

# Impact of model-based iterative reconstruction on low-contrast lesion detection and image quality in abdominal CT: a 12-reader-based comparative phantom study with filtered back projection at different tube voltages

André Euler<sup>1</sup> · Bram Stieltjes<sup>1</sup> · Zsolt Szucs-Farkas<sup>2</sup> · Reto Eichenberger<sup>1</sup> · Clemens Reisinger<sup>1</sup> · Anna Hirschmann<sup>1</sup> · Caroline Zaehringer<sup>1</sup> · Achim Kircher<sup>1</sup> · Matthias Streif<sup>1</sup> · Sabine Bucher<sup>1</sup> · David Buegler<sup>1</sup> · Luigia D'Errico<sup>1</sup> · Sebastián Kopp<sup>1</sup> · Markus Wilhelm<sup>1</sup> · Sebastian T. Schindera<sup>1,3</sup>

Received: 24 August 2016 / Revised: 8 February 2017 / Accepted: 20 March 2017 / Published online: 3 April 2017  
© European Society of Radiology 2017, corrected publication August 2017

## Abstract

**Objectives** To evaluate the impact of model-based iterative reconstruction (MBIR) on image quality and low-contrast lesion detection compared with filtered back projection (FBP) in abdominal computed tomography (CT) of simulated medium and large patients at different tube voltages.

**Methods** A phantom with 45 hypoattenuating lesions was placed in two water containers and scanned at 70, 80, 100, and 120 kVp. The 120-kVp protocol served as reference, and the volume CT dose index (CTDI<sub>vol</sub>) was kept constant for all protocols. The datasets were reconstructed with MBIR and FBP. Image noise and contrast-to-noise-ratio (CNR) were assessed. Low-contrast lesion detectability was evaluated by 12 radiologists.

**Results** MBIR decreased the image noise by 24% and 27%, and increased the CNR by 30% and 29% for the medium and large phantoms, respectively. Lower tube voltages increased the CNR by 58%, 46%, and 16% at 70, 80, and 100 kVp, respectively, compared with 120 kVp in the medium phantom and by 9%, 18% and 12% in the large phantom. No significant

difference in lesion detection rate was observed (medium: 79–82%; large: 57–65%;  $P > 0.37$ ).

**Conclusions** Although MBIR improved quantitative image quality compared with FBP, it did not result in increased low-contrast lesion detection in abdominal CT at different tube voltages in simulated medium and large patients.

## Key Points

- MBIR improved quantitative image quality but not lesion detection compared with FBP.
- Increased CNR by low tube voltages did not improve lesion detection.
- Changes in image noise and CNR do not directly influence diagnostic accuracy.

**Keywords** Multidetector computed tomography · Radiological phantom · Model-based iterative reconstruction · Filtered back projection · Low-contrast detection

## Introduction

In the last decade, substantial efforts have been made to reduce patients' exposure to ionising radiation from computed tomography (CT). However, the diagnostic accuracy should not be impeded by this process. Lesion detection is a common task in daily practise and particularly important to assess metastatic disease in oncological patients [1]. Alterations in image quality parameters of a CT scan, including image noise and lesion-to-background contrast-to-noise ratio (CNR), potentially influence lesion detectability [2]. This phenomenon has been especially observed for the detection of low-contrast

The original version of this article was revised: The name of the 12th author has been corrected to Luigia D'Errico.

✉ Sebastian T. Schindera  
sschindera@aol.com

<sup>1</sup> Clinic of Radiology and Nuclear Medicine, University Hospital Basel, Basel, Switzerland

<sup>2</sup> Institute of Radiology, Hospital Centre of Biel, Biel, Switzerland

<sup>3</sup> Institute of Radiology, Cantonal Hospital Aarau, Aarau, Switzerland

lesions, which is one of the most difficult tasks of diagnostic radiology in routine practice [2, 3].

Radiologists generally have two available options to improve lesion-to-background CNR: the application of iterative reconstruction algorithms (IRs) and imaging at low tube voltages. Several studies demonstrated that IRs reduce the image noise and consequently improve the CNR compared with conventional filtered back projection (FBP) [4–8]. Reducing the tube voltage increases the attenuation of iodinated contrast material due to the augmented photoelectric effect [9]. Increased attenuation may lead to improved lesion delineation between two tissue types with a different iodine uptake, as confirmed in former studies [10, 11]. The main drawback of low tube voltages is the increased image noise caused by the reduced photon flux. Especially in larger patients, this increase in image noise may outweigh the improved lesion contrast and result in an overall decrease of the lesion-to-background CNR. To increase the photon flux, the tube current time product of the CT scanner must be increased.

Some studies have demonstrated an improvement of image quality parameters and lesion detection rate using IRs [12, 13], whereas others demonstrated no change in diagnostic accuracy despite improvements in image noise and CNR [14, 15]. Model-based iterative reconstruction algorithms (MBIR) are the third generation of IRs. Their complex statistical model takes into account X-ray physics and CT scanner optics during the iteration process. MBIRs have shown to improve image quality parameters and potentially reduce radiation dose [16–18]. A phantom study, evaluating the impact of a recent MBIR on low-contrast lesion detectability, demonstrated significantly improved detectability using MBIR compared with FBP at 120 kV [19]. However, the results of the study were limited by the assessment at a single tube voltage as well as by the phantom's small diameter (16.5 cm) which does not reflect the average abdominal diameter of an adult patient in the Western world.

