Physics and Imaging in Radiation Oncology 33 (2025) 100743

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

Uncertainty estimation in female pelvic synthetic computed tomography generated from iterative reconstructed cone-beam computed tomography

Yvonne J.M. de Hond *, Paul M.A. van Haaren, Rob H.N. Tijssen, Coen W. Hurkmans

Department of Radiation Oncology, Catharina Hospital Eindhoven, the Netherlands

ARTICLE INFO

Keywords:
Iterative reconstruction
Cone-beam computed tomography
Synthetic CT
Robustness
Parameter variations
Aleatoric uncertainty

ABSTRACT

Background and Purpose: Iterative reconstruction (IR) can be used to improve cone-beam computed tomography (CBCT) image quality and from such iterative reconstructed (iCBCT) images, synthetic CT (sCT) images can be generated to enable accurate dose calculations. The aim of this study was to evaluate the uncertainty in generating sCT from iCBCT using vendor-supplied software for online adaptive radiotherapy.

Materials and Methods: Projection data from 20 female pelvic CBCTs were used to reconstruct iCBCT images. The process was repeated with 128 different IR parameter combinations. From these iCBCTs, sCTs were generated. Voxel value variation in the 128 iCBCT and 128 sCT images per patient was quantified by the standard deviation (STD). Additional sub-analysis was performed per parameter category.

Results: Generated sCTs had significantly higher maximum STD-values, median of 438 HU, compared to input iCBCT, median of 198 HU, indicating limited robustness to parameter changes. The highest STD-values of sCTs were within bone and soft-tissue compared to air. Variations in sCT numbers were parameter dependent. Scatter correction produced the highest variance in sCTs (median: 358 HU) despite no visible changes in iCBCTs, whereas total variation regularization resulted in the lowest variance in sCTs (median: 233 HU) despite increased iCBCT blurriness.

Conclusions: Variations in iCBCT reconstruction parameters affected the CT number representation in the sCT. The sCT variance depended on the parameter category, with subtle iCBCT changes leading to significant density alterations in sCT. Therefore, it is recommended to evaluate both iCBCT and sCT generation, especially when updating software or settings.

1. Introduction

In image guided radiotherapy, cone-beam computed tomography (CBCT) is often used for patient positioning and verification. The poor image quality of the CBCT, however, is limiting the usage of CBCT for online adaptive replanning. To improve the image quality, iterative reconstruction (IR) can be incorporated in commercial CBCT reconstruction packages [1,2]. Although the image quality is improved in iterative reconstructed CBCT (iCBCT) images, the iCBCT numbers are not calibrated Hounsfield units (HU). To adapt the values to HU and enable dose calculations, bulk density CT or synthetic CT (sCT) images may be generated from the iCBCT images [3–6]. For online replanning, calibrated CT numbers are required for dose calculations and the anatomy on the CBCT needs to be preserved on the sCT.

Although the sCT models often enhance the image quality and improve CBCT CT number accuracy, the anatomical preservation may not always be sufficient [7]. In addition to anatomical preservation differences in sCT generation models, density differences could occur due to uncertainties in the sCT generation model [8]. In addressing deep learning (DL)-model uncertainties, methods such as Laplace approximation, Monte Carlo dropout, and network output discrepancies are commonly used [8–13]. These methods are incorporated into the model and, therefore, not applicable for retrospective evaluation of fully trained, vendor-supplied models.

To be able to retrospectively evaluate fully trained sCT models, aleatoric uncertainty can be estimated with test-time augmentation (TTA) [14–16]. TTA is a known method, which is a data augmentation technique applied during the testing phase. Multiple augmentations of the input image are commonly used, such as rotation, flipping or scaling and jittering [14–16]. However, these augmentation methods do not address uncertainties arising from potential variations in iCBCT reconstruction. Variations in sCT generation may arise from CBCT

* Corresponding author at: Department of Radiation Oncology, Catharina Hospital Eindhoven, Michelangelolaan 2, 5623 EJ Eindhoven, the Netherlands.
E-mail address: yvonne.d.hond@catharinaziekenhuis.nl (Y.J.M. de Hond).

https://doi.org/10.1016/j.phro.2025.100743
Received 19 September 2024; Received in revised form 19 February 2025; Accepted 26 February 2025
Available online 5 March 2025

