CT for evaluation of urolithiasis: image quality of ultralow dose CT with model-based iterative reconstruction and diagnostic performance of low dose CT with statistical iterative reconstruction

Poster No.: C-0782
Congress: ECR 2015
Type: Scientific Exhibit
Keywords: Abdomen, CT, Radiation safety
DOI: 10.1594/ecr2015/C-0782

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method is strictly prohibited.

You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.

www.myESR.org
Aims and objectives

PURPOSE: To compare radiation dose and image quality in regular, low, and ultralow dose CT protocols, and to evaluate diagnostic performance of low dose CT for urolithiasis.

Introduction

Urolithiasis is a common health problem, and its estimated lifetime risk is 5-10% in Europe and the United States [1-3]. An increase in incidence and prevalence of the disease has been reported worldwide in both children and adults [4-7]. For these patients, multi-detector computed tomography (MDCT) provides a rapid and accurate examination with high sensitivity, specificity and accuracy (100%, 96% and 98%, respectively) [8]. Furthermore, because of the low detection rate of other modalities including intravenous pyelography or ultrasonography, MDCT has emerged as the frontline diagnostic modality [9, 10].

As MDCT usage becomes more widespread, the importance of radiation exposure is becoming a concern. There were considerable increases of total effective radiation doses and the proportion of medical sources in the past decades. Among the man-made sources, MDCT is the largest proportion, accounting for about one-half [11]. Although debatable, there are increasing concerns that even a single typical abdomen CT examination may increase the risk of carcinogenesis [12]. Because of multiple MDCT examinations during treatment and the relatively high recurrence rate of urolithiasis, the patient's lifetime cumulative radiation exposure can be higher, especially in young patients. For these reasons, many investigators have made an effort to decrease the effective dose of MDCT for urolithiasis, using a low-dose approach that results in an estimated dose from 0.7 to 4.2 millisieverts (mSv) [13]. Lowering the effective dose decreases the quality of the image due to the excessive image noise, potentially causing a decrease in diagnostic performance and confidence. Therefore, the as low as reasonably achievable (ALARA) principle is used to achieve the lowest radiation dose possible while maintaining an optimal image quality in MDCT examinations [14].

The iterative reconstruction (IR) algorithms have been introduced and many studies have reported improved image quality and radiation dose reduction while maintaining diagnostic performance compared to filtered back-projection (FBP) [15]. There are several statistical IR algorithms commercially available and a few knowledge or model-based IR algorithms have investigated and introduced more recently. Due to the differences in each algorithm and the reconstructed images, there are different advantages and disadvantages.
The purpose of this study was to compare the radiation dose, and the objective and subjective image qualities in regular dose, low dose, and ultralow dose MDCT protocols reconstructed by FBP, statistical IR and knowledge-based IR. We also evaluated the differences in diagnostic performance between regular dose and low dose protocols for urolithiasis.
Methods and materials

This study was approved by the institutional review board and informed consent was obtained from all of the patients after providing the study’s details, including information on the additional radiation dose.

Patient Population and Study Design

Between February and May 2013, 336 consecutive patients who were suspected of having urolithiasis and scheduled to undergo non-enhanced CT were enrolled in this study. Instead of using specific eligibility criteria, we relied on the assessments carried out by the outpatient or emergency room physicians that led to the clinical suspicion of urolithiasis. Before the exam, we informed the patients about this prospective study and acquired consent. Patients who did not want to undergo additional scans or didn’t have urolithiasis were excluded (Figure 1).

CT Protocol

All studies were performed with a 256-MDCT scanner (Brilliance iCT, Philips Healthcare, Cleveland, OH, USA). First, all patients underwent a regular dose (RD) protocol scan from the proximal aspect of the T12 vertebra to the distal aspect of symphysis pubis in the supine position. If the patient had a urolithiasis in the RD scan and consented to this study, additional scans using the low dose (LD) and ultralow dose (ULD) protocols were performed. The scan range of LD protocol was the same as the RD protocol, but that of the ULD protocol was limited to the area where the urinary stones were expected. To minimize positional differences, the time between each scan was kept to a minimum. The RD and LD protocols were acquired at manually fixed peak tube voltage of 120 kVp and 100 kVp, respectively, with automated Z-axis dose modulation by the scout image (DoseRight, Philips Healthcare, Cleveland, OH, USA). The maximal tube current of the RD and LD protocols was limited to 150 mAs and 100 mAs, respectively. The ULD protocol was acquired at manually fixed peak tube voltage and tube current of 100 kVp and 20 mAs. The remaining scanning parameters were as follows: detector configuration, 128x0.625; pitch, 0.915; beam collimation, 80 mm; rotation time, 0.4 sec; and helical acquisition.
**Image Reconstruction**

