

# Image quality assessment of deep learning image reconstruction in computed tomography using tube current modulation: a phantom study

kazuhiro takeuchi (✉ [takeuchi.kazuhiro@kagawa-u.ac.jp](mailto:takeuchi.kazuhiro@kagawa-u.ac.jp))

Kagawa University Hospital: Kagawa Daigaku Igakubu Fuzoku Byoin <https://orcid.org/0000-0002-0383-2799>

**Yasuhiro Ide**

Kagawa University Hospital: Kagawa Daigaku Igakubu Fuzoku Byoin

**Yuichiro Mori**

Kagawa University Hospital: Kagawa Daigaku Igakubu Fuzoku Byoin

**Yusuke Uehara**

Kagawa University Hospital: Kagawa Daigaku Igakubu Fuzoku Byoin

**Hiroshi Sukeishi**

Kagawa University Hospital: Kagawa Daigaku Igakubu Fuzoku Byoin

**Sachiko Goto**

Okayama University Faculty of Health Sciences Graduate School of Health Sciences: Okayama Daigaku Igakubu Hoken Gakka Daigakuin Hokengaku Kenkyuka

---

## Research Article

**Keywords:** computed tomography, deep learning, image reconstruction, tube current modulation, phantom size.

**Posted Date:** June 25th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-591163/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

The novel deep learning image reconstruction (DLIR) is known to change its image quality characteristics according to object contrast and image noise. In clinical practice, computed tomography (CT) image noise is usually controlled by tube current modulation (TCM) to accommodate changes in object size. This study aimed to evaluate the image quality characteristics of DLIR for different object sizes when in-plane noise is controlled by TCM. We used Mercury 4.0 phantoms with different object sizes. Phantom image acquisition was performed on a GE Revolution CT system to investigate the impact of the DLIR algorithm compared to standard reconstructions: filtered back projection (FBP) and hybrid iterative reconstruction (hybrid-IR). For image quality evaluation, the noise power spectrum (NPS), task-based transfer function (TTF), and detectability index ( $d'$ ) were determined. The NPS of DLIR was very similar to that of FBP, and the information in the high-frequency region was maintained. In terms of TTF, DLIR showed higher resolution than hybrid-IR at low- to medium-contrast ( $\Delta 50$ ,  $\Delta 90$ HU), but not necessarily higher than FBP. At the simulated contrast and lesion size, DLIR showed higher detectability than hybrid-IR, regardless of the phantom size. In this study, we evaluated a novel DLIR algorithm by reproducing clinical behaviors. The findings indicate that DLIR produces higher image quality than hybrid-IR regardless of the phantom size, although it depends on the reconstruction strength.

## Introduction

Computed tomography (CT) examinations have been introduced in several medical institutions because of their high diagnostic imaging ability, high throughput, and few restrictions. Due to the increase in the number and the high diversity of examinations, the proportion of CT in medical exposure is also increasing, and optimization of CT exposure dose is an essential condition [1, 2]. Filtered back projection (FBP), which is the basic image reconstruction algorithm of CT, has a linear relationship between its dose and image quality, although the relationship between image quality and CT dose is a trade-off, making it difficult to reduce image noise. Therefore, hybrid (hybrid-IR) and model-based iterative reconstruction (MBIR) have been developed to maintain image quality and reduce image noise. Several studies on CT using iterative reconstruction (IR) have been reported and used in clinical practice [3, 4].

However, several problems have been reported with the IR algorithm, and it has been recognized that the noise level of the image and the reconstruction strengths affect the texture when analyzing image noise reduction [5]. Thus, it has been reported that it is difficult to reduce the dose compared to FBP [6, 7]. In contrast, deep learning image reconstruction (DLIR), which was newly developed by the deep neural network (DNN), has been reported to reduce noise and dose without affecting the texture in the image compared to IR [8–11, 12]. The novel DLIR is a nonlinear image reconstruction algorithm that depends on image noise and contrast and is affected by the subject size in clinical practice. In clinical applications, tube current modulation (TCM) is applied to optimize the incident dose according to the subject size [13]. TCM is a radiation exposure reduction technique that optimizes the incident dose for each patient by simulating the body thickness size from scout images and regulating the X-ray dose by modulating the tube current to fit the irregular patient size.

Solomon et al. [14] used the Mercury Phantom V3.0 to evaluate the image quality of DLIR and evaluated different object sizes using a fixed tube current (FTC), which was necessary to maintain measurement accuracy. However, the Mercury Phantom is useful only for clinical evaluation of TCM [15]. Thus, in this study, we assumed a clinical CT of the torso, and the task of image quality assessment targeted two low- to medium-contrast inserts of polystyrene ( $\approx \Delta 50$ HU) and water ( $\approx \Delta 90$ HU) in the Mercury Phantom. For advanced reconstruction methods such as DLIR, high-contrast objects may be easily distinguished from noise and edges based on previous physical and clinical evaluations [8, 11].

