in the image domain. Our results indicate that TF decreases noise magnitude to a level similar to ASiR-V 60 %, while minimizing noise texture changes when compared to FBP. This enables a sharper image appearance for TF on one hand and a higher dose reduction potential on the other hand.

Many studies still only use contrast-to-noise ratio (CNR) or conventional Fourier-based metrics (i.e. Modulation Transfer Function) to assess image quality, but overestimate potential dose reduction due to the lack of lesion size, contrast or noise texture consideration when dealing with IR [12,43–45]. To overcome these limitations, our study used a CHO model observer in the image domain, to compute an area under the ROC curve as surrogate of detectability index. These model observer demonstrate a stronger correlation with human observer performance compared to conventional CNR or standard Fourier metrics

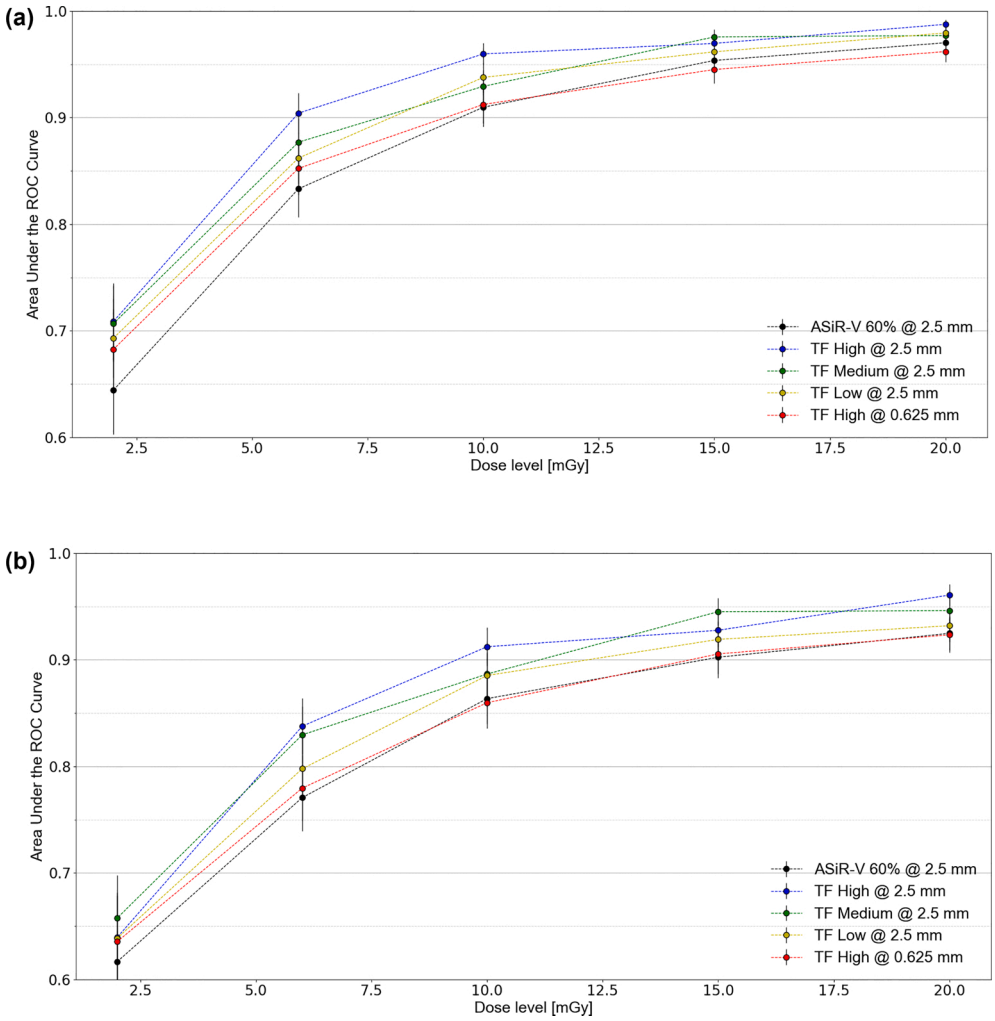

Fig. 5. Area under the ROC curve as a function of the $CTDI_{vol}$ and algorithms investigated for 8-mm-diameter (a) and 6-mm-diameter (b) lesions for QRM L. Overall, TF with a slice thickness of 2.5 mm outperformed ASiR-V 60 % and increasing DLIR levels substantially increased detectability. For visualization purpose the AUC scale started at 0.6.

