PHILIPS

Computed Tomography

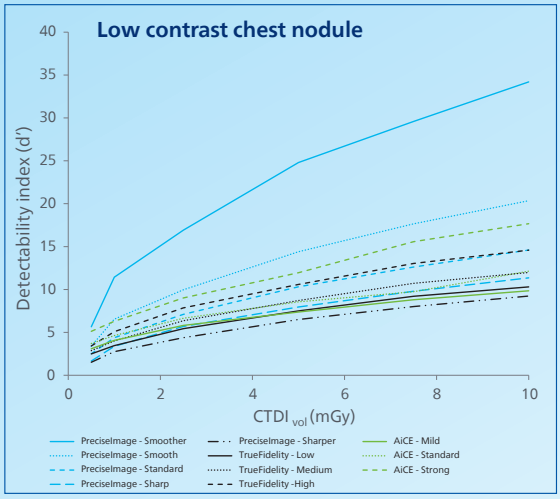
White paper

Precise Image: A multi-vendor task-based image quality comparison

Niels van der Werf, PhD - CT Clinical Scientist Europe - Philips Healthcare

Overview

This white paper objectively quantifies the image quality of deep-learning reconstruction (DLR) of Philips Precise Image and compares these AI reconstruction results with the DLR image quality offered by other CT manufacturers. Precise Image significantly improved image quality in comparison to filtered back projection (FBP) and iterative reconstruction. In addition, Precise Image outperformed the AI reconstruction of competitors in terms of detectability indices for low contrast chest nodules.



Introduction

Recently, a deep-learning reconstruction (DLR) technique called Precise Image was introduced to overcome the limitations of traditional CT reconstruction methods. While conventional filtered back projection (FBP) often results in high-noise images and iterative reconstruction (IR) techniques reduce noise at the expense of image smoothness with unfamiliar textures, Precise Image offers the best of both worlds.[1] It delivers low-noise images with a familiar FBP-like appearance, without the drawbacks of FBP or IR.

When it comes to evaluating image quality across various reconstruction techniques, traditional metrics such as image noise, signal-to-noise ratio (SNR), and contrast-to-noise ratio (CNR) fall short.[2] Therefore, other more advanced objective image-quality metrics have been previously proposed to better describe the potential image quality in clinical practice. One of these metrics is the detectability index (d'), which assesses the radiologist's ability to accurately detect lesions by considering factors such as noise, resolution and lesion-specific characteristics. For this, the noise-power-spectrum (NPS) and a task-specific transfer function (TTF) are considered. The NPS analysis helps quantify not just the amount of noise but also its texture and granularity, while the TTF evaluates image resolution for specific clinical tasks. Together, these metrics ensure that Precise Image meets the highest standards of clinical excellence, enhancing diagnostic confidence and improving patient outcomes.

Phantom setup

To objectively assess the image quality of Precise Image, an ACR 464 (CT ACR 464 Phantom, Gammex Technology) was used, in line with the methodology from two recent papers of Greffier, et al (**Figure 1**).[3,4] With this phantom, advanced quantitative image quality properties for noise and resolution were evaluated. Noise properties were measured using the NPS in 20 consecutive slices of a uniform region of the phantom. Resolution properties were estimated with a TTF in ten consecutive slices for various task-specific contrasts, i.e., air (-1000 Hounsfield units (HU)) and polyethylene (-95 HU).

Raw data was acquired on a Philips Incisive CT system with software version 5.01 (Philips, Best) using a standard helical chest CT protocol (collimation 64 x 0.625 mm; rotation time 0.35 s; pitch 1) with a tube potential of 120 kVp.[4] Six dose levels were used: 10, 7.5, 5, 2.5, 1, and 0.5 mGy. For this, fixed tube current (mA) values were used. Raw data was reconstructed on a 512 x 512 element matrix with a 250 mm field-of-view using all three available reconstruction methods: FBP, IR (iDose[4], Philips), and DLR (Precise Image, Philips). For both FBP and IR, the B and YA filter were used. For IR specifically, both levels 4 and 7 were used. For the Philips DLR--Precise Image--raw data was reconstructed with all available levels (Smoother, Smooth, Standard, Sharp, Sharper), with both soft-tissue and lung settings. In accordance with Greffier, et al., a slice thickness of 1 mm with an overlapping slice increment of 0.5 mm was used for all reconstructions.[4]

