

## A comparative study based on image quality and clinical task performance for CT reconstruction algorithms in radiotherapy

Hua Li,<sup>1a</sup> Steven Dolly,<sup>1</sup> Hsin-Chen Chen,<sup>1</sup> Mark A. Anastasio,<sup>2</sup> Daniel A. Low,<sup>3</sup> Harold H. Li,<sup>1</sup> Jeff M. Michalski,<sup>1</sup> Wade L. Thorstad,<sup>1</sup> Hiram Gay,<sup>1</sup> and Sasa Mutic<sup>1</sup>

*Department of Radiation Oncology,<sup>1</sup> Washington University, St. Louis MO; Department of Biomedical Engineering,<sup>2</sup> Washington University, St. Louis MO; Department of Radiation Oncology,<sup>3</sup> University of California Los Angeles, CA, USA*  
huli@radonc.wustl.edu

Received 20 May, 2015; accepted 6 March, 2016

CT image reconstruction is typically evaluated based on the ability to reduce the radiation dose to as-low-as-reasonably-achievable (ALARA) while maintaining acceptable image quality. However, the determination of common image quality metrics, such as noise, contrast, and contrast-to-noise ratio, is often insufficient for describing clinical radiotherapy task performance. In this study we designed and implemented a new comparative analysis method associating image quality, radiation dose, and patient size with radiotherapy task performance, with the purpose of guiding the clinical radiotherapy usage of CT reconstruction algorithms. The iDose<sup>4</sup> iterative reconstruction algorithm was selected as the target for comparison, wherein filtered back-projection (FBP) reconstruction was regarded as the baseline. Both phantom and patient images were analyzed. A layer-adjustable anthropomorphic pelvis phantom capable of mimicking 38–58 cm lateral diameter-sized patients was imaged and reconstructed by the FBP and iDose<sup>4</sup> algorithms with varying noise-reduction-levels, respectively. The resulting image sets were quantitatively assessed by two image quality indices, noise and contrast-to-noise ratio, and two clinical task-based indices, target CT Hounsfield number (for electron density determination) and structure contouring accuracy (for dose-volume calculations). Additionally, CT images of 34 patients reconstructed with iDose<sup>4</sup> with six noise reduction levels were qualitatively evaluated by two radiation oncologists using a five-point scoring mechanism. For the phantom experiments, iDose<sup>4</sup> achieved noise reduction up to 66.1% and CNR improvement up to 53.2%, compared to FBP without considering the changes of spatial resolution among images and the clinical acceptance of reconstructed images. Such improvements consistently appeared across different iDose<sup>4</sup> noise reduction levels, exhibiting limited interlevel noise (< 5 HU) and target CT number variations (< 1 HU). The radiation dose required to achieve similar contouring accuracy decreased when using iDose<sup>4</sup> in place of FBP, up to 32%. Contouring accuracy improvement for iDose<sup>4</sup> images, when compared to FBP, was greater in larger patients than smaller-sized patients. Overall, the iDose<sup>4</sup> algorithm provided superior radiation dose control while maintaining or improving task performance, when compared to FBP. The reader study on image quality improvement of patient cases shows that physicians preferred iDose<sup>4</sup>-reconstructed images on all cases compared to those from FBP algorithm with overall quality score: 1.21 vs. 3.15,  $p = 0.0022$ . However, qualitative evaluation

<sup>a</sup> Corresponding author: Hua Li, Department of Radiation Oncology, Washington University School of Medicine, 4921 Parkview Place, Saint Louis, MO 63110, USA: phone: (314) 362 0129; fax: (314) 362 8521; email: huli@radonc.wustl.edu

strongly indicated that the radiation oncologists chose iDose<sup>4</sup> noise reduction levels of 3–4 with additional consideration of task performance, instead of image quality metrics alone. Although higher iDose<sup>4</sup> noise reduction levels improved the CNR through the further reduction of noise, there was pixelization of anatomical/tumor structures. Very-low-dose scans yielded severe photon starvation artifacts, which decreased target visualization on both FBP and iDose<sup>4</sup> reconstructions, especially for the 58 cm phantom size. The iDose<sup>4</sup> algorithm with a moderate noise reduction level is hence suggested for CT simulation and treatment planning. Quantitative task-based image quality metrics should be further investigated to accommodate additional clinical applications.

PACS number(s): 87.57.C-, 87.57.Q-

Key words: ALARA principle, task-based image quality, iDose<sup>4</sup> reconstruction, simulation CT, radiotherapy

## I. INTRODUCTION

The filtered back-projection (FBP) algorithm has been widely utilized as the standard image reconstruction algorithm for commercial CT scanners.<sup>(1)</sup> Although it is computationally efficient, the FBP algorithm does not account for the stochastic nature of projection data and can produce streaking artifacts, especially for projection data with severe noise contamination. Moreover, such conditions are likely to produce inaccurate estimates of CT Hounsfield numbers. Recently, iterative reconstruction techniques have been developed to more effectively mitigate image artifacts and reduce image noise.<sup>(2,3)</sup> In radiotherapy, CT simulation images have been frequently utilized in specific tasks, including determination of tissue electron density (via CT Hounsfield numbers) and delineation of tumor and normal organ anatomy (by either manual techniques or increasingly by the use of semiautomated or automated techniques).<sup>(4-8)</sup> However, the relationship between the image quality improvements of iterative reconstruction and the performance of specific radiotherapy tasks has not been fully investigated as yet.

The iDose<sup>4</sup> iterative reconstruction algorithm (Philips Medical, Cleveland, OH) has been commercially released as an artifact and radiation dose reduction tool, and mainly is used to improve image quality combined with dose reduction capabilities for diagnostic imaging purpose, while following the as-low-as-reasonably-achievable (ALARA) principle.<sup>(3,9-11)</sup> It has been reported from routine works of our clinical institutions that traditional diagnostic image quality definitions may not be realistic enough to determine the performance of the radiotherapy tasks, which certainly influence treatment planning accuracy. Inspired by a recent paradigm shift of image quality assessment that image quality should be defined by the ability of a user to perform medically or scientifically relevant tasks with the image data,<sup>(12,13)</sup> the purpose of this study is: 1) to conduct new comparative experiments of the FBP and iDose<sup>4</sup> algorithms, with varying noise reduction levels, by associating the ALARA principles with radiotherapy task performance, and 2) to provide suggestions for clinical radiotherapy usage of the iDose<sup>4</sup> algorithm. This study is expected to fill the gap between image quality assessment and the clinical usage of CT simulation images in radiotherapy.