The purpose of our study was to assess the impact of MBIR and low-kVp imaging on image quality and low-contrast detectability in a contrast-detail CT phantom in abdominal CT of simulated medium and large patients.

## Materials and methods

### Phantom design

An iodine-containing contrast-detail phantom with 45 hypodense, low-contrast lesions (QRM, Moehrendorf, Germany) was used in this study. The background parenchyma of the phantom was hyperattenuating, simulating parenchyma in arterial phase. The lesions had three different diameters (5, 10, and 15 mm) and three different lesion-to-background contrast values (10, 25, and 50 HU). The

phantom was placed in two separate water-filled cylindrical plastic containers with diameters of 30 cm and 40 cm, mimicking the abdominal cross-sectional dimensions of a medium and a large patient (estimated body weights: 72–85 kg for the medium phantom and 118–142 kg for the large phantom [20]).

### Scanning protocol

Both phantoms were scanned using a third-generation dual-source CT scanner (SOMATOM Force; Siemens Healthineers, Forchheim, Germany) at 70, 80, 100, and 120 kVp. Our institutional abdominal protocol at 120 kVp, 150 ref. mAs, collimation of  $192 \times 0.6$  mm, gantry rotation time of 0.5 s, pitch of 0.8 and automatic tube current modulation with an average strength curve (CareDose4D; Siemens Healthineers) served as the reference protocol, and the tube current time products of the other three protocols (70, 80, 100 kVp) were adjusted to keep the volume CT dose index ( $CTDI_{vol}$ ) constant within each phantom size. The  $CTDI_{vol}$  was 8 mGy and 19 mGy for the medium and large phantoms, respectively. The effective mAs values measured 685, 417, 200, and 120 for the medium phantom and 1616, 985, 472, and 283 for the large phantom at 70, 80, 100, and 120 kVp, respectively.

Each scan was reconstructed with a recent MBIR, known as advanced modelled iterative reconstruction (ADMIRE; Siemens Healthineers) at a strength of 3 and with FBP [21]. The I40- and Br40-kernels were applied for MBIR and FBP, respectively. The 5-mm-thick axial images with an increment of 2.5 mm were used for evaluation. Eight datasets, with a total of 816 axial images, were generated for each phantom size.

### Objective image quality

CT numbers were measured in the background parenchyma and in the centre of the 15-mm lesions by a third-year radiology resident (A.E.), by placing circular regions of interest (ROIs). The 15-mm lesions were chosen to minimise measurement errors due to partial volume effects. All measurements were performed thrice on a vendor-specific software solution (Syngo.via, version VB10A; Siemens Healthineers). The standard deviation of the attenuation of the background parenchyma served as image noise. The lesion-to-background CNR was calculated as follows:

$$[ROI(B) - ROI(L)]/N,$$

where  $ROI(B)$  = mean attenuation of the background parenchyma;  $ROI(L)$  = mean attenuation of the lesion;  $N$  = mean image noise. The CNR values of the 15-mm lesions were averaged to create a mean CNR value for each dataset.

## Low-contrast lesion detection

Six radiologists (three board-certified radiologists and three residents with 12, 9, 7, 4, 3, and 1 years of experience in abdominal CT imaging, respectively) analysed the eight CT datasets of the medium phantom, and an additional six different radiologists (four board-certified radiologists and two residents with 10, 8, 7, 7, 4, and 4 years of experience in abdominal CT imaging, respectively) analysed the eight datasets of the large phantom. We chose two different reader groups to reduce the recall bias due to the large number of similar datasets. They had to scroll through the dataset and mark the position and grade of conspicuity of every lesion on a reading sheet. The conspicuity was graded according to the following three categories: 1 = perhaps present; 2 = most likely present; 3 = definitely present. The readers were free to change the window width and level for analysis.

Each reading session was separated by 1–3 weeks. To reduce recall bias, the succession of the datasets was randomised and the geometrical orientation of the images was changed after every reading: (1) dataset reading from first to last image; (2) dataset flipped vertically; (3) dataset rotated 90° clockwise; (4) dataset reading from last to first image.

## Subjective image quality

Ten different slices of each CT dataset were anonymised and randomly presented to the same 12 readers. Thus, every reader had to evaluate 80 different CT images. Each reader had to grade subjective image noise (grade 1 = unacceptable; grade 2 = above average; grade 3 = average; grade 4 = below average; grade 5 = absent) and image quality (grade 1 = bad, no diagnosis possible; grade 2 = poor, diagnostic confidence substantially reduced; grade 3 = moderate but sufficient for diagnosis; grade 4 = good; grade 5 = excellent) on a five-point scale for each image.

## Statistics

Marks made by the readers were compared with the construction plan of the phantom and were classified as true positive (TP) or false positive (FP). Missed lesions were regarded as false negatives (FNs). The data from six readers with each phantom were averaged, and the mean values were used for further statistics. Lesion detection rates and numbers of FPs with various combinations of peak kilovoltage (kVp) and reconstruction algorithms were compared using Fisher's exact test. The number of FPs was rather low; thus, we did not use statistical methods suitable for free response datasets. Inter-observer agreement was assessed by calculating Cohen's kappa value. Subjective image quality and subjective noise as rated by the readers with the various CT protocols were compared using Wilcoxon's matched pairs test. All of the analyses

were performed using the Statistica software package (Statsoft, Tulsa, OK, USA). The level of statistical significance was set at  $P < 0.05$ .

