Iterative Model Reconstruction (IMR) in MDCT Below 2 mSv for the Detection of Urinary Calculi: Diagnostic Accuracy and Image Quality in Comparison to Filtered Back-Projection and 4th Generation Iterative Reconstruction (iDose4)

Iterative Model Reconstruction (IMR) in der MDCT unter 2 mSv zur Detektion einer Urolithiasis: Diagnostische Genauigkeit und Bildqualität im Vergleich zur Filterrückprojektion und der iterativen Rekonstruktion der vierten Generation (iDose4)

Authors
Jakob Schmidt-Holtz¹, Azien Laqmani¹, Sebastian Butscheidt¹, Max Kurfürst¹, Maxim Avanesov¹, Cyrus Behzadi¹, Clemens Spink¹, Simon Veldhoen², Hans Dieter Nagel², Gerhard Adam¹, Marc Regier¹

Affiliations
1 Department for Diagnostic and Interventional Radiology, Endoscopy and Nuclear Medicine, University Hospital Hamburg Eppendorf, Hamburg, Germany
2 Department for Diagnostic and Interventional Radiology, University Medical Center Würzburg, Germany
3 Research & Technic for Radiology, Dr. HD Nagel, Buchholz, Germany

Key words
abdomen, image manipulation/reconstruction, CT, ureter, urinary tract, urolithiasis

received 22.03.2017
accepted 21.12.2017

Bibliography
DOI https://doi.org/10.1055/s-0044-100724
Published online: 2018
Fortschr Röntgenstr 2018; 190: 630–636
© Georg Thieme Verlag KG, Stuttgart · New York
ISSN 1438-9029

Correspondence
Jakob Schmidt-Holtz
Department of Diagnostic and Interventional Radiology, University Medical Center Hamburg-Eppendorf, Harvestehuder Weg 55, 20149 Hamburg, Germany
Tel.: ++ 49/0 17/24 06 55 49
jak.schmidt-holtz@uke.de

ZUSAMMENFASSUNG
Ziel
Das Ziel dieser Studie war es, die Auswirkungen der iterativen Modellrekonstruktion (IMR) auf die Sicherheit des Radiologen bei der Steindetektion und auf die Bildqualität im Vergleich zu gefilterten Rückprojektionen (FBP) und iDose Level 4 (iDose4) im Stein-MDCT mit Strahlendosen unter 2 mSv zu beurteilen.

Materialien und Methoden
Für 32 aufeinanderfolgende Patienten mit Verdacht auf ein Harnleiterkonkrement wurden die Rohdaten der nativen Scans (256 Slice MDCT, 120 kV, 40 Referenz mAs, mittleres CTDIvol: 2.7 ± 0.8 mGy, mittleres DLP: 126 ± 38 mGy × cm) unter Verwendung einer Prototypversion von IMR (Level 1–3), iDose4 (Level 4) und FBP mit einer Schichtdicke von 3 mm rekonstruiert. Die Bildanalyse wurde jeweils unabhängig von zwei geblindeten Radiologen durchgeführt. Die diagnostische Sicherheit des Radiologen bei der Steindetektion erfolgte auf der Grundlage einer 5-Punkte-Skala (1 – sicherer Ausschluss, 5 – Konkrement definitiv vorhanden) wie auch für die Auswertung der Bildqualität in Bezug auf die Darstellung von anatomischen Details (1 = mangelhaft; 5 = Ausgezeichnet). Ein klinischer Referenzstandard für die Steindetektion wurde nicht erhoben. Die statistische Auswertung beinhaltet die weighted kappa-analysis und den Wilcoxon-Test.