## II. MATERIALS AND METHODS

### A. Description of the phantom

An anthropomorphic human male pelvis phantom (CIRS Inc., Norfolk, VA), as shown in Fig. 1, with additional layers of attenuation, was employed to simulate adult male patient lateral diameters of 38, 43, 48, 53, and 58 cm. The pelvis phantom was constructed from proprietary epoxy materials that mimic human tissue density and radiation attenuation properties of human tissue within 1% from 50 keV to 25 MeV. The overall phantom dimensions were 38 cm wide, 26 cm thick, and 37 cm tall. The phantom contained anatomically precise bone, cartilage, spinal cord, vertebral disks, muscle, intestines, bladder, prostate, rectum, and interstitial fat. All anatomical dimensions of this phantom were based on the visible human project datasets that serve as a reference for the study of human anatomy and are available through the National Library of Medicine. The additional attenuation layers were provided by tissue-equivalent bolus material (Mick Radio-Nuclear Instruments, Inc. Mount Vernon, NY), which was a synthetic oil gel with a specific gravity of 1.02, and based on vinyl plastic containing a large amount of diisodecyl phthalate.

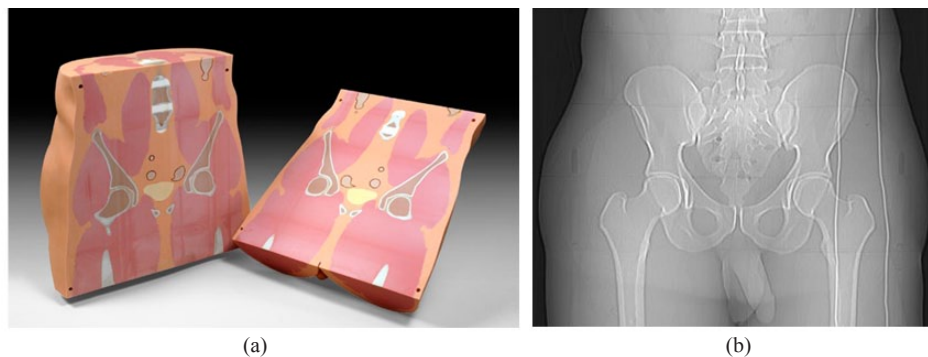


FIG. 1. Anthropomorphic pelvis phantom (a) (courtesy to CIRS Inc., Norfolk, VA); (b) CT anterior–posterior survey image of the phantom.

### B. CT Acquisitions of the phantom

CT scans were acquired with a Philips Brilliance 64 slice CT simulator (Philips Medical, Cleveland, OH). To assure that the scans did not exceed the tube current limits of the scanner, we employed the commercial 4D CT protocol, which reduced the pitch to 0.06 and provided effective mAs up to three to four times the standard helical protocol. The untagged 4D CT scans were reconstructed to simulate higher dose level helical CT scans for the analysis performed in this study. Tube potentials lower than 120 kVp were not suitable for scanning the smallest phantom size of 38 cm, and certainly not any larger sizes due to higher tissue attenuation.<sup>(14,15)</sup> Each sized phantom was scanned with two tube potential settings, 120 and 140 kVp. The tube current modulation function was disabled in order to evaluate the image quality under the full range of effective mAs. Standard B filter (one of the image reconstruction kernels) was used because it was calibrated by the manufacturer to provide accurate CT Hounsfield numbers and is suitable for routine abdomen, pelvis, and CT angiography scans. It is a recommended filter for clinical routine use. The detailed scanning parameters for this study were summarized in the Table 1.

TABLE 1. The detailed CT scanning parameters utilized in this study.

<i>Adjustable Scanning Parameters</i>	<i>Values</i>
kVp	120 and 140
Collimator setting (mm)	64 × 0.625
Rotation time (s)	0.5
Pitch	0.06
Reconstruction Filter	Standard B
Reconstruction slice thickness (mm)	3 mm
mAs/slice (effective mAs)	50, 100, 150, 200, 250, 500, 1000, 1500, 2000, 2500, and 3333 (for 120 kVp) or 3000 (for 140 kVp)
Corresponding CTDI <sub>vol</sub> (mGy) for 120 kVp scans	3.3, 6.6, 9.8, 13.1, 16.7, 32.7, 65.4, 98.1, 130.8, 168.0, and 216.9
Corresponding CTDI <sub>vol</sub> (mGy) for 140 kVp scans	4.9, 9.7, 14.5, 19.3, 24.7, 48.5, 97.0, 145.4, 193.9, 241.6, and 289.2
Reconstruction Algorithms	FBP and iDose <sup>4</sup> with noise reduction levels 1, 3, & 6
Phantom Diameters (cm)	38, 43, 48, 53, and 58

### C. CT Acquisitions of clinical patients

In addition to the CT scans of the phantom, CT image data from 14 head and neck cancer patients, 10 prostate cancer patients, and 10 gynecology (GYN) cancer patients were collected based on an approved IRB protocol to evaluate the visibility of anatomical structures. The iDose<sup>4</sup> reconstruction algorithm was retrospectively applied to the collected patient cases for further evaluation. Each scan originally reconstructed with FBP was reconstructed using iDose<sup>4</sup> algorithm with noise reduction levels 1–6. The iDose<sup>4</sup>-based reconstructed images were compared to the images reconstructed with the traditional FBP technique by two radiation oncologists and using the five-point scores described in Materials & Methods section F below.

### D. iDose<sup>4</sup> iterative reconstruction algorithm

The iDose<sup>4</sup> algorithm, which improves image quality while reducing radiation dose, is achieved through an iterative process as briefly described below. An adaptive linear filter is used on noisy projections in the projection domain followed by quantum mottle noise reduction in the image domain. The dual-domain processing steps are helpful for suppressing streak artifacts and noise, as well as visualizing underlying anatomical information. The noise filtering in the projection domain by the iDose<sup>4</sup> algorithm begins with identification of noisy point measurements in each projection. The noise model utilizes photon statistics to identify these data points, which are given a weight penalty through an iterative edge-preserving diffusion process. In the image domain, the quantum mottle noise distribution within the image volume is estimated and then subtracted using a best-fit structural model, chosen by the user, as a template. Multifrequency noise subtraction is implemented to maintain noise power spectrum constancy. The suppression effect of the quantum mottle noise in an image can be controlled by the iDose<sup>4</sup> reconstruction noise reduction levels (1–6), corresponding to a varied reduction range from low to high. By means of the dual-domain processing, the iDose<sup>4</sup> reconstruction algorithm can correct bias artifacts and maintain noise power spectrum constancy, while preserving spatial resolution. More details regarding the iDose<sup>4</sup> reconstruction algorithm can be found in the technical paper provided by Philips and in the literature.<sup>(3,9-11)</sup>