## Results

### Objective image quality

The CT numbers of the background parenchyma and the lesions were comparable for both reconstruction algorithms ( $P = 0.53$ ) (Table 1). MBIR decreased the image noise by 23.5%, 21.8%, 26.3%, and 23.5% in the medium phantom and by 26.9%, 25.8%, 27.9%, and 29.1% in the large phantom compared with FBP at 70, 80, 100, and 120 kVp, respectively. In the large phantom, the image noise was comparable between 120 and 100 kVp, but it was substantially increased at 80 kVp and 70 kVp within both reconstruction algorithms. MBIR increased the CNR by 32.4%, 28.1%, 32.0%, and 27.3% in the medium phantom and by 23.5%, 27.8%, 29.4%, and 33.3% in the large phantom compared with FBP at 70, 80, 100, and 120 kVp, respectively.

In the medium phantom, the reduction of the tube voltage increased the CNR by 60.7%, 46.4%, and 17.9% at 70, 80, and 100 kVp, respectively, compared with 120 kVp with MBIR and by 54.5%, 45.5%, and 13.6% at 70, 80, and 100 kVp, respectively, compared with 120 kVp with FBP. In the large phantom, the CNR increased by 5%, 15%, and 10% at 70, 80, and 100 kVp, respectively, compared with 120 kVp with MBIR and by 13.3%, 20%, and 13.3% at 70, 80, and 100 kVp, respectively, compared with 120 kVp with FBP (Fig. 1).

### Subjective image quality

In the medium phantom, MBIR resulted in significantly reduced image noise and increased image quality compared with FBP at all four tube voltages ( $P = 0.001$ , respectively) (Table 2).

In the large phantom, MBIR also resulted in significantly improved image noise based on pooled data of all four tube voltages ( $P = 0.029$ ). However, there was no significant difference in image quality at any of the four tube voltages ( $P = 0.79$ ). We must note that the subjective image quality parameters were graded higher in the large phantom compared with the medium phantom, likely due to different preferences regarding image quality between the two reader groups. The lesion conspicuity was graded slightly higher with FBP compared with MBIR in the medium phantom ( $2.38 \pm 0.7$  vs  $2.30 \pm 0.7$ , respectively,  $P = 0.024$ ) and similarly with both algorithms in the large phantom ( $2.15 \pm 0.8$  vs  $2.13 \pm 0.7$ ,  $P = 0.44$ ).

**Table 1** Objective image quality

	70 kVp		80 kVp		100 kVp		120 kVp	
	M	L	M	L	M	L	M	L
Backgr. (HU)								
FBP	186.3 ± 0.6	173.3 ± 0.6	167.7 ± 0.6	158.3 ± 0.6	145.0 ± 1.0	136.7 ± 1.2	132.3 ± 1.2	124.7 ± 0.6
MBIR	186.3 ± 1.2	172.0 ± 1.0	167.7 ± 0.6	157.0 ± 1.0	144.7 ± 0.6	135.7 ± 0.6	131.7 ± 0.6	124.7 ± 0.6
Lesion (HU)								
FBP	78.7 ± 0.6	68.0 ± 1.0	76.7 ± 0.6	71.3 ± 1.2	75.7 ± 0.6	72.0 ± 1.0	74.7 ± 0.6	66.3 ± 1.2
MBIR	80.7 ± 0.6	72.0 ± 1.0	77.7 ± 0.6	73.0 ± 1.0	76.0 ± 1.0	72.3 ± 1.2	75.3 ± 0.6	67.7 ± 0.6
Noise (HU)								
FBP	18.7 ± 0.6	38.7 ± 0.6	17.0 ± 1.0	28.7 ± 0.6	16.7 ± 0.6	24.0 ± 0.0	17.0 ± 0.0	23.0 ± 1.0
MBIR	14.3 ± 0.6	28.3 ± 0.6	13.3 ± 0.6	21.3 ± 0.6	12.3 ± 0.6	17.3 ± 0.6	13.0 ± 0.0	16.3 ± 0.6
CNR								
FBP	3.4 ± 2.3	1.7 ± 1.0	3.2 ± 2.1	1.8 ± 1.1	2.5 ± 1.6	1.7 ± 1.0	2.2 ± 1.2	1.5 ± 0.9
MBIR	4.5 ± 2.9	2.1 ± 1.3	4.1 ± 2.6	2.3 ± 1.5	3.3 ± 2.1	2.2 ± 1.4	2.8 ± 1.6	2.0 ± 1.4

The data are the mean values with standard deviations

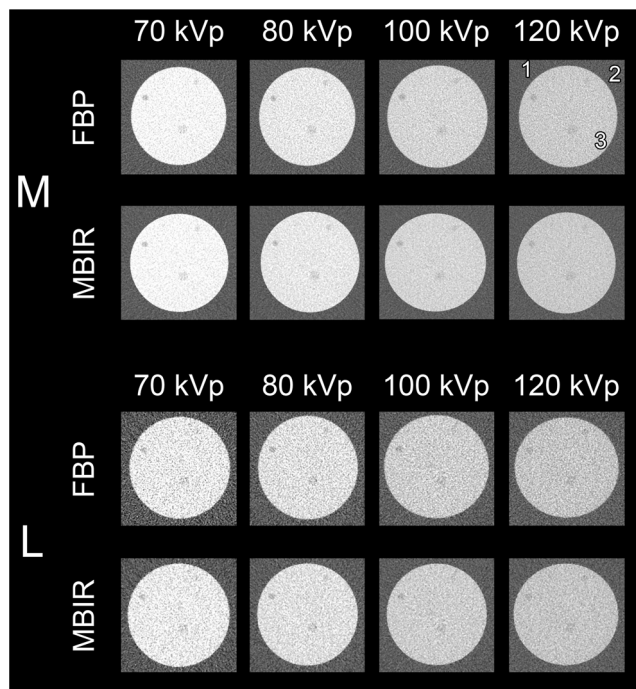
The values for the lesion represent a lesion with a diameter of 15 mm and a lesion-to-background contrast value of 50 HU