The raw data of RD protocol was reconstructed into axial and coronal images using the FBP algorithm and that of LD protocol was reconstructed using iDose level 5 (Philips Medical Systems, Best, the Netherlands) which is a commercial statistical IR algorithm. The raw data of the ULD protocol was reconstructed to another three axial images using FBP, iDose level 5, and iterative model reconstruction (IMR) Soft Tissue level 3 which is a pre-released knowledge-based IR prototype system. All reconstructed section thicknesses and intervals were 3mm. Thus five reconstructed imaging data sets were ultimately obtained: RD-FBP, LD-iDose, ULD-FBP, ULD-iDose and ULD-IMR (Figure 2).

**Radiation Dose**

To evaluate the radiation dose, the effective radiation doses of each protocol were calculated as mSv from the recorded dose-length product (DLP) reported by the CT scanner with the conversion factor (0.015) \[16\]. The effective radiation dose of a full scan range in the ULD protocol was assumed by the number of slices compared to the RD and LD scans because of the limited scan range in its protocol. The reduction in radiation dose was compared between the three scans.

**Image Quality Assessment**

*Objective image noise assessment*

During the objective noise measurement, the radiology resident (one of the two reviewers of diagnostic performance) recorded the standard deviation (SD) of Hounsfield units (HU) by placing a circular region of interest (ROI) of 80-110 mm\(^2\) in the skeletal muscle and the subcutaneous fat area at the same level where the largest urinary stone was present. The objective image noise was reflected by the SD of attenuation in the circular ROI.

*Subjective image assessment*

To evaluate the subjective image, the staff radiologist (one of the two reviewers of diagnostic performance) who was blinded to the detailed technical scanning parameters used and subjectively graded the image's quality, noise, and the reviewer's confidence in each image set according to the previously reported 3 or 5 point scale \[13\]. All of
the images were displayed in a random fashion and the reviewer was permitted to change the window level and width. Consequently, the subjective image quality, with respect to the depiction of urinary stones, was rated on a 5-point scale (1 = poor, not diagnostically acceptable for interpretation; 2 = suboptimal, worse than acceptable quality; 3 = acceptable, diagnostic interpretation possible; 4 = good; and 5 = excellent). Subjective image noise was rated on a 3-point scale using the graininess of the image or pixel to pixel variation (1 = minimal; 2 = acceptable; 3 = excessive, rendering diagnostic interpretation impossible). In addition, the confidence level of radiologist with regard to the stone depiction was rated on a 3-point scale (1 = no confidence; 2 = confidence with reservations; and 3 = highly confident).

Diagnostic Performance and Inter-observer Agreement

All of the reconstructed image data sets were transmitted to the picture archiving and communication system (PACS) for image analysis. A consensus panel was formed to establish the reference diagnosis of stones based on RD-FBP and consisted of two abdominal staff radiologists who did not evaluate the diagnostic performance in the LD images. For the diagnostic performance, two reviewers (a radiology resident, 3rd year and a staff radiologist with 12 years of experience) interpreted the number, location, and size (< 3mm or larger) of the urolithiasis independently in LD image sets. The concordant rate between RD and LD image sets was calculated according to the diagnostic performance of both reviewers, and the inter-observer agreement was evaluated. All comparisons were performed after dividing the stones into 3 groups; all sizes, # 3 mm, and < 3 mm.

Statistical Analysis

Statistical analysis was performed using IBM SPSS 20 (IBM Software Inc.). All the recorded data was presented as mean ± SD. For the radiation dose and objective image noise, a paired-t test was performed to compare each image set. The subjective image quality was compared by using the Wilcoxon signed rank test. Inter-observer agreement between the radiology resident and staff radiologist for urolithiasis in LD image set was calculated using Cohen’s kappa values and all kappa values were interpreted according to following: poor for a k value of 0.19 or lower; fair for a k value of 0.20-0.39; moderate for a k value of 0.40-0.59; substantial for a k value of 0.60-0.79; and almost perfect for a k value of 0.80-1.00. The p value < 0.05 was considered to show a statistically significant difference.
Fig. 1: Flow chart illustrating the study design for patient selection.

© Radiology, Chung-Ang University Hospital - Seoul/KR
Fig. 2: Five reconstructed axial image sets of a 40-year-old man with a left ureterovesical junction stone. The effective radiation doses of each protocol were 6.0 mSv (RD), 1.4 mSv (LD) and 0.8 mSv (assumed ULD). In contrast to the FBP or iDose, IMR shows a less noisy image and the objective image noises on fat were as followings: 13.97 (RD, A), 25.21 (LD, B), 110.90 (ULD-FBP, C), 32.21 (ULD-iDose, D), and 7.42 (ULD-IMR, E).

