Six iterative reconstruction algorithms in brain CT- A phantom study on image quality at different radiation doses.

Löve, Askell; Aurumskjöld, Marie-Louise; Siemund, Roger; Stålhammar, Fredrik; Björkman-Burtscher, Isabella; Söderberg, Marcus

Published in:
British Journal of Radiology

DOI:
10.1259/bjr.20130388

2013

BJR

Received: 30 June 2013 | Revised: 4 September 2013 | Accepted: 16 September 2013

doi: 10.1259/bjr.20130388

Cite this article as:
Löve A, Olsson M-L, Siemund R, Stålhammar F, Björkman-Burtscher IM, Söderberg M. Six iterative reconstruction algorithms in brain CT: a phantom study on image quality at different radiation dose levels. Br J Radiol 2013;86:20130388.

FULL PAPER

Six iterative reconstruction algorithms in brain CT: a phantom study on image quality at different radiation dose levels

[1]A LÖVE, MD, [2]M-L OLSSON, MSc, [1]R SIEMUND, MD, PhD, [1]F STÅLHAMMAR, MD, [1,3]I M BJÖRKMAN-BURTSCHER, MD, PhD and [2]M SÖDERBERG, PhD

[1]Department of Neuroradiology, Skåne University Hospital, Lund University, Lund, Sweden
[2]Medical Radiation Physics Malmö, Skåne University Hospital, Lund University, Lund, Sweden
[3]Lund University Bioimaging Centre, Lund University, Lund, Sweden

Address correspondence to: Dr Askell Löve
E-mail: *askell.love@skane.se*

Objective: To evaluate the image quality produced by six different iterative reconstruction (IR) algorithms in four CT systems in the setting of brain CT, using different radiation dose levels and iterative image optimisation levels.
Methods: An image quality phantom, supplied with a bone mimicking annulus, was examined using four CT systems from different vendors and four radiation dose levels. Acquisitions were reconstructed using conventional filtered back-projection (FBP), three levels of statistical IR and, when available, a model-based IR algorithm. The evaluated image quality parameters were CT numbers, uniformity, noise, noise-power spectra, low-contrast resolution and spatial resolution.
Results: Compared with FBP, noise reduction was achieved by all six IR algorithms at all radiation dose levels, with further improvement seen at higher IR levels. Noise-power spectra revealed changes in noise distribution relative to the FBP for most statistical IR algorithms, especially the two model-based IR algorithms. Compared with FBP, variable degrees of improvements were seen in both objective and subjective low-contrast resolutions for all IR algorithms. Spatial resolution was improved with both model-based IR algorithms and one of the statistical IR algorithms.
Conclusion: The four statistical IR algorithms evaluated in the study all improved the general image quality compared with FBP, with improvement seen for most or all evaluated quality criteria. Further improvement was achieved with one of the model-based IR algorithms.
Advances in knowledge: The six evaluated IR algorithms all improve the image quality in brain CT but show different strengths and weaknesses.

Iterative reconstruction (IR) algorithms are one of the most recent advances in CT. Since the introduction of the first IR algorithm in 2008 [1], multiple clinical studies have shown the potential of such algorithms to improve the image quality and allow for the reduction of radiation dose while maintaining diagnostic acceptability [2–7].

Although all IR algorithms perform iterative image optimisation at some point in the CT image generation process, there are considerable technical differences between the available IR solutions. Furthermore, some vendors even offer more than one type of IR algorithm in their product range. Although detailed mechanisms of the current algorithms remain undisclosed, they can be classified into two basic categories [8,9] (Table 1): (A) statistical iterative optimisation based on photon statistics, assuming an ideal system and (B) model-based iterative optimisation that in addition attempts to model the system and the acquisition process, including system optics. Both of these optimise the image date in the projection domain and the image domain. A disadvantage of the model-based algorithms is the increased reconstruction time that makes them unsuitable at present for acute clinical situations.

With a few exceptions [10,11], studies on IR from the literature have compared IR algorithms with filtered back-projection (FBP) reconstruction from the same vendor. As the IR algorithms can be expected to have different strengths and weaknesses, side-by-side assessment of their performance should be of interest. Such evaluation is best carried out in a phantom under standardised conditions.

The purpose of this phantom study was to objectively and subjectively evaluate the image quality produced by six different IR algorithms in four CT systems from different vendors, using a variety of radiation dose levels and iterative

Table 1. Classification of major iterative reconstruction algorithms based on mechanism of function, including acronyms, product names and vendors (in alphabetical order)

Algorithm	Acronym	Vendor
Statistical iterative optimisation		
ASIR	Adaptive Statistical Iterative Reconstruction	GE Healthcare, Milwaukee, MI
iDOSE[4]	Product name, not acronym	Philips Medical Systems, Best, Netherlands
SAFIRE	Sinogram Affirmed Iterative Reconstruction	Siemens Healthcare, Forchheim, Germany
AIDR 3D	Adaptive Iterative Dose Reduction 3D	Toshiba Medical Systems, Tokyo, Japan
Model-based iterative optimisation		
Veo	Product name, not acronym	GE Healthcare
IMR	Iterative Model Reconstruction	Philips Medical Systems

image optimisation levels. The study was designed to simulate the demanding conditions of brain CT, with emphasis on noise and low-contrast resolution.

MATERIALS AND METHODS

CT examinations

Examinations were conducted using four different CT systems, Discovery™ CT750 HD (GE Healthcare, Milwaukee, MI), Brilliance iCT (Philips Medical Systems, Best, Netherlands), Definition Flash (Siemens Healthcare, Forchheim, Germany) and Aquilion ONE (Toshiba Medical Systems, Tokyo, Japan). All scans were performed in a helical mode using a 20-cm-diameter image quality phantom (Catphan® 600; The Phantom Laboratory, Greenwich, NY), with an external bone-mimicking ring (Teflon® Annulus CTP299; The Phantom Laboratory) to simulate the beam hardening effects created by the skull in clinical brain CT. The phantom includes multiple modules for assessment of CT imaging performance such as CT numbers, spatial resolution, low-contrast resolution and noise/uniformity.

Examination parameters were chosen according to the recommendations for adult brain CT published by the American Association of Physicists in Medicine (AAPM) [12], except for tube current (mA), which was adjusted to give a fixed baseline CT dose index by volume ($CTDI_{vol}$) of 120 mGy for all the four CT systems. Because the AAPM recommendations did not specifically address the reconstruction filters for the IR algorithms, we used the same filters for IR and FBP if available; otherwise, a default brain reconstruction filter was used. Automatic exposure control was disabled.