Ergebnisse
17 Nierenbecken- und Uretersteine wurden gefunden. 11 weitere Konkremente wurden im Harnblasenstrom oder der Harnblase selber lokalisiert. Bei der Anwendung von IMR wurde mit jeder Iterationsstufe eine deutliche Verbesserung der Bildqualität beobachtet (Mittelwert für FBP, 2,0, iDose4, 2,9, IMR L1, bis L3, 3,9, alle p < 0,001). Mit den höheren IMR-Stufen L2 und L3 wurde ein gewisses Maß an sogenanntem „blotchiness“ (Unscharfe) anatomischer Konturen beobachtet. Die Sicherheit des Radiologen wurde so unabhängig vom IMR-Level signifikant verbessert (sichere Steindetektion FBP, 69%; iDose4, 81%; IMR L1 bis L3, 95%, alle p > 0,001). Mit zunehmenden IMR-Werten wurde die Reduktion von Streifenartefakten durch eine Abnahme des Bildrauschens quantifiziert. Ein Verlust anatomischer Informationen wurde nicht beobachtet. Die Sensitivität der Steindetektion waren äquivalent für alle MDCTs, die...
Over the last decade, the number of MDCT scans performed for Schmidt-Holtz J et al. Iterative Model Reconstruction specificity (96 %) in the detection of urinary stone disease [1] its high diagnostic value by achieving a high sensitivity (98 %) and Multidetector computed tomography (MDCT) has already proven the portion of young patients affected is above-average [4]. With a lifetime-estimated recurrence rate of approximately over 50 % [3], patients suffering from urolithiasis are often repeatedly referred for cross-sectional imaging. The portion of young patients affected is above-average [4]. In combination with the growing accessibility and usage of MDCT in these cases, the scans have evolved to become an increasing part of population radiation exposure. As a result, development of low-dose protocols was needed. After broad implementation, these have proven to be efficient and further accepted as a standard procedure for the diagnosis of urolithiasis [5–7]. Early ethical concerns [8, 9] have been invalidated [10].

However, with this development growing image noise proved to be the trade-off for lowering radiation dose levels. Iterative reconstruction (IR) tools have been shown to be successful in compensating for this loss in image quality at the existing dose reduction levels in the past [11–13]. With ongoing improvement of IR tools and its ability to decrease even higher image noise ratios, the radiation dose reduction potential changes and needs to be re-evaluated each time.

Most recently, a new generation of IR tools called Iterative Model Reconstruction (IMR®; Philips Healthcare, Best, the Netherlands) was introduced. With the underlying target being to once again lower radiation dose exposure “as low as reasonably achievable” (ALARA), the new potential capability of IMR to allow for further lowering has to be evaluated.
Thus, the purpose of this study is to assess the impact of IMR on reader confidence with respect to stone detection and on image quality in comparison to FBP and iDose4 in abdominal MDCT with radiation doses below 2 mSv.

Materials and Methods

Patient collective

In this retrospective study the data from 32 consecutive patients (21 men; mean age: 53.5 years; range: 20 – 82 years) with suspected urinary stone disease were analyzed.

The study was reviewed and approved by the local ethics committee. Due to the retrospective nature of the study protocol, the ethical review board waived the need for written informed consent.

MDCT image acquisition

MDCT examinations were performed using a 256-slice MDCT scanner (Brilliance ICT, Philips Healthcare, Best, The Netherlands). All patients underwent non-contrast MDCT of the upper urinary tract performed in craniocaudal direction from the top of the kidneys to the bottom of the urinary bladder. As only a prototype version of IMR was available in our study, which was neither FDA- nor CE-cleared, regular diagnosis based on images reconstructed with IMR was not allowed. Hence, the dose settings that were appropriate for iDose, which was used for regular diagnosis, had to be applied, using the following parameters: tube voltage of 120 KV; reference tube current time product of 40 mAs, corresponding to a CTDIvol of 2.7 mGy for a standard-sized patient; gantry rotation time of 0.4s; pitch of 0.922; collimation of 64 × 1.25 mm. Automatic exposure control system (ACS) in combination with z-axis dose modulation (Z-DOM) was activated.

MDCT image reconstruction:

The raw data of all MDCTs were reconstructed with FBP, iDose4 (level 4) and three different IMR levels (IMR level 1 (IMR L1), level 2 (IMR L2) and level 3 (IMR L3)) allowing for an intraindividual comparison between all reconstruction methods.

IMR reconstruction was performed with the “Body Routine” setting. The different levels of IMR refer to the degree of noise reduction provided by the IMR software. In the case of FBP and iDose, the standard body kernel (B) was applied. A slice thickness of 3 mm with 50% overlap was used for all data sets. The recommended fixed setting “BodyRoutine” was used for IMR.