### E. Metrics of image quality and radiotherapy task performance

This comparative study was conducted to quantitatively compare the iDose<sup>4</sup> iterative reconstruction algorithm to the traditional FBP algorithm based on two image quality-based metrics, noise and contrast-to-noise ratio (CNR), and two radiotherapy task-based metrics, target CT Hounsfield number and target contouring accuracy (TCA). Considering that the prostate is a critical organ in radiotherapy, the prostate of the anthropomorphic phantom was selected as the target of analysis here. Three adjacent cross-sectional slices of the prostate were selected, and the same slices were chosen throughout all the reconstructed image sets for fair comparison. The contours of the prostate on the three slices were manually outlined by a trained medical physicist and reviewed by an experienced dosimetrist; radiation oncologists reviewed some selected datasets, as well. The average CT Hounsfield number and noise  $\sigma$  (CT Hounsfield number standard deviation) for each contoured prostate region were measured.  $\sigma_{target}$ , which describes the fluctuations in the target signal that may affect target delineation, was defined as

$$\sigma_{target} = \sqrt{\frac{1}{(F-1)} \sum_{f=1, f \in target}^F (CT\#_f - \overline{CT\#_{target}})^2} \quad (1)$$

where  $CT\#_f$  represents the CT number of each voxel  $f$ ,  $\overline{CT\#_{target}}$  is the average of CT numbers of all the voxels within the target, and  $F$  is the total voxel numbers within the target. Additionally, the average CT Hounsfield number and noise at the surrounding region outside the prostate were measured in the same way, and was used as the background values. The contrast-to-noise ratio (CNR), which describes the ability of the target to be detected against its background, was defined as

$$CNR = \frac{|\overline{CT\#_{target}} - \overline{CT\#_{background}}|}{\sigma_{background}} \quad (2)$$

where  $\overline{CT\#_{background}}$  is the average of CT Hounsfield numbers of a narrow band of three voxels' width surrounding and outside the prostate, and  $\sigma_{background}$  as the background noise, can be calculated using Eq. (1). The CT number, which is used to determine tissue electron density and therefore affects the dose calculation results, was simply defined as the average of CT numbers of the target, and its variations can be evaluated from its changes along with various scan protocols. The TCA of each reconstructed image set was defined as

$$TCA = (C^{target} \cap C_{ref}^{target}) / |C^{target} \cup C_{ref}^{target}| \quad (3)$$

where  $C^{target}$  is the set of target voxels on the image to be evaluated, and  $C_{ref}^{target}$  is the set of target voxels on reference image. This TCA was defined to consider the target coverage accuracy since it is the most important parameters to evaluate the accuracy of radiation treatment. Confirmed with both radiation oncologists who performed image quality evaluation in this study, the CT image of the 38 cm original phantom scanned with 120 kVp and the greatest available CTDI<sub>vol</sub> of 216.9 mGy and reconstructed with FBP algorithm was defined as the reference image. The reference ground truth contour was defined as the manual prostate contour  $C_{ref}$  delineated on this reference image. A value of one for TCA indicates a complete overlap between the contouring result and the reference contour, while a zero value represents no overlap between two contours. The relationships between the radiation dose, image quality metrics, and task-based indices were analyzed from the CT images acquired in Materials & Methods section C above.

### F. Qualitative five-point scoring criteria

We determined a five-point scoring criteria based on physicians' suggestions for qualitative evaluation of clinical patient images reconstructed by FBP and iDose<sup>4</sup>, respectively. For each case, the FBP-based and iDose<sup>4</sup>-based images were reviewed side-by-side. The radiation oncologists were allowed to browse through each image series, zoom in and out, and adjust the window settings. They ranked the overall image quality based on the visual conspicuity of critical structures using a five-point score: 0 = totally obscured, no structures identifiable; 1 = marked artifacts, questionable recognition; 2 = faint anatomic recognition; 3 = anatomical recognition with low confidence; 4 = anatomical recognition with medium confidence; and 5 = anatomical recognition with high confidence for segmentation. All the patients' images were evaluated by two radiation oncologists using the five-point scores. The preferred dataset and the related iDose<sup>4</sup> noise-reduction level (if iDose<sup>4</sup>-based image was preferred) were also indicated. The final ranking was obtained by averaging the ranks from two radiation oncologists. A Wilcoxon signed-rank test was performed to analyze the quality ranking and to determine significant differences between the FBP and iDose<sup>4</sup> reconstructed images. A p-value of less than 0.05 indicated a statistically significant difference.

## III. RESULTS

### A. Phantom CT simulation data

#### A.1 Noise analysis

Figure 2 illustrates the relationship of image noise to the reconstruction algorithms and scanning protocols of varying radiation dose levels. Here each reported noise value is the average of measurements on three continuous slices, and the variations are within five HU for all the

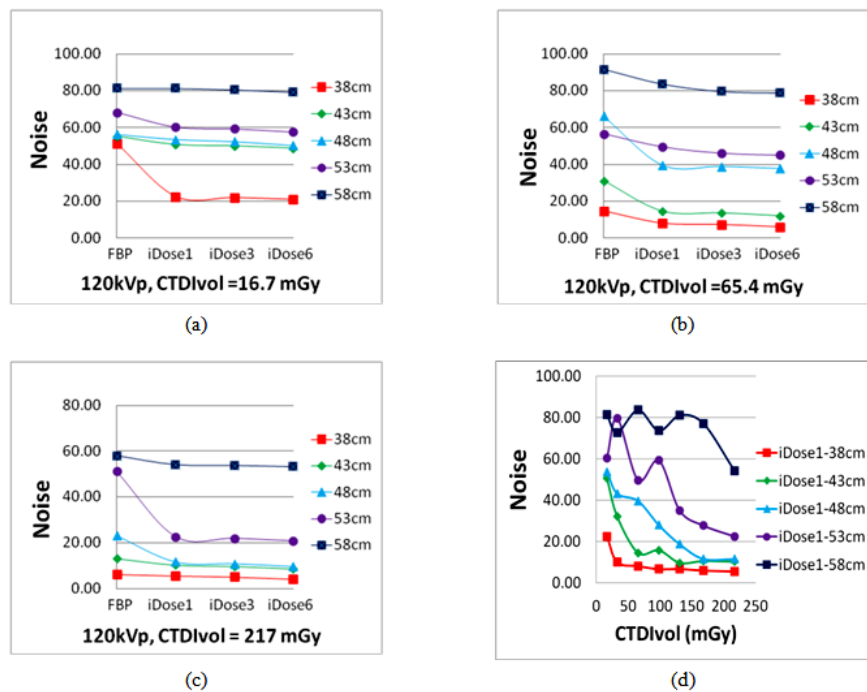


FIG. 2. Noise comparison for the prostate; (a)–(c) image noise plots with respect to different reconstruction algorithms under various radiation doses; (d) image noise plot obtained by the iDose<sup>4</sup> with noise reduction level 1 algorithm while altering both the phantom thickness and radiation dose, using a 120 kVp setting.