Currently, there have been few reports on the physical evaluation of DLIR related to the measurement of low contrast of 50HU. Furthermore, image quality characteristics using IR at low contrast show different characteristics among the apparatus [16]. A 50HU contrast in the abdominal region is important for clinical evaluation in CT to assess faint shadows such as liver tumors and intravenous thrombus [17, 18, 19]. Thus, this study aimed to evaluate the image quality characteristics of different subject sizes reconstructed by DLIR using images acquired under controlled in-plane noise using TCM.

## Methods

## Deep learning image reconstruction

Developed by GE, DLIR [true fidelity imaging (TFI); GE Healthcare, Waukesha, WI, USA] is a DNN-based reconstructive algorithm that uses high-quality FBP data for teacher data in the learning process. DLIR has low, medium, and high set strengths, each reconstructed from a different DNN model. The traditionally used hybrid-IR (ASiR-V; GE Healthcare, Waukesha, WI, USA) differs such that the strength is adjusted by blending FBP and IR [9]. In addition, the advanced intelligent clear-IQ engine (AiCE; Canon Medical Systems, Tokyo, Japan), which has been commercialized as reconstruction using deep learning, differs such that it uses MBIR images for the learning process [12].

## Object size modulation phantom

Mercury 4.0 Phantom (Gammex, Middleton, WI, USA) was developed to evaluate the TCM installed in a CT system [20, 21]. This phantom was approved by the American Association of Physicists in Medicine (AAPM) Task Group (TG) 233 [21]. Mercury 4.0 Phantom has five sections with different diameters (16, 21, 26, 31, and 36 cm) made of polyethylene (Fig. 1). The phantom has a uniform layer for evaluating noise characteristics and five cylinders of water, bone, polystyrene, 10 mg/mL iodine, and air at a constant distance from the center as contrast inserts for evaluating the resolution. By scanning this phantom with TCM, it is possible to acquire images under clinical conditions, and by evaluating the image quality at each phantom size, comprehensive image quality assessment can be performed.

## Data acquisition and image reconstruction

The data acquisition in this study used a 256-slice Revolution CT (GE Healthcare, Waukesha, WI, USA) with hybrid-IR and DLIR as the image reconstruction algorithms. The X-ray tube voltage was 120 kV, while the X-ray tube current used was the TCM, which was controlled with a noise index (NI). The current was modulated at 10–720 mA by the phantom diameter and NI. NI is determined by the operator, and the imaging dose is controlled to have a constant standard deviation (SD) in the central region of the phantom, regardless of patient size [22]. In this study, two NIs (13.7 and 22.4) were set, in which the tube current varied without saturation at each phantom size within the range of the upper and lower limits of the TCM (10–720 mA). Considering the guidance level of the adult abdomen announced by the International Atomic Energy Agency (IAEA), NI = 13.7 is the standard dose level, and NI = 22.4 is the low dose level [23]. The volume CT dose index (CTDIvol) reported by the scanner console was recorded in a DICOM radiation dose structured report file after each scan. The mean CTDIvol recorded from the apparatus was 11.6 mGy at standard dose level (NI = 13.7) and 3.07 mGy at low dose level (NI = 22.4). Nominal CTDIvols with phantom diameters of 16, 21, 26, 31, and 36 cm were 2.0, 3.3, 7.4, 17.5, and 28.8 mGy, respectively, at standard doses; at low doses, it was 0.73, 1.18, 2.45, 4.81, and 6.49 mGy, respectively. The detector configuration was 0.625 mm × 256 rows, and the rotation time was 0.6 s/rotation. When NI = 13.7, the focal spot size was XL, and when NI = 22.4, the focal spot size was S. The images were subjected to FBP, hybrid IR, and DLIR. Hybrid-IR was evaluated for the ASiR-V50 (IR50; a 50 × 50 combination of FBP and IR), which is frequently used in clinical practice, and IR100, which shows strong performance. DLIR was evaluated for three strengths: low (DL-L), medium (DL-M), and high (DL-H). Images were reconstructed with a standard kernel: slice thickness, 1.25 mm (gap less); field of view, 400 mm; and matrix size, 512 × 512 pixels.

## Image quality assessment

ImQuest, developed by the clinical imaging physics group at Duke University, was used to analyze the acquired images. ImQuest is an open-source software that uses the technology described in TG233 of the AAPM and is compatible with the Mercury 4.0 Phantom. To assess the image quality characteristics, the noise power spectrum (NPS), task-based modulation transfer function (TTF), and task-based detectability index ( $d'$ ) were calculated using ImQuest. The quality and quantity of in-plane noise were evaluated using NPS. As shown in Fig. 2a, we set 40 × 40 pixel ROIs and obtained two-dimensional NPS from 60 consecutive axial slices (240 ROIs) in each section with uniform layers [21]. As shown in Fig. 2b, the resolution was measured for two low-to-medium contrast inserts: polystyrene ( $\approx \Delta 50\text{HU}$ ) and water ( $\approx \Delta 90\text{HU}$ ). A circular ROI was set for each insert, and the TTF was evaluated by the circular edge technique by averaging 80 consecutive axial slices in each section [21, 24, 25]. On the other hand,  $d'$  is an index that can be virtually simulated as the detectability by a radiologist using the measured NPS and TTF and by setting the detection task that is clinically required using the following equation:

$$d'^2 = \frac{\{\iint |W(u, v)|^2 \cdot TTF(u, v)^2 \cdot E(u, v)^2 dudv\}^2}{\iint |W(u, v)|^2 \cdot TTF(u, v)^2 \cdot NPS(u, v)^2 \cdot E(u, v)^4 dudv}$$

where  $u$  and  $v$  are the spatial frequencies of the  $x$ - and  $y$ -coordinates, respectively.  $NPS(u, v)$  and  $TTF(u, v)$  are measured from the phantom, while  $W(u, v)$  is a task function found from arbitrary tumor sizes, assuming clinical practice. On the other hand,  $E(u, v)$  is an eye filter that models the sensitivity of the human visual system to various spatial frequencies and requires the setting of observation conditions [26]. In this study, we simulated target tumors in the abdominal region and used two low-to-medium TTFs, NPS obtained from different doses, and phantom sizes. Then, the calculation condition of  $d'$  was based on a non-prewhitening matched filter with an eye filter (NPWE), a reconstructed field of view of 400 mm, a viewing distance of 450 mm, a zoom factor of 1.5, and a simulated tumor expressed in a Gaussian field of 8 mm [17, 21, 26, 27]. From the obtained  $d'$ , the rate of increase ( $d\%$ ) of each algorithm for FBP was calculated.

## Results

Figure 3 shows a portion of the NPS results at low doses, and Table 1 shows the results of the NPS peak and NPS average at each dose level and phantom size. NPSpeak showed similar values depending on the TCM with the set NIs; however, it was larger when the phantom size was 16 cm and smaller when the phantom size was 36 cm, and fluctuated slightly compared to 21, 26, and 31 cm diameters. This trend was also observed for different doses. In all conditions, FBP had the highest noise level, and DL-H had the lowest noise level. The NPSaverage frequency of DLIR compared with FBP was almost similar ( $0.01$ – $0.04 \text{ mm}^{-1}$  at low dose and  $0$ – $0.03 \text{ mm}^{-1}$  at normal dose), and the difference between FBP and IR50 was approximately  $0.04 \text{ mm}^{-1}$  regardless of dose. The difference between FBP and IR50 was approximately  $0.04 \text{ mm}^{-1}$  regardless of dose, but the difference between FBP and IR100 was very large (approximately  $0.18 \text{ mm}^{-1}$  at low dose and approximately  $0.17 \text{ mm}^{-1}$  at normal dose). Phantom size has no specific effect on the NPS.

Figure 4 shows some of the TTF results calculated for polystyrene ( $\approx \Delta 50\text{HU}$ ) and water ( $\approx \Delta 90\text{HU}$ ), and Table 2 shows the TTF50% results for each dose level and phantom size. The average deviation for each dose and phantom size was 4.2%. In TTF polystyrene ( $\approx \Delta 50\text{HU}$ ), FBP was the highest, while DL-M and IR50 were similar. Compared with FBP, both DLIR and ASIR-V decreased the TTF50% as the reconstruction strength increased. The rate of decrease of TTF50% with respect to FBP also showed the largest decrease rate of DL-H, with a maximum of 44.4% (36 cm in diameter, NI = 22.4). In hybrid-IR, the rate of decrease in IR100 was the highest, with a maximum of 59.1% (21 cm in diameter, NI = 22.4). In TTF water ( $\approx \Delta 90\text{HU}$ ), FBP and DLIR were similar, and DLIR showed favorable results in the low spatial frequency region. The results obtained from water were similar to those of polystyrene, and compared with FBP, both DLIR and ASIR-V decreased TTF50% as the reconstruction strength increased. In DLIR, the rate of decrease of TTF50% with respect to FBP had the largest decrease rate of DL-H, with a maximum of 16.3% (16 cm in diameter, NI = 22.4). In hybrid-IR, the rate of decrease in IR100 was the highest, with a maximum of 44.4% (36 cm in diameter, NI = 22.4).

Figure 5 shows the results of  $d'$  (low dose) for each phantom size, and Table 3 shows the results of the increase in  $d\%$  at each dose level and phantom size. The trend of the results was similar regardless of the variation in dose, and FBP was the least detectable under all conditions. Compared with IR50, the detectability of DL-M and DL-H was high, and the detectability of DL-H was high under all conditions. The smaller the phantom size, the higher the detectability. However, there was no difference in  $d'$  by phantom size in the 21, 26, and 31 cm diameters, where the SD variation was small.

## Discussion

The image quality of the DLIR was evaluated using TCM with the Mercury 4.0 Phantom. The SD varied when the phantom size was 16 cm and 36 cm relative to the set NI. The NPSpeak values under each condition also varied accordingly, possibly due to the bowtie filter of the CT system [28]. The bowtie filter intends to equalize the in-plane dose distribution of the human body. In general, the size of the bowtie filter varies according to the size of the subject, but it cannot be changed during the scan. Therefore, we believe that it affected the control by TCM and caused the SD to vary.