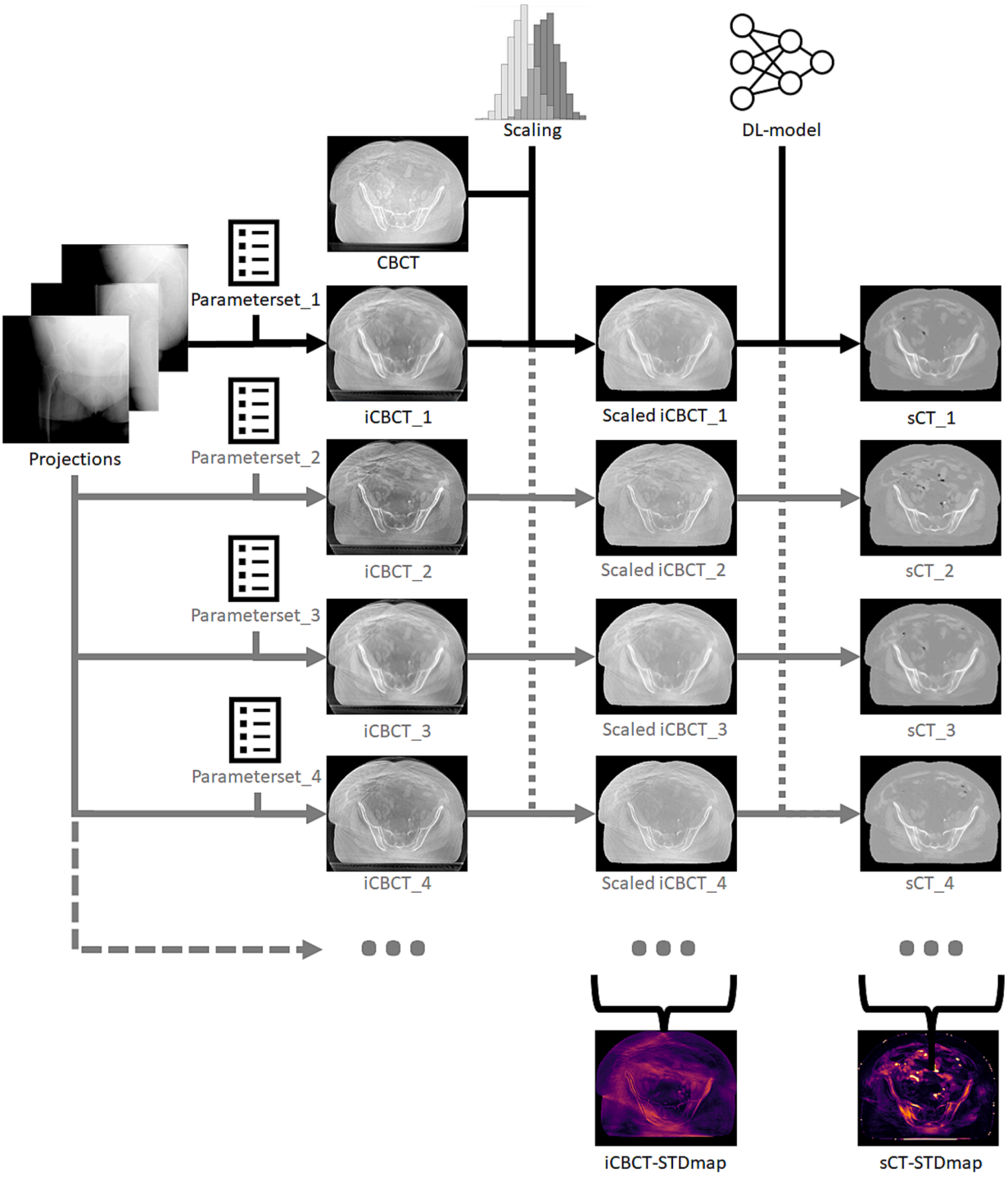


Fig. 1. Workflow to generate the stdmap of a sct and icbct. To generate these maps, 6 parameters were varied in a total of 128 different parameter sets.

reconstruction software updates. For example, parameter settings in IR programs can affect the iCBCT image, impacting both image quality and anatomical correctness of the generated sCT. Additionally, these variations in iCBCT reconstructions may influence the uncertainty of sCT predictions. Therefore, iCBCT reconstruction parameter variation is introduced as a type of TTA. The aim of this study was to evaluate the uncertainty in generating synthetic CT (sCT) from iterative cone-beam CT (iCBCT) using vendor-supplied software for online adaptive radiotherapy.

2. Materials and methods

2.1. Data

This study retrospectively included data of 20 female cervical and rectal cancer patients treated at Catharina Hospital, Eindhoven, the Netherlands between July 2022 and September 2023. Patients with metal orthopedic implants and patients who had undergone hysterectomy were excluded. CBCT images were acquired on multiple Xray Volumetric Imaging (XVI) v5.0.6 and v5.0.7 systems (Elekta AB,

Table 1
a) Varied parameters of FRESCO. b) Varied parameters for influence of each parameter category analysis. The analyzed parameter category was varied while the default setting was used for the other parameters.

a) Per patient analysis

Varied parameter category	Parameter name	Default setting	Variation	Number of variations
All combined	Noise variance scaling	1	0.5 and 1	2
	Filter size	1: 3*3 pixels	1 & 2: 3*3 and 5*5 pixels	2
	Scatter correction method	Soft	Soft and scaled	2
	Scatter correction threshold	Soft:0.6	Soft:0.06 and 0.7 Scaled:1e-5 and 1e-4	2
	Number of Iterations	100	100 and 200	2
	Regularization TV lambda	0.03	0, 0.01, 0.02 and 0.03	4
	Total number of combinations:			128

b) Additional sub-analysis: influence of each parameter category

Varied parameter category	Parameter name	Default setting	Variation	Number of variations
Filtering projections	Noise variance scaling	1	8 steps between 0.5 and 1	8
	Filter size	1: 3*3 pixels	1 & 2: 3*3 and 5*5 pixels	2
	Total number of combinations:			16
Estimated scatter	Scatter correction method	Soft	Soft and scaled	2
	Scatter correction threshold	0.7	If soft: 8 steps between 0.7 and 0.9 If scaled: 8 steps between 1e and 4 and 1e-5	8
	Total number of combinations:			16
Iterations	Number of Iterations	100	16 steps between 100 and 200	16
	Total number of combinations:			16
Regularization	Regularization TV Lambda	0.03	16 steps between 0 and 0.03	16
	Total number of combinations:			16

Sweden). CBCT imaging settings were: medium Field-of-View, bowtie filter, 64 mA and 40 ms per projection, 5.5 frames/sec in a 360° arc. Reconstructed CBCT dimensions were 41 cm x 26.4 cm x 41 cm (x, y, z), with a voxel size of 1 mm x 1 mm x 1 mm. Clinical CBCT scans were reconstructed using the standard XVI Feldkamp-Davis-Kress (FDK) algorithm with uncalibrated CT numbers. First fraction CBCT images were used in this study. This research, based on anonymized patient data, was approved under the Dutch non-WMO legislation and received ethical approval on August 27, 2021 (reference number nWMO-2020.235).

2.2. Uncertainty estimation workflow

To assess the robustness of the sCT model to variations in iCBCT reconstruction parameter-settings, uncertainty maps were generated using the workflow illustrated in Fig. 1. The different steps are explained in more detail in the subsections below. First, the raw projections of the CBCT scan were used to reconstruct multiple iCBCT images per patient, all with different IR parameters. Second, the histograms of the iCBCTs were scaled to the histogram of the corresponding clinical CBCT image to minimize the iCBCT number differences between different iCBCT images generated with varying parameter-settings. Third, the iCBCT images were used to generate sCT images. Finally, the variation per voxel in the different sCT images was calculated by the standard deviation of the different sCT images.