The CNR value is the mean of all of the measurements of the 15-mm lesions in a dataset

**Low-contrast lesion detection**

No significant difference in overall low-contrast lesion detection rate was noted regardless of the tube voltage or the reconstruction algorithm (detection rate of 76.7-80.7% for the

medium phantom and of 56.7-65.2% for the large phantom; *P* value range of 0.37-1) (Table 3). The average detection rate was 78.7% and 62.5% for the medium and large phantom, respectively. The inter-observer agreement was excellent for both sizes (overall  $\kappa$  range, 0.82-0.89). False-positive findings were substantially increased with MBIR compared with FBP for the medium phantom (*P* = 0.052). A sub-analysis of the lesion sizes showed, on average, comparable detection rates for both reconstruction algorithms (Table 4, Fig. 2a and b).



**Fig. 1** Example of three lesions: one lesion with a diameter of 10 mm and a lesion-to-liver contrast value of 25 HU (1); one lesion with a diameter of 10 mm and a lesion-to-liver contrast value of 10 HU (2); and one lesion with a diameter of 15 mm and a lesion-to-liver-contrast value of 10 HU (3). Despite the increased iodine attenuation of the liver parenchyma at lower tube voltages, the different lesions were equally detectable with MBIR and FBP at all four tube voltages. Note the higher image noise and decreased conspicuity of lesion (2) in the large phantom

**Discussion**

Despite an improvement in image quality parameters using MBIR compared with FBP, low-contrast detectability was not significantly improved. The improvement of image quality parameters by MBIR is consistent with the findings of recent studies [22–28]. However, our results are contradictory to those of the only other study evaluating the impact of the same MBIR (ADMIRE; Siemens Healthineers) on low-contrast detectability [19]. The authors reported a significant improvement in the detection rate of low-contrast lesions at 120 kVp using MBIR at a strength of 3 compared with FBP. Compared with our study, there are two major methodical differences in study design. First, a relatively small phantom (diameter of 16.5 cm) mimicking very lean or paediatric patients was used. In our study, we simulated medium and large patients (diameter of 30 cm and 40 cm), and our results revealed that patient size had a substantial impact on low-contrast detectability. We observed a substantial decrease in the average detection rate between the medium (78.7%) and large phantom (62.5%). The second main difference was that lesion detection accuracy was evaluated by a two-alternative forced choice approach and by scoring the total number of

**Table 2** Data for subjective evaluation of image noise and image quality

	70 kVp		80 kVp		100 kVp		120 kVp	
	M	L	M	L	M	L	M	L
<b>Noise</b>								
FBP	2.7 ± 0.7 (2.7   2.0   3.2)	2.3 ± 0.7 (2.0   1.9   3.2)	2.7 ± 0.6 (2.6   2.3   3.3)	2.8 ± 0.7 (2.8   2.1   3.3)	2.7 ± 0.6 (2.9   2.1   3.3)	3.2 ± 0.6 (3.4   3.0   3.5)	2.3 ± 0.5 (2.4   2.0   2.6)	3.2 ± 0.7 (3.5   2.8   3.7)
MBIR	3.1 ± 0.2 (3.1   3.0   3.2)	2.5 ± 0.8 (2.2   1.9   3.3)	3.2 ± 0.4 (3.2   2.8   3.6)	2.8 ± 0.6 (2.8   2.4   3.3)	3.3 ± 0.5 (3.3   2.8   3.6)	3.4 ± 0.6 (3.5   3.4   3.6)	3.0 ± 0.4 (3.0   2.6   3.3)	3.4 ± 0.8 (3.4   3.3   3.7)
<b>Quality</b>								
FBP	2.7 ± 0.7 (2.7   2.0   3.2)	2.8 ± 0.5 (2.8   2.6   3.2)	2.7 ± 0.6 (2.6   2.3   3.3)	3.0 ± 0.5 (2.9   2.5   3.1)	2.7 ± 0.6 (2.9   2.1   3.3)	3.2 ± 0.4 (3.2   2.9   3.4)	2.3 ± 0.5 (2.4   2.0   2.6)	3.3 ± 0.4 (3.4   3.2   3.6)
MBIR	3.2 ± 0.4 (3.1   2.9   3.6)	2.7 ± 0.5 (2.7   2.4   3.0)	3.5 ± 0.3 (3.6   3.2   3.8)	3.0 ± 0.3 (3.0   2.8   3.2)	3.4 ± 0.4 (3.4   3.1   3.7)	3.3 ± 0.4 (3.3   3.0   3.6)	3.1 ± 0.4 (3.1   2.7   3.4)	3.4 ± 0.7 (3.6   2.8   3.9)

Image noise was rated on a five-point scale (1 = major, 5 = absent)

Image quality was also rated on a five-point scale (1 = bad, no diagnosis possible; 5 = excellent)

Numbers in brackets are as follows: (median | lower quartile | upper quartile)

visible object groups. In contrast, we chose to simulate a typical clinical scenario by performing the reading on a full volumetric stack of images, as is common with clinical CT datasets. Twelve clinical radiologists with broad ranges of experience participated in our study. We hypothesise that these major differences in study design contributed to the discrepancy in low-contrast detection rate.