According to the NPS average frequency and shape of the NPS, DLIR was able to maintain the shape and reduce the amount of image noise without any shift in the NPS average frequency compared to FBP. The ideal noise reduction, where the noise characteristics are improved by increasing the dose, is when the shape of the NPS is maintained, and the amount of shift is reduced; the noise reduction method of the DLIR algorithm can be considered to reduce noise without changing the details of the original image. On the other hand, the NPS average frequency of IR50 did not show a large variation; however, similar to IR100, it showed a characteristic shape as if it had been smoothed in the high-frequency region. The effect of smoothing on IR100 was strong, and the NPS average frequency also shifted significantly to the low-frequency side. The IR100 image also shows the image quality influenced by high-frequency components (Fig. 6), which is consistent with previous research [5]. Regarding the TTF, the focus size was automatically changed by the NI setting, and the effect on TTF was expected, although there was almost no effect [29]. In previous studies on the TTF of DLIR, it has been reported that the resolution characteristics are improved compared to the FBP, which is not necessarily the case in this experiment [8, 10]. Hara et al. [30] shows the effect of the ROI position on the measured value of the MTF measurement. The effect of the selected phantom on the measurement results cannot be ignored. However, it is considered that DLIR easily recognizes the contrast at a high contrast, such as 300 HU, as reported in a previous study; at a low contrast of approximately 50 HU, as in this study, the resolution characteristics were lower than FBP.

Regarding  $d'$ , FBP was less detectable at all doses and phantom sizes. In hybrid-IR and DLIR, compared with IR50, the detectability of the image reconstructed by DLIR was high under all conditions. DL-H showed the highest detectability, with an improvement in reconstruction strength. Regarding TTF, DLIR did not always show high-resolution characteristics due to contrast; however, NPS showed high noise characteristics (low image noise), and the simulated tumor showed high  $d'$  in reconstruction by DLIR. Moreover, Urikura et al. [19] reported that the effect of the resolution characteristics was small for such low-contrast visibility, and considered that the high noise reduction effect of DLIR was highly useful in clinical practice. For phantom sizes of 16 and 36 cm,  $d'$  varies greatly, indicating the effect of noise characteristics.

The present study has some limitations. First, the tube voltage, which is the scan condition, was only 120 kVp, and all scans were performed only with helical scans. Second, the analysis using different phantom sizes was based on the adult size, and smaller phantom sizes do not represent children. The phantom is also circular and not an ellipse, which imitates a real person. In this study, the arrangement of the measured ROIs was the same regardless of the phantom size when measuring the NPS and TTF, which was necessary to maintain the accuracy of the experiment, although previous studies have described the effect of the measurement position. Finally, the finding of the present study has not been verified using actual clinical images. Under clinically assumed conditions, the subject size may change significantly, and in such cases, the detectability may change because of equipment hardware problems such as bowtie filters and output restrictions.

## Conclusions

In this study, we evaluated the overall image quality assessment of DLIR when TCM was used, targeting low to medium contrasts and varying phantom sizes and NIs. The results of noise control by TCM showed that DLIR has high detectability due to its high noise reduction capability, regardless of the phantom size. However, as shown in the phantom experiment, some phantom sizes may be overcorrected because of the output and hardware limitations of the CT system, and the same phenomenon may be observed in actual clinical practice depending on the size of the patient. Furthermore, when DLIR is used, which has a larger change in detectability than hybrid-IR, the operator needs to tune it according to the purpose.

## Declarations

### Funding

The authors did not receive support from any organization for the submitted work.

### Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

## Availability of data and materials

Not applicable

## Code availability

ImQuest Version 7.1 image quality evaluation software

## References

- [1] OECD (2015) Health at a Glance 2015: OECD Indicators. OECD Publishing, Paris
- [2] Brenner DJ, Hall EJ (2007) Computed tomography-an increasing source of radiation exposure. *N Engl J Med* 357: 2277-2284
- [3] Geyer LL, Schoepf UJ, Meinel FG et al. (2015) State of the Art: Iterative CT Reconstruction Techniques. *Radiology* 276:339-357. <https://doi.org/10.1148/radiol.2015132766>
- [4] Löve A, Olsson M, Siemund R (2013) Six iterative reconstruction algorithms in brain CT: a phantom study on image quality at different radiation dose levels. *Br J Radiol* 86(1031):20130388. <https://doi.org/10.1259/bjr.20130388>
- [5] McCollough CH, Yu L, Kofler JM et al (2015) Degradation of CT Low-Contrast Spatial Resolution Due to the Use of Iterative Reconstruction and Reduced Dose Levels. *Radiology* 276(2):499-506. <https://doi.org/10.1148/radiol.15142047>
- [6] Mileto A, Guimaraes LS, McCollough CH et al (2019) State of the Art in Abdominal CT: The Limits of Iterative Reconstruction Algorithms. *Radiology* 293(3):491-503. <https://doi.org/10.1148/radiol.2019191422>
- [7] Urikura A, Hara T, Ichikawa K et al (2016) Objective assessment of low-contrast computed tomography images with iterative reconstruction. *Phys Med* 32(8):992-998. <https://doi.org/10.1016/j.ejmp.2016.07.003>
- [8] Greffier J, Hamard A, Pereira F et al (2020) Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for CT: a phantom study. *Eur Radiol* 30:3951-3959. <https://doi.org/10.1007/s00330-020-06724-w>
- [9] Hsieh J, Liu E, Nett B, Tang J, Thibault JB, Sahney S (2019) A new era of image reconstruction: TrueFidelity™. Technical white paper on deep learning image reconstruction. GE Healthcare
- [10] Racine D, Becce F, Viry A et al (2020) Task-based characterization of a deep learning image reconstruction and comparison with filtered back-projection and a partial model-based iterative reconstruction in abdominal CT: A phantom study. *Phys Med* 76:28-37. <https://doi.org/10.1016/j.ejmp.2020.06.004>
- [11] Benz DC, Benetos G, Rampidis G et al (2020) Validation of deep-learning image reconstruction for coronary computed tomography angiography: Impact on noise, image quality and diagnostic accuracy. *J Cardiovasc Comput Tomogr* 14(5):444-451. <https://doi.org/10.1016/j.jcct.2020.01.002>
- [12] Higaki T, Nakamura Y, Zhou J, et al (2020) Deep Learning Reconstruction at CT: Phantom Study of the Image Characteristics. *Acad Radiol* 27(1):82-87. <https://doi.org/10.1016/j.acra.2019.09.008>
- [13] Solomon JB, Li X, Samei E (2013) Relating noise to image quality indicators in CT examinations with tube current modulation. *AJR Am J Roentgenol* 200:592-600. <https://doi.org/10.2214/AJR.12.8580>
- [14] Solomon J, Lyu P, Marin D, Samei E (2020) Noise and spatial resolution properties of a commercially available deep learning-based CT reconstruction algorithm. *Med Phys* 47(9):3961-3971. <https://doi.org/10.1002/mp.14319>
- [15] Ria F, Solomon JB, Wilson JM, Samei E (2020) Technical Note: Validation of TG 233 phantom methodology to characterize noise and dose in patient CT data. *Med Phys* 47(4):1633-1639. <https://doi.org/10.1002/mp.14089>

- [16] Takata T, Ichikawa K, Mitsui W et al (2017) Object shape dependency of in-plane resolution for iterative reconstruction of computed tomography. *Phys Med* 33:146-151. <https://doi.org/10.1016/j.ejmp.2017.01.001>
- [17] Furlan A, Marin D, Vanzulli A et al (2011) Hepatocellular carcinoma in cirrhotic patients at multidetector CT: hepatic venous phase versus delayed phase for the detection of tumour washout. *Br J Radiol* 84(1001):403-412. <https://doi.org/10.1259/bjr/18329080>
- [18] Cham MD, Yankelevitz DF, Shaham D et al (2000) Deep venous thrombosis: detection by using indirect CT venography. The Pulmonary Angiography-Indirect CT Venography Cooperative Group. *Radiology* 226(3):744–751. <https://doi.org/10.1148/radiology.216.3.r00se44744>
- [19] Urikura A, Ichikawa K, Hara T et al (2014). Spatial resolution measurement for iterative reconstruction by use of image-averaging techniques in computed tomography. *Radiol Phys Technol* 7:358–366. <https://doi.org/10.1007/s12194-014-0273-2>
- [20] Ria F, Solomon JB, Wilson JM et al (2020) Technical Note: Validation of TG 233 phantom methodology to characterize noise and dose in patient CT data. *Med Phys* 47(4):1633-1639. <https://doi.org/10.1002/mp.14089>
- [21] Samei E, Bakalyar D, Boedeker KL et al (2019) Performance evaluation of computed tomography systems: Summary of AAPM Task Group 233. *Med Phys* 46(11):e735-e756. <https://doi.org/10.1002/mp.13763>
- [22] Gies M, Kalender WA, Wolf H, Suess C (1999) Dose reduction in CT by anatomically adapted tube current modulation. I. Simulation studies. *Med Phys* 26 (11), 2235–2247. <https://doi.org/10.1118/1.598779>
- [23] IAEA (1994) IAEA Safety Series No. 1154, International Basic Safety Standards for Protection of Radiation Sources. IAEA, Vienna
- [24] Richard S, Husarik DB, Yadava G et al (2012) Towards task-based assessment of CT performance: system and object MTF across different reconstruction algorithms. *Med Phys*; 39(7):4115-4122
- [25] Yu L, Vrieze TJ, Leng S et al (2015) Technical Note: Measuring contrast- and noise-dependent spatial resolution of an iterative reconstruction method in CT using ensemble averaging. *Med Phys* 42(5):2261-2267. <https://doi.org/10.1118/1.4916802>
- [26] Solomon J, Samei E (2016) Correlation between human detection accuracy and observer model-based image quality metrics in computed tomography. *J Med Imaging (Bellingham)* 3(3):035506. <https://doi.org/10.1117/1.Jmi.3.3.035506>
- [27] Chen YL, Ko CJ, Chien SY et al (2011) Tumor size as a prognostic factor in resected small hepatocellular carcinoma: a controversy revisited. *J Gastroenterol Hepatol* 26(5):851-857.
- [28] Gomez-Cardona D, Cruz-Bastida JP, Li K, Budde A et al (2016) Impact of bowtie filter and object position on the two-dimensional noise power spectrum of a clinical MDCT system. *Med Phys* 43(8):4495-4506. <https://doi.org/10.1118/1.4954848>
- [29] Grimes J, Duan X, Yu L et al (2015) The influence of focal spot blooming on high-contrast spatial resolution in CT imaging. *Med Phys* 42(10):6011-6020. <https://doi.org/10.1118/1.4931053>
- [30] Hara T, Ichikawa K, Sanada S, Ida Y (2010) Image quality dependence on in-plane positions and directions for MDCT images. *Eur J Radiol* 75(1):114-21. <https://doi.org/10.1016/j.ejrad.2009.03.060>