2.2.1. Iterative CBCT-images

The iCBCT was reconstructed with Framework for Reconstruction and Simulation of Conebeam images V24.2.1 (FRESCO) (Elekta AB). In this framework, an iterative polyenergetic quantitative (polyquant) reconstruction method was used, which is an iterative statistical reconstruction algorithm that subtracts the estimated Poisson noise in the CBCT image [17]. The algorithm attempts to minimize the difference between measured projection values and forward-projected values from an iteratively updated model of the patient using a polychromatic integrator.

Several parameters influence iCBCT reconstruction. Six different parameters were varied to create 128 different iCBCT images per CBCT scan (Table1a). The parameters were selected based on the ability to control the noise in the CBCT images, affecting both noticeable and subtle changes. All possible combinations of the six different parameter variations were applied to generate the 128 images per patient.

The first two parameters adjust Wiener filter denoising: magnitude (noise variance scaling) and filter size. Increasing these parameters reduces the noise in the image. The second two parameters control the estimated scatter, which is subtracted from the projection images before each FDK reconstruction. The scatter correction method and threshold parameters control the iterations for the polyquant problem and total variation (TV) regularization. Although more iterations sharpen the iCBCT image, ring and streak artifacts become more pronounced. The regularization TV parameter reduces the noise in the reconstructed volume. Increasing the regularization TV parameter increases blurriness but reduces detail. Four regularization TV parameters were used, including no regularization (Lambda = 0).

The iCBCT reconstructed with the research version did not have realistic CBCT numbers, whereas the sCT model was trained on CBCT data with realistic CBCT numbers. To compensate for this difference, the histogram of the iCBCT image was matched to the histogram of the clinical CBCT. Only values within the body contour were used for matching the histograms with the scikit-image toolbox in python. The values outside the body contour of the iCBCT reconstructions were set to −1000.

An additional in-depth analysis was performed to analyze whether the influence of some parameters was larger than other parameters. The analysis was performed four times, varying the parameters for filtering projections, estimated scatter, regularization or the number of iterations. The analyzed parameter category was varied while the default setting was used for the other parameters. Per parameter category evaluation, 16 different parameter-settings were used (Table 1b). This number was based on preliminary experiments (supplementary materials A1).

2.2.2. sCT generation

The different iCBCTs were used as augmented input for the Advanced Medical Image Registration Engine research V3.47 (ADMIRE) (Elekta AB). The used 2D Cycle Generative Adversarial Network (cycle-GAN) model in ADMIRE was EKT CBCT-sCT M + F Pelvis V5.0. This model was trained by the vendor on paired CBCT and planning CT pelvic data from different hospitals. The first 40 and last 40 slices of 1 mm in the iCBCT and sCT images, which on average accounted for 2083 cm^3 [range: 810 cm^3 – 4061 cm^3] and represented 13 % [range: 5 % – 23 %] of the imaged body contour, were excluded to minimize influence of limited CBCT reconstruction FOV artefacts in the iCBCT.