[46,47].

Physical characterization of TF has recently been outlined in other studies [48–50]. We found similar results regarding noise texture, also with an eye filter to simulate human noise texture visualisation. The noise texture in TF images was comparable to FBP in terms of spatial frequency content (RMSD) and spatial frequency shift (PFD) whereas ASiR-V provided images with low frequency content resulting in blurry images which makes low-contrast detections tasks challenging [11, 51–53].

Greffier et al. found it possible to reduce the dose by 48 % and Racine et al. 61 %. However, we estimated an approximately 30 % dose reduction potential for TF without impairing lesion detectability in comparison to ASiR-V. Various points can explain these different results. First, we have used a CHO model observer in the image domain instead of NPWE model observer in Fourier domain. The NPWE model observer is strongly affected by noise magnitude especially in planar (2D) formulation without managing influence of longitudinal variations [54, 55]. Second, the lesion size and contrast are not the same between studies. Greffier studied large lesion size (10 mm) and Racine studied a 50 HU in contrast lesion: both are simulated lesions. In other phantom-based studies, liver lesions were generally simulated with contrasts in the range of −12 to −40 HU and sizes from 5 mm to 20 mm, which is consistent with our study [12,43,56]. In our study, we used spherical real low contrast lesions of 8 and 6 mm in diameter with a contrast of 20 HU relative to the background and corresponding to human low contrast lesions, not easily detectable in order to highlight the benefits of the TF algorithm in the most challenging conditions.

The potential of TF to increase image quality and, as a consequence, to enable an additional dose reduction has recently been also outlined in few retrospective studies. Jensen et al. found that overall image quality and overall lesion diagnostic confidence evaluated by two radiologists were significantly higher for TF than ASiR-V in abdominal CT for an overweight population (mean body mass index of 27 ± 5 (range, 20–41)) [53]. They suggested that compared to FBP, TF could allow more radiation dose reduction that the 25% to 40% dose reduction range found with ASIR-V. For another clinical task, Benz et al. found that TF provided superior qualitative image quality on 43 patients undergoing coronary CT angiography compared with ASiR-V which should contribute to a further reduction in the dose of CT radiation [57]. Our phantom study quantifies this potential dose reduction in simulating real low contrast lesion detectability according the strength of TF used, in order to determine a clinical applicable approach.

Despite these appealing results, some limitations need to be mentioned.

First, this study used phantom images. As TF was trained on phantom and patient images with anatomical noise, it is reasonable to expect that its performance evaluated on a semi anthropomorphic phantom with uniform background will further need a dedicated image quality study. 3D printing of a phantom created from real patient images shall be conducted in the near future, but we expect these future results to