We also evaluated the impact of different tube voltages on image quality and low-contrast lesion detection. We were particularly interested to analyse lower tube voltages because the

**Table 3** Data for the detection of 45 simulated hypodense lesions for both phantom sizes

Size	Conspicuity		No. of TP		No. of FP		Sensitivity (%)	
	M	L	M	L	M	L	M	L
<b>70 kV</b>								
FBP	2.35	2.11	36.0	27.3	1.8	1.8	80.0	60.7
MBIR	2.32	2.19	36.3	25.5	2.7	3.2	80.7	56.7
<b>80 kV</b>								
FBP	2.39	2.24	35.8	29.2	3.8	4.2	79.6	64.8
MBIR	2.35	2.09	35.7	28.8	5.8	2.2	79.3	64.1
<b>100 kV</b>								
FBP	2.44	2.17	35.2	29.3	1.2	3.0	78.2	65.2
MBIR	2.38	2.19	36.0	28.3	6.5	2.2	80.0	63.0
<b>120 kV</b>								
FBP	2.33	2.06	33.7	27.7	2.2	2.0	74.8	61.5
MBIR	2.17	2.06	34.5	28.8	5.7	3.7	76.7	64.1

Data are the means of results from 12 independent readers for all 45 lesions

The conspicuity of the tumours was rated on a three-point scale: 1 = might be present; 2 = most likely present; 3 = definitely present

TP true-positive findings, FP false-positive findings

third-generation dual-source CT scanner used in this study potentially enabled imaging of medium and large patients at tube voltages of 70 and 80 kVp. In former scanner generations, these tube voltages were typically restricted to small patients. Changes in tube voltage could greatly influence image noise and CNR, which was an observation confirmed in our study. Our phantom design, with a higher iodine content of the background compared with the hypoattenuating lesions, augmented the beneficial effect of low tube voltages on lesion-to-background CNR. Lower tube voltages of 70, 80, and 100 kVp resulted in increased CNR for both phantom sizes compared with 120 kVp (16–58% in the medium phantom and 9–18% in the large phantom, respectively). Image noise was relatively constant within each reconstruction algorithm in the medium phantom. However, comparing data at low kVp to those at high kVp, the notably larger size of the large phantom (40 cm) caused a marked increase in the phantom attenuation at lower kVp settings, leading to a proportionate increase in image noise. This observation confirmed that imaging at very low kVp settings could still increase image noise disproportionately in large patients despite recent access to the very high tube current output by the third-generation dual-source CT used in this study.

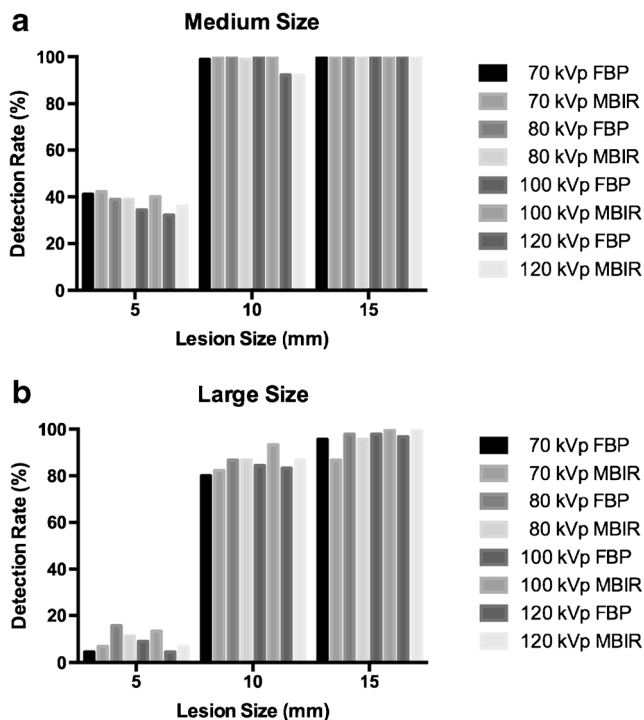
In the scientific literature, only a few studies evaluated the effect of reducing the tube voltage on diagnostic accuracy in abdominal CT. Beneficial effects of low tube voltages on lesion detection or conspicuity of hypovascular and hypervascular lesions have been reported in phantom and clinical studies [10, 11, 29–32]. Despite the overall improvement of CNR by reducing the tube voltage, the low-contrast lesion detection was not significantly altered in our study. Our findings might be explained by the low iodine concentration of the hypodense lesions in our phantom. Contrast enhancement at a particular tube voltage is proportional to iodine concentration