© Radiology, Chung-Ang University Hospital - Seoul/KR
Results

Patient Population

Ultimately, 46 men and 19 women (65 total patients) were included in this study. The mean age of patients was 49.6 years and the range of the age was 19-83 years. The mean body mass index (BMI) was 24.3 kg/m$^2$ and the range of BMI was 16.2 to 36.4.

Radiation Dose

The effective radiation doses of each protocol were 6.29 mSv ± 1.76 (RD), 1.48 mSv ± 0.44 (LD), and the assumed effective dose of the full scan range (same range from the other two protocols) in the ULD protocol was 0.64 mSv ± 0.12 (ULD) (Figure 3). There were statistically significant differences (p < 0.001) between the effective doses of each protocol. The estimated dose reductions compared to RD were 76.4 % in LD and 89.8% in the assumed ULD.

Image Quality Assessment

Objective image noise assessment

The objective image noise was significantly lower in the ULD-IMR set (both measured on muscle and fat) than in any other image sets, including RD. The other results were significantly different from each other even between the RD and LD image sets (Table 1 and Figure 4).

Subjective image assessment

The subjective image assessments are summarized in Table 2 and Figure 5. There was no statistically significant difference in the subjective noise and diagnostic confidence between the RD and LD image sets, but the quality, noise, and diagnostic confidence of ULD-IMR were lower than RD and LD (p < 0.05). In case of noise, the scores of ULD-IMR, LD-iDOSE and RD-FBP were 4.2, 4.8 and 5 respectively (Good - Excellent). Among the three image sets reconstructed from the ULD protocol, IMR (ULD-IMR) showed better
subjective quality, noise, and diagnostic confidence than iDose (ULD-iDose) and FBP (ULD-FBP).

Diagnostic Performance and Interobserver Agreement

A total of 178 stones, including 16 right ureter stones (9.0%), 29 left ureter stones (16.3%), 58 right kidney stones (32.6%), and 75 left kidney stones (42.1%) were diagnosed in the RD image sets as a standard reference. Among the stones, 98 (55.1%) were smaller than 3mm. The concordant rates of the two reviewers examining the LD images compared to the standard reference are summarized in Table 3. The overall concordant rates of two reviewers were 88.2% and 84.3%, respectively. In cases of stones ≤ 3mm, the staff radiologist detected 79 of 80 (98.8%) and the resident radiologist detected 78 of 80 (97.5%). However, in cases of stones < 3mm, the staff radiologist detected 78 of 98 (79.6%) and the resident radiologist detected only 72 of 98 (73.5%) (Figure 6). For each location, the staff radiologist and radiology resident achieved concordant rates in the ureter of 97.8% and 91.1%, respectively (Figure 7) and in kidney of 85.0% and 81.2%, respectively. For the ureter, both reviewers had 100% concordant rates in case of stones ≤ 3mm. There were four false positive cases of the staff radiologist's results that were all very tiny or subtle. The inter-observer agreement was substantial (kappa value = 0.61).
Fig. 3: Graph shows effective doses in each protocol. In contrast to the RD protocols, LD and ULD protocols showed statistically significant low doses and the estimated dose reductions were 76.4% (LD) and 89.8% (assumed ULD).

© Radiology, Chung-Ang University Hospital - Seoul/KR
Fig. 4: Graph shows the mean objective image noise of each image set. ULD-IMR showed a statistically significant lower objective image noise than any other image set in both muscle and fat.

Table 1: Objective image noise of each image set.

<table>
<thead>
<tr>
<th></th>
<th>RD</th>
<th>LD</th>
<th>ULD-FBP</th>
<th>ULD-iDose</th>
<th>ULD-IMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muscle</td>
<td>17.7±3.5</td>
<td>24.6±4.2</td>
<td>99.2±23.9</td>
<td>37.6±7.2</td>
<td>10.5±3.7</td>
</tr>
<tr>
<td>Fat</td>
<td>15.8±2.8</td>
<td>21.5±5.2</td>
<td>84.6±21.8</td>
<td>33.7±7.6</td>
<td>11.1±7.0</td>
</tr>
</tbody>
</table>