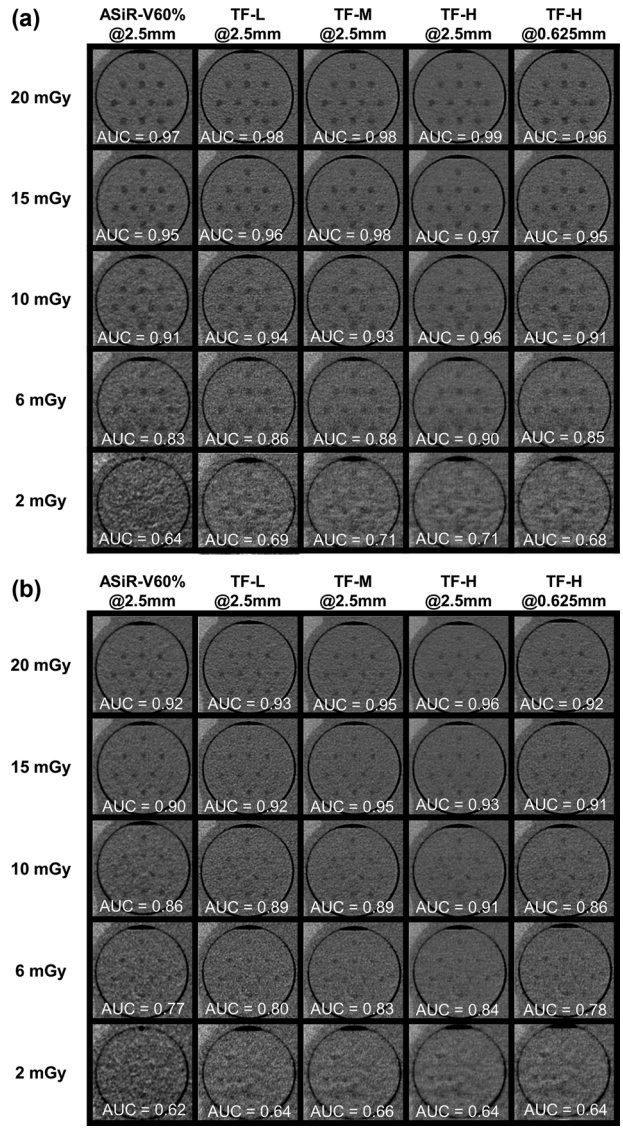

Fig. 6. On axial CT images, a 3.5 × 3.5 cm2 region of interest of the cylindrical module containing spherical lesions of 8 (a) and 6 mm (b) in diameter for QRM L is reconstructed with ASiR-V at 60 % and TF at the low, medium and high strengths as function of the dose level. Areas under the ROC curve result obtained with CHO model observer for each category are indicated in each square. Note that for visualization purpose images are obtained by averaging 10 series to increase signal to noise ratio.

correlate with those found in this phantom study. Second, some lesions in the phantom would have been more similar to clinical routine if they contained iodine material, to assess the impact of kVp on detectability but this study used 100 kVp only, as in clinical practice. Third, the reduction in dose potential was only calculated with ASiR-V level of 60 %. Since the detectability at low contrast is strongly affected by the amplitude of the noise but also by the texture of the noise, it will be interesting to study the reduction of the dose potential at different ASiR-V levels, especially at a high level. However several studies have shown that it is difficult to use high levels of IR in clinical practice due to excessive image smoothing and artificial appearance of the image [7,9,14]. Finally, in order to keep this study focused and concise, we considered out of scope to study the image quality in the other two reconstruction planes, which would deserve further dedicated work.

In conclusion, TrueFidelity deep learning image reconstruction algorithm improves low contrast detectability without changing noise texture whatever the strength or dose level in comparison to ASiR-V. For low-contrast diagnostic tasks in abdominal CT examinations, TF high strength has the potential to enable an additional dose optimization of approximately 30 % compared to current IR techniques. The introduction of DLIR algorithms in clinical routine marks the beginning of a new era for CT dose optimization. Future studies using a task based paradigm assessment performed using other anatomical localization and therefore with a different anatomical texture will be carried out to confirm the impact of DLIR algorithms.

Table 3
Potential dose reduction with TF compared to ASiR-V 60 % (ΔD) [±95 % CI]. Dose reduction is calculated for an AUC value obtained with ASiR-V 60 % at 5.1 mGy for QRM medium and 8.4 mGy for QRM large corresponding to P50 local clinical DRLs for liver tumour assessment. *p < 0.05 (significant).

Dose reduction potential (ΔD)				
	QRM M		QRM L	
Reconstruction algorithms compared / sphere dimension	6 mm	8 mm	6 mm	8 mm
TF High 2.5 mm vs ASiR-V 60 %	−33%* [−66 % ; −1%]	−25% [−54 % ; +4%]	−33%* [−58 ; −9%]	−40% * [−60 % ; −19 %]
TF medium 2.5 mm vs ASiR-V 60 %	−26% [−67 % ; 1%]	−21% [−53 % ; +10 %]	−29%* [−56 ; −1%]	−23% [−51 %;+4%]
TF low 2.5 vs ASiR-V 60 %	+2% [−48 % ; +52 %]	+3% [−40 % ; +47 %]	−16% [−49 ; +17 %]	−18% [−48 %;+10 %]
TF high 0.625 mm vs ASiR-V 60 %	−2% [−51 % ; +51 %]	+10 % [−39 % ; +60 %]	−1%* [−42 ; −38 %]	−6% [−44 %;+31 %]

Funding source

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Guarantor

The scientific guarantor of this publication is Hugues Brat, Medical Doctor.