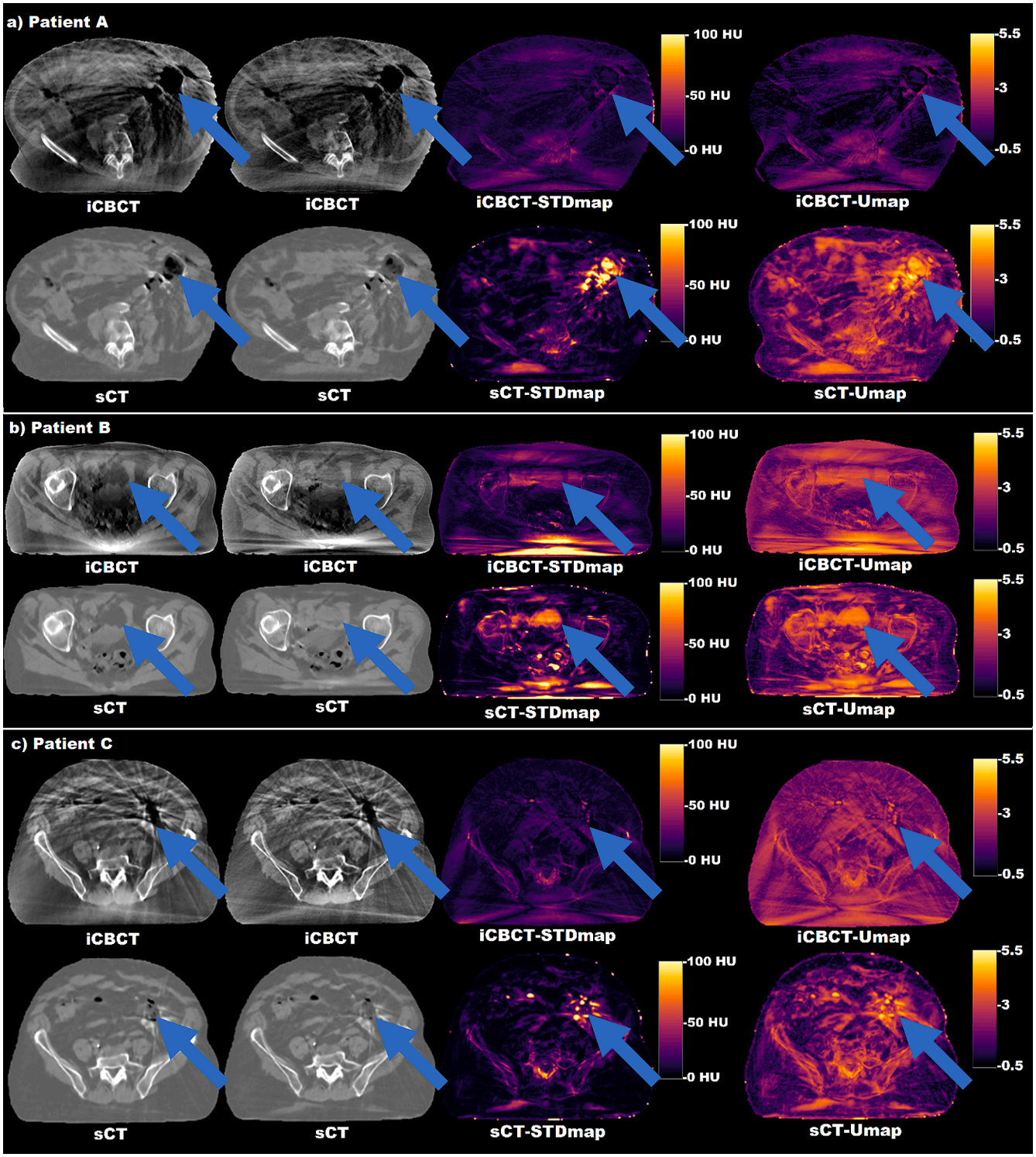


Fig. 2. Typical examples of iCBCT, sCT and their corresponding STDmaps and Umaps of three patients. In the first two columns, two iCBCTs per patient, which were reconstructed with the default parameter-setting (first column) and opposite side of parameter range* (second column), and below their associated sCT. In the third column, the iCBCT-STDmaps between 128 different iCBCTs and underneath the sCT-STDmaps of 128 sCTs. In the fourth column, the Umaps, calculated by the natural logarithm of the STDmaps. Window and level were 400/-200 for the iCBCT, 1000/0 for the sCT images. * Regularization TV Lambda 0 instead of 0.03, and 200 iterations instead of 100 iterations, etc.

2.2.3. *Uncertainty maps*

The uncertainty maps (Umaps) were computed for each individual CBCT projection dataset. For each patient, the uncertainty was assessed voxel-wise by the standard deviation (STD) across multiple sCTs resulting in one STD map (STDmap). The Umap was assessed voxel-wise by the natural logarithm of the STDmap across the multiple sCTs according to the following formula [15,16]:

$$Umap\,[x,y,z] = LOG(STDmap[x,y,z] + 1) \quad (1)$$

In which the x,y,z are voxel coordinates. This resulted in one sCT-Umap per patient. The calculation of STDmap and Umap was similarly computed among the different iCBCT images for evaluation of the

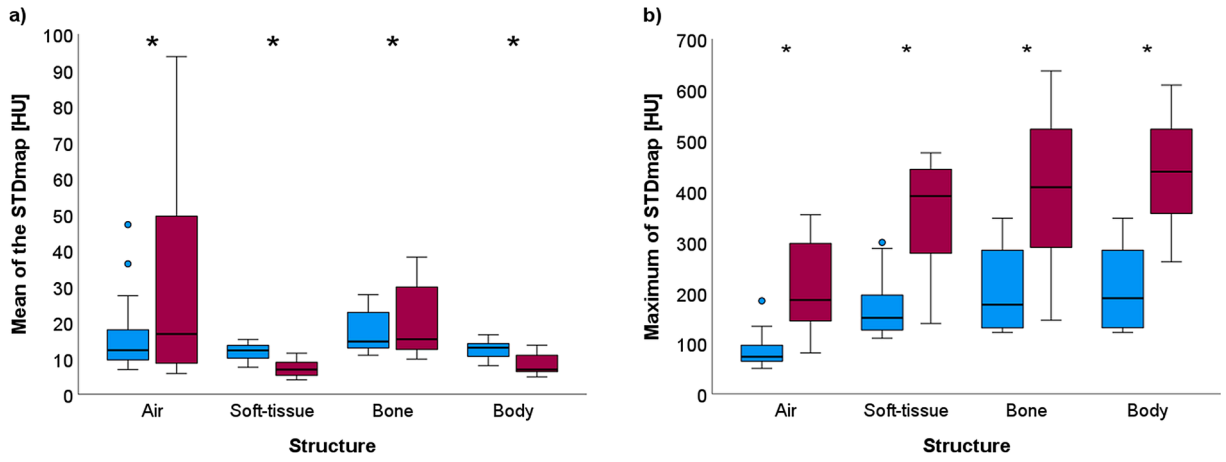


Fig. 3. boxplot of the stdmap values generated from iCBCTS (blue) or sCTS (red) per structure. a) mean and b) maximum values of the stdmap. A star indicates a significant difference between the iCBCT and sCT. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

amount of introduced input uncertainty.

In summary, this uncertainty estimation method was applied separately in three different scenarios: (1) on iCBCT data with 128 different reconstruction parameters, (2) on sCT data based on the 128 iCBCT images, (3) on the per-parameter sub-analysis.

2.3. Evaluation

The sCT-Umaps (sCT-Umaps) and the iCBCT-Umaps (iCBCT-Umaps) were generated to analyze the impact of iCBCT image variation on sCT generation. Umaps were evaluated within different types of tissue: air, bone and soft-tissue to determine if uncertainty was different across different tissue types and to quantitatively assess the uncertainty dependence on specific parameters. These structures were segmented on the clinical CBCT images in Raystation V12a (Raysearch Laboratories, Stockholm, Sweden). The body contour was segmented using the body segmentation option. The air within the body was segmented using HU-level thresholding. The thresholding level was based on the image histogram, specifically at the highest HU-level before the number of voxels per bin starts increasing (supplementary materials A2). Bones were segmented using the medium bone segmentation option in Raystation. This method translated threshold values, bone $>$ 200 HU and non-bone $<$ 100 HU, into the CBCT image specific values using the CBCT density table. All segmentations were verified and adapted manually if necessary. The soft-tissue segmentations were generated by extracting the bone and air segmentation. Since the image space of the CBCT, iCBCT and sCT is comparable, the structures on the CBCT were also used for the STDmaps of the iCBCT and sCT.

Visual inspection was conducted to qualitatively assess reconstruction parameter effects on iCBCT, sCT, STDmaps, and Umaps. The STDmaps were evaluated on average STD and maximum STD-values within a structure. A Wilcoxon signed rank test was used to statistically test if the sCT-STDmaps were different compared to the iCBCT-STDmaps, which was the variation in input to the sCT generation model [18]. A p-value below 0.05 was considered significant.