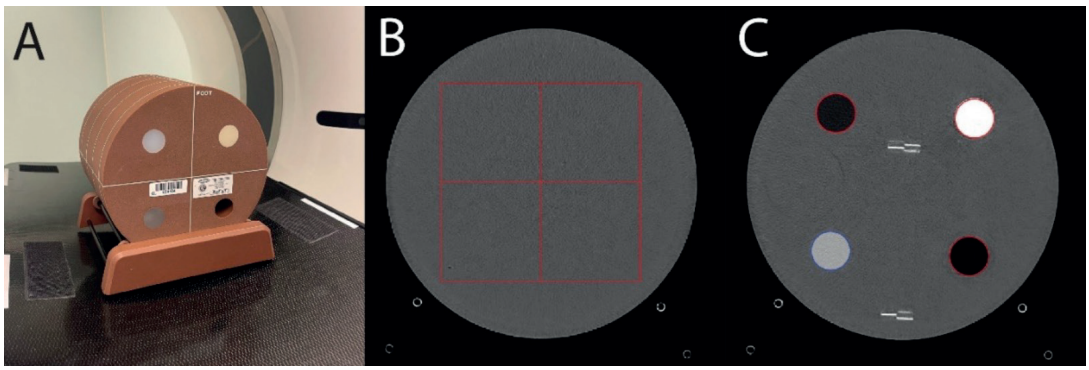


Figure 1: Overview of the phantom setup (A), noise-power-spectrum slice (B), and task-transfer function slice (C).

Task-based image-quality assessment

Three tasks, identical to the tasks defined by Greffier, et al., were defined to objectively assess image quality using the iQMetrix-CT software.[3,4] These tasks were defined as the detection of chest lesions of 5 mm in diameter for a low-contrast soft tissue chest nodule, a ground-glass opacity and a high-contrast pulmonary lesion with a contrast of 50, 200, and 950 HU, respectively. For each detection task, a non-pre-whitening observer model with eye filter was used to calculate the detectability index.[5–7] Polyethylene TTF results were used for the two lowest contrast tasks, while the air insert TTF result was used for the high-contrast task. All tasks were defined on a 512 x 512 elements matrix with a pixel size of 0.488 x 0.488 mm^2. Interpretation conditions were a zoom factor of 1.5, a 500 mm viewing distance and a 500 mm field of view. For each task, Precise Image results were compared with the results from Greffier, et al., for DLR from Canon (AiCE) and GE (TrueFidelity™) as extracted from this publication.[3]

For a full interpretation of the results, it is important to understand the relationship between the detectability index (d'), and the ROC curve (receiver operating characteristic) that underlies it.[8] The ROC curve tracks the probability of true positive detections of the object or lesion under study, against false positive detections (cases when the observer decides that the object is present, when in truth it is not present), as a decision threshold is varied over a range of possible values. A d' number converts this 2-D ROC curve into a 1-D scalar value, assuming Gaussian distributions of the decision variable. A d'=3.5 corresponds to an AUC (area under the ROC curve) = 0.9932, meaning that the observer correctly detects the object 99.32% of the time. The worst AUC value, 0.5, means that the observer does no better than random guessing, and this corresponds to a d' = 0. Thus, any d' values above ~3.5 do not truly reflect a significant increase in detectability (since object detection is already almost perfect at d' ~ 3.5). d' values greater than 3.5 most often reflect a noise reduction, which may offer benefits in terms of shape discrimination or diagnostic confidence but are not in fact better in terms of strict detectability.

Image quality results

Image noise

As expected, noise magnitude (standard deviation in HU) decreased with increasing radiation dose, increased IR levels, and smoother Precise Image reconstruction levels (**Figure 2** and **Figure 3**). In comparison to FBP and IR, noise magnitude was reduced up to 81% and 73% for Precise Image, respectively (**Table 1** and **Table 2**). For all Precise Image levels, increasing the radiation dose from 0.5 mGy to 10 mGy resulted in an average decrease of noise magnitude of approximately 3.7 times.

The average spatial frequency of the noise increased for sharper Precise Image reconstruction level (**Figure 4** and **Figure 5**). For each level, the average spatial frequency was constant (< 3% deviation) for all radiation doses for the Standard, Sharp, and Sharper level (**Table 1** and **Table 2**). For the Smooth and Smoother level, a maximum deviation of 12% was found. Average NPS spatial frequencies were higher for all Precise Image levels with the Lung setting, in comparison to the Soft Tissue setting.