Statistics and biometry

No complex statistical methods were necessary for this paper.

Informed consent

N/A.

Ethical approval

Institutional Review Board approval was not required for this study because of its phantom-only experimental design.

Study subjects or cohorts overlap

N/A.

Methodology

- prospective
- experimental
- performed at one institution

CRediT authorship contribution statement

D. Racine: Conceptualization, Methodology, Software, Writing - original draft. **H.G. Brat:** Conceptualization, Methodology, Validation, Writing - review & editing. **B. Dufour:** Methodology, Validation, Writing - review & editing. **J.M. Steity:** Resources, Writing - review & editing. **M. Hussenot:** Methodology, Resources, Writing - review & editing. **B. Rizk:** Validation, Writing - review & editing. **D. Fournier:** Resources, Writing - review & editing. **F. Zanca:** Conceptualization, Methodology, Validation, Writing - review & editing.

Declaration of Competing Interest

One of the authors (Michel Hussenot) is an employee of GE Healthcare. He had no control over data acquisition, analysis or interpretation. The other authors state that they have no competing interests to declare.

References

[1] D.J. Brenner, E.J. Hall, Computed tomography — an increasing source of radiation exposure, N. Engl. J. Med. 357 (2007) 2277–2284, https://doi.org/10.1056/NEJMra072149.

[2] D.J. Brenner, Radiation risks potentially associated with low-dose CT screening of adult smokers for lung cancer, Radiology 231 (2004) 440–445, https://doi.org/10.1148/radiol.2312030880.

[3] R. Le Coultre, J. Bize, M. Champendal, D. Wittwer, N. Ryckx, A. Aroua, P. Trueb, F. R. Verdun, Exposure of the Swiss population by radiodiagnostics: 2013 review, Radiat. Prot. Dosimetry 169 (2016) 221–224, https://doi.org/10.1093/rpd/ncv462.

[4] L. Yu, H. Li, J.G. Fletcher, C.H. McCollough, Automatic selection of tube potential for radiation dose reduction in CT: a general strategy, Med. Phys. 37 (2010) 234–243, https://doi.org/10.1118/1.3264614.

[5] A.E. Papadakis, J. Damilakis, Automatic tube current modulation and tube voltage selection in pediatric computed tomography, Invest. Radiol. 54 (2019) 265–272, https://doi.org/10.1097/RLI.0000000000000537.

[6] M.K. Kalra, M.M. Maher, T.L. Toth, B. Schmidt, B.L. Westerman, H.T. Morgan, S. Saini, Techniques and applications of automatic tube current modulation for CT, Radiology 233 (2004) 649–657, https://doi.org/10.1148/radiol.2333031150.

[7] A. Mileto, L.S. Guimaraes, C.H. McCollough, J.G. Fletcher, L. Yu, State of the art in abdominal CT: the limits of iterative reconstruction algorithms, Radiology 293 (2019) 491–503, https://doi.org/10.1148/radiol.2019191422.

[8] M.J. Willemink, T. Leiner, P.A. de Jong, L.M. de Heer, R.A.J. Nievelstein, A.M. R. Schilham, R.P.J. Budde, Iterative reconstruction techniques for computed tomography part 2: initial results in dose reduction and image quality, Eur. Radiol. 23 (2013) 1632–1642, https://doi.org/10.1007/s00330-012-2764-z.

[9] J. Greffier, J. Frandon, A. Larbi, J.P. Beregi, F. Pereira, CT iterative reconstruction algorithms: a task-based image quality assessment, Eur. Radiol. 30 (2020) 487–500, https://doi.org/10.1007/s00330-019-06359-6.

[10] M. Katsura, I. Matsuda, M. Akahane, J. Sato, H. Akai, K. Yasaka, A. Kunimatsu, K. Ohtomo, Model-based iterative reconstruction technique for radiation dose reduction in chest CT: comparison with the adaptive statistical iterative reconstruction technique, Eur. Radiol. 22 (2012) 1613–1623, https://doi.org/10.1007/s00330-012-2452-z.