## Tables

Table 1. NPSpeak and NPSaverage values with different phantom sizes for each dose condition.

NPS peak (HU<sup>2</sup> mm<sup>2</sup>)

phantom size (cm)	NI 13.7						NI 22.4					
	FBP	DL-L	DL-M	DL-H	IR50	IR100	FBP	DL-L	DL-M	DL-H	IR50	IR100
16	370	134	82	40	147	68	1088	340	188	73	428	182
21	563	209	127	64	221	105	1573	503	283	116	662	282
26	557	212	132	67	224	104	1642	561	316	132	707	359
31	520	198	126	65	219	104	1732	613	352	150	795	407
36	750	282	178	91	319	148	2646	926	543	258	1375	755

NPS average spatial frequency (mm<sup>-1</sup>)

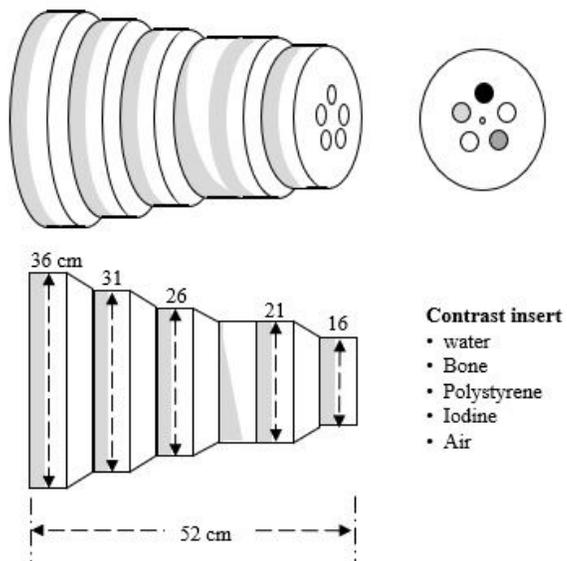
phantom size (cm)	NI 13.7						NI 22.4					
	FBP	DL-L	DL-M	DL-H	IR50	IR100	FBP	DL-L	DL-M	DL-H	IR50	IR100
16	0.33	0.32	0.31	0.30	0.29	0.14	0.33	0.32	0.32	0.30	0.29	0.14
21	0.33	0.31	0.30	0.29	0.29	0.14	0.32	0.32	0.31	0.30	0.28	0.14
26	0.32	0.31	0.30	0.28	0.28	0.14	0.31	0.31	0.30	0.28	0.27	0.13
31	0.31	0.30	0.29	0.27	0.27	0.14	0.30	0.30	0.29	0.28	0.26	0.13
36	0.31	0.30	0.29	0.27	0.27	0.14	0.28	0.29	0.28	0.26	0.24	0.13

Table 2. Target transfer function at 50% (TTF50%) for polystyrene and water inserts at different dose conditions and for each phantom size.

phantom size (diameter)		TTF 50% (mm <sup>-1</sup> )											
		NI: 13					NI: 22.4						
		FBP	DL-L	DL-M	DL-H	IR50	IR100	FBP	DL-L	DL-M	DL-H	IR50	IR100
polystyrene ( $\approx\Delta 50\text{HU}$ )	16	0.38	0.32	0.31	0.29	0.29	0.22	0.44	0.38	0.31	0.28	0.3	0.2
	21	0.41	0.32	0.3	0.28	0.28	0.21	0.44	0.34	0.3	0.25	0.28	0.18
	26	0.37	0.31	0.29	0.26	0.28	0.19	0.39	0.32	0.28	0.22	0.26	0.18
	31	0.35	0.3	0.28	0.26	0.26	0.19	0.31	0.27	0.24	0.22	0.23	0.18
	36	0.37	0.31	0.29	0.27	0.27	0.21	0.36	0.28	0.25	0.2	0.24	0.17
water ( $\approx\Delta 90\text{HU}$ )	16	0.38	0.39	0.38	0.38	0.37	0.35	0.43	0.39	0.38	0.36	0.35	0.28
	21	0.38	0.38	0.37	0.36	0.33	0.29	0.37	0.36	0.35	0.33	0.3	0.22
	26	0.36	0.36	0.35	0.34	0.31	0.27	0.35	0.35	0.34	0.32	0.27	0.21
	31	0.34	0.34	0.33	0.33	0.29	0.26	0.37	0.33	0.32	0.31	0.28	0.21
	36	0.35	0.35	0.35	0.34	0.3	0.26	0.36	0.36	0.32	0.31	0.24	0.2