Radiation dose

The acquisitions were carried out on all four CT systems using four different radiation dose levels: 120 mGy (100%, baseline), 84 (70%), 48 (40%) and 12 mGy (10%). Radiation dose reduction was accomplished through reduction in tube current (mA) only. Exact $CTDI_{vol}$ was achieved by direct measurement of $CTDI_{vol}$ in a 16-cm-diameter polymethyl methacrylate CTDI phantom using a CT Dose Profiler connected to Piranha multifunction meter (all from RTI Electronics, Gothenburg, Sweden). The tube current was adjusted until the correct dose level was reached. The baseline dose of 120 mGy for the 20-cm phantom was chosen, as it roughly equals 60 mGy in a standard 16-cm phantom, which is the recommended dose for head CT imaging according to the European guidelines on quality criteria for computed tomography [13].

Image reconstruction

All reconstructions were created using standard commercially available algorithms, except for the Philips model-based reconstructions that were created using a pre-release iterative model reconstruction (IMR) prototype (v. 1.2.0.0 R06), which is not approved for clinical use.

Raw data from all 16 combinations of CT systems and dose levels were reconstructed with the respective FBP and IR algorithms, using different IR levels where possible, resulting in 4–5 reconstructions per CT system and a total of 72 reconstructions. The IR levels were chosen to reflect the complete range of available levels: low (IR1), medium (IR2) and high (IR3). Of the model-based IR algorithms, GE Veo™ did not offer any adjustable reconstruction parameters, whereas for Philips IMR, a brain reconstruction (low contrast) was chosen with a medium IR level (Level 2 of 3).

An overview of the evaluated CT systems, IR algorithms, IR levels, reconstruction filters and key examination parameters is presented in Table 2.

Image quality assessment

The program AutoQA Lite™ (v. 3.01, 2010; Iris QA, LLC, Frederick, MD) was used for automatic evaluation of the image quality parameters in the phantom images, including assessment of CT numbers, noise, uniformity and spatial resolution. The program automatically detects the phantom orientation, performs appropriate measurements in the different phantom modules and provides summary reports with descriptive statistical data. CT numbers (Figure 1a) were measured for seven different materials encompassing a wide range of Hounsfield units (HU): air (−1000 HU), polymethylpentene (−200 HU), low-density polyethylene (−100 HU), polystyrene (−35 HU), acrylic (120 HU), Delrin (340 HU) and Teflon (990). In the

Table 2. Overview of evaluated CT systems, IR-algorithms and radiation dose levels

Parameters	GE	Philips	Siemens	Toshiba
	Discovery CT750 HD	Brilliance iCT	Definition Flash	Aquilion ONE
Algorithms	FBP	FBP	FBP	FBP
IR1	ASIR 10%	iDose^4 Level 1	SAFIRE Strength 1	AIDR 3D mild
IR2	ASIR 50%	iDose^4 Level 3	SAFIRE Strength 3	AIDR 3D standard
IR3	ASIR 90%	iDose^4 Level 5	SAFIRE Strength 5	AIDR 3D strong
Model-based IR	Veo	IMR low-contrast L2	None	None
Tube current (mA)				
120 mGy	685	646	432	400
84 mGy	490	452	302	280
48 mGy	280	258	173	170
12 mGy	70	65	43	45
Tube voltage (kV)	120	120	120	120
Collimation (mm)	32×0.625	64×0.625	128×0.6	32×0.5
Pitch	0.531	0.391	0.55	0.656
Rotation time (s)	0.5	0.5	1.0	0.75
Display FOV (mm)	240	240	240	240
Reconstruction filters	Standard Plus (FBP/IR)	Brain UB (FBP/IR)	H31s (FBP)	Head Brain: FC26
	Default (Veo)	Low-contrast L2 (IMR)	J30s (SAFIRE)	
Thickness (mm)	5	5	5	5
Increment (mm)	5	5	5	5

AIDR 3D, adaptive iterative dose reduction–three dimensional; ASIR, adaptive statistical iterative reconstruction; FBP, filtered back-projection; FOV, field of view; IMR, iterative model reconstruction; IR, iterative reconstruction; SAFIRE, sinogram affirmed iterative reconstruction.
Scan parameters were adapted directly from the recommendation of American Association of Physicists in Medicine [12], with some exceptions as explained in Materials and Methods. Discovery CT750 HD obtained from GE Healthcare, Milwaukee, MI; Brilliance iCT, Philips Medical Systems, Best, Netherlands; Definition Flash, Siemens Healthcare, Forchheim, Germany; Aquilion ONE, Toshiba Medical Systems, Tokyo, Japan.

image uniformity module of the phantom (Figure 1d), mean attenuation (HU) and noise [standard deviation (SD)] were measured in five identical regions of interest (ROIs), one centrally located and four peripherally located. Uniformity was calculated as the maximal difference in the CT numbers between the central ROI and a peripheral ROI. For evaluation of signal-to-noise ratios (SNR), measurements from all five ROIs were pooled.

Noise-power spectra (NPS) were calculated in the reconstructions from the uniform module of the phantom as:

$$\mathrm{NPS}(u,v)=N_xN_y\Delta_x\Delta_y\left\{\left|\mathrm{FT}[\Delta I(x,y)]\right|^2\right\},$$

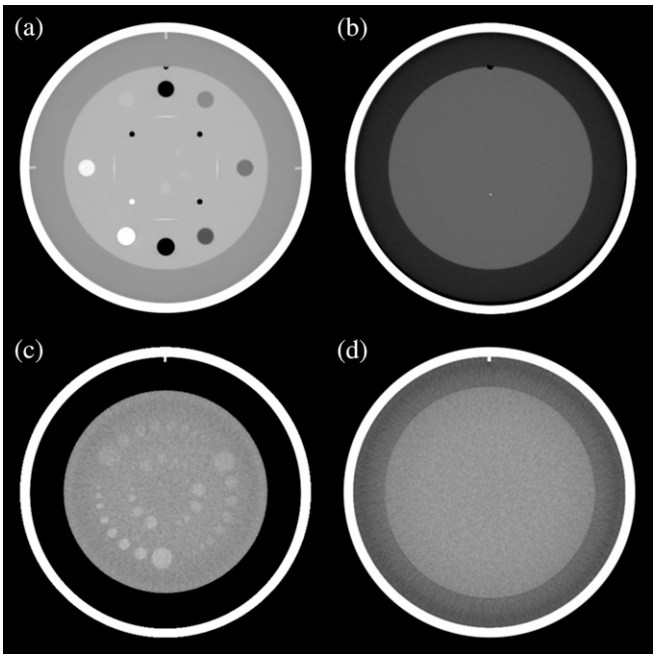
where (u, v) and (x, y) are a Fourier transform pair, ΔI is the deviation from the mean of a noise image (flat-field image), N_x and N_y are the number of pixels and Δ_x and Δ_y are the pixel sizes in the x and y direction, respectively [14]. NPS was calculated for ROIs of 92.8 cm^2 (205×205 pixels, 0.47-mm pixel size) located at the phantom centre using the image processing program, ImageJ (v. 1.46; National Institutes of Health, Bethesda, MD) and the plugin Radial Profile Plot (by Paul Baggethun, v. 14, May 2009; Pittsburgh, PA). The resulting NPS curves show the distribution of noise power in the frequency space.