[11] D.C. Rotzinger, D. Racine, C. Beigelman-Aubry, K.M. Alfudhili, N. Keller, P. Monnin, F.R. Verdun, F. Becce, Task-based model observer assessment of a partial model-based iterative reconstruction algorithm in thoracic oncologic multidetector CT, Sci. Rep. 8 (2018) 17734, https://doi.org/10.1038/s41598-018-36045-4.

[12] S.T. Schindera, D. Odedra, S.A. Raza, T.K. Kim, H.-J. Jang, Z. Szucs-Farkas, P. Rogalla, Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? Radiology 269 (2013) 511–518, https://doi.org/10.1148/radiol.13122349.

[13] A. Mileto, D.A. Zamora, A.M. Alessio, C. Pereira, J. Liu, P. Bhargava, J. Carnell, S. M. Cowan, M.K. Dighe, M.L. Gunn, S. Kim, O. Kolokythas, J.H. Lee, J.H. Maki, M. Moshiri, A. Nasrullah, R.B. O'Malley, U.P. Schmiedl, E.V. Soloff, G.V. Toia, C. L. Wang, K.M. Kanal, CT detectability of small low-contrast hypoattenuating focal lesions: iterative reconstructions versus filtered back projection, Radiology 289 (2018) 443–454, https://doi.org/10.1148/radiol.2018180137.

[14] L.L. Geyer, U.J. Schoepf, F.G. Meinel, J.W. Nance, G. Bastarrika, J.A. Leipsic, N. S. Paul, M. Rengo, A. Laghi, C.N. De Cecco, State of the art: iterative CT reconstruction techniques, Radiology 276 (2015) 339–357, https://doi.org/10.1148/radiol.2015132766.

[15] R. Singh, S.R. Digumarthy, V.V. Muse, A.R. Kambadakone, M.A. Blake, A. Tabari, Y. Hoi, N. Akino, E. Angel, R. Madan, M.K. Kalra, Image quality and lesion detection on deep learning reconstruction and iterative reconstruction of submillisievert chest and abdominal CT, Am. J. Roentgenol. 214 (2020) 566–573, https://doi.org/10.2214/AJR.19.21809.

[16] J.G. Fletcher, L. Yu, J.L. Fidler, D.L. Levin, D.R. DeLone, D.M. Hough, N. Takahashi, S.K. Venkatesh, A.-M.G. Sykes, D. White, R.M. Lindell, A.L. Kotsenas, N.G. Campeau, V.T. Lehman, A.C. Bartley, S. Leng, D.R. Holmes, A.Y. Toledano, R. E. Carter, C.H. McCollough, Estimation of observer performance for reduced radiation dose levels in CT: eliminating reduced dose levels that are too low is the first step, Acad. Radiol. 24 (2017) 876–890, https://doi.org/10.1016/j.acra.2016.12.017.

[17] M.J. Willemink, W.A. Koszek, C. Hardell, J. Wu, D. Fleischmann, H. Harvey, L. R. Folio, R.M. Summers, D.L. Rubin, M.P. Lungren, Preparing medical imaging data for machine learning, Radiology (2020), 192224, https://doi.org/10.1148/radiol.2020192224.

[18] M.J. Willemink, P.B. Noël, The evolution of image reconstruction for CT—from filtered back projection to artificial intelligence, Eur. Radiol. 29 (2019) 2185–2195, https://doi.org/10.1007/s00330-018-5810-7.

[19] J. Hsieh, E. Liu, B. Nett, J. Tang, J.-B. Thibault, S. Sahney, A New Era of Image Reconstruction: TrueFidelity, GE Healthc. White paper, 2019. JB68676XX.

[20] R.W. Bouwman, M. Goffi, R.E. van Engen, M.J.M. Broeders, D.R. Dance, K. C. Young, W.J.H. Veldkamp, Can the channelized Hotelling observer including aspects of the human visual system predict human observer performance in mammography? Phys. Medica PM Int. J. Devoted Appl. Phys. Med. Biol. Off. J. Ital. Assoc. Biomed. Phys. AIFB 33 (2017) 95–105, https://doi.org/10.1016/j.ejmp.2016.12.015.