Figure 2: Precise Image noise magnitude for all reconstruction levels and radiation dose levels for the two lowest contrast tasks, based on the Soft Tissue setting.

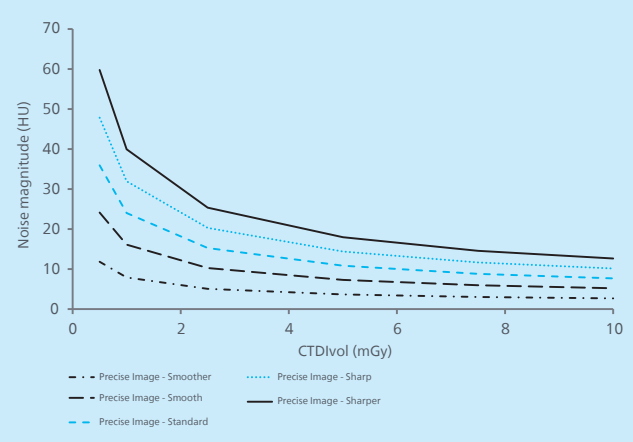


Figure 3: Precise Image noise magnitude for all reconstruction levels and radiation dose levels for the high contrast task, based on the Lung setting.

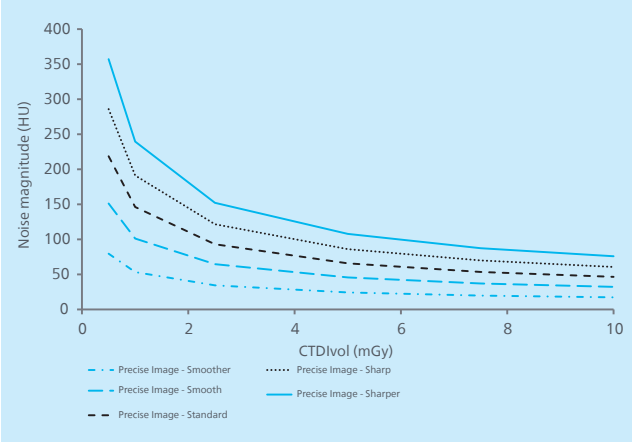


Figure 4: Precise Image average spatial frequency (fav,) of the noise for all reconstruction levels and radiation dose levels for the two lowest contrast tasks, based on the Soft Tissue setting.

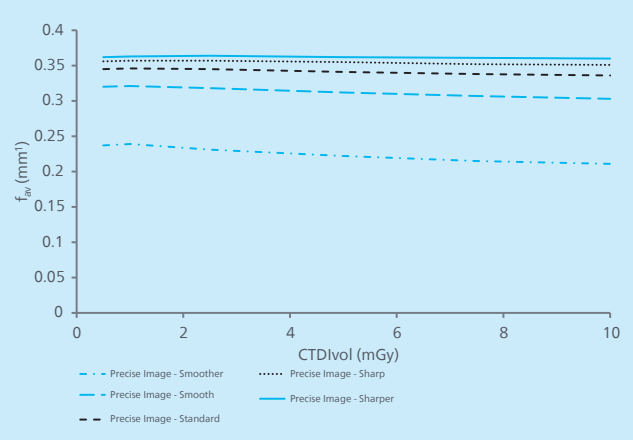
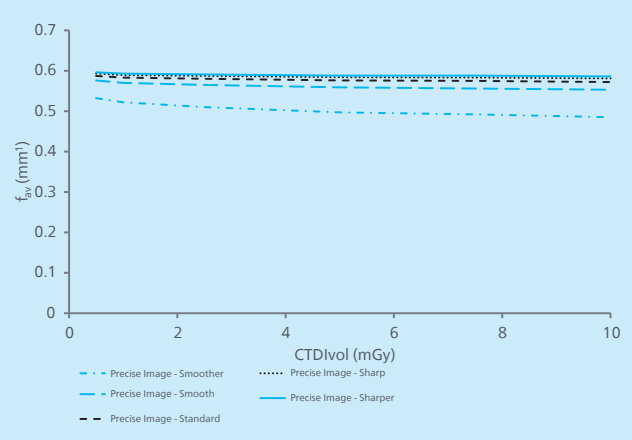


Figure 5: Precise Image average spatial frequency (fav,) of the noise for all reconstruction levels and radiation dose levels for the high contrast task, based on the Lung setting.