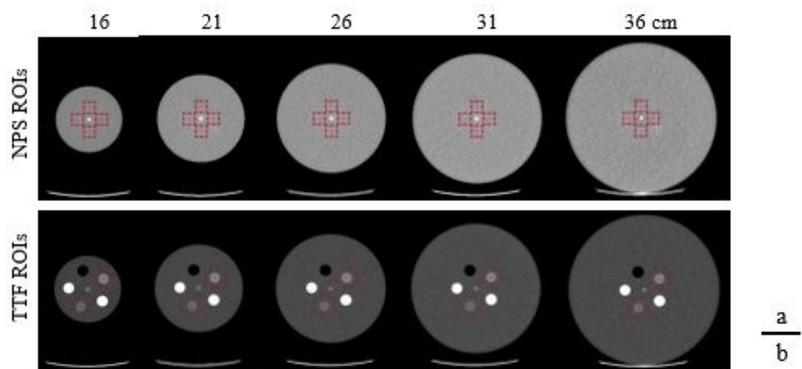
Table 3. Rate of increase in detectability index ( $d'$ ) with each phantom size for different contrasts (polystyrene and water) at each dose condition.

phantom size (diameter)		rate of increase in $d'$ (%)									
		NI: 13					NI: 22.4				
		DL-L	DL-M	DL-H	IR50	IR100	DL-L	DL-M	DL-H	IR50	IR100
polystyrene ( $\approx\Delta 50\text{HU}$ )	16	40.0	76.8	124.2	20.5	48.7	30.2	69.6	170.2	20.2	55.7
	21	43.9	73.1	125.1	23.7	59.7	64.4	103.3	163.0	27.5	66.1
	26	38.0	60.2	103.3	17.2	44.7	47.4	84.9	143.8	23.8	54.7
	31	28.2	56.5	90.8	14.5	36.6	48.7	78.2	123.4	18.7	42.5
	36	21.9	62.5	100.0	20.6	45.9	50.2	89.5	136.7	18.7	44.7
water ( $\approx\Delta 90\text{HU}$ )	16	45.5	86.7	139.6	24.3	58.1	63.6	109.1	196.2	23.6	64.2
	21	41.7	72.4	125.6	21.1	57.8	69.3	108.0	172.5	23.6	54.4
	26	45.5	70.9	114.4	22.8	55.3	59.6	103.4	171.1	26.5	55.3
	31	32.3	65.8	103.1	15.7	41.5	50.5	106.3	169.0	22.8	52.8
	36	41.9	61.4	98.8	18.5	48.9	63.2	94.4	166.2	23.6	53.3