Subjective image evaluation

Multiplanar reconstructions were created for every examination. A commercially available PACS system (Centricity PACS-IW, GE Healthcare, Chicago, Illinois, United States) was used for image interpretation. Two radiologists with 5 and 10 years of abdominal MDCT experience independently evaluated all data sets in a blinded fashion.

All reconstructions were randomized and six separate review sessions were performed in order to achieve a maximum level of blinding.

The subjective image quality was rated based on a 5-point Likert scale defined as follows:
1. poor image quality, major noise, with poorly defined anatomic details.
2. reduced image quality, substantial noise, with limitations in anatomic details.
3. acceptable image quality, moderate noise.
4. good image quality, minor noise, with clear anatomic details.
5. excellent image quality, no noise, with distinct anatomic details.

Ratings were separately performed for the renal pelvis, the urinary bladder, and the proximal, middle and distal third of the ureteral course. The rating of anatomical structures was mainly influenced by the ability to clearly outline the displayed anatomical borders with particular regard to the urinary and renal tract.

For the determination of image appearance and the known blurring effect of IR techniques, the 5-point Likert scale targeted the “blottiness” of the entire MDCT data set influencing the detailed depiction of anatomical details only:
1. major blotchy appearance.
2. increased blotchy appearance.
3. moderate blotchy appearance.
4. minor blotchy appearance.
5. no blotchy appearance.

Reader confidence with respect to diagnosing urolithiasis was graded dividing the data into the sections identical to those performed for the analysis of the subjective image quality (renal pelvis, proximal, middle and distal ureteral third and the urinary bladder, respectively). Ratings were performed for both the left and the right side separately, applying the following definition of calculus score:
1. definitely no calculus
2. probably no calculus
3. indeterminate
4. probably calculus
5. definitely calculus

Finally, consensus readout was performed and the reference standard was defined. To determine the diagnostic value, definite stone detection or exclusion groups were determined. A relation to possible findings was set and a ‘confidence’ index was generated, which is defined as

\[
\text{confidence index} = \frac{\text{number of cases with calculus score 1 or 5}}{\text{total number of cases}}
\]

Thus, reader confidence with respect to diagnosis was evaluated. A subsequent correlation of the radiologist’s diagnosis and clinical findings of urolithiasis or the lack thereof was not carried out. Thus, a corresponding reference standard was not established.
Objective image evaluation

Radiation dose and noise levels

For every MDCT scan a dose report was generated. All relevant dose parameters including CT dose index volume (CTDIs), dose-length product (DLP) and tube voltage (KV) were documented.

For the quantitative analysis the Hounsfield Units (HU) and the standard deviations (SD) were determined by placing circular regions of interest at the following locations: the subcutaneous fat at the dorsolateral back; the right subcutaneous gluteal fat; the aortic lumen; the internal obturator muscle and the urinary bladder. These measurements were performed on three adjacent slices and mean values were obtained.

Statistical analysis

Continuous variables were expressed as mean ± SD. Categorical variables are presented as mean + 95% confidence interval. Inter-group comparisons between ≥ 3 related groups were assessed applying the Friedman-test. Post-hoc analysis between 2 related groups was performed using the paired Student’s t-test for normally distributed data and the Wilcoxon signed-rank test for skewed data samples. Interobserver agreement between the two reviewers was assessed by weighted kappa. Statistical significance was defined for 2-tailed p-values of less than 0.05. Data collection and statistical calculations were performed using SPSS Statistics 22 software (IBM Inc. SPSS Statistics, Chicago, IL).

Results

Subjective image quality

The subjective image quality results for each part of the urinary tract are displayed in Fig. 1. First, a mean score for the image quality of the upper urinary tract was generated for all segments. The scores substantially increased from FBP to iDose4 as well as from iDose4 to IMR. FBP scored an average of 2.0 points on the Likert scale. iDose4 accomplished an average score of 2.9. IMR level 1 was rated to perform with a substantial increase in the subjective image quality (4.2) and a decrease of image noise. With further noise reduction by the IMR software at the higher levels 2 and 3, the mean score for image quality slightly decreased to an average of 4.0 and 3.9, respectively.