	$CTDI_{vol}$ (mGy)	FBP	IR level		Precise Image				
			4	7	Smoother	Smooth	Standard	Sharp	Sharper
Noise magnitude (HU)	0.5	61.8	44.5	30.1	11.8	24.1	35.9	47.8	59.8
	1	41.3	29.7	20.1	7.9	16.1	24.0	32.0	39.9
	2.5	26.1	18.8	12.7	5.1	10.3	15.3	20.3	25.4
	5	18.4	13.3	9.0	3.7	7.3	10.8	14.4	18.0
	7.5	14.9	10.8	7.3	3.0	6.0	8.8	11.7	14.6
	10	12.9	9.3	6.4	2.7	5.2	7.7	10.2	12.7
Average NPS spatial frequency (mm^{-1})	0.5	0.318	0.292	0.248	0.237	0.320	0.345	0.356	0.362
	1	0.321	0.294	0.250	0.239	0.321	0.346	0.357	0.363
	2.5	0.322	0.295	0.250	0.231	0.318	0.345	0.357	0.364
	5	0.321	0.293	0.247	0.222	0.312	0.341	0.355	0.362
	7.5	0.320	0.292	0.245	0.215	0.307	0.338	0.352	0.361
	10	0.320	0.290	0.242	0.211	0.303	0.336	0.351	0.360

Table 1: Noise magnitude (Hounsfield Units (HU)) and average noise power spectrum (NPS) spatial frequencies for all radiation dose levels and reconstructions for the two lowest contrast tasks (B filter for FBP and IR, Soft tissue setting for Precise Image).

	$CTDI_{vol}$ (mGy)	FBP	IR level		Precise Image				
			4	7	Smoother	Smooth	Standard	Sharp	Sharper
Noise magnitude (HU)	0.5	257.9	183.3	117.6	79.5	151.1	218.5	285.8	357.0
	1	172.6	122.2	77.7	53.3	101.1	146.1	191.1	239.5
	2.5	109.5	77.5	49.3	34.3	64.5	93.0	121.5	152.2
	5	77.5	54.8	34.8	24.2	45.6	65.8	86.0	107.7
	7.5	62.9	44.5	28.3	19.7	37.0	53.4	69.8	87.4
	10	54.6	38.6	24.5	17.2	32.2	46.4	60.6	75.9
Average NPS spatial frequency (mm^{-1})	0.5	0.523	0.519	0.515	0.532	0.576	0.587	0.593	0.596
	1	0.522	0.518	0.513	0.522	0.570	0.583	0.589	0.593
	2.5	0.522	0.518	0.513	0.510	0.565	0.580	0.587	0.591
	5	0.520	0.516	0.512	0.497	0.559	0.576	0.584	0.588
	7.5	0.519	0.515	0.511	0.492	0.556	0.575	0.583	0.588
	10	0.519	0.515	0.510	0.485	0.553	0.572	0.581	0.586

Table 2: Noise magnitude (Hounsfield Units (HU)) and average noise power spectrum (NPS) spatial frequencies for all radiation dose levels and reconstructions for the high contrast task (YA filter for FBP and IR, Lung setting for Precise Image).

Image resolution: task transfer function

In comparison to FBP and IR, image resolution was increased for all Precise Image levels with increased $TTF_{50\%}$ values for the air insert (**Table 3**). For the polyethylene insert, equal results were shown for all dose levels above 0.5 mGy.

	$CTDI_{vol}$ (mGy)	FBP	IR level		Precise Image				
			4	7	Smoother	Smooth	Standard	Sharp	Sharper
$TTF_{50\%}$ - Air (mm^{-1})	0.5	0.499	0.496	0.492	0.600	0.606	0.643	0.768	0.780
	1	0.586	0.483	0.496	0.715	0.729	0.603	0.595	0.692
	2.5	0.501	0.514	0.527	0.680	0.696	0.688	0.673	0.659
	5	0.557	0.578	0.602	0.687	0.773	0.735	0.802	0.789
	7.5	0.574	0.500	0.511	0.728	0.719	0.723	0.717	0.717
	10	0.584	0.597	0.536	0.701	0.704	0.695	0.685	0.742
$TTF_{50\%}$ - Polyethylene (mm^{-1})	0.5	0.254	0.238	0.218	0.209	0.212	0.228	0.173	0.212
	1	0.314	0.283	0.257	0.329	0.332	0.351	0.354	0.357
	2.5	0.306	0.268	0.294	0.323	0.327	0.327	0.369	0.356
	5	0.319	0.300	0.280	0.381	0.369	0.376	0.369	0.373
	7.5	0.317	0.304	0.297	0.363	0.368	0.365	0.362	0.365
	10	0.327	0.319	0.287	0.380	0.380	0.375	0.375	0.373