**Table 4** Number of true-positive findings by diameter and contrast value of the simulated tumours

Size	Tumour diameter (mm)						Lesion-to-liver contrast (HU)					
	5		10		15		10		25		50	
	M	L	M	L	M	L	M	L	M	L	M	L
70 kV												
FBP	6.2 (41.3)	0.7 (4.6)	14.8 (98.7)	12.0 (80.0)	15.0 (100)	14.3 (95.3)	10.0 (66.7)	7.3 (48.7)	11.2 (74.7)	9.0 (60.0)	14.8 (98.7)	10.7 (71.3)
MBIR	6.3 (42.0)	1.0 (6.7)	15.0 (100)	12.3 (82.0)	15.0 (100)	13.0 (86.7)	10.3 (68.7)	7.3 (48.7)	11.3 (75.3)	8.0 (53.3)	14.7 (98.0)	11.0 (73.3)
80 kV												
FBP	5.8 (38.7)	2.3 (15.3)	15.0 (100)	13.0 (86.7)	15.0 (100)	14.7 (98.0)	10.0 (66.7)	7.7 (51.3)	11.2 (74.7)	10.7 (71.3)	14.7 (98.0)	11.7 (78.0)
MBIR	5.8 (38.7)	1.7 (11.3)	14.8 (98.7)	13.0 (86.7)	15.0 (100)	14.3 (95.3)	9.8 (65.3)	8.7 (58.0)	11.7 (78.0)	10.0 (66.7)	14.2 (94.7)	10.3 (68.7)
100 kV												
FBP	5.2 (34.7)	1.3 (8.7)	15.0 (100)	12.7 (84.7)	15.0 (100)	14.7 (98.0)	10.0 (66.7)	7.3 (48.7)	11.2 (74.7)	10.3 (68.7)	14.0 (93.3)	11.0 (73.3)
MBIR	6.0 (40.0)	2.0 (13.3)	15.0 (100)	14.0 (93.3)	15.0 (100)	15.0 (100)	10.0 (66.7)	9.3 (62.0)	12.2 (81.3)	9.7 (64.7)	13.8 (92.0)	12.0 (80.0)
120 kV												
FBP	4.8 (32.0)	0.7 (4.7)	13.8 (92.0)	12.5 (83.3)	15.0 (100)	14.5 (96.7)	9.2 (61.3)	7.2 (48.0)	10.8 (72.0)	10.0 (66.7)	13.7 (91.3)	10.5 (70.0)
MBIR	5.4 (36.0)	1.0 (6.7)	13.8 (92.0)	13.0 (86.7)	15.0 (100)	15.0 (100)	9.0 (60.0)	8.0 (53.3)	11.2 (74.7)	10.0 (66.7)	14.0 (93.3)	11.0 (73.3)

Data are the means of all datasets assessed by six independent and blinded readers for each phantom size  
 The mean sensitivity for each subgroup is presented in parentheses  
 The total number of simulated tumours for each diameter and each contrast value was 15 in each dataset

[33]. Although the CT numbers of the background



**Fig. 2** The two graphs show the percentage detection rate for each dataset depending on the lesion size in the medium (a) and the large (b) phantom. Note the substantial lower detection rate of 5-mm lesions in the large phantom

parenchyma increased at lower tube voltages, the attenuation of the hypodense lesions differed only slightly. However, the low iodine concentration of our lesions simulates a realistic physiological scenario given that these lesions are typically hypovascular and exhibit low iodine uptake in vivo. Further studies must determine the thresholds of difference in iodine concentration between lesions and background that are needed to make imaging at low tube voltages beneficial.

Another interesting result of our study is that there might also be potential disadvantages to the use of iterative reconstruction algorithms. Although MBIR significantly improved subjective image noise for both sizes and image quality in the medium phantom, the overall lesion conspicuity was graded slightly lower compared with FBP. We observed that the number of false-positive findings was substantially increased using MBIR in the medium phantom (on average five vs two FPs with MBIR compared with FBP). This finding was close to statistical significance ( $P = 0.052$ ), and we strongly believe that significance would have been attained with a higher lesion count.

The results of our study are consistent with some studies, but also in contrast to several other studies evaluating the impact of IRs on diagnostic accuracy [14, 15, 34]. These controversial results are a consequence of the wide differences in applied study designs. In particular, phantom studies differ substantially in regard to lesion diameter, lesion-to-background contrast values, phantom shape and size and in

regard to the investigated clinical task. This raises the question as to whether direct comparisons among these studies are still meaningful. In addition, a comprehensive framework that fully encompasses image quality and diagnostic accuracy in a clinical context is currently lacking. As noted by Fletcher et al. [35], image quality parameters alone, such as image noise, are not necessarily correlated with the diagnostic accuracy of radiologists.

To date, no standardised, objective and reproducible method is routinely used to assess the impact of novel radiation dose reduction techniques, such as IR technique, on diagnostic accuracy. The implementation of model-observer algorithms could overcome this problem in future. Model-observer algorithms simulate human observer performance and therefore potentially enable standardised, reproducible evaluation of image quality in a clinical context [36–38].

Our study had several limitations. First, our phantom had a cylindrical shape, which did not precisely reflect the geometry of the abdomen of humans. Second, we only evaluated one specific type of iterative reconstruction algorithm of one vendor. We did not perform a comparison among different generations of iterative reconstruction algorithms, and we only evaluated a single strength. The investigation of additional reconstruction strength levels would not have been practicable due to recall bias. Third, there was possible recall bias of the readers with regard to the location of the lesions. We minimised this bias by changing the geometrical orientation of the datasets and by separating the reading sessions by 1–3 weeks.