Objective image quality

The CT numbers remained constant for each anatomical structure. No substantial differences between iteration tools or levels were found. No pattern of deviation could be related. The observed extent of variation ranged from 0.5% to 2.0%. The mean SD determined for all anatomical landmarks showed a considerable decrease when applying iDose4 and IMR in comparison to FBP. Using iDose4 with level 4, image noise was reduced by 45% on average. IMR L1 allowed for an even greater reduction of 70%. With IMR L3 the highest noise reduction of 79% was achieved. The constancy of the mean CT numbers as well as the reduction in mean SD were calculated for each anatomical structure as shown in Fig. 2.

Reader confidence with respect to stone detection

The results show an increase in reader confidence with respect to definite stone detection and exclusion from FBP to iDose4 and to IMR. FBP achieved a confidence index of 66%, iDose4 of 81% and IMR of 96% independent of the level (Fig. 3). Both radiologists reported a high influence of the extent of distinction of anatomical structures on their confidence regarding stone detection. This particularly affected evaluation in the area of the lower ureter and bladder.

A drop of only 2% with IMR level 1 and 3% with levels 2 and 3 from the upper to the lower urinary tract was observed. Subse-
quently, a nearly consistent high level of reliability with IMR level 1 (96%) and level 2 (97%) was shown. With FBP and iDose4, however, a drop of 26% (FBP) and 16% (iDose4) was shown.

**Image appearance and noise**

No blotchy appearance was found on the images reconstructed using FBP (mean value of 5). Alterations in the image appearance were noted only with the introduction of iDose4 or IMR. In iDose4 the blotchiness was assessed to be between minor and no blotchy appearance (mean value of 4.6). With further influence of the IMR reconstruction technique, increasing blotchiness was observed (mean values of 3.7 (IMR L1), 3.2 (IMR L2) and 2.8 (IMR L3)), resulting in a moderate blotchy image appearance.

**Radiation dose**

The mean effective tube current time product was 41 ± 12 mAs (range 19 – 80 mAs) with a fixed tube voltage of 120 KV. The mean CTDIvol was 2.7 ± 0.8 mGy (range 1.29 – 5.44 mGy). On average, the DLP amounted to 126 ± 38 mSv × cm (range from 49 – 253 mSv × cm). By applying the conversion factor of $k = 0.015 \text{ mSv/mGy} \times \text{cm}$ from AAPM report 96, this resulted in an overall mean effective dose of 1.9 ± 0.6 mSv.

**Discussion:**

MDCT has been shown to be highly reliable in the diagnosis of urinary stone disease [14]. However, the challenge to diminish radiation dose exposure remains. With the initial introduction of IR techniques, a continuous reduction of MDCT dose settings with a simultaneous positive influence on image quality was achieved in the past.

Several studies have evaluated the impact of IR on image quality parameters as well as on the reader’s diagnostic confidence with regard to pinpointing pathological findings. Not only the detection of urolithiasis but also various other types of CT examinations and findings were investigated [15 – 17]. With the very recent introduction of the next generation IR tool IMR, this study aimed to evaluate the changes of these effects in comparison to its predecessors FBP and iDose4.

Our results show that IMR reduces image noise and increases reader confidence with respect to stone detection to a large extent. This is reflected in substantially reduced noise levels (reduction of 70% to 80% in comparison to FBP and of 50% to 60% in comparison to iDose4 (level 4)) and improved reader confidence with regard to stone detection (from 66% with FBP to 81% with iDose4 and 96% with each IMR level (Fig. 3)). Thus, an increase of up to 30% from FBP to IMR and of 15% from iDose4 to IMR was observed, which represents a significant improvement.

However, with growing influence of the IMR algorithm in higher IMR levels, especially IMR level 3, a noticeable altering of image appearance can be observed. This alteration, in other studies already described as "blotchiness" or a "waxy" appearance, refers to blurring with a loss of sharpness of anatomical contours [14, 18]. This visual phenomenon accounts for the reduction of subjective image quality as well as the image appearance score despite an ongoing noise reduction from IMR level 1 to level 3 (Fig. 4).