Table 3: Task-based transfer function at 50% (TTF50%) for the air and polyethylene inserts for FBP, iterative reconstruction (IR) and all Precise Image levels for all radiation dose levels. For the air insert, the results are based on the YA-filter for FBP and IR, and on the Lung setting for Precise Image. For the polyethylene insert, the results are based on the B-filter for FBP and IR, and on the Soft Tissue setting for Precise Image.

Detectability index

For Precise Image, detectability index for all simulated lesions increased with increased smoothing and radiation dose (**Figure 6, Figure 7** and **Figure 8**). For the two lowest contrast tasks, the detectability index based on the Precise Image reconstructions was higher than the conventional FBP and IR reconstructions (**Figure 6** and **Figure 7**).
The Smoother level resulted in the highest detectability index, independent of radiation dose. For the high contrast task, with inherent high detection rate, the detectability index did not improve based on the Precise Image reconstruction (**Figure 8**).

Figure 6: Detectability index (d') for filtered-back-projection (FBP), iterative reconstruction (iDose levels 4 and 7), and all available Precise Image levels for the detection of the low contrast soft tissue chest nodule (50 Hounsfield Units (HU) contrast) based on the noise-power-spectrum and polyethylene insert task-transfer-function, both with the Soft Tissue setting.

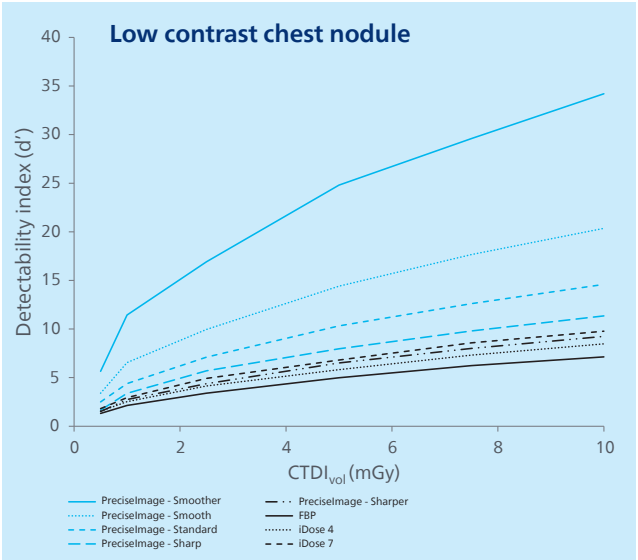


Figure 7: Detectability index (d') for filtered-back-projection (FBP), iterative reconstruction (iDose levels 4 and 7), and all available Precise Image levels for the detection of the ground glass opacity (200 Hounsfield Units (HU) contrast) based on the noise-power-spectrum and polyethylene insert task-transfer-function, both with the Soft Tissue setting.