Low-contrast resolution was assessed in two steps using the low-contrast module of the phantom (Figure 1c). The module contains three sets of outer supraslice cylinders, with a z-axis dimension of 40 mm, which according to the manufacturer have a nominal contrast of 1.0% (10 HU), 0.5% (5 HU) and 0.3% (3 HU). Each set consists of nine cylinders with diameters ranging from 2 to 15 mm. Objective assessment was done by measuring the CT numbers (HU) and noise (SD) in identical ROIs placed in the largest cylinder in each of the three sets and in the background. Contrast-to-noise ratios (CNRs) were calculated using the formula:

$$\mathrm{CNR}=\frac{\mathrm{HU}_{\mathrm{Object}}-\mathrm{HU}_{\mathrm{Background}}}{\mathrm{SD}_{\mathrm{Background}}},$$

Subjective assessment of all 72 reconstructions was carried out in consensus by 5 observers, 3 radiologists and 2 medical physicists with 5–22 years of experience in reading CT images. Two image quality criteria were assessed separately for each of the three sets of supraslice cylinders (Figure 1c): (A) smallest discernible

Figure 1. Images, as examples, from the Catphan® phantom (The Phantom Laboratory, Greenwich, NY) with bone-mimicking annulus showing modules for assessment of CT numbers (a), spatial resolution (b), low-contrast resolution (c) and noise and uniformity (d). To ensure an adequate reproduction of the internal structure of the phantom, each image has optimised window levels and window widths.



cylinder and (B) smallest sharply defined cylinder. The images were presented individually in a randomised order on a picture archiving and communication system workstation using the viewing and scoring software Viewer for Digital Evaluation of X-ray images (ViewDex; The Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden), v. 2.0 [15]. As absolute CT numbers varied slightly between reconstructions, the window level was adjusted for each reconstruction to match the mean attenuation in the homogeneous background, while the window width was fixed at 80 HU, as is common for clinical brain CT.

Finally, spatial resolution [modulation transfer function (MTF)] was calculated from the discrete Fourier transform of the average vertical and horizontal line spread functions of the point spread function from the scan of a small tungsten carbide bead (Figure 1b).

RESULTS

CT numbers

As expected, because of different beam qualities, there were differences in the absolute values of the CT number for specific materials between the CT systems, ranging from 11 HU (polymethylpentene at 120 mGy; range: −174 HU to −185 HU) to 197 HU (Teflon at 12 mGy; range: 903 HU to 1100 HU). Within CT systems, however, the differences in CT number between FBP and IR were in most cases negligible (≤4 HU) irrespective of the radiation dose level, although with some exceptions: for GE Veo, a deviation of 19–20 HU (Delrin) and 31–33 HU (Teflon) was observed at all dose levels, for Siemens IR1–IR3, 6–9 HU deviation for Teflon at 12 mGy and for Toshiba IR1–IR2, 5–7 HU deviation for polymethylpentene at 12 mGy.

Uniformity

The uniformity of CT number between the central and peripheral ROIs was ≤3 HU for all CT systems at all radiation doses, except for the Siemens and Toshiba systems at 12 mGy, where there were deviations of 4.8–5.0 and 3.6–5.2 HU, respectively.

Noise

The noise measurement results are shown in Table 3. Within each CT system, noise was reduced and SNR improved with IR compared with FBP at all radiation dose levels, with only two exceptions: for GE IR1, a larger reduction in signal (HU) than in noise (SD) at all dose levels resulted in lower SNR (0.76–2.33) than with FBP (0.8–2.67). Same mechanism resulted in lower SNR for GE Veo (1.24–1.85) than with FBP (1.59–3.67) at 48 120-mGy dose levels.

For all IR algorithms and dose levels, a higher IR level resulted in reduced noise and improved SNR, except for Toshiba at 120 mGy, where the peak noise reduction was reached with IR2 already. For the statistical IR algorithms, the same IR level resulted in a similar noise reduction (ΔSD%=3–13%), whereas the model-based IR algorithms progressively reduced noise with lower doses: GE Veo reduced noise by 2–33% (ΔSD%=31%), with notable noise reduction (11–33%) seen only at 12–48 mGy, whereas Philips IMR reduced noise by 31–56% (ΔSD%=25%) at all dose levels, with the greatest noise reduction at 12 mGy.

For FBP, the absolute noise levels (SD) were lower at all dose levels for Philips (2.3–7.2 HU) than for the other three systems (3.3–12.9 HU). For IR, the lowest absolute noise levels at all dose levels were produced by GE IR3 (1.4–5.0 HU), Philips IR3 (1.5–4.6 HU) and Philips IMR (1.6–3.1 HU). Compared with FBP from the same vendor, the greatest relative noise reduction (SD%) was achieved with GE IR3 (54–58%), followed by Philips IMR (31–56%) and Siemens IR3 (41–47%).

Noise power spectrum

NPS for all reconstruction algorithms at 84 and 12 mGy are shown in Figure 2. The greatest downward shift of the NPS curves (noise reduction) between FBP and statistical IR algorithms was seen for the GE system. Overall, the shape of the NPS curves (noise distribution) was visually most consistent between the FBP and the statistical IR algorithms for the Philips system. The model-based IR algorithms Philips IMR and GE Veo had a different NPS curve form than the respective FBP and IR algorithms.

Objective low-contrast resolution (contrast-to-noise ratio)

Cumulative CNR results are shown in Figure 3. Within each CT system, the CNR generally improved with increased radiation dose and higher IR level for all statistical IR algorithms, except for GE where the peak CNR was reached at 84 mGy. Of the model-based IR algorithms, Philips IMR had 7–50% higher cumulative CNR than Philips IR3 at all dose levels except 84 mGy, where it was 17% lower, and GE Veo had 24–41% lower

Table 3. Measured mean CT numbers (HU), noise (SD), percentage noise compared with FBP (SD%) and signal-to-noise ratios (SNR) for all combinations of CT systems, reconstruction algorithms and radiation dose levels