The decrease in image appearance from IMR L1 to L3, however, had no negative impact on reader confidence regarding stone detection. This is due to the fact that the high density of urinary calculi allows easy identification in non-enhanced abdominal MDCT independent of the reconstruction technique (Fig. 5).

Especially in IMR L3, discrete alterations of anatomical details had no significant impact on reader confidence. Still it has to be taken into account that no clinical reference standard was defined. Instead, the problem reported by both readers was not the identification but in some cases the reliable assignment of possible calculus to part of the urinary tract. This problem was mainly caused by higher noise levels arising partly from lower noise reduction potential, especially by FBP, and partly from anatomical regions often related to higher noise levels and artifacts. Referring to the urinary tract, these regions are the middle and lower third of the ureter as well as the urinary bladder and its ostia (Fig. 1, 5). In many cases this problem is essential for differentiation between urolithiasis and nonspecific calcifications such as phleboliths or calcified lymph nodes.
In this study reader confidence regarding stone detection varied greatly from the upper urinary tract to the lower. This change in reader confidence, however, directly depends on the IR tool being used. With FBP the confidence index dropped by 23% (from 69% to 43%) and with iDose4 by 16% (from 84% to 68%) from the upper to the lower urinary tract, whereas a drop of only a few percent was observed with IMR (L1 from 96% to 94%, L2 and L3 from 97% to 94%). This leads to the conclusion that even though higher IMR levels alter image appearance and even very fine anatomical structures can potentially be lost [19], the diagnostic value for this specific task substantially improves. Nevertheless, the discussion of missing a possible differential diagnosis with altering image appearance as stated by Veldhoen et al. [14] remains.

However, CT examinations for urolithiasis especially concern young patients who often undergo multiple CT scans over the years [14]. The obligation to consider radiation dose-related cancer risks remains unchanged [4, 20, 21]. Based on our results, as well as similar results [15], the newly introduced radiation dose reduction potential offered by IMR has to be discussed.

Our study achieved an average effective dose of 1.9 mSv without altering the already implemented standard dose settings. In consideration of the significant improvements in reader confidence, a further dose reduction should be possible with IMR. This was shown by Park et al. as well as other studies [22–25], which achieved average effective dose levels as low as 0.68 mSv. Future studies need to evaluate the full dose reduction potential, taking the loss of a possible differential diagnosis and the alteration of image appearance into account. The trade-off needs to be evaluated and optimal dose settings have to be identified.

This study had the following limitations:

As already mentioned, the prototype version of IMR used in this study had no FDA or CE clearance. To ensure regular diagnosis without IMR, the impact of reductions below the level that is appropriate with iDose could not be investigated.

Based on the fact that a clinical correlation of the radiologist’s interpretations did not take place, the possible radiological diagnoses of urolithiasis could not be confirmed or disproved. Thus, a clinical reference standard is not defined.

Even though the radiologists were blinded, an ideal blinding could not be assured in all cases due to the visible differences between each iterative reconstruction technique (FBP, iDose4 and IMR), such as noise levels and sharpness of anatomical structures.

A further limitation was that possible differential diagnoses were not investigated. IMR proved to have a positive effect on reader confidence regarding the diagnosis or exclusion of urolithiasis. On the other hand, the loss of a possible differential diagnosis due to an alteration in image appearance by the IMR algorithm, especially in IMR L3, remains uncertain.

Furthermore, the size of possible stones and a final confirmation of determined stones were not established in this study.

In conclusion, IMR substantially exceeds the capability of its predecessors iDose4 and FBP to improve image noise levels as well as subjective image quality for non-enhanced low-dose MDCT of the urinary tract. Thus, on the basis of missing clinical correlation, considerably better subjective reader confidence was achieved.

The results of the presented study allude to a further potential to reduce radiation dose exceeding the dose settings applied. However, the extent of this potential needs to be clarified in future studies.

Conflict of Interest

The authors declare that they have no conflict of interest.

References