measurements. For a given phantom size and a given scanning protocol, the iDose<sup>4</sup> reconstruction algorithm always produced images with lower noise levels than the FBP algorithm, reducing image noise up to 66.1% when compared with the FBP algorithm without considering the clinical acceptance of the images and the changes on special resolutions of images reconstructed by iDose<sup>4</sup> and FBP reconstruction algorithms, respectively. This highest noise reduction happened on the images that were reconstructed with iDose<sup>4</sup> reconstruction with noise reduction level 6 on the projection data scanned with the protocol of CTDI<sub>vol</sub> of 16.7 mGy on 38 cm sized phantom. As the radiation dose increases, the iDose<sup>4</sup> algorithm reduces noise more effectively for the larger phantoms. As shown in Figs. 2(a) to 2(c), for the largest phantom size employed in this study (58 cm), the iDose<sup>4</sup> algorithm was unable to reduce the noise, when compared with FBP, with a 120 kVp scan parameter. This phenomenon occurs even for the highest CTDI<sub>vol</sub> (217 mGy) used in this study, as the noise within the prostate was only reduced by 6.6%.

### A.2 Contrast-to-noise ratio (CNR) analysis

Figure 3 illustrates the comparison of the CNR with different scanning doses and image reconstruction algorithms for different phantom thicknesses. The iDose<sup>4</sup> reconstruction algorithm produced images with a higher CNR than the FBP algorithm. In general, the largest improvement in CNR occurred at noise reduction level 6 for a given phantom size and scan dose. The maximal improvement of the iDose<sup>4</sup> algorithm over the FBP algorithm occurred for the 38 cm sized phantom. As shown in Fig. 3(a), the CNR is 1.65 when using the FBP algorithm, but increases by 369.5% to 7.74 when using iDose<sup>4</sup> with noise reduction level 1. For a particular phantom size and scan parameters, the CNR increases with an increasing iDose<sup>4</sup> noise reduction level. For the largest phantom size used in this study (58 cm), limited CNR improvement was observed when comparing the iDose<sup>4</sup> algorithm with FBP, for 120 kVp scan parameters. A similar trend was observed from our experiments of the 140 kVp scans, which were not illustrated here.

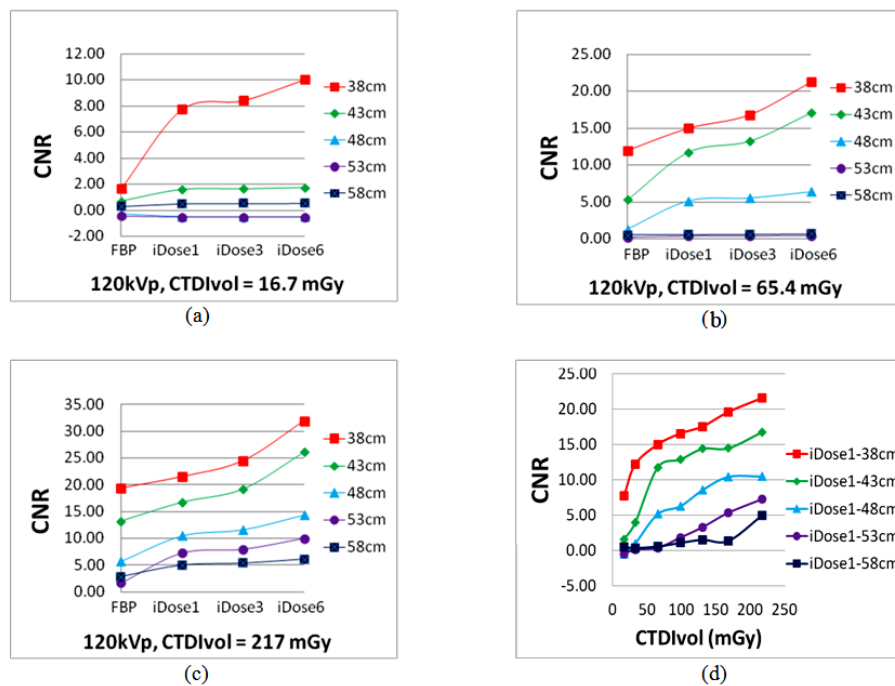


Fig. 3. CNR comparison ((a) to (c)) of various reconstruction algorithms; (d) CNRs of a single reconstruction under changes of the phantom thickness and the radiation dose.

### A.3 Target CT number variation analysis

Accurate and consistent CT Hounsfield numbers are required in order to properly determine the tissue electron density in dose calculations, an important task in radiotherapy. A comparison for the average CT Hounsfield number of the target obtained by different imaging protocols was performed, as illustrated in Fig. 4. Given the same phantom size and radiation dose, the CT number obtained by using the FBP algorithm was different from those acquired by the iDose<sup>4</sup> algorithm. While looking into the iDose<sup>4</sup> algorithm, for a particular CTDI<sub>vol</sub> value and phantom size, a nearly constant CT number was maintained, indicating that iDose<sup>4</sup> noise reduction level does not affect the consistency of the CT number. Taking the example of the 58 cm sized phantom, the maximum difference of the average CT numbers between the iDose<sup>4</sup> reconstructions with noise reduction levels 1, 3, and 6 was only a few of Hounsfield units. However, the CT number can vary given different phantom sizes. For a particular reconstruction method, the variations of the CT numbers caused by the phantom size decreased as the CTDI<sub>vol</sub> increased. This phenomenon can be observed based on Fig. 4(d), wherein the CT numbers obtained by applying the iDose<sup>4</sup> reconstruction with noise reduction level 1 for various sizes of the phantom converge progressively as the CTDI<sub>vol</sub> increased.