3. Results

3.1. Visual

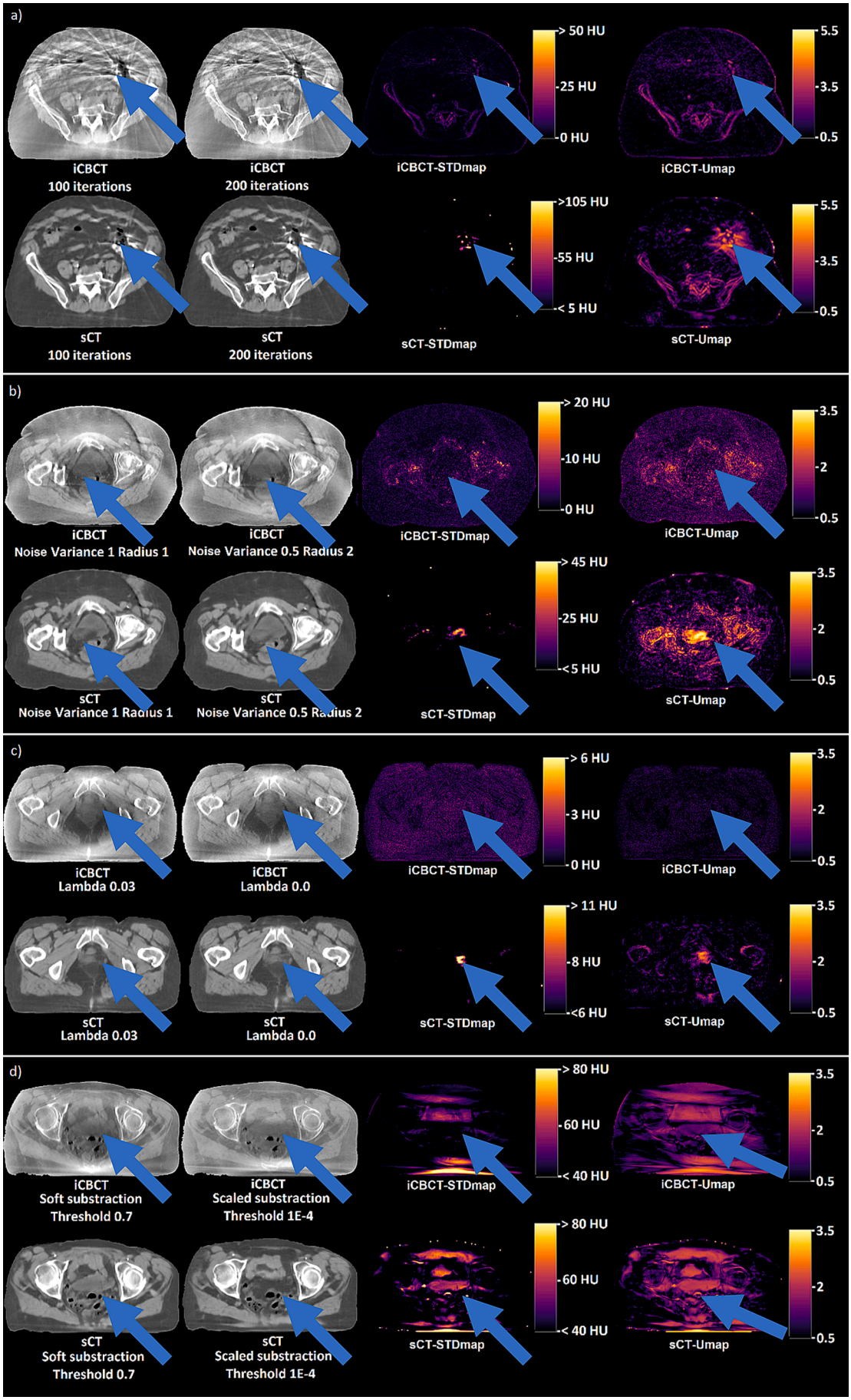
Typical examples of iCBCT, sCT,STDmaps and the Umaps for three patients are shown in Fig. 2. These patients were partly selected on the diversity in artifact types and partly at random to represent typical examples. In all three cases, the iCBCT in the second column showed a higher resolution compared to the iCBCT in the first column. However, the iCBCT images in the first column contained visually less noise and artefacts. The sCT images generated from these iCBCT images showed variations in synthesized tissue. For instance, in patient A, the air bubble present in sCT in the first column and in both iCBCTs images, was partly synthetized with higher values, which were more comparable to fat, in the second sCT. This region had higher STD-values compared to other areas, as visualized in the sCT-STDmaps. These regions were not highlighted in the iCBCT-STDmap. In patient B, the values comparable to fat tissue in sCT in the first column and in both iCBCTs images were, however, lower in the other sCT, which were more comparable to water than to fat. This region with differing types of tissue in the two sCTs was highlighted only in the sCT-Umaps. In patient C, differences in air-pockets and high values, more comparable to bone, were observed within the bowel. The regions with differing types of tissue in the two sCTs were highlighted only in the sCT-Umaps.

3.2. Standard deviation sCT compared to iCBCT

Although the mean values of the sCT-STDmaps within soft-tissue structure ($p < 0.001$) and body contour ($p < 0.001$) were significantly lower compared to iCBCT-STDmaps, the maximum STD-values in soft-tissue structure ($p < 0.001$) and body structure ($p < 0.001$) were significantly higher in sCT compared to iCBCT (Fig. 3). Within the body structure, the mean STD-value was lower in sCT-STDmap, with a median of 7 HU (range: 5 HU – 14 HU), compared to iCBCT-STDmap, 13 HU (range: 8 HU – 17 HU). However, the maximum STD-value was larger in sCT-STDmap, with a median of 438 HU (range: 260 HU – 609 HU), compared to iCBCT-STDmap, 198 HU (range: 121 HU – 346 HU). Within the air structure, the sCT-STDmaps had a significant higher mean ($p = 0.014$) as well as maximum values ($p < 0.001$) compared to the iCBCT-STDmap.