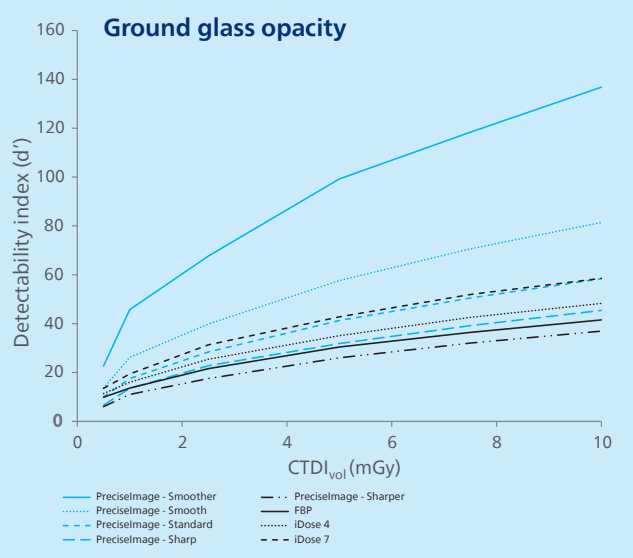
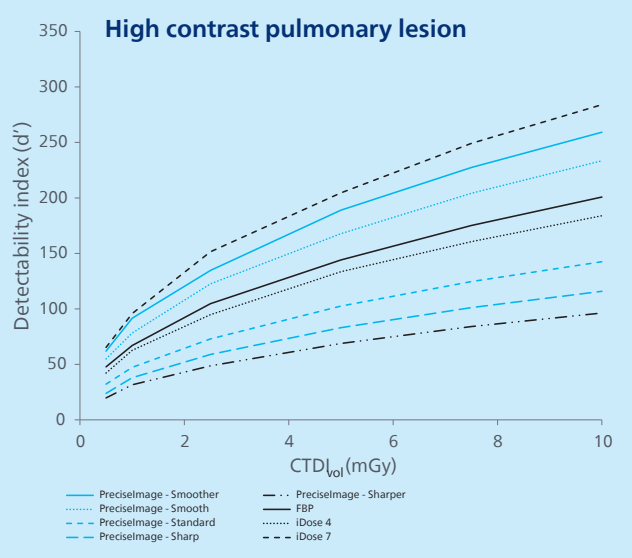


Figure 8: Detectability index (d') for filtered-back-projection (FBP), iterative reconstruction (iDose levels 4 and 7), and all available Precise Image levels for the detection of the high contrast pulmonary lesion (50 Hounsfield Units (HU) contrast) based on the noise-power-spectrum and air insert task-transfer-function, both with the Lung setting.



Multi-vendor comparison

In comparison to the previously published analysis on DLR from two other vendors, noise levels of Precise Image with the Soft Tissue setting, and the Smoother, Smooth, and Standard level were comparable.[3] The Sharp and Sharper reconstruction setting, and all levels based on the Lung setting provided increased image noise. For the Soft Tissue setting, average NPS spatial frequencies were also comparable to the previously published DLR analysis. For the Lung setting, increased average NPS spatial frequencies were shown for Precise Image.

For the polyethylene insert, image resolution (based on the TTF at 50%) for Precise Image was comparable to the previously reported values for the other vendors.[3] For the air insert, Precise Image showed increased image resolution, especially for the sharper reconstruction levels.

Finally, independent of radiation dose, Precise Image (level Smoother) d' values were higher than both other DLR reconstruction techniques (**Figure 9** and **Figure 10**). For dose levels of 1 mGy and beyond, this was true for both the Smoother and Smooth reconstruction level. For the high contrast pulmonary lesion with inherent high image contrast, the highest detectability index was reached for AiCE, irrespective of the radiation dose (**Figure 11**).

Figure 9: Detectability index (d') for all available deep learning reconstruction algorithms for the detection of the low contrast soft tissue chest nodule (50 Hounsfield Units (HU) contrast).

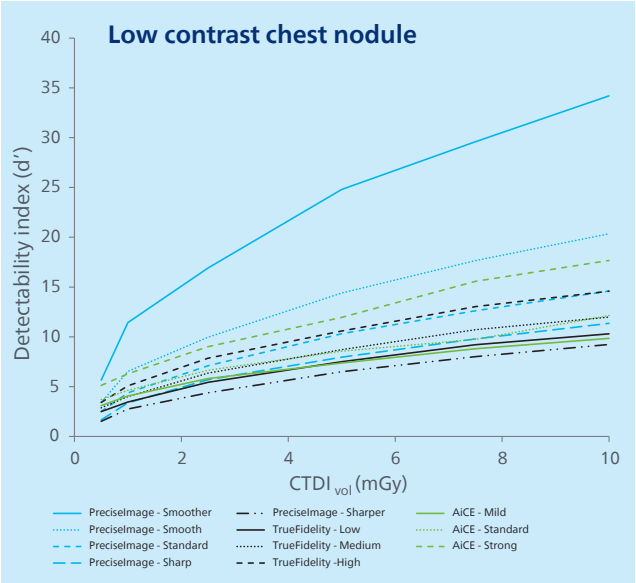


Figure 10: Detectability index (d') for all available deep learning reconstruction algorithms for the detection of the ground glass opacity (200 HU contrast).