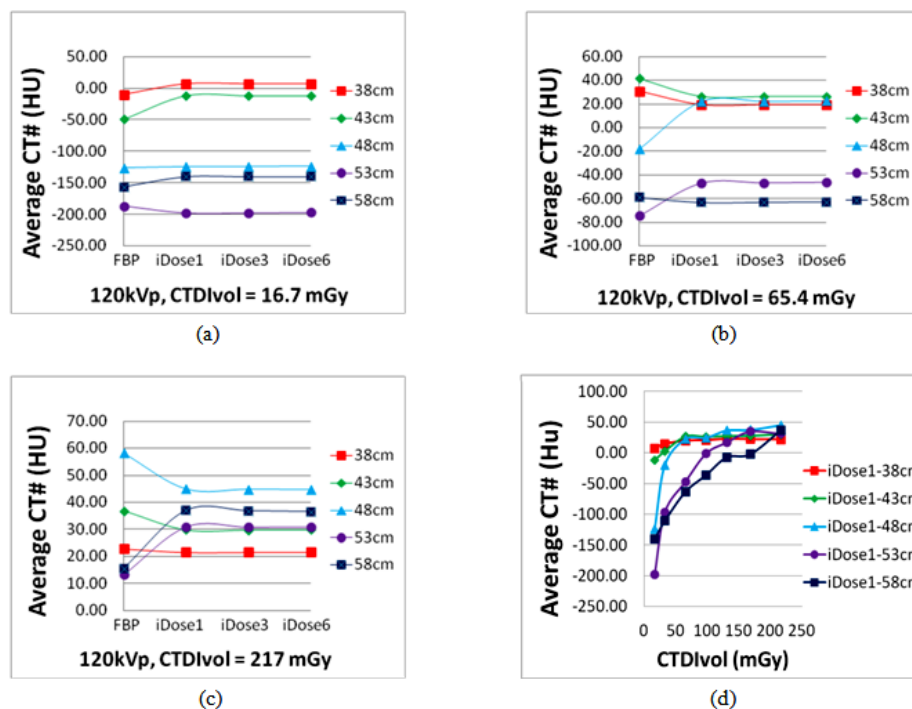


FIG. 4. Average CT Hounsfield number comparison; (a)-(c) plots of the CT number with respect to different imaging parameters for the FBP, the iDose<sup>4</sup> reconstruction with noise reduction levels 1, 3, and 6; (d) CT number variations obtained by using the iDose<sup>4</sup> reconstruction with noise reduction level 1 and altering the phantom thickness and the radiation dose.

### A.4 Target contouring accuracy (TCA) analysis

Other than CT number variations, the target contouring accuracy has been considered as another important task performance measure since it directly influences the accuracy of the dose-volume relationship of a particular structure. Figure 5 illustrates the relationship of the TCA for manual contouring with respect to the radiation dose based on the FBP and iDose<sup>4</sup> reconstruction with noise reduction level 3 for the prostate of the phantom. The TCA for both the FBP and iDose<sup>4</sup> algorithms depended greatly on both the phantom size and the radiation dose. When using the



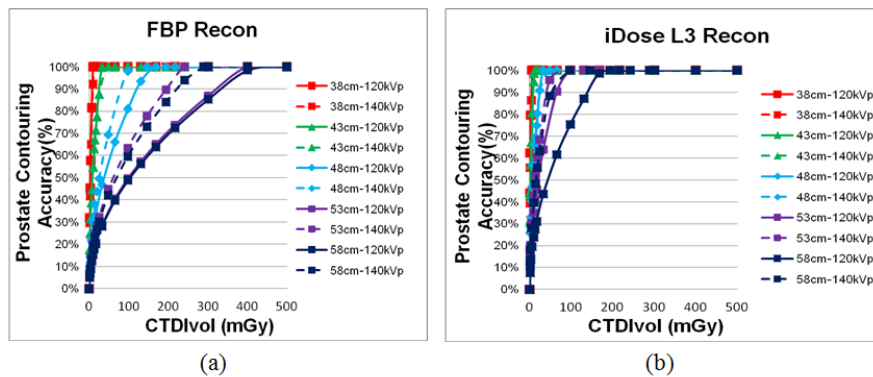


Fig. 5. Relationship between the manual prostate contouring accuracy and the CTDI<sub>vol</sub> based on the FBP (a) and iDose<sup>4</sup> reconstruction with noise reduction level 3 (b) under various phantom sizes and kVp settings.

120 kVp scan protocol, 100% TCA could not be achieved for the large-size phantoms of 53 and 58 cm diameter, even when a very high CTDI<sub>vol</sub> was used in this study (217 mGy). Generally, using the iDose<sup>4</sup> algorithm can achieve 100% TCA with a lower CTDI<sub>vol</sub> than the FBP algorithm requires. It was additionally observed from this analysis that higher radiation doses (i.e., CTDI<sub>vol</sub>) were needed for larger sized patients in order to achieve the same contouring accuracy as that performed on smaller sized patients. More specifically, for the smallest phantom diameter (38 cm), the prostate could be delineated with 100% accuracy for the lowest-dose scan (i.e., a CTDI<sub>vol</sub> of 16.7 mGy) for both the FBP and iDose<sup>4</sup> reconstruction with noise reduction level 3.

## B. Patient simulation CT data

The image data for 34 patients, reconstructed by the iDose<sup>4</sup> algorithm with noise reduction levels 1 to 6, were qualitatively evaluated by two radiation oncologists using the five-point scoring criteria. Table 2 lists the evaluation results. For all cases, the oncologists preferred to use the simulation CT images reconstructed by the iDose<sup>4</sup> with moderate noise reduction levels 3–5 for treatment planning. Overall, the iDose<sup>4</sup> algorithm provided superior radiation dose control while maintaining or improving task performance when compared to FBP. The reader study on patient cases shows that physicians prefer iDose<sup>4</sup> reconstructed images compared to those from FBP algorithm with overall quality score: 1.21 vs. 3.15,  $p = 0.0022$ . The qualitative evaluation also indicated that the radiation oncologists chose an optimal and moderate iDose<sup>4</sup> noise reduction level (3 or 4) instead of the highest noise reduction level 6.