In conclusion, the improvement of quantitative image quality parameters by MBIR compared with FBP at different tube voltages did not result in significantly improved low-contrast lesion detectability in simulated abdominal CT of medium and large patients. Although image noise and CNR are well accepted parameters for grading quantitative image quality, they tend to not fully encompass the influencing factors that determine diagnostic accuracy, thus limiting their value as a fundamental basis for dose optimisation in CT. Model-observer algorithms could replace crude measurements of quantitative image quality parameters in the future by enabling benchmarking of image quality and diagnostic accuracy in a clinical context.

#### Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Sebastian T. Schindera.

**Conflict of interest** Sebastian T. Schindera received a research grant by Siemens Healthcare and Bayer Healthcare

**Funding** The authors state that this work has not received any funding.

**Statistics and biometry** No complex statistical methods were necessary for this paper.

**Ethical approval** Institutional Review Board approval was not required because of the design as a phantom study.

#### Methodology

- prospective
- experimental
- performed at one institution

#### References

1. Tirumani SH, Kim KW, Nishino M, Howard SA, Krajewski KM, Jagannathan JP et al (2014) Update on the role of imaging in management of metastatic colorectal cancer. *Radiographics* 34:1908–1928
2. Kanal KM, Chung JH, Wang J, Bhargava P, Kohr JR, Shuman WP et al (2011) Image noise and liver lesion detection with MDCT: a phantom study. *AJR Am J Roentgenol* 197:437–441
3. Goenka AH, Herts BR, Obuchowski NA, Primak AN, Dong F, Karim W et al (2014) 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* 272:154–163
4. Deak Z, Grimm JM, Treitl M, Geyer LL, Linsenmaier U, Komer M et al (2013) Filtered back projection, adaptive statistical iterative reconstruction, and a model-based iterative reconstruction in abdominal CT: an experimental clinical study. *Radiology* 266:197–206
5. Aurumskjold ML, Ydstrom K, Tingberg A, Soderberg M (2017) Improvements to image quality using hybrid and model-based iterative reconstructions: a phantom study. *Acta Radiol* 58:53–61
6. Holmquist F, Nyman U, Siemund R, Geijer M, Soderberg M (2016) Impact of iterative reconstructions on image noise and low-contrast object detection in low kVp simulated abdominal CT: a phantom study. *Acta Radiol* 57:1079–1088
7. Kim JH, Choo KS, Moon TY, Lee JW, Jeon UB, Kim TU et al (2016) Comparison of the image qualities of filtered back-projection, adaptive statistical iterative reconstruction, and model-based iterative reconstruction for CT venography at 80 kVp. *Eur Radiol* 26:2055–2063
8. Song JS, Lee JM, Sohn JY, Yoon JH, Han JK, Choi BI (2015) Hybrid iterative reconstruction technique for liver CT scans for image noise reduction and image quality improvement: evaluation of the optimal iterative reconstruction strengths. *Radiol Med* 120: 259–267
9. Kalva SP, Sahani DV, Hahn PF, Saini S (2006) Using the K-edge to improve contrast conspicuity and to lower radiation dose with a 16-MDCT: a phantom and human study. *J Comput Assist Tomogr* 30: 391–397
10. Schindera ST, Nelson RC, Mukundan S Jr, Paulson EK, Jaffe TA, Miller CM et al (2008) Hypervascular liver tumors: low tube voltage, high tube current multi-detector row CT for enhanced detection—phantom study. *Radiology* 246:125–132
11. Robinson E, Babb J, Chandarana H, Macari M (2010) Dual source dual energy MDCT: comparison of 80 kVp and weighted average 120 kVp data for conspicuity of hypo-vascular liver metastases. *Invest Radiol* 45:413–418
12. Mievil FA, Gudinchet F, Brunelle F, Bochud FO, Verdun FR (2013) Iterative reconstruction methods in two different MDCT scanners: physical metrics and 4-alternative forced-choice detectability experiments—a phantom approach. *Phys Med* 29:99–110