System	Radiation doses											
	120 mGy			84 mGy			48 mGy			12 mGy		
	HU/SD	SD%	SNR	HU/SD	SD%	SNR	HU/SD	SD%	SNR	HU/SD	SD%	SNR
GE												
FBP	8.7/3.3	100	2.7	9.0/4.0	100	2.2	8.8/5.6	100	1.6	8.7/10.8	100	0.8
IR1	6.5/2.8	86	2.3	6.9/3.5	87	2.0	7.1/5.0	89	1.4	7.7/10.2	94	0.8
IR2	6.3/2.0	62	3.1	6.7/2.6	63	2.6	6.8/3.6	66	1.9	7.3/7.4	68	1.0
IR3	6.2/1.4	42	4.5	6.5/1.7	43	3.7	6.5/2.6	46	2.5	6.9/5.0	46	1.4
Veo	5.9/3.2	97	1.9	6.2/4.0	98	1.6	6.1/4.9	89	1.2	6.0/7.3	67	0.8
Philips												
FBP	20.4/2.3	100	8.9	20.4/2.7	100	7.6	20.4/4.1	100	5.0	20.9/7.2	100	2.9
IR1	20.5/2.2	97	9.2	20.4/2.5	94	8.1	20.4/3.7	91	5.5	20.8/6.3	88	3.3
IR2	20.5/1.8	79	11.3	20.4/2.1	77	9.8	20.4/3.3	80	6.3	20.8/5.5	77	3.8
IR3	20.5/1.5	66	13.5	20.4/1.7	63	12.0	20.4/2.7	67	7.5	20.8/4.6	64	4.5
IMR	20.0/1.6	69	12.6	20.0/1.6	59	12.7	20.1/2.1	52	9.4	20.5/3.1	44	6.6
Siemens												
FBP	16.9/3.3	100	5.1	16.3/3.8	100	4.3	16.2/5.6	100	2.9	16.0/10.0	100	1.6
IR1	16.9/2.8	85	6.0	16.4/3.3	85	5.0	16.3/4.9	87	3.4	16.0/8.4	83	1.9
IR2	16.9/2.4	71	7.1	16.5/2.7	71	6.0	16.3/4.1	74	4.0	16.2/6.8	67	2.4
IR3	17.0/2.0	59	8.7	16.6/2.2	58	7.5	16.5/3.3	60	4.9	16.3/5.3	53	3.1
Toshiba												
FBP	4.3/3.5	100	1.2	4.2/4.0	100	1.0	4.5/5.2	100	0.9	4.2/12.9	100	0.3
IR1	4.0/2.7	77	1.5	4.0/3.1	77	1.3	4.4/4.0	78	1.1	4.1/8.4	65	0.5
IR2	3.9/2.3	66	1.7	3.9/2.7	66	1.5	4.2/3.2	62	1.3	4.1/7.4	57	0.6
IR3	3.9/2.4	67	1.7	3.8/2.5	61	1.5	4.2/3.0	58	1.4	4.2/7.0	54	0.6

FBP, filtered back-projection; HU, Hounsfield units; IR1–3, iterative reconstruction, levels 1–3; SD, standard deviation; Veo, a model-based IR algorithm. See Table 2 for system names and full manufacturer details.

cumulative CNR than GE IR3 at all dose levels, except for 12 mGy, where it was 25% higher.

A comparison of cumulative CNR between CT systems revealed highest overall CNR for Philips IMR (10.2 at 120 mGy) and Toshiba IR3 (10.1 at 120 mGy). For all dose levels, the highest average CNR at 1.0% nominal contrast was seen for Philips ($CNR_{1.0\%}$=3.6), whereas Toshiba showed the highest nominal contrast at 0.3% ($CNR_{0.3\%}$ = 1.3).

Subjective low-contrast resolution (contrast-to-noise ratio)

As seen in Figure 4, trends toward increased subjective object visibility were seen for both quality criteria with increasing radiation dose and IR level. For the GE system, the number of discernible objects increased from 5–6 at 12 mGy to 19–20 at 48 mGy, with no further increase at higher dose levels. A more constant increase in the number of discernible objects was seen with increasing dose for the other three CT systems. For all CT systems, the number of sharply defined objects increased with higher dose levels. This, however, was slow for the Toshiba system, with a sudden increase in the number of sharply defined objects from 2 to 7 seen between 84 and 120 mGy. The model-based IR algorithm IMR from Philips reproduced 1–4 more discernible objects than the statistical IR at the same dose level, and no improvement was shown for GE Veo.

Evaluation of CT systems and IR algorithms with respect to the number of discernible objects (Figure 4) shows that Philips reproduced more objects than the other systems at 12 mGy (9–14 objects), whereas the same is true for GE at 48 mGy (15–20 objects). At higher dose levels, all systems perform similarly (>15 objects), except for Toshiba at 84 mGy, where 13 discernible objects were reproduced maximally. The mean number of discernible objects for all dose levels and algorithms was greatest for Philips (16.2), followed by GE (15.3),

Figure 2. Noise power spectra (NPS) for all CT systems and reconstruction algorithms at 84 and 12 mGy radiation dose. The curve form represents the distribution of noise (*y*-axis) as a function of spatial frequency (*x*-axis). Higher curves implicate more noise. FBP, filtered back-projection; IMR, iterative model reconstruction; IR1–3, iterative reconstruction, Levels 1–3; Veo, a model-based IR algorithm. See Table 2 for system names and full manufacturer details.