As shown in Fig. 6, although the iDose<sup>4</sup> algorithm with a higher noise reduction level can improve the CNR, the performed iterative diffusion process (as described in Materials & Methods section D above) achieves edge-preserving smoothness only from a numerical point of view and lacks consideration of anatomical knowledge and geometrical priors, resulting in unrealistic degradation of spatial resolution (as indicated by the circles) and sharper but jagged edges (as indicated by the arrow), as described in image processing literature.<sup>(16,17)</sup>

TABLE 2. Qualitative evaluation results on clinical patient cases.

Treatment Site	Number of the Evaluated Patient Cases	Which Reconstruction Algorithm Is Preferred by Physicians? (FBP or iDose <sup>4</sup> )	If iDose <sup>4</sup> Algorithm Was Preferred, Which Noise Reduction Level Was Preferred?
Head & Neck	14	iDose <sup>4</sup>	4 to 5
Prostate patients	10	iDose <sup>4</sup>	4 to 5
GYN patients	10	iDose <sup>4</sup>	3 or 4

Figure 7 shows an example of photon starvation artifacts on images of large-sized phantom scanned with low-dose protocol. The diameter size of the phantom is 53 cm. The phantom was scanned with low-dose protocol ( $\text{CTDI}_{\text{vol}} = 29.93 \text{ mGy}$ ). The images were reconstructed with FBP and iDose<sup>4</sup> algorithms, respectively. It is clear that the photon starvation artifacts exist on both images.

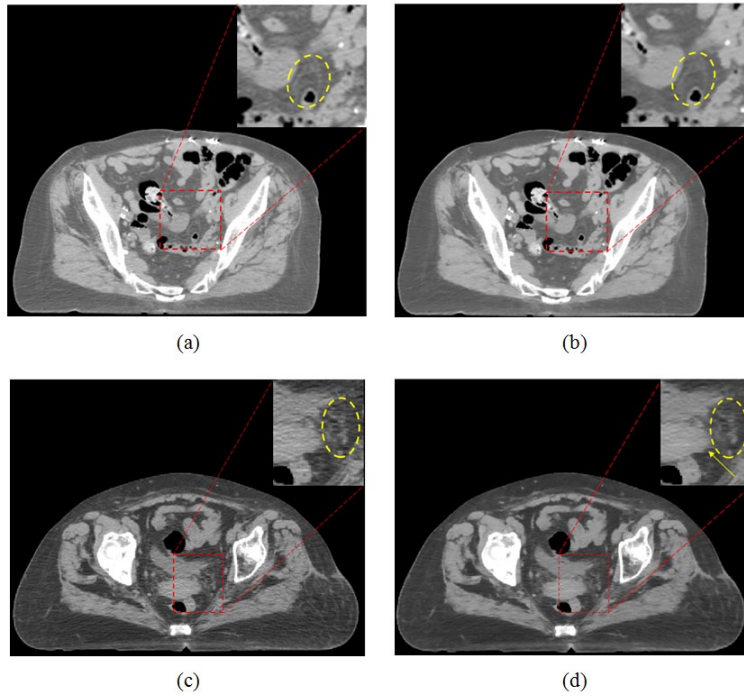


FIG. 6. Assessment of anatomy conspicuity and realism for the CT images produced by the iDose<sup>4</sup> algorithm with (a) and (c) noise-reduction-level 1, and (b) and (d) noise reduction level 6.

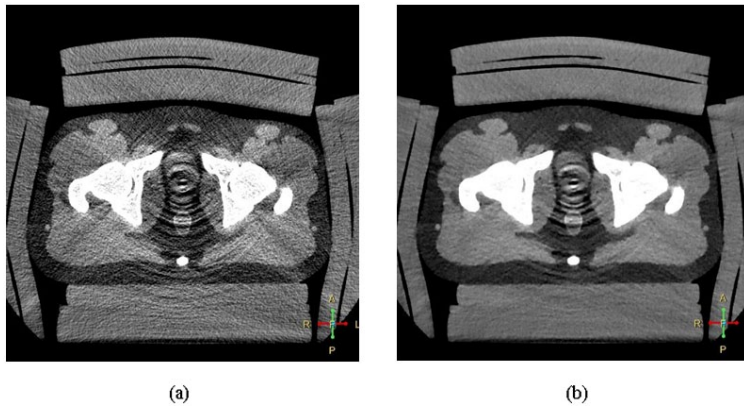


FIG. 7. The photon starvation artifacts are shown in images of 53 cm sized phantom scanned with same low-dose protocol of  $\text{CTDI}_{\text{vol}} = 29.93 \text{ mGy}$ , and reconstructed with different reconstruction algorithms: (a) the image is reconstructed with FBP; (b) the image is reconstructed with iDose<sup>4</sup> and noise reduction level 6.

#### IV. DISCUSSION

This comparative study has illustrated that the overall clinical efficacy of the FBP and the iDose<sup>4</sup> reconstruction algorithms are dependent on both the patient size and the CT scan parameters. For smaller sized patients, comparably accurate structure contours can be obtained from the images reconstructed by both the FBP and the iDose<sup>4</sup> algorithms. However, the image noise was lower and the contrast-to-noise ratio was improved in the images obtained by using iDose<sup>4</sup> when compared to the FBP algorithm (see Fig. 2). The noise reduction comparison shown in Fig. 2 is performed to compare only the noise level changes between iDose<sup>4</sup> and FBP reconstruction algorithms without considering the changes of spatial resolution for these two reconstruction algorithms. In this study, the scanning protocols were varied only on effective mAs and reconstruction algorithms while keeping other scanning parameters, such as slice thickness and reconstruction filter, the same. We have used these protocols to scan the line-pair phantom provided by the manufacturer and measured the MTF of each images. Interestingly, all the images yielded the same full width at half maximum (FWHM) of line profile as 2.6 lp/cm, and no spatial (high contrast) resolution differences were observed between FBP and iDose<sup>4</sup> reconstructed images in this study. One of our on-going works is to provide complete investigation on the relation among all scanning parameters to noise level, and spatial and contrast resolution of varied reconstruction algorithms in order to provide guidance on the improvement of contouring accuracy of CT simulations in radiation therapy.