[21] R.W. Bouwman, A. Mackenzie, R.E. van Engen, M.J.M. Broeders, K.C. Young, D. R. Dance, G.J. den Heeten, W.J.H. Veldkamp, Toward image quality assessment in mammography using model observers: detection of a calcification-like object, Med. Phys. 44 (2017) 5726–5739, https://doi.org/10.1002/mp.12532.

[22] S. Young, P.R. Bakic, K.J. Myers, R.J. Jennings, S. Park, A virtual trial framework for quantifying the detectability of masses in breast tomosynthesis projection data, Med. Phys. 40 (2013), 051914, https://doi.org/10.1118/1.4800501.

[23] C.K. Abbey, F.W. Samuelson, R. Zeng, J.M. Boone, M.P. Eckstein, K. Myers, Classification images for localization performance in ramp-spectrum noise, Med. Phys. 45 (2018) 1970–1984, https://doi.org/10.1002/mp.12857.

[24] M. Han, J. Baek, A performance comparison of anthropomorphic model observers for breast cone beam CT images: a single-slice and multislice study, Med. Phys. 46 (2019) 3431–3441, https://doi.org/10.1002/mp.13598.

[25] C. Franck, A. Snoeckx, M. Spinhoven, H. El Addouli, S. Nicolay, A. Van Hoyweghen, P. Deak, F. Zanca, Pulmonary nodule detection in chest Ct using a deep learning-based reconstruction algorithm, Radiat. Prot. Dosimetry (2021), https://doi.org/10.1093/rpd/ncab025.

[26] S.K.N. Dilger, S. Leng, B. Chen, R.E. Carter, C.P. Favazza, J.G. Fletcher, C. H. McCollough, L. Yu, Localization of liver lesions in abdominal CT imaging: II. Mathematical model observer performance correlates with human observer performance for localization of liver lesions in abdominal CT imaging, Phys. Med. Biol. 64 (2019) 105012, https://doi.org/10.1088/1361-6560/ab1a62.

[27] D. Racine, A.H. Ba, J.G. Ott, F.O. Bochud, F.R. Verdun, Objective assessment of low contrast detectability in computed tomography with Channelized Hotelling Observer, Phys. Medica Eur. J. Med. Phys. 32 (2016) 76–83, https://doi.org/10.1016/j.ejmp.2015.09.011.

[28] A. Ba, M.P. Eckstein, D. Racine, J.G. Ott, F. Verdun, S. Kobbe-Schmidt, F. O. Bochud, Anthropomorphic model observer performance in three-dimensional detection task for low-contrast computed tomography, J. Med. Imaging Bellingham Wash. 3 (2016), 011009, https://doi.org/10.1117/1.JMI.3.1.011009.

[29] J.G. Ott, A. Ba, D. Racine, A. Viry, F.O. Bochud, F.R. Verdun, Assessment of low contrast detection in CT using model observers: developing a clinically-relevant tool for characterising adaptive statistical and model-based iterative reconstruction, Z. Med. Phys. 27 (2017) 86–97, https://doi.org/10.1016/j.zemedi.2016.04.002.

[30] L. Yu, B. Chen, J.M. Kofler, C.P. Favazza, S. Leng, M.A. Kupinski, C.H. McCollough, Correlation between a 2D channelized Hotelling observer and human observers in a low-contrast detection task with multislice reading in CT, Med. Phys. 44 (2017) 3990–3999, https://doi.org/10.1002/mp.12380.

[31] C.P. Favazza, A. Ferrero, L. Yu, S. Leng, K.L. McMillan, C.H. McCollough, Use of a channelized Hotelling observer to assess CT image quality and optimize dose reduction for iteratively reconstructed images, J. Med. Imaging Bellingham Wash. 4 (2017), 031213, https://doi.org/10.1117/1.JMI.4.3.031213.