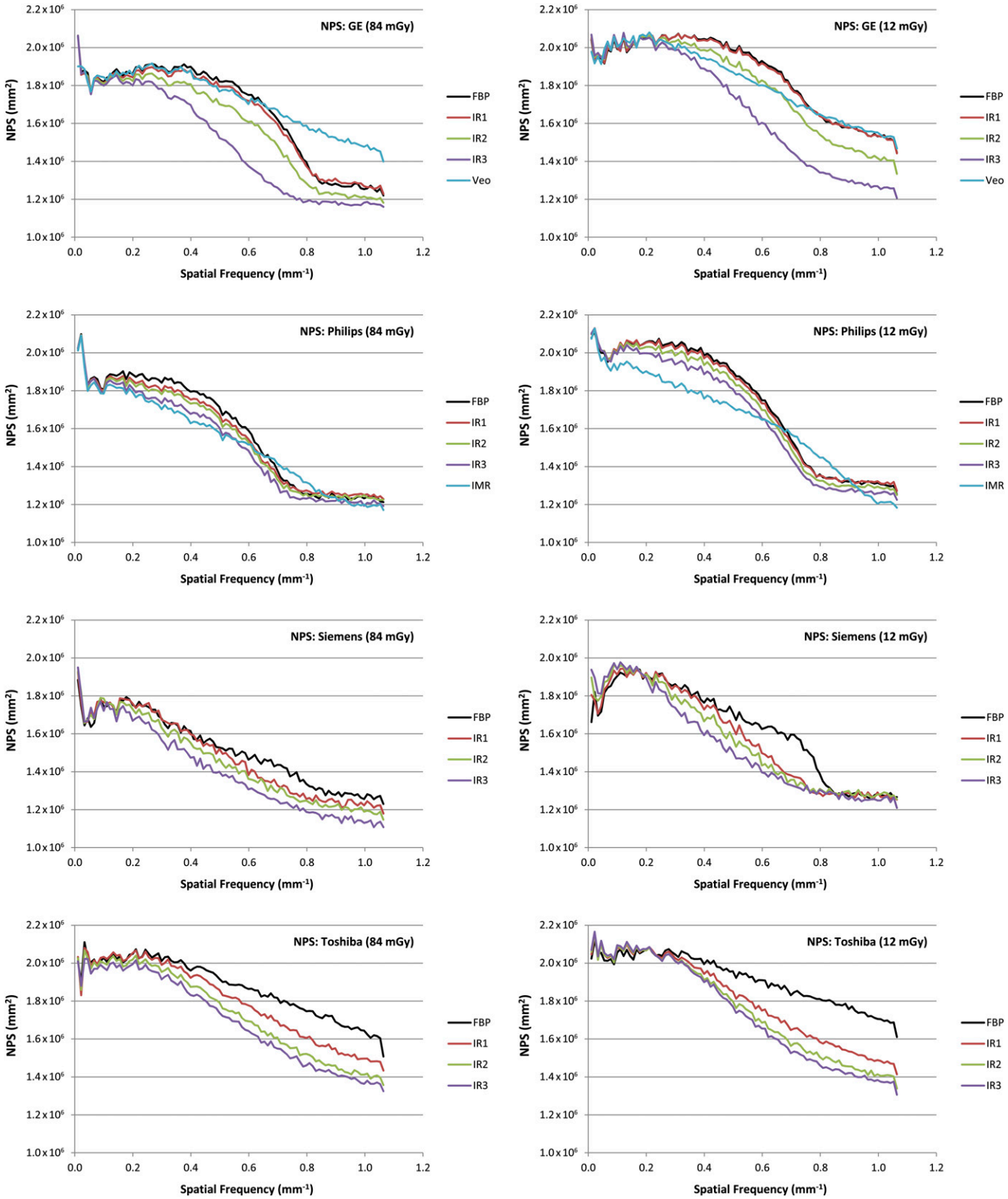
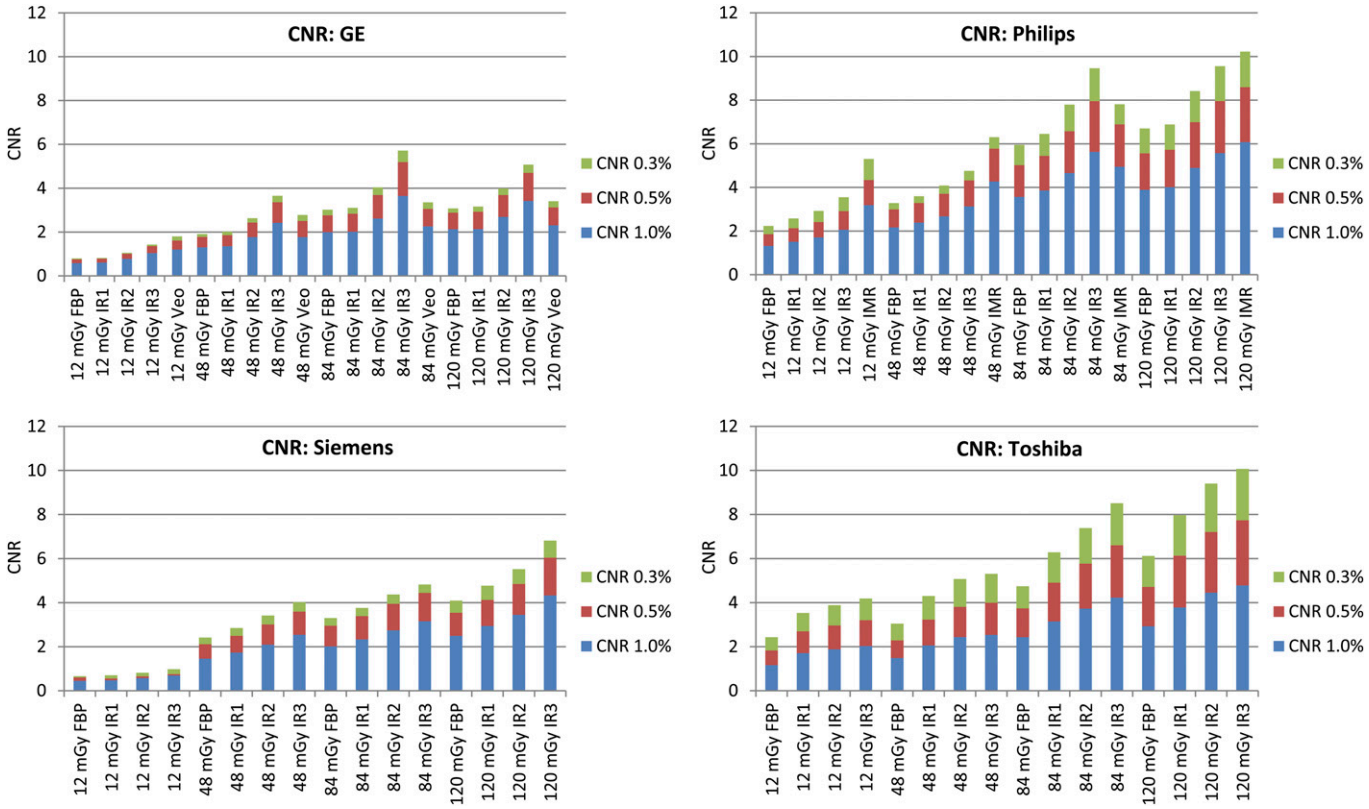


Figure 3. Objective evaluation of low-contrast resolution—cumulative contrast-to-noise ratios (CNRs) from all three nominal contrast levels (0.3%, 0.5% and 1.0%), for all combinations of radiation doses and reconstruction algorithms. FBP, filtered back-projection; IMR, iterative model reconstruction; IR1–3, iterative reconstruction, Levels 1–3; Veo, a model-based IR algorithm. See Table 2 for system names and full manufacturer details.



Siemens (14.2) and Toshiba (12.6). The mean number of sharply defined objects for all dose levels and algorithms was greatest for Philips (5.2), followed by Siemens (4.3), GE (4.1) and Toshiba (2.1).

Figure 5 contains samples from the images used for subjective and objective evaluation of low-contrast resolution.

Spatial resolution

Results for spatial resolution are shown in Table 4. Of the statistical IR algorithms, only GE ASIR improved the spatial resolution compared with FBP, with up to 5–6% improvement of MTF seen at 84 mGy. Spatial resolution was unaffected with Philips iDOSE[4], but slightly reduced with both Siemens SAFIRE and Toshiba AIDR 3D. Both model-based IR algorithms improved the spatial resolution compared with FBP.