3.3. Additional sub analysis: Influence of each parameter category

In the varied parameter category, the artifact magnitude remained consistent across all 20 cases, while their locations varied among individuals. In the first typical example, increasing the number of iterations from 100 to 200 sharpened the iCBCT image, though the streaking artefacts became more pronounced (Fig. 4a). The sCTs showed variability in synthesized air bubbles, and high-density areas were synthesized in the bowel. These areas were highlighted on the Umap, indicating higher uncertainty compared to the iCBCT-Umap. Although Wiener filtering variations resulted in visually comparable iCBCT images, sCTs differed in synthesized tissue (Fig. 4b). For example, a lower density was synthesized in a part of the bladder in the left sCT, which



(caption on next page)

Fig. 4. Typical examples of iCBCT, sCT and corresponding STDmaps and Umaps per varied parameter. The following parameters were varied: a) the number of iterations, b) wiener filtering parameters, c) regularization parameter and d) the scatter correction parameters. In the first and second columns, two iCBCTs per patient, which were reconstructed with largest different parameter setting used in this study, and underneath their associated sCT. In the third two column, the iCBCT-STDmaps between 16 different iCBCTs and underneath the sCT-STDmaps of 16 sCTs. In the fourt colm, the iCBCT-Umaps and underneath the sCT-Umaps Window and level were 400/-300 for the iCBCT, 500/0 for the sCT images. For better visibility, the scale is different for each STDmap. The Umap scales of varied number of iterations are different than the other parameter categories. While the magnitude of artifacts observed was generally consistent across all 20 cases within the varied parameter category, the location of the artifacts varied between individuals.

was only highlighted on the sCT-Umap. Similar effects were observed with varying TV regularization, however, the uncertainty was lower than with Wiener filtering (Fig. 4c). Although varying scatter correction parameters resulted in larger uncertainties among the iCBCT images compared to varying other parameters, the differences were subtle in the iCBCT images (Fig. 4d). In the sCT images, CT number variations in the CTV and abdominal muscles, appearing as fat, were highlighted on the sCT-Umap.

Variance in scatter correction resulted in the largest mean STD-values across all structures compared to variance of the other parameters (Fig. 5). Although the mean STD-values in soft-tissue were significantly larger in iCBCT compared to sCT, all values remained below 15 HU ($p < 0.001$) (Fig. 5b). In contrast, the maximum STD-values were significantly higher in sCT compared to iCBCT, except for TV regularization within the bone structure (supplementary materials A3). The magnitude of discrepancies varied depending on the varied parameter category. While scatter correction resulted in the largest mean STD-values, the maximum STD-values were most influenced by both scatter correction and the number of iterations. Within the body contour, for example, scatter correction resulted in the highest variance in sCTs (median: 358 HU), whereas TV regularization resulted in the lowest variance in sCTs (median: 233 HU). The highest maximum STD-values were observed in bone, with a median of 290 HU for scatter correction.

4. Discussion

In this study, the robustness of sCT generation based on iterative CBCT reconstruction was evaluated. The variance in sCT densities was significantly higher compared to iCBCT densities, indicating that the sCT model was not robust to parameter changes in the IR. Variations in specific parameters had more influence on sCT-Umap than other parameters of the iCBCT reconstruction.

The mean STD-values were highest within the air structures of sCT images and significantly higher in sCT images compared to iCBCT images. For example, air-pockets in the bowel had an increased density, comparable to fat or tissue. This aligns with van Harten et al., in which air-filled sinuses on the MRI were synthesized as tissue-like densities on the sCT [9]. While the current study focused on female pelvic scans,