[32] H. Brat, F. Zanca, S. Montandon, D. Racine, B. Rizk, E. Meicher, D. Fournier, Local clinical diagnostic reference levels for chest and abdomen CT examinations in adults as a function of body mass index and clinical indication: a prospective multicenter study, Eur. Radiol. 29 (2019) 6794–6804, https://doi.org/10.1007/s00330-019-06257-x.

[33] I.E. Commission, Medical Electrical Equipment-Part 2-44. Particular Requirements for the Safety of X-Ray Equipment for Computed Tomography, IEC 60601-2-44, 2002 (accessed February 3, 2020), https://ci.nii.ac.jp/naid/20000855289/.

[34] L.-H. Chen, C. Jin, J.-Y. Li, G.-L. Wang, Y.-J. Jia, H.-F. Duan, N. Pan, J. Guo, Image quality comparison of two adaptive statistical iterative reconstruction (ASiR, ASiR-V) algorithms and filtered back projection in routine liver CT, Br. J. Radiol. 91 (2018), 20170655, https://doi.org/10.1259/bjr.20170655.

[35] M.H. Goodenberger, N.A. Wagner-Bartak, S. Gupta, X. Liu, R.Q. Yap, J. Sun, E. P. Tamm, C.T. Jensen, 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. 42 (2018) 184–190, https://doi.org/10.1097/RCT.0000000000000666.

[36] S. Lee, H. Kwon, J. Cho, The detection of focal liver lesions using abdominal CT: a comparison of image quality between adaptive statistical iterative reconstruction V and adaptive statistical iterative reconstruction, Acad. Radiol. 23 (2016) 1532–1538, https://doi.org/10.1016/j.acra.2016.08.013.