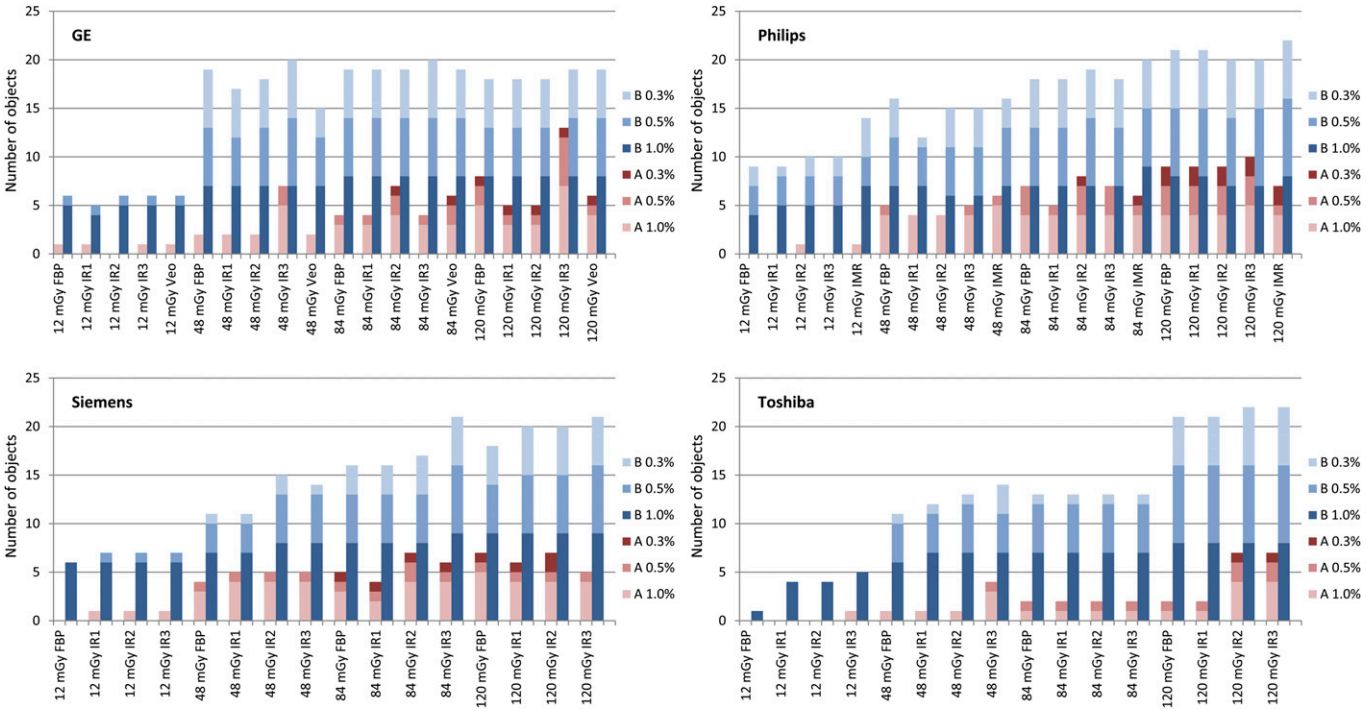
DISCUSSION

Since their introduction, IR algorithms have been extensively studied using both phantoms and clinical data. To our knowledge, this is the first study on IR algorithms from more than two vendors. Furthermore, none of the IR algorithms included in our study has been evaluated before in phantoms using brain CT protocols and a bone-mimicking annulus.

CT image quality is the product of a complicated interplay of many different parameters. Although many of these parameters are system specific and fixed, some are user adjustable. Side-by-side evaluation of the performance of different CT systems is difficult because of the inherent differences in hardware, scanning parameters, reconstruction kernels and absolute CT numbers. There is no way to completely compensate for these differences. Basic standardisation can be achieved using independently recommended scan protocols for clinical CT examinations, such as the AAPM recommendations for adult head CT [12] used in this study. In light of this, it must be stressed that the study results apply to the specific CT systems and examination parameters only and are not necessarily valid in any other context. Furthermore, it should be highlighted that surrogate markers of image quality as measured in this study do not necessarily reflect the diagnostic value of the images.

According to our tests, both reconstruction algorithms and software programs for the automatic evaluation of the image quality used in this study produce identical results when applied repeatedly on the same raw data and images. The only expected variation between repeated scans with identical scan parameters and phantom position is thus caused by the inconsistency in the signal chain. However, these inconsistencies are very small in modern CT systems, and the evaluation of this is beyond the scope of this article. The study results show that measurement inconsistency is not large enough to affect the clear relationship between the radiation dose and the key image quality parameters.

Figure 4. Subjective evaluation of low-contrast resolution—cumulative representation of the number of sharply defined (A) and discernible (B) cylinders (Figure 1c) at all three nominal contrast levels (0.3%, 0.5% and 1.0%), for all combinations of doses and reconstruction algorithms. FBP, filtered back-projection; IMR, iterative model reconstruction; IR1–3, iterative reconstruction, Levels 1–3; Veo, a model-based IR algorithm. See Table 2 for system names and full manufacturer details.



Noise

Noise reduction is a key feature of the IR algorithms and has previously been demonstrated in brain CT for both GE ASIR [4,16,17] and for Siemens first iterative algorithm (IRIS) [18]. Noise is a major determinant of the image quality in brain CT because it negatively affects the low-contrast resolution that is required for adequate reproduction of internal brain anatomy.

Noise in FBP reconstructions is usually quantified by the SD of CT numbers in a ROI, assuming normal distribution. As the SD of noise is not strongly dependent on the absolute CT number values, it could be an ideal parameter for a direct comparison between different CT systems. However, as noise distribution in extensively optimised IR might differ from FBP [19], SD is unreliable for direct comparison of noise levels between FBP and

Figure 5. Combined image showing one set of supraslice cylinders from the low-contrast phantom module (Figure 1c) for all 72 combinations of radiation doses and reconstruction algorithms. Each set consists of 9 cylinders of varying sizes, with 1.0% nominal contrast. Note that the small size of the sample images, in combination with limited background context, would produce different results for subjective visibility assessment—with tendencies towards increased visibility. FBP, filtered back-projection; IR1–3, iterative reconstruction, Levels 1–3; MBIR, model-based iterative reconstruction. See Table 2 for system names and full manufacturer details.