The results shown in Fig. 3 illustrate that, with the ability to reduce image noise while retaining boundary information, the iDose<sup>4</sup> reconstruction algorithm produces images with a superior CNR when compared with images generated using the FBP method. The results strongly suggested that the CNR was insufficient in distinguishing the performance of CT reconstruction algorithms for large-size patients. In these cases, task-based quality assessment, which was stated in the last paragraph of this section, would be required. Our experimental results also indicated that the use of iDose<sup>4</sup> had no obvious improvement of the clinical task performance (i.e., accuracy of contouring a structure) over FBP for patients with near or below 38 cm diameter (see Fig. 5) even with the lowest-dose level scanning protocol in this study ( $CTDI_{vol} = 3.3$  mGy). Yet, as the patient size increased, the image quality becomes more dependent upon the selected reconstruction algorithm, which directly influences image characteristics: noise and CNR. In these cases, the iDose<sup>4</sup> reconstruction can provide more clear appearance of anatomical structures through noise reduction and CNR improvement. Overall, the iDose<sup>4</sup> algorithm provided more accurate structure delineation, compared to the FBP method, for the majority of scan protocols and patient sizes in this study.

In addition, the results shown in Fig. 5 illustrate that an important clinical suggestion can be made based on the proposed task performance-based metrics rather than the standard image quality-based metrics such as CNR. For small sized patients it may not be necessary to increase the  $CTDI_{vol}$  in order to achieve improved TCA, even though the CNR may improve, because the TCA is sufficient for a radiotherapy treatment; in that case a radiation dose increase is very likely to cause unnecessary radiation exposure. However, a higher  $CTDI_{vol}$  may be needed for larger size patients in order to achieve 100% TCA.

As shown in Fig. 6, the iDose<sup>4</sup> reconstruction with noise reduction level 6 may misrepresent delicate signs of pathology and target tissues, from the perception of the oncologists. According to our investigation, this was the major reason the iDose<sup>4</sup> with noise reduction level 6 was not recommended for use in the clinic. This qualitative study reinforces the assertion that standard image quality characteristics, including the noise and contrast, should not be used as the sole criteria for assessing image quality when considering radiotherapy tasks.

While further considering the effect of CT scan parameters on image quality, the photon starvation artifacts that are caused by a deficit of X-ray photons in certain areas cannot always be eliminated by increasing the radiation dose, especially for larger sized patients. As shown in Fig. 2(d), for a 38 cm diameter phantom, the image noise can be effectively reduced by increasing

the radiation dose from 16.7 mGy to 34 mGy. However, it is impractical to do so for a 58 cm size phantom (and unsafe for patients) because of the effects of photon starvation artifact. In particular, given a low  $\text{CTDI}_{\text{vol}}$  (e.g., 16.7, 32.7, and 65.4 mGy in this study), the increase of the structure contouring confidence achieved by the iDose<sup>4</sup> algorithm was very limited for larger sized patient data (i.e., 48 cm diameter or greater in this study). In addition, as shown in Fig. 2(d), the noise level from the 53 cm and 58 cm phantom sizes is not monotonically decreasing. The reason is also due to the irregular photon starvation artifact streaks existing in the images. Low-dose scanning protocols that inherently cause severe photon starvation artifacts should be avoided in CT simulations, especially of larger sized patients. These results also demonstrated that both reconstruction algorithms cannot physically resolve the photon starvation artifacts, and therefore the radiation dose needs to be increased in the scanning protocol of larger sized patients in order to achieve desirable image quality for contouring purpose.

In order to further analyze the change in CT number and its affect on dose calculation, we performed the quantitative comparison to ascertain whether the treatment plans generated using different radiation dose scans and reconstruction algorithms were clinically equivalent.<sup>(11,12)</sup> For the 53 cm sized phantom, five pairs of different dose images (tube potential: 120 kVp;  $\text{CTDI}_{\text{vol}}$ : 32.7, 65.4, 98.1, 130.8, and 168.0 mGy) were reconstructed through the FBP and the iDose<sup>4</sup> L3 algorithms, respectively. The prostate, bladder, and rectum contours in the reference image (in Material & Methods section E) were firstly mapped to each image dataset. The clinical IMRT treatment plans for prostate cancer were optimized based on a consistent prescription. It was originally determined based on the reference image and reference contours, and recomputed on these five image pairs. The  $\gamma$  dose distribution comparison tool<sup>(18,19)</sup> was employed to compute the associated dose distributions. Dose areas with a  $\gamma$  value greater than 0.95 with a 3%/3 mm DTA criterion were considered clinically equivalent. After comparing the reference and test treatment plans, all plans achieved a passing rate of  $\geq 99.9\%$ . Thus, while different radiation doses and reconstruction algorithms produced images with varying CT number, this variation did not significantly alter the calculated dose distribution.

The dosimetry study explained above indicates that the dose distribution differences were mostly within 1% (at most 3%) of the prescription dose, and were therefore considered as dosimetric equivalent for these five CT image datasets acquired with different dose-level protocols. This finding can be explained by the Compton interaction of megavoltage therapeutic beams with human body tissues, and was described in one of our previous studies.<sup>(20)</sup> As well-known in radiation treatment planning, the attenuation of a megavoltage beam in water is less than 3% per cm penetration, which is much less than that of a kilovoltage beam.<sup>(21)</sup> In this study, for large sized phantom of 53 cm and 58 cm, the CT number differences on images acquired with the scanning protocols of moderate dose ( $\text{CTDI}_{\text{vol}}$  of 65.4 mGy) and the higher dose protocol ( $\text{CTDI}_{\text{vol}}$  of 216.9 mGy) was less than 100 HU, which translates to a 1 mm uncertainty per cm depth. Without loss of generality, we can assume that the effective prostate region has a size of less than 5 cm, the uncertainty in depth determination is therefore less than 0.5 cm, and it results in a dose calculation uncertainty of  $3\%/\text{cm} \times 0.5 \text{ cm} = 1.5\%$  at most.<sup>(20)</sup> As such, the dose distributions are considered as dosimetric-equivalent no matter which CT image dataset was used for dose computation as long as the same contour was used. As shown in Fig. 4, there was an improvement in CT Hounsfield number consistency when using iDose<sup>4</sup> reconstruction algorithm with moderate radiation dose-based scanning protocol. However, as shown in Fig. 7, iDose<sup>4</sup> reconstruction cannot reduce photon starvation artifacts; very-low-dose scanning protocols should be avoided for both FBP and iDose<sup>4</sup> reconstructions to eliminate or reduce photon starvation artifacts and improve the contouring accuracy. The moderate dose level scanning protocol should be used and adjusted based on patient sizes for CT simulations in radiation therapy.