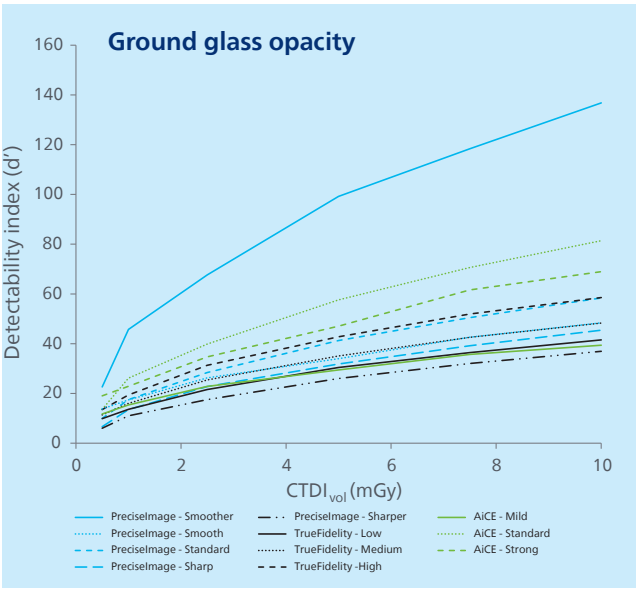
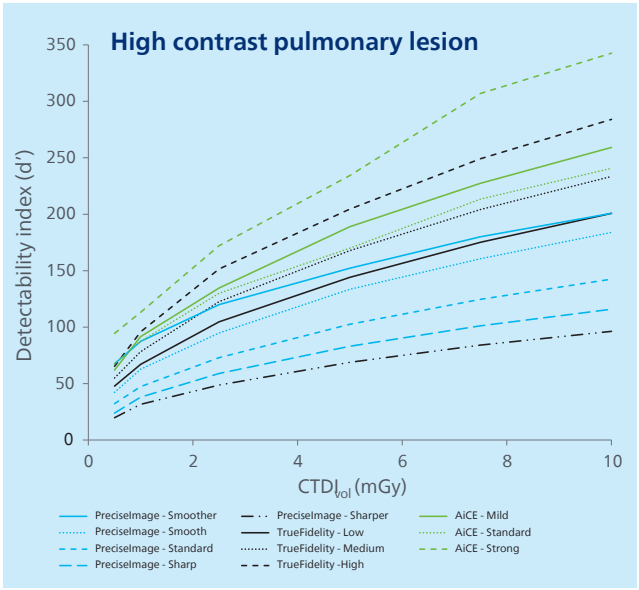


Figure 11: Detectability index (d') for all available deep learning reconstruction algorithms for the detection of the high contrast pulmonary lesion (950 HU contrast).



Conclusion

Precise Image significantly improved detectability indices for the three simulated lesions studied, in comparison to conventional FBP and IR reconstructions. Also, Precise Image showed superior detectability indices compared to competition for low-contrast chest nodules and ground glass opacities. For the high-contrast pulmonary lesions, image quality was comparable among all vendors, with inherent high detectability index overall.

1. Willemink MJ, Noël PB. The evolution of image reconstruction for CT-from filtered back projection to artificial intelligence. Eur Radiol. 2019;29:2185–2195.
2. Smith TB, Solomon J, Samei E. Estimating detectability index in vivo: development and validation of an automated methodology. Journal of Medical Imaging. 2017;5:1.
3. Greffier J, Frandon J, Si-Mohamed S, et al. Comparison of two deep learning image reconstruction algorithms in chest CT images: A task-based image quality assessment on phantom data. Diagn Interv Imaging. 2022;103:21–30.
4. Greffier J, Si-Mohamed S, Frandon J, et al. Impact of an artificial intelligence deep-learning reconstruction algorithm for CT on image quality and potential dose reduction: A phantom study. Med Phys. 2022;49:5052–5063.
5. Greffier J, Boccalini S, Beregi JP, et al. CT dose optimization for the detection of pulmonary arteriovenous malformation (PAVM): A phantom study. Diagn Interv Imaging. 2020;101:289–297.
6. Greffier J, Hamard A, Pereira F, et al. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. Eur Radiol. 2020;30:3951–3959.
7. Solomon J, Samei E. Correlation between human detection accuracy and observer model-based image quality metrics in computed tomography. J Med Imag. 2016;3:35506.
8. Barrett HH, Myers KJ. Foundations of Image Science. Hoboken, NJ: John Wiley & Sons, 2004

4522 991 88461 * NOV 2024

How to reach us
Please visit www.philips.com