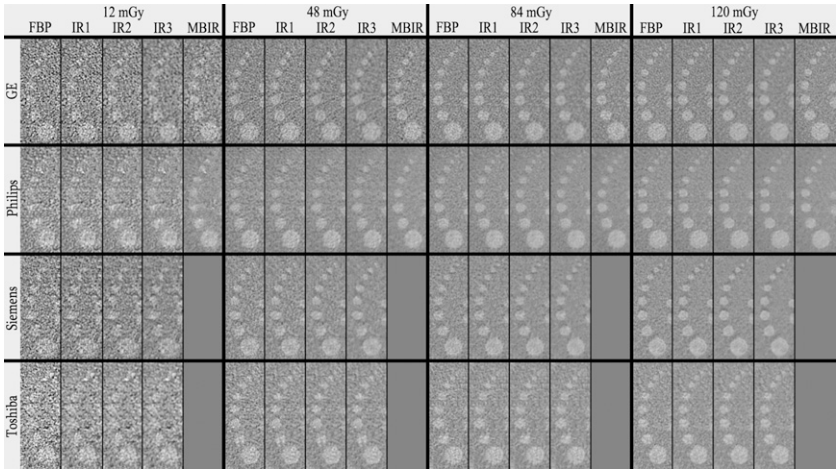


Table 4. Spatial resolution (MTF) for all CT systems and IR algorithms at 84 mGy, measured by calculating the point spread function from the scan of a small tungsten carbide bead

System	84 mGy	
	$MTF_{50\%}$	$MTF_{10\%}$
GE		
FBP	3.7 (100)	6.6 (100)
IR1	3.7 (100)	6.6 (101)
IR2	3.8 (102)	6.8 (104)
IR3	3.9 (105)	7.0 (106)
Veo	4.0 (107)	9.3 (141)
Philips		
FBP	3.6 (100)	5.8 (100)
IR1	3.6 (99)	5.8 (100)
IR2	3.6 (99)	5.8 (100)
IR3	3.5 (99)	5.8 (100)
IMR	4.1 (114)	6.8 (117)
Siemens		
FBP	3.3 (100)	6.9 (100)
IR1	3.2 (97)	6.0 (87)
IR2	3.3 (100)	6.1 (89)
IR3	3.4 (102)	6.2 (89)
Toshiba		
FBP	3.7 (100)	6.3 (100)
IR1	3.5 (96)	6.0 (95)
IR2	3.5 (96)	5.9 (95)
IR3	3.3 (91)	5.7 (91)

IR1–3, iterative reconstruction, Levels 1–3; MTF, modulation transfer function. Results for the other dose levels were equivalent, except for 12 mGy where the low dose level resulted in unreliable MTF-measurements. Values represent frequency (cycles per centimetre) at an MTF level of 50% and 10%, with percent spatial resolution compared with filtered back-projection (FBP) in parenthesis. Higher values indicate improved spatial resolution.
See Table 2 for system names and full manufacturer details.

IR. This might in part explain some of our results, such as the variable noise reduction (SD) seen for different IR algorithms and why the GE Veo algorithm does not seem to reduce noise (SD) at higher dose levels. Furthermore, our results indicate that the model-based IR algorithms do not exhibit a fixed dose/noise relationship like FBP. Altogether these considerations may limit the usability of conventional objective image quality parameters and increase the importance of subjective evaluation steps.

Of the statistical IR algorithms, GE IR3 produced the greatest relative noise reduction (SD%), followed by Philips IR3. However, as the baseline noise for Philips FBP was considerably lower than for GE FBP, the absolute noise levels (SD) for the two IR algorithms were identical. Although the statistical IR algorithms showed relatively constant relative noise reduction (SD%) irrespective of the dose level, the model-based algorithms showed increasing noise reduction with lower dose levels, with the greatest effect seen at 12 mGy. This indicates that the new model-based IR algorithms may allow radiation dose reduction beyond what has been shown for the statistical algorithms. The poor noise reduction of GE Veo at higher dose levels (84–120 mGy) was unexpected and inconsistent with previous studies on paediatric phantoms and body phantoms [11,20]. However, results from these studies are not directly comparable, as the brain CT approach in our study involves much higher radiation doses, resulting in lower baseline noise levels.

Although SNRs were calculated for all reconstructions (Table 3), they were of limited use as they are based on absolute CT values that vary considerably, especially between CT systems and also between GE ASIR and Veo. Furthermore, the use of SD of noise can also be questioned here, as discussed earlier.

Noise power spectrum

To compensate for the limitations of SD and SNR for the assessment of noise, NPS was calculated. NPS provides information about noise spatial characteristics (texture) and is therefore ideal for the detection of changes in noise distribution between different reconstructions [19].

As the IR algorithms modulate noise, the shape of the NPS curves can be expected to differ somewhat between FBP and IR. This is probably one of the factors behind reports of "unnatural" image texture in IR compared with FBP [16,21]. Compared with FBP, the shape of the NPS curves was least affected by Philips iDOSE^4 of all the evaluated algorithms.

Both model-based IR algorithms in the study had considerably different NPS curves than their FBP and IR counterparts, reflecting more extensive noise manipulation and highlighting the inadequateness of comparison based solely on the SD of noise. These results are consistent with a previously published study on GE ASIR/Veo and Philips iDOSE^4 in paediatric body CT [11].

Objective low-contrast resolution (contrast-to-noise ratio)

As expected, given the results from noise measurements, CNR increased proportionally with both radiation dose and IR level for all CT systems. This has previously been shown in brain CT for GE ASIR [16].

Despite having equally low absolute noise levels (SD) for IR3 at all dose levels, as measured in the noise/uniformity phantom module, CNR was much higher for Philips than for GE in the low-contrast module. There are three reasons for this: (1) different mean absolute values for CT number, measuring 47.7 HU (GE) and 59.9 HU (Philips) in the background of the low-contrast module, (2), different mean absolute differences of the CT number between the object and its adjacent background, measuring 8.8 HU (GE) and 10.6 HU (Philips) between the background and 1% nominal contrast cylinders, (3) different mean noise (SD), measuring 6.0 HU (GE) and 3.4 HU (Philips)

in the background of the low-contrast module. Similar mechanisms cause the relatively low CNR for Siemens.

Subjective low-contrast resolution (contrast-to-noise ratio)

As CNR is insensitive to image artefacts/imperfections other than noise, low-contrast resolution was also evaluated subjectively. The subjective evaluation provided a rough estimate of the image quality through evaluation of only two image quality criteria. Unavoidably, subjective evaluation introduces various observer biases [22,23]. Therefore, the two evaluation methods, subjective and objective, complement each other well.

Interestingly, both the GE and Siemens systems performed better in the subjective evaluation of low-contrast resolution than in the objective measurements. In the case of the Siemens system, this could in part be explained by a subtle circular "shadow" artefact seen around many of the low-contrast cylinders (Figure 5), most obvious at higher dose levels, and that might account for increased subjective visibility. No other artefacts were noted for any of the CT systems.

The GE system outperformed all other systems with regard to discernible cylinders at 48 mGy. The same was true for the Toshiba system at 120 mGy, whereas it was inferior to all other systems with respect to sharply defined objects at 84 mGy.

Spatial resolution

One of the major technical challenges in CT imaging is to reduce noise while maintaining good spatial resolution. Poor spatial resolution is a sideeffect of noise suppression and explains the need for a separate reconstruction for evaluation of skeletal structures in head CT. In our study, we assessed how MTF, the ability of a CT system to reproduce spatial information of an object into an image, was affected by radiation dose levels and reconstruction algorithms.