This study has shown that the iDose<sup>4</sup> reconstruction algorithm, as compared with the FBP reconstruction algorithm, has the potential to produce CT images which improve radiotherapy task performance. It can reduce noise while maintaining boundary information, maintain CT

Hounsfield number consistency, and increase anatomical structure conspicuity. However, the accuracy of structure contouring could decrease when the iDose<sup>4</sup> noise reduction level increased from 3 to 6, which may lead to a distortion in the image texture of the target. This strongly suggested that improvement of standard image quality characteristics (e.g., the noise and CNR in this study) does not necessarily bring clinical benefits for radiotherapy tasks. As such, either the noise or CNR should not be used as sole criteria for guiding the usage of iDose<sup>4</sup> reconstruction for CT simulations; instead, the scanning protocol and certain level of anatomical conspicuity (or realism) that fulfills radiotherapy task requirements must be jointly considered.<sup>(22)</sup> On the other hand, for the patient cases analyzed in this study, the radiation dose was chosen by means of the automatic tube current modulation. This comparative study suggested that, even if the scanning dose is reduced to a certain level (depending upon the patient size), as long as the iDose<sup>4</sup> algorithm is employed to keep the image quality, the radiotherapy tasks could still be accomplished. Our future research will focus on the investigation of task-based image quality assessment frameworks so as to provide an improved quantitative image selection platform for radiotherapy applications.

## V. CONCLUSIONS

This study evaluated a commercially available Philips iDose<sup>4</sup> iterative reconstruction technique in radiotherapy environment and compared it to the traditional filter back-projection (FBP) reconstruction. The iDose<sup>4</sup> algorithm demonstrated a superior ability over the FBP algorithm in reducing image noise, improving the CNR, maintaining the CT number consistency, and improving target delineation accuracy for both energy levels. We suggest that the iDose<sup>4</sup> algorithm can be used safely and effectively for CT simulation and the subsequent treatment planning. However, the usage of iDose<sup>4</sup> reconstruction algorithm should be determined by the radiotherapy tasks (e.g., target contour conspicuity and CT number consistency for accurate treatment planning) accompanied with patient-specific factors (e.g., body size), instead of the common image quality metrics only, such as noise, contrast, and/or CNR. Particularly, our experiments suggested that a moderate noise reduction level of iDose<sup>4</sup> reconstruction was preferred for the radiotherapy tasks, as an overly high noise reduction level is most likely to cause unrealistically smooth CT images with mosaic artifacts.

## ACKNOWLEDGMENTS

We would like to especially thank Dr. Lifeng Yu from the Mayo Clinic for his helpful discussion and insight regarding task-based image quality assessment.

## COPYRIGHT

This work is licensed under a [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/).

## REFERENCES

1. Hsieh J. Computed tomography: principles, design, artifacts, and recent advances, 2nd edition. Hoboken, NJ: Wiley; 2009.
2. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. *Phys Med*. 2012;28(2):94–108.
3. Scibelli A. iDose4 iterative reconstruction technique. Amsterdam: Philips Healthcare System; 2012. Available from: <http://acceptance.netforum.healthcare.philips.com/Clinical/global/Explore/White-Papers/CT/iDose4-iterative-reconstruction-technique>

4. Purdy JA. Intensity-modulated radiation therapy. *Int J Radiat Oncol Biol Phys.* 1996;35(4):845–46.
5. Ma L, Geis PB, Boyer AL. Quality assurance for dynamic multileaf collimator modulated fields using a fast beam imaging system. *Med Phys.* 1997;24(8):1213–20.
6. Boyer AL and Yu CX. Intensity-modulated radiation therapy with dynamic multileaf collimators. *Semin Radiat Oncol.* 1999;9(1):48–59.
7. Michalski JM, Purdy JA, Winter K, et al. Preliminary report of toxicity following 3D radiation therapy for prostate cancer on 3DOG/RTOG 9406. *Int J Radiat Oncol Biol Phys.* 2000;46(2):391–402.
8. Perez CA, Michalski JM, Purdy JA, Lockett MA. New trends in prostatic cancer research. Three-dimensional conformal radiation therapy (3-D CRT), brachytherapy, and new therapeutic modalities. *Rays.* 2000;25(3):331–43.
9. Arapakis I, Efsthathopoulos E, Tsitsia V, et al. Using iDose(4) iterative reconstruction algorithm in adults' chest–abdomen–pelvis CT examinations: effect on image quality in relation to patient radiation exposure. *Br J Radiol.* 2014;87(1036).
10. Laqmani A, Buhk JH, Henes FO, et al. Impact of a 4th generation iterative reconstruction technique on image quality in low-dose computed tomography of the chest in immunocompromised patients. *RoFo.* 2013;185(8):749–57.
11. Olsson ML and Norrgren K. An investigation of the iterative reconstruction method iDose<sup>4</sup> on a Philips CT Brilliance 64 using a Catphan 600 phantom. *Proc. SPIE 8313 Medical Imaging.* 2012.
12. Caucci L and Barrett HH. Objective assessment of image quality. V. Photon-counting detectors and list-mode data. *J Opt Soc Am A Opt Image Sci Vis.* 2012;29(6):1003–16.
13. Barrett HH and Myers K. *Foundations of image science.* Hoboken, NJ: Wiley-Interscience; 2004.
14. Yu L, Bruesewitz MR, Thomas KB, Fletcher JG, Kofler JM, McCollough CH. Optimal tube potential for radiation dose reduction in pediatric CT: principles, clinical implementations, and pitfalls. *Radiographics.* 2011;31(3):835–48.
15. Yu L, Li H, Fletcher JG, McCollough CH. Automatic selection of tube potential for radiation dose reduction in CT: a general strategy. *Med Phys.* 2010;37(1):234–43.
16. Fu S, Ruan Q, Wang W, Chen J. Feature-oriented fuzzy shock-diffusion equation for adaptive image resolution enhancement. *Lect Notes Comput Sci.* 2006;4223:874–82.
17. Shanmugam K and Wahidabanu RS. Condensed anisotropic diffusion for speckle reduction and enhancement in ultrasonography. *Eurasip J Image Video Process.* 2012;12.
18. Low DA and Dempsey JF. Evaluation of the gamma dose distribution comparison method. *Med Phys.* 2003;30(9):2455–64.
19. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys.* 1998;25(5):656–61.
20. Li H, Noel C, Chen H, et al. Clinical evaluation of a commercial orthopedic metal artifact reduction tool for CT simulations in radiation therapy. *Med Phys.* 2012;39(12): 7507–17.
21. Khan FM. *The physics of radiation therapy.* Philadelphia, PA: Lippincott Williams Wilkins; 2009.
22. Li H, Yu L, Anastasio MA, et al. Automatic CT simulation optimization for radiation therapy: a general strategy. *Med Phys.* 2014;41(3):031913.