Mean SD of HU ± SD
Table 2: Subjective image assessment scores of each image set

© Radiology, Chung-Ang University Hospital - Seoul/KR

<table>
<thead>
<tr>
<th></th>
<th>RD</th>
<th>LD</th>
<th>ULD-FBP</th>
<th>ULD-iDose</th>
<th>ULD-IMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality</td>
<td>5.0±0.1</td>
<td>4.8±0.4</td>
<td>1.4±0.5</td>
<td>3.0±0.4</td>
<td>4.2±0.5</td>
</tr>
<tr>
<td>Noise</td>
<td>1.0±0.0</td>
<td>1.0±0.1</td>
<td>3.0±0.2</td>
<td>2.0±0.3</td>
<td>1.3±0.5</td>
</tr>
<tr>
<td>Confidence</td>
<td>3.0±0.0</td>
<td>3.0±0.1</td>
<td>1.2±0.4</td>
<td>2.1±0.3</td>
<td>2.7±0.5</td>
</tr>
</tbody>
</table>

Quality scale: 1, poor; 2, suboptimal; 3, acceptable; 4, good; 5, excellent
Noise scale: 1, minimal; 2, acceptable; 3, excessive and impossible interpretation
Diagnostic Confidence scale: 1, no confidence; 2, confidence with reservation; 3, highly confidence

Fig. 5: Graph shows the mean subjective image assessment scores of each image set. There were statistically significant differences between the subjective image scores (quality, noise, and diagnostic confidence) of ULD-IMR and RD or LD image sets. However, IMR showed a better image in the same ULD protocol than FBP or iDose. There was no statistically significant difference in the subjective noise and diagnostic confidence between the RD and LD image sets.
Fig. 6: Axial CT images of 34-year-old man suspected of having urolithiasis. Consensus panel interpreted that a true stone less than 3mm in the left kidney on the RD image. The staff radiologist detected the stone but the radiology resident missed it in an LD image.

© Radiology, Chung-Ang University Hospital - Seoul/KR

Fig. 7: Axial CT images of 66-year-old woman suspected of having urolithiasis. Consensus panel interpreted that multiple stones, including a true stone of less than 3mm in the left mid ureter in a RD image. Both the staff radiologist and the radiology resident detected this lesion on the LD image.

© Radiology, Chung-Ang University Hospital - Seoul/KR
Table 3: Concordant rates of LD images compared to RD images as standard references

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>≥3mm</th>
<th>&lt;3mm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Staff (%)</td>
<td>Resident (%)</td>
</tr>
<tr>
<td>Ureter</td>
<td>45</td>
<td>44 (97.8)</td>
<td>41 (91.1)</td>
</tr>
<tr>
<td>Rt. Ureter</td>
<td>16</td>
<td>16 (100)</td>
<td>15 (93.8)</td>
</tr>
<tr>
<td>Lt. Ureter</td>
<td>29</td>
<td>28 (96.6)</td>
<td>26 (89.7)</td>
</tr>
<tr>
<td>Kidney</td>
<td>133</td>
<td>113 (85.0)</td>
<td>108 (81.2)</td>
</tr>
<tr>
<td>Rt. Kidney</td>
<td>58</td>
<td>46 (79.3)</td>
<td>45 (77.6)</td>
</tr>
<tr>
<td>Lt. Kidney</td>
<td>75</td>
<td>67 (89.3)</td>
<td>63 (84.0)</td>
</tr>
<tr>
<td>Overall</td>
<td>178</td>
<td>157 (88.2)</td>
<td>150 (84.3)</td>
</tr>
</tbody>
</table>

© Radiology, Chung-Ang University Hospital - Seoul/KR
Conclusion

CT is the best and the first line imaging examination for urolithiasis. As increased fear of the potential cancer risk of radiation exposure several studies to reduce the effective dose in urolithiasis have been published, and they reported an effective dose ranging from 0.7 to 2.8 mSv [17, 18]. Our study showed an average of 1.48 mSv for an effective dose in the LD protocol, similar to previous reports. Despite the low effective dose, a previous meta-analysis study of low-dose urolithiasis CT showed high sensitivity and specificity [18]. The effective dose of intravenous pyelography is reported to range between 0.7 to 3.7 mSv, but the diagnostic performance was significantly lower than CT [10, 19]. Thus, LD protocols have strengths including a higher diagnostic performance and no need for contrast material in similar effective dose. For the ULD protocols, the assumed effective dose was 0.64 mSv and it was very low, even comparable to conventional single plain radiography of the kidneys, ureters and bladder [19]. During the diagnosis and treatment of urolithiasis, the average of four imaging studies including 1.7 CT is performed in the year after initial event [20]. Moreover, there have been increasing trends showing a 1.4 % increase in incidence and a relatively high recurrence rate for urolithiasis, so increased cumulative doses can raise the risk of cancer in the patient with urolithiasis [3, 7, 17]. Thus, LD or ULD protocols can potentially contribute to reduce lifetime cancer risk.