Although all the statistical IR algorithms accomplished considerable noise reduction, only GE ASIR and Philips iDOSE^4 managed to maintain or improve spatial resolution ($MTF_{50\%}$ and $MTF_{10\%}$). This is in agreement with an earlier study on GE ASIR in brain CT [16]. For all IR algorithms, spatial resolution was largely unaffected by the IR level. Both model-based algorithms clearly improve spatial resolution. This is particularly remarkable for Philips IMR, which at the same time greatly reduces noise.

CONCLUSIONS

The evaluated IR algorithms have different strengths and weaknesses. The four statistical IR algorithms, all improved the general image quality compared with conventional FBP reconstruction from the same vendor, with improvements seen for most or all evaluated quality criteria. Additional improvement was achieved with one of the model-based IR algorithms, especially at lower dose levels, indicating further dose reduction potential.

ACKNOWLEDGMENTS

Professor Mats Nilsson, for assistance with NPS evaluation, and Lars Herrnsdorf, for assistance with radiation dose measurements.

FUNDING

Supported by Skåne University Hospital Lund and Region Skåne.

REFERENCES

1. Hara AK, Paden RG, Silva AC, Kujak JL, Lawder HJ, Pavlicek W. Iterative reconstruction technique for reducing body radiation dose at CT: feasibility study. AJR Am J Roentgenol 2009;193:764–71. doi: 10.2214/AJR.09.2397
2. Kalra MK, Woisetschlager M, Dahlstrom N, Singh S, Lindblom M, Choy G, et al. Radiation dose reduction with Sinogram Affirmed Iterative Reconstruction technique for abdominal computed tomography. J Comput Assist Tomogr 2012;36:339–46. doi: 10.1097/RCT.0b013e31825586c0
3. Pontana F, Duhamel A, Pagniez J, Flohr T, Faivre JB, Hachulla AL, et al. Chest computed tomography using iterative reconstruction vs filtered back projection (part 2): image quality of low-dose CT examinations in 80 patients. Eur Radiol 2011;21:636–43.
4. Ren Q, Dewan SK, Li M, Li J, Mao D, Wang Z, et al. Comparison of adaptive statistical iterative and filtered back projection reconstruction techniques in brain CT. Eur J Radiol 2012;81:2597–601. doi: 10.1016/j.ejrad.2011.12.041
5. Yamada Y, Jinzaki M, Hosokawa T, Tanami Y, Sugiura H, Abe T, et al. Dose reduction in chest CT: comparison of the adaptive iterative dose reduction 3D, adaptive iterative dose reduction, and filtered back projection reconstruction techniques. Eur J Radiol 2012;81:4185–95. doi: 10.1016/j.ejrad.2012.07.013
6. Hou Y, Liu X, Xv S, Guo W, Guo Q. Comparisons of image quality and radiation dose between iterative reconstruction and filtered back projection reconstruction algorithms in 256-MDCT coronary angiography. Am J Roentgenol 2012;199:588–94. doi: 10.2214/AJR.11.7557
7. Katsura M, Sato J, Akahane M, Matsuda I, Ishida M, Yasaka K, et al. Comparison of pure and hybrid iterative reconstruction techniques with conventional filtered back projection: Image quality assessment in the cervicothoracic region. Eur J Radiol 2013;82:356–60. doi: 10.1016/j.ejrad.2012.11.004
8. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. Phys Med 2012;28:94–108. doi: 10.1016/j.ejmp.2012.01.003
9. Willemink MJ, de Jong PA, Leiner T, de Heer LM, Nievelstein RA, Budde RP, et al. Iterative reconstruction techniques for computed tomography part 1: technical principles. Eur Radiol 2013;23:1623–31.
10. Desai GS, Thabet A, Elias AY, Sahani DV. Comparative assessment of three image reconstruction techniques for image quality and radiation dose in patients undergoing abdominopelvic multidetector CT examinations. Br J Radiol 2013;86:20120161.
11. Mieville FA, Gudinchet F, Brunelle F, Bochud FO, Verdun FR. Iterative reconstruction methods in two different MDCT scanners: physical metrics and 4-alternative forced-choice detectability experiments—a phantom

approach. Phys Med 2013;29:99–110. doi: 10.1016/j.ejmp.2011.12.004
12. The American Association of Physicists in Medicine (AAPM). Adult routine head CT protocols. Version 1.1. 2012 [updated 1 June, 2012; cited 5 January, 2013] Available from http://aapm.org/pubs/CTProtocols/documents/AdultRoutineHeadCT.pdf
13. Menzel HG, Schilbilla B, Teunen D (eds). European guidelines on quality criteria for computed tomography (EUR 16262). Luxembourg: European Commission; 1999.
14. Dobbins JT 3rd. Image quality metrics for digital systems. In: Handbook of medical imaging. Vol. I. Physics and psychophysics. Beutel J, Kundel HL, Van Metter RL (eds). Bellingham, WA: SPIE Press; 2000.
15. Borjesson S, Hakansson M, Bath M, Kheddache S, Svensson S, Tingberg A, et al. A software tool for increased efficiency in observer performance studies in radiology. Radiat Prot Dosimetry 2005;114:45–52. doi: 10.1093/rpd/nch550
16. Rapalino O, Kamalian S, Payabvash S, Souza LC, Zhang D, Mukta J, et al. Cranial CT with adaptive statistical iterative reconstruction: improved image quality with concomitant radiation dose reduction. Am J Neuroradiol 2012;33:609–15. doi: 10.3174/ajnr.A2826
17. Kilic K, Erbas G, Guryildirim M, Arac M, Ilgit E, Coskun B. Lowering the dose in head CT using adaptive statistical iterative reconstruction. Am J Neuroradiol 2011;32:1578–82. doi: 10.3174/ajnr.A2585
18. Korn A, Fenchel M, Bender B, Danz S, Hauser TK, Ketelsen D, et al. Iterative reconstruction in head CT: image quality of routine and low-dose protocols in comparison with standard filtered back-projection. Am J Neuroradiol 2012;33:218–24. doi: 10.3174/ajnr.A2749
19. Hsieh J. Computed tomography: principles, design, artifacts and recent advances. 2nd edn. Bellingham, WA: SPIE Press; 2009.
20. Vardhanabhuti V, Loader R, Roobottom CA. Assessment of image quality on effects of varying tube voltage and automatic tube current modulation with hybrid and pure iterative reconstruction techniques in abdominal/pelvic CT: a phantom study. Invest Radiol 2013;48:167–74. doi: 10.1097/RLI.0b013e31827b8f61
21. Niu YT, Mehta D, Zhang ZR, Zhang YX, Liu YF, Kang TL, et al. Radiation dose reduction in temporal bone CT with iterative reconstruction technique. Am J Neuroradiol 2012;33:1020–6. doi: 10.3174/ajnr.A2941
22. Sackett DL. Bias in analytic research. J Chronic Dis 1979;32:51–63.
23. Sica GT. Bias in research studies. Radiology 2006;238:780–9. doi: 10.1148/radiol.2383041109