13. Yu MH, Lee JM, Yoon JH, Baek JH, Han JK, Choi BI et al (2013) Low tube voltage intermediate tube current liver MDCT: sinogram-affirmed iterative reconstruction algorithm for detection of hypervascular hepatocellular carcinoma. *AJR Am J Roentgenol* 201:23–32
14. Schindera ST, Odedra D, Raza SA, Kim TK, Jang HJ, Szucs-Farkas Z et al (2013) Iterative reconstruction algorithm for CT: can radiation dose be decreased while low-contrast detectability is preserved? *Radiology* 269:511–518
15. Schindera ST, Odedra D, Mercer D, Thiphavong S, Chou P, Szucs-Farkas Z et al (2014) Hybrid iterative reconstruction technique for abdominal CT protocols in obese patients: assessment of image quality, radiation dose, and low-contrast detectability in a phantom. *AJR Am J Roentgenol* 202:W146–W152
16. Chang W, Lee JM, Lee K, Yoon JH, Yu MH, Han JK et al (2013) Assessment of a model-based, iterative reconstruction algorithm (MBIR) regarding image quality and dose reduction in liver computed tomography. *Invest Radiol* 48:598–606
17. Volders D, Bols A, Haspeslagh M, Coenegrachts K (2013) Model-based iterative reconstruction and adaptive statistical iterative reconstruction techniques in abdominal CT: comparison of image quality in the detection of colorectal liver metastases. *Radiology* 269:469–474
18. Fontarensky M, Alfidja A, Perignon R, Schoenig A, Perrier C, Mulliez A et al (2015) Reduced radiation dose with model-based iterative reconstruction versus standard dose with adaptive statistical iterative reconstruction in abdominal CT for diagnosis of acute renal colic. *Radiology* 276:156–166
19. Solomon J, Mileto A, Ramirez-Giraldo JC, Samei E (2015) Diagnostic performance of an advanced modeled iterative reconstruction algorithm for low-contrast detectability with a third-generation dual-source multidetector CT scanner: potential for radiation dose reduction in a multireader study. *Radiology* 275:735–745
20. Menke J (2005) Comparison of different body size parameters for individual dose adaptation in body CT of adults. *Radiology* 236:565–571
21. Gordic S, Morsbach F, Schmidt B, Allmendinger T, Flohr T, Husarik D et al (2014) Ultralow-dose chest computed tomography for pulmonary nodule detection: first performance evaluation of single energy scanning with spectral shaping. *Invest Radiol* 49:465–473
22. Gordic S, Desbiolles L, Stolzmann P, Gantner L, Leschka S, Husarik DB et al (2014) Advanced modelled iterative reconstruction for abdominal CT: qualitative and quantitative evaluation. *Clin Radiol* 69:e497–e504
23. Botsikas D, Stefanelli S, Boudabbous S, Toso S, Becker CD, Montet X (2014) Model-based iterative reconstruction versus adaptive statistical iterative reconstruction in low-dose abdominal CT for urolithiasis. *AJR Am J Roentgenol* 203:336–340
24. Patino M, Fuentes JM, Hayano K, Kambadakone AR, Uyeda JW, Sahani DV (2015) A quantitative comparison of noise reduction across five commercial (hybrid and model-based) iterative reconstruction techniques: an anthropomorphic phantom study. *AJR Am J Roentgenol* 204:W176–W183
25. Scholtz JE, Wichmann JL, Husers K, Albrecht MH, Beeres M, Bauer RW et al (2016) Third-generation dual-source CT of the neck using automated tube voltage adaptation in combination with advanced modeled iterative reconstruction: evaluation of image quality and radiation dose. *Eur Radiol* 26:2623–2631
26. Schaller F, Sedlmair M, Raupach R, Uder M, Lell M (2016) Noise reduction in abdominal computed tomography applying iterative reconstruction (ADMIRE). *Acad Radiol* 23:1230–1238
27. de Margerie-Mellon C, de Bazelaire C, Montlahuc C, Lambert J, Martineau A, Coulon P et al (2016) Reducing radiation dose at chest CT: comparison among model-based type iterative reconstruction, hybrid iterative reconstruction, and filtered back projection. *Acad Radiol* 23:1246–1254
28. Park HJ, Lee JM, Park SB, Lee JB, Jeong YK, Yoon JH (2016) Comparison of knowledge-based iterative model reconstruction and hybrid reconstruction techniques for liver CT evaluation of hypervascular hepatocellular carcinoma. *J Comput Assist Tomogr* 40:863–871
29. Holm J, Loizou L, Albiin N, Kartalis N, Leidner B, Sundin A (2012) Low tube voltage CT for improved detection of pancreatic cancer: detection threshold for small, simulated lesions. *BMC Med Imaging* 12:20
30. Hernandez-Giron I, Calzado A, Geleijns J, Joemai RM, Veldkamp WJ (2015) Low contrast detectability performance of model observers based on CT phantom images: kVp influence. *Phys Med* 31:798–807
31. Marin D, Nelson RC, Bamhart H, Schindera ST, Ho LM, Jaffe TA et al (2010) Detection of pancreatic tumors, image quality, and radiation dose during the pancreatic parenchymal phase: effect of a low-tube-voltage, high-tube-current CT technique—preliminary results. *Radiology* 256:450–459
32. Lv P, Liu J, Zhang R, Jia Y, Gao J (2015) Combined use of automatic tube voltage selection and current modulation with iterative reconstruction for CT evaluation of small hypervascular hepatocellular carcinomas: effect on lesion conspicuity and image quality. *Korean J Radiol* 16:531–540
33. Bae KT (2010) Intravenous contrast medium administration and scan timing at CT: considerations and approaches. *Radiology* 256:32–61
34. Dodge CT, Tamm EP, Cody DD, Liu X, Jensen CT, Wei W et al (2016) Performance evaluation of iterative reconstruction algorithms for achieving CT radiation dose reduction—a phantom study. *J Appl Clin Med Phys* 17:5709
35. Fletcher JG, Yu L, Li Z, Manduca A, Blezek DJ, Hough DM et al (2015) Observer performance in the detection and classification of malignant hepatic nodules and masses with CT image-space denoising and iterative reconstruction. *Radiology* 276:465–478
36. Tseng HW, Fan J, Kupinski MA, Sainath P, Hsieh J (2014) Assessing image quality and dose reduction of a new X-ray computed tomography iterative reconstruction algorithm using model observers. *Med Phys* 41:071910
37. Ott JG, Ba A, Racine D, Viry A, Bochud FO, Verdun FR (2016) 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*. doi:10.1016/j.zemedi.2016.04.002
38. Racine D, Ryckx N, Ba A, Ott JG, Bochud FO, Verdun FR (2016) Benchmarking of CT for patient exposure optimisation. *Radiat Prot Dosimetry* 169:78–83