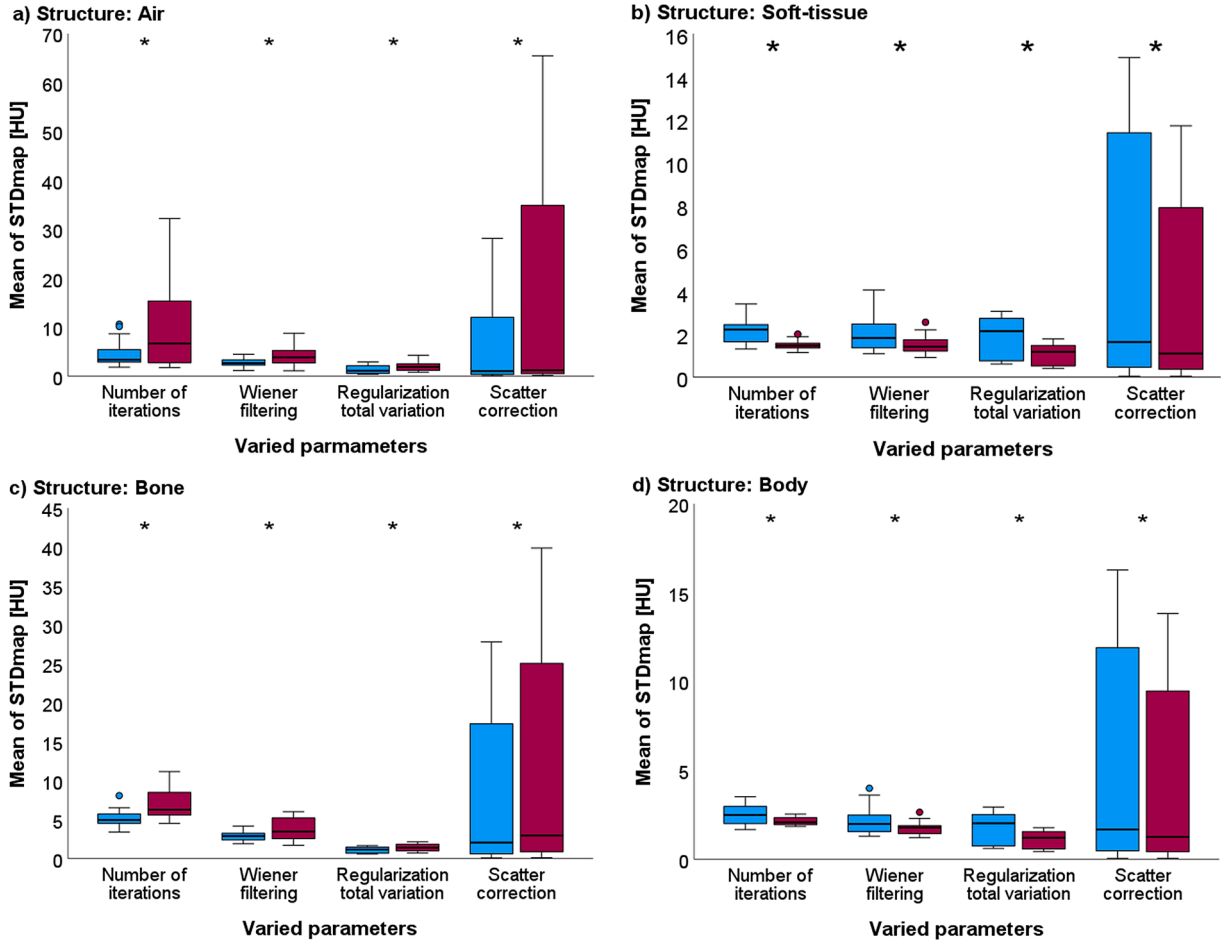


Fig. 5. boxplot of the STDmap values (mean) generated from iCBCTs (blue) or sCTs (red) per parameter setting for air, soft-tissue, bone and body. A star indicates a significant difference between the iCBCT and sCT. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

similar density variance due to iCBCT reconstruction differences is expected in male scans as well. Rusanov et al, observed CT-density variance in sCTs generated from prostate cancer CBCT images [16]. Since the variance in the current study occurred across all tissue types, these findings will probably apply to CBCT scans of other body parts.

In the additional sub-analysis, variations in specific parameters affected the sCT-Umap more than other iCBCT reconstruction parameters. For example, scatter correction introduced the highest variance in sCTs, while changes in iCBCTs were visually subtle, despite the observed values in iCBCT-STDmap. Conversely, variations in TV regularization resulted in the lowest STD-values in sCT images, despite increased iCBCT blurriness. Therefore, changes in the iCBCT did not necessarily predict their effect on sCT. Imperceptible alterations in the input image can lead to incorrect predictions by DL- networks, as observed in various non-radiotherapy fields [19–21]. Similarly, changes in the iCBCT reconstruction, even when not directly observable in the iCBCT image, can still influence sCT generation. Therefore, it is recommended that the image preparation workflow includes an evaluation of both the iCBCT and the generated sCT, particularly if software versions or settings of either one is updated. To assess the impact of software updates and ensure the model is used on data it was trained on, transparency regarding model characteristics and the training dataset is essential.

The maximum STD-values were significantly larger in sCT compared to iCBCT, with the highest STD-values in bone and soft-tissue structures and lowest in air structure. These variations would result in clinically irrelevant dose differences, as indicated in supplementary materials A4 and consistent with the evaluation of air-pocket differences in Lemus et al. [22]. However, the interpretation of the CT-values to anatomical structures or tissue types on the sCT image can differ due to the smaller range of CT-values representing soft-tissue compared to bone tissue. Most maximum STD-values in sCT images exceeded the HU range of soft-tissue, ranging from −120 HU to 70 HU [23]. However, most maximum STD-values were lower than the HU range for both cancellous and cortical bone, ranging from 300 to 900 HU [24]. In the first case it could convert tissue into fat or even into lung tissue, while in the second case, the CT-values could still represent bone [25]. This tissue-to-fat or fat-to-lung conversion can impact plan selection or segmentation, with changes in this range having a larger effect than variations around −1000 HU or 500 HU.

IR variations could also influence DL-based auto-segmentation prediction. Variations in IR algorithms for CT imaging affected the model-based auto-segmentations demonstrated by Miller et al. [26]. The average center of mass shift between rectum segmentations was 4.5 mm, which is comparable to the clinical planning target volume (PTV) margin [26,27]. Therefore, variations in IR could have a major clinical impact.

In addition to the magnitude, the location and volume of regions with high STD-values are also critical factors. Specifically, areas located near or within the PTV could have a larger clinical impact than areas located at the image boundaries. Further research is required to assess the clinical relevance of these variances in sCT images.

In conclusion, variations in parameter-settings of the iCBCT reconstruction significantly affected the representation of the anatomy on the sCT. The variance in sCTs was depending on the parameter category. Subtle alterations in the iCBCT leaded to density and anatomical changes in the sCT. Therefore, it is recommended that the image preparation workflow includes an evaluation of both the iCBCT and the generated sCT, particularly if software versions or settings of either one is updated.

Credit author statement

All authors contributed to the conception of the work, analysis and interpretation of data. All authors contributed to and approved the final version of the manuscript. All authors agreed to be accountable for the work in ensuring that questions related to the accuracy or integrity of the manuscript are appropriately investigated and resolved.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used chat gpt 4.0 in order to check written sentences on language errors and to reduce the number of words in a section. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Yvonne de Hond was financially supported by a research grant by Elekta (grant number SOW_20210426, Elekta Ltd., Crawley, UK). All remaining authors have declared no conflicts of interest.

Acknowledgments

Yvonne de Hond was financially supported by a research grant by Elekta (grant number SOW_20210426, Elekta Ltd., Crawley, UK).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2025.100743.

References

[1] Mao W, Liu C, Gardner SJ, Siddiqui F, Snyder KC, Kumarasiri A, et al. Evaluation and clinical application of a commercially available iterative reconstruction algorithm for CBCT-based IGRT. Technol Cancer Res Treat 2019;18:1533033818823054. https://doi.org/10.1177/1533033818823054.

[2] Gardner SJ, Mao W, Liu C, Aref I, Elshaikh M, Lee JK, et al. Improvements in CBCT image quality using a novel iterative reconstruction algorithm: a clinical evaluation. Adv Radiat Oncol 2019;4:390–400. https://doi.org/10.1016/j.adro.2018.12.003.

[3] O'Hara CJ, Bird D, Al-Qaisieh B, Speight R. Assessment of CBCT-based synthetic CT generation accuracy for adaptive radiotherapy planning. J Appl Clin Med Phys 2022;23:e13737. https://doi.org/10.1002/acm2.13737.

[4] Groot Koerkamp ML, Bol GH, Kroon PS, Krikke LL, Harderwijk T, Zoetelief AJ, et al. Bringing online adaptive radiotherapy to a standard C-arm linac. Phys Imaging Radiat Oncol 2024;31:100597. https://doi.org/10.1016/j.phro.2024.100597.