[37] J.B. Solomon, O. Christianson, E. Samei, Quantitative comparison of noise texture across CT scanners from different manufacturers, Med. Phys. 39 (2012) 6048–6055, https://doi.org/10.1118/1.4752209.
[38] W. Vennart, ICRU Report 54: medical imaging—the assessment of image quality, Radiography 3 (1996) 243–244, https://doi.org/10.1016/S1078-8174(97)90038-9.
[39] International Commission on Radiation Units and Measurements, ICRU Report No. 87: radiation dose and image-quality assessment in computed tomography, J. ICRU 12 (2012) 1–149, https://doi.org/10.1093/jicru/ndt007.
[40] A. Ba, C.K. Abbey, J. Baek, M. Han, R.W. Bouwman, C. Balta, J. Brankov, F. Massanes, H.C. Gifford, I. Hernandez-Giron, W.J.H. Veldkamp, D. Petrov, N. Marshall, F.W. Samuelson, R. Zeng, J.B. Solomon, E. Samei, P. Timberg, H. Förnvik, I. Reiser, L. Yu, H. Gong, F.O. Bochud, Inter-laboratory comparison of channelized hotelling observer computation, Med. Phys. 45 (2018) 3019–3030, https://doi.org/10.1002/mp.12940.
[41] J.G. Brankov, Evaluation of channelized hotelling observer with internal-noise model in a train-test paradigm for cardiac SPECT defect detection, Phys. Med. Biol. 58 (2013) 7159–7182, https://doi.org/10.1088/0031-9155/58/20/7159.
[42] A. Wunderlich, F. Noo, B.D. Gallas, M.E. Heilbrun, Exact confidence intervals for channelized hotelling observer performance in image quality studies, IEEE Trans. Med. Imaging 34 (2015) 453–464, https://doi.org/10.1109/TMI.2014.2360496.
[43] J. Solomon, D. Marin, K. Roy Choudhury, B. Patel, E. Samei, 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 284 (2017) 777–787, https://doi.org/10.1148/radiol.2017161736.
[44] F.R. Verdun, D. Racine, J.G. Ott, M.J. Tapiovaara, P. Toroi, F.O. Bochud, W.J. H. Veldkamp, A. Schegerer, R.W. Bouwman, I.H. Giron, N.W. Marshall, S. Edyvean, Image quality in CT: from physical measurements to model observers, Phys. Med. 31 (2015) 823–843, https://doi.org/10.1016/j.ejmp.2015.08.007.
[45] J.G. Ott, F. Becce, P. Monnin, S. Schmidt, F.O. Bochud, F.R. Verdun, Update on the non-prewhitening model observer in computed tomography for the assessment of the adaptive statistical and model-based iterative reconstruction algorithms, Phys. Med. Biol. 59 (2014) 4047–4064, https://doi.org/10.1088/0031-9155/59/4/4047.
[46] B. Chen, L. Yu, S. Leng, J. Kofler, C. Favazza, T. Vrieze, C. McCollough, Predicting detection performance with model observers: fourier domain or spatial domain? Proc. SPIE. Int. Soc. Opt. Eng. 9783 (2016) https://doi.org/10.1117/12.2216962.
[47] J. Solomon, E. Samei, Correlation between human detection accuracy and observer model-based image quality metrics in computed tomography, J. Med. Imaging Bellingham Wash. 3 (2016), 035506, https://doi.org/10.1117/1.JMI.3.3.035506.
[48] J. Greffier, A. Hamard, F. Pereira, C. Barrau, H. Pasquier, J.P. Beregi, J. Frandon, Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study, Eur. Radiol. 30 (2020) 3951–3959, https://doi.org/10.1007/s00330-020-06724-w.
[49] D. Racine, F. Becce, A. Viry, P. Monnin, B. Thomsen, F.R. Verdun, D.C. Rotzinger, 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. 76 (2020) 28–37, https://doi.org/10.1016/j.ejmp.2020.06.004.
[50] J. Solomon, P. Lyu, D. Marin, E. Samei, Noise and spatial resolution properties of a commercially available deep learning-based CT reconstruction algorithm, Med. Phys. n/a (n.d.). https://doi.org/10.1002/mp.14319.
[51] C.H. McCollough, L. Yu, J.M. Kofler, S. Leng, Y. Zhang, Z. Li, R.E. Carter, Degradation of CT low-contrast spatial resolution due to the use of iterative reconstruction and reduced dose levels, Radiology 276 (2015) 499–506, https://doi.org/10.1148/radiol.15142047.
[52] C.T. Jensen, N.A. Wagner-Bartak, L.N. Vu, X. Liu, B. Raval, D. Martinez, W. Wei, Y. Cheng, E. Samei, S. Gupta, Detection of colorectal hepatic metastases is superior at standard radiation dose CT versus reduced dose CT, Radiology 290 (2019) 400–409, https://doi.org/10.1148/radiol.2018181657.
[53] C.T. Jensen, X. Liu, E.P. Tamm, A.G. Chandler, J. Sun, A.C. Morani, S. Javadi, N. A. Wagner-Bartak, Image quality assessment of abdominal CT by use of new deep learning image reconstruction: initial experience, Am. J. Roentgenol. 215 (2020) 50–57, https://doi.org/10.2214/AJR.19.22332.
[54] S.S. Hsieh, N.J. Pelc, Improvements in low contrast detectability with iterative reconstruction and the effect of slice thickness. Med. Imaging 2017 Phys. Med. Imaging, International Society for Optics and Photonics, 2017, p. 1013253, https://doi.org/10.1117/12.2253937.
[55] P. Monnin, A. Viry, F.R. Verdun, D. Racine, Slice NEQ and system DQE to assess CT imaging performance, Phys. Med. Biol. 65 (2020) 105009, https://doi.org/10.1088/1361-6560/ab807a.
[56] A.H. Goenka, B.R. Herts, N.A. Obuchowski, A.N. Primak, F. Dong, W. Karim, M. E. Baker, Effect of reduced radiation exposure and iterative reconstruction on detection of low-contrast low-attenuation lesions in an anthropomorphic liver phantom: an 18-reader study, Radiology 272 (2014) 154–163, https://doi.org/10.1148/radiol.14131928.
[57] D.C. Benz, G. Benetos, G. Rampidis, E. von Felten, A. Bakula, A. Sustar, K. Kudura, M. Messerli, T.A. Fuchs, C. Gebhard, A.P. Pazhenkottil, P.A. Kaufmann, R. R. Buechel, Validation of deep-learning image reconstruction for coronary computed tomography angiography: impact on noise, image quality and diagnostic accuracy, J. Cardiovasc. Comput. Tomogr. (2020), https://doi.org/10.1016/j.jcct.2020.01.002.