The beam energy is determined by the peak kilovoltage of the tube and has an exponential relationship to the effective radiation dose. The kilovoltage of the tube has a complex effect on image noise and tissue contrast, so maintaining constant image quality is one of the purposes in radiation dose reduction [14]. Several ways to obtain better images have been researched and applied in current clinical studies. In these manners, lowering the peak kilovoltage, increasing the milliampere-second, and using automated tube current modulation are regarded as the best strategies for radiation dose reduction and several studies have reported [14, 18]. Furthermore, depending on the development of model-based IR algorithms, several investigators have published reports using the algorithms and they found it to be useful in image quality with more effective dose reduction [15, 21-24]. However, it is more time consuming during the reconstruction phase, taking between 10 and 90 minutes depending on the number of slices, making model-based IR to inappropriate for urgent examinations [25, 26]. Nevertheless, new computer technique has reached a tolerable rate of speed using knowledge-based IR [27].

The image reconstructed with knowledge or model-based IR shows an improvement in spatial resolution and image noise especially in low-dose or ultralow-dose scans, but the scans appear differently, often described as "waxy" or "plastic" as compared with FBP [26]. In our study, IMR represents the best objective image noise, despite of ULD protocols like other previous reports. Though IMR did not show a comparable assessment
to the RD or LD protocols, statistically significant improvements were seen compared to FBP and iDose using ULD protocol and the subjective assessment scores were still higher. Based on this, it can provide better image quality in terms of subjective image quality, noise, and diagnostic confidence with proper dose modulation.

In the diagnostic performance of LD CT for urolithiasis, our study showed overall concordant rates of 88.2% and 84.3% (staff radiologist and radiology resident, respectively) compared to standard reference. Although relatively lower concordant rates of 79.6% and 73.5% in case of renal stones < 3mm, additional diagnostic or therapeutic procedures may not be required clinically because stones < 5mm pass spontaneously up to 98% of the time [1]. For the ureter stones, the overall concordant rates were 97.8% and 91.1%, but in cases of stones # 3mm, both the staff radiologist and the radiology resident achieved a concordant rate of 100%. In cases of ureter stones < 3mm, the concordant rate of the radiology resident was substantial of 63.6%. However, because of the lower clinical significance in stones < 5mm, LD image can provide an image strong enough for good diagnostic performance even in less-experienced interpreters [1].

There were several limitations in our study. First, additional radiation was required for the study, which can be an ethical problem. However, the ULD protocol scan was only performed over the a limited range, and the actual total effective radiation dose of each patient (ranged 4.3 to 15.9 mSv) did not exceed the range of regular dose CT protocols previously reported [18]. Second, a reference standard was not confirmed with the extraction or excretion of stones. In particular, it is debatable whether subtle and small stones in RD are true stones or not. However, we tried to reduce this limitation by using a consensus panel. Third, personal variation, such as BMI, which can vary the radiation dose or image quality, was not considered in this study. Moreover, the relatively small number of patients involved in this study within one institution may not be sufficient to show statistical differences. Although the reviewer was blinded, experienced staff radiologists may be able to distinguish between 5 image sets due to the difference in image texture. Thus, bias may be present in the subjective image assessment. Finally, LD and ULD acquisitions were performed only in cases of positive initial CT. Thus, reviewers that interpreted the LD series were also biased.

However, we believe that our study offers a number of positive features. The different CT protocols were performed at the same time to allow for the exact comparison of virtually identical stones, whereas most of the previous studies had time interval. Furthermore, our study compared RD using FBP reconstruction and LD using IR and evaluated diagnostic performance. There can be a corresponding diagnostic performance of ULD using IMR because of the lowest objective image noise.

In conclusion, knowledge-based IR can provide better objective image noise and subjective image assessment of a condition using the same radiation dose as other
reconstruction algorithms. Furthermore, LD protocols showed a comparable diagnostic performance to RD protocols. Although knowledge-based IR didn't show better image quality compared to RD or LD protocols, its potential strength can be expected in LD protocols. To decrease effective radiation dose, further studies are needed to confirm the diagnostic performance of ULD protocols using knowledge-based IR.
Personal information

1Department of Radiology, Chung-Ang University Hospital, Chung-Ang University College of Medicine, 102, Heukseok-ro, Dongjak-gu, Seoul, 156-755, Korea

Ji In Choi, MD, Sung Bin Park, MD, Joonho Hur, MD, Jong Beum Lee, MD,

Hyun Jeong Park, MD, Yang Soo Kim, MD
References