[5] Liu Y, Lei Y, Wang T, Fu Y, Tang X, Curran WJ, et al. CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy. Med Phys 2020;47:2472–83. https://doi.org/10.1002/mp.14121.

[6] Sherwani MK, Gopalakrishnan S. A systematic literature review: deep learning techniques for synthetic medical image generation and their applications in radiotherapy. Front Radiol 2024;4:1385742. https://doi.org/10.3389/fradi.2024.1385742. eCollection@2024.:1385742.

[7] De Hond YJM, Kerckhaert CEM, van Eijnatten MAJM, Van Haaren PMA, Hurkmans CW, Tijssen RHN. Anatomical evaluation of deep-learning synthetic computed tomography images generated from male pelvis cone-beam computed tomography. Phys Imaging Radiat Oncol 2023;25:100416. https://doi.org/10.1016/j.phro.2023.100416.

[8] Galapon Jr AV, Thummerer A, Langendijk JA, Wagenaar D, Both S. Feasibility of Monte Carlo dropout-based uncertainty maps to evaluate deep learning-based synthetic CTs for adaptive proton therapy. Med Phys 2024;51:2499–509. https://doi.org/10.1002/mp.16838.

[9] van Harten LD, Wolterink JM, Verhoeff JC, Išgum I. Automatic Online Quality Control of Synthetic CTs. arXiv (Cornell University); 2019.

[10] Zhang X, Sisniega A, Zbijewski WB, Lee J, Jones CK, Wu P, et al. Combining physics-based models with deep learning image synthesis and uncertainty in intraoperative cone-beam CT of the brain. Med Phys 2023;50:2607–24. https://doi.org/10.1002/mp.16351.

[11] Kläser K, Borges P, Shaw R, Ranzini M, Modat M, Atkinson D, et al. A multi-channel uncertainty-aware multi-resolution network for MR to CT synthesis. Appl Sci (Basel) 2021;11:1667. https://doi.org/10.3390/app11041667.

[12] Reinhold JC, He Y, Han S, Chen Y, Gao DY, Lee J, et al. Validating Uncertainty in Medical Image Translation. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI); 2020. p. 95–8. https://doi.org/10.1109/isbi45749.2020.9098543.

[13] van den Berg CAT, Meliadiadó EF. Uncertainty assessment for deep learning radiotherapy applications. Semin Radiat Oncol 2022;32:304–18. https://doi.org/10.1016/j.semradonc.2022.06.001.
[14] Iwana BK, Uchida S. An empirical survey of data augmentation for time series classification with neural networks. PLoS One 2021;16:e0254841. https://doi.org/10.1371/journal.pone.0254841.
[15] Wang G, Li W, Aertsen M, Deprest J, Ourselin S and Vercauteren T. Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. Neurocomputing. (Amst) 2019; 335:34-45. doi: 10.1016/j.neucom.2019.01.103. Epub@2019 Feb 7.:34-45.
[16] Rusanov B, Hassan GM, Reynolds M, Sabet M, Rowshanfarzad P, Bucknell N, et al. Transformer CycleGAN with uncertainty estimation for CBCT based synthetic CT in adaptive radiotherapy. Phys Med Biol 2024;69:10. https://doi.org/10.1088/1361-6560/ad1cfc.
[17] Mason JH, Perelli A, Nailon WH, Davies ME. Polyquant CT: direct electron and mass density reconstruction from a single polyenergetic source. Phys Med Biol 2017;62:8739. https://doi.org/10.1088/1361-6560/aa9162.
[18] Wilcoxon F. Individual comparisons by ranking methods. Biometrics 1945;1:80–3. https://www.doi.org/10.2307/3001968.
[19] Bortsova G, González-Gonzalo C, Wetstein SC, Dubost F, Katramados I, Hogeweg L, et al. Adversarial attack vulnerability of medical image analysis systems: unexplored factors. Med Image Anal 2021;73:102141. https://doi.org/10.1016/j.media.2021.102141.
[20] Goodfellow IJ, Shlens J, Szegedy C. Explaining and harnessing adversarial examples. CoRR 2014. https://doi.org/10.48550/arXiv.1412.6572. abs/1412.6572.
[21] Cao K, Liu M, Su H, Wu J, Zhu J, Liu S. Analyzing the noise robustness of deep neural networks. IEEE Trans Vis Comput Graph 2021;27:3289–304. https://doi.org/10.1109/TVCG.2020.2969185.
[22] Lemus OMD, Tanny S, Cummings M, Webster M, Wancura J, Jung H, et al. Influence of air mapping errors on the dosimetric accuracy of prostate CBCT-guided online adaptive radiation therapy. J Appl Clin Med Phys 2023;24:e14057. https://doi.org/10.1002/acm2.14057.
[23] Spratt JD, Salkowski LR, Weir J, Abrahams PH. Imaging Atlas of Human Anatomy 2010.
[24] Lim Fat D, Kennedy J, Galvin R, O'Brien F, Mc Grath F, Mullett H. The Hounsfield value for cortical bone geometry in the proximal humerus—an in vitro study. Skeletal Radiol 2012;41:557–68. https://doi.org/10.1007/s00256-011-1255-7.
[25] Elsayed O, Mahar K, Kholief M, Khater HA. Automatic detection of the pulmonary nodules from CT images. SAI Intelligent Systems Conference (IntelliSys) 2015; 2015:742–6. https://doi.org/10.1109/IntelliSys.2015.7361223.
[26] Miller C, Mittelstaedt D, Black N, Klahr P, Nejad-Davarani S, Schulz H, et al. Impact of CT reconstruction algorithm on auto-segmentation performance. J Appl Clin Med Phys 2019;20:95–103. https://doi.org/10.1002/acm2.12710.
[27] Pötter R, Tanderup K, Kirisits C, de LA, Kirchheiner K, Nout R, et al. The EMBRACE II study: the outcome and prospect of two decades of evolution within the GEC-ESTRO GYN working group and the EMBRACE studies. Clin Transl Radiat Oncol 2018;9:48–60. https://doi.org/10.1016/j.ctro.2018.01.001.