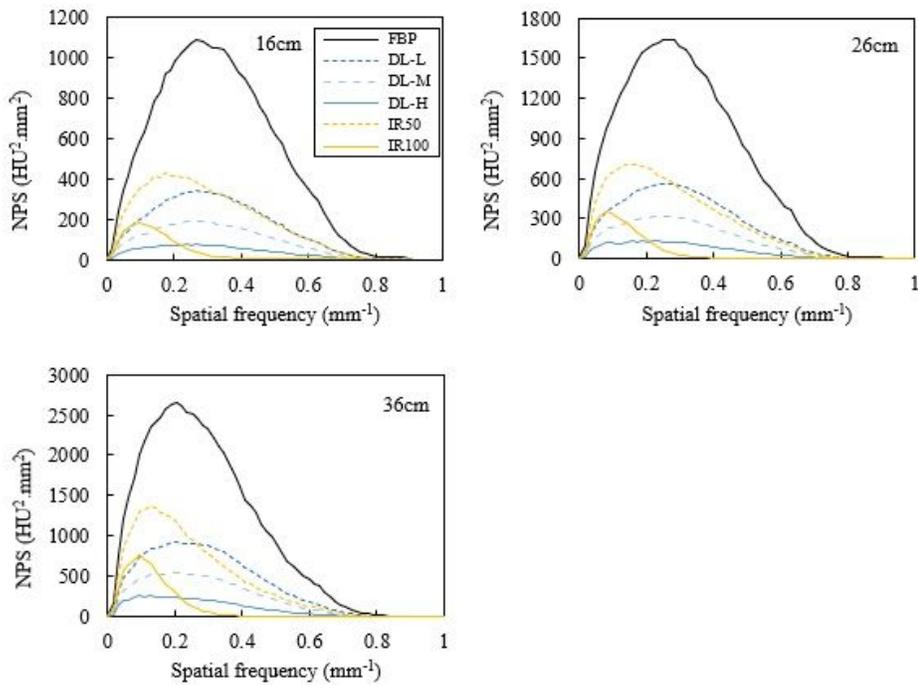
**Figure 1**

Composition of Mercury Phantom used in this study



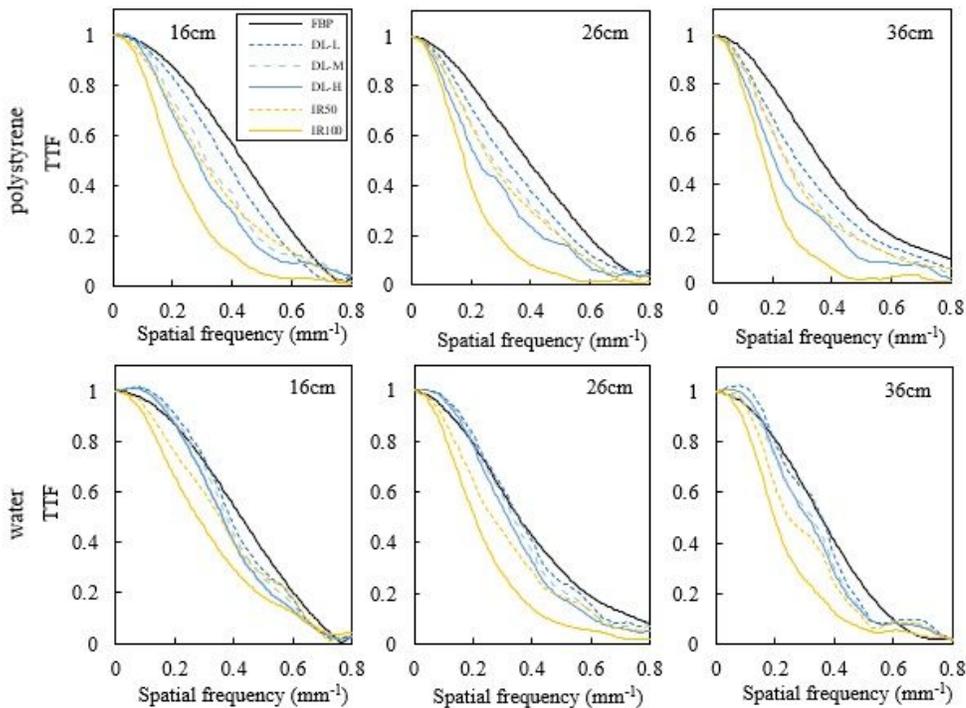
**Figure 2**

a) Regions of interest (ROIs) located for the noise power spectrum (NPS) measurements b) ROIs located for task-based modulation transfer function (TTF) measurements with polystyrene and water inserts



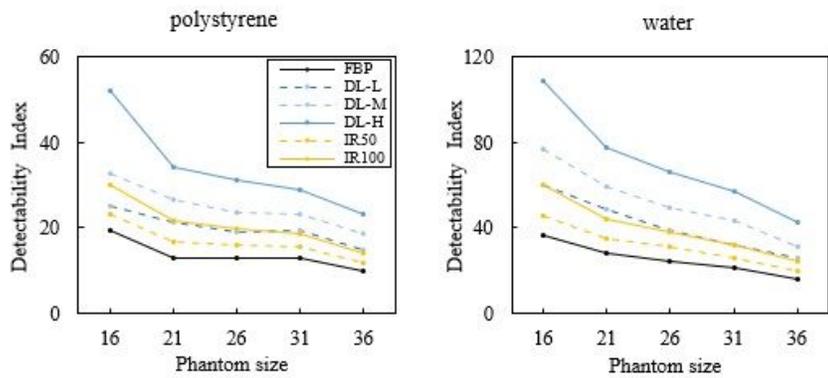
**Figure 3**

Noise power spectrum (NPS) at different phantom sizes measured under low dose conditions; 21 cm and 31 cm were equivalent to 26 cm



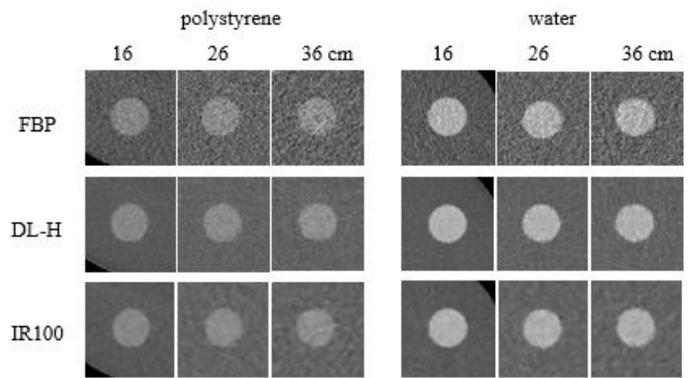
**Figure 4**

The TTFs of polyethylene and water with different phantom sizes at low dose conditions; 21 and 31 cm, were less affected by phantom size, and the trend with irradiation dose was similar.



**Figure 5**

Detectability index ( $d'$ ) by phantom size for different contrasts (polystyrene and water) at low dose conditions



**Figure 6**

Images of each contrast insert with phantom size at low dose conditions