

Original Research Article

Comparing methods to improve cone-beam computed tomography for dose calculations in adaptive proton therapy



Casper Dueholm Vestergaard ^{a,b,*¹}, Nadine Vatterodt ^{a,b,1}, Ulrik Vindelev Elstrøm ^a, Kenneth Jensen ^a, Ole Nørrevang ^a, Ludvig Paul Muren ^{a,b,*}, Stine Sofia Korreman ^{a,b}, Vicki Trier Taasti ^{a,b}

^a Danish Centre for Particle Therapy, Aarhus University Hospital, Aarhus, Denmark

^b Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

ARTICLE INFO

Keywords:

Adaptive proton therapy
Cone-beam computed tomography
Synthetic computed tomography
Head-and-neck cancer
Proton dose calculation

ABSTRACT

Background and purpose: Proton therapy requires dose monitoring, often performed based on repeated computed tomography (reCT) scans. However, reCT scans may not accurately reflect the internal anatomy and patient positioning during treatment. In-room cone-beam CT (CBCT) offers a potential alternative, but its low image quality limits proton dose calculation accuracy. This study therefore evaluated different methods for quality-improvement of CBCTs (synthetic CTs; sCTs) for use in adaptive proton therapy of head-and-neck cancer patients.

Materials and methods: Thirty-five CBCTs from twenty-four head-and-neck cancer patients were used to assess four sCT generation methods: an intensity-correction method, two deformable image registration methods, and a deep learning-based method. The sCTs were evaluated against same-day reCTs for CT number accuracy, proton range accuracy through single-spot plans, and dose recalculation accuracy of clinical plans via dose-volume-histogram (DVH) parameters.

Results: All four methods generated sCTs with improved image quality while preserving the anatomy relative to the CBCT. The differences in absolute median proton range between sCT methods were small and generally less than the difference between sCT and reCT, which had median differences of 1.0–1.1 mm. Similarly, differences in DVH parameters were generally small between the sCT methods. While outliers were identified for all four methods, these outliers were often consistent for all sCT methods and could be attributed to anatomical and/or positional discrepancies between the CBCT and reCT.

Conclusions: All four sCT methods enabled accurate proton dose calculation and preserved the anatomy, making them of value for adaptive proton therapy.

1. Introduction

The favorable dose distribution of protons, characterized by the Bragg peak, allows precise targeting of the tumor, while reducing the dose to surrounding healthy tissues and thereby the risk of radiation-induced side effects [1–3]. However, the characteristics of the Bragg peak make proton therapy sensitive to anatomical changes and tissue density variations, which can lead to under-dosage of the tumor and/or over-dosage of healthy tissues [4,5]. Tumor shrinkage, weight loss, or positional variations during the treatment for head-and-neck cancer patients necessitate careful dose monitoring to maintain optimal dose delivery [6–8].

Cone-beam computed tomography (CBCT) scans are often used to verify patient positioning and monitor anatomical changes, yet CBCT image artifacts, noise, and CT number instability can hinder accurate proton dose calculations [9]. Patients are therefore often rescanned on a CT scanner (repeated CT; reCT) to ensure that the delivered dose distribution aligns with the planned dose. However, anatomical and positional differences between the patient setup at the CT scanner and on the treatment couch can compromise the accuracy of the reCT dose evaluation, making it less representative of the actual delivered dose distribution.

The need for dose evaluation on the patient's anatomy during treatment has driven the development of methods to generate quality-

* Corresponding authors.

E-mail addresses: Casper.Dueholm.Vestergaard@rm.dk (C.D. Vestergaard), ludvmure@rm.dk (L.P. Muren).

¹ Shared first authors.

improved CBCT images (collectively referred to as synthetic CTs (sCTs) in the following) for more accurate proton dose calculations based on the CBCT scans. Traditional approaches improve image quality through intensity correction, including bulk density override [10,11] or histogram matching [12,13]. Other solutions apply deformable image registration (DIR) to deform the planning CT (pCT) to the CBCT [14–17]. These methods may be used individually or in combination [18,19], with the extent of manual effort depending on available software. More recently, deep learning (DL) has shown promising results for sCT generation, enabling fast sCT generation with minimal manual input [20–22].

In this study, we compared four sCT generation methods (both commercial and in-house), spanning workflows from manual to fully automated approaches. Rather than proposing a new technique, the aim was to evaluate existing methods in terms of their suitability to assess the need for adaptation in proton therapy of head-and-neck cancer patients, and their potential for daily dose monitoring based on the patient anatomy seen at the treatment.

2. Methods and materials

2.1. Patient data

The study included 24 head-and-neck cancer patients. The use of data was approved by an internal review board and the Central Denmark Region Committees on Health Research Ethics (case number: 2119125). The patients were treated with intensity-modulated proton therapy (IMPT) with a prescribed dose of 66/68/68.4 Gy in 33/34/38 fractions to the high-risk clinical target volume (CTV, denoted CTV1), 60 Gy to the intermediate-risk CTV (CTV2), and 50 Gy to the elective volumes (CTV3), delivered with a simultaneous integrated boost according to the DAHANCA guidelines [23]. The clinical plans were created in Eclipse v16.1 (Varian – a Siemens Healthineers company, Palo Alto, CA, US) with multi-field robust optimization using combined ± 4 mm setup and $\pm 3.5\%$ range uncertainty (details on beam arrangements and dose calculation settings in [Supplementary material S1](#)).

A total of 35 CBCTs were selected from the 24 patients. The CBCTs were selected to minimize anatomical mismatch to the same-day reCT, while having a varying degree of mismatch between the pCT and the CBCT to cover a range of difficulty levels in sCT generation. This assessment was done through slice-wise visual inspection by an experienced medical physicist. The generated sCTs were compared to the same-day reCTs, with the reCTs as reference. Details on the CT and CBCT protocols can be found in the [Supplementary material S2](#).

2.2. Synthetic CT generation

Four different sCTs were generated based on each of the 35 CBCTs. Additional input to the sCT generation methods included the corresponding pCT and the rigid registration between the pCT and CBCT created as part of clinical routine. The methods are briefly described in the following, while additional details can be found in [Supplementary material S3](#).

CBCT correction method (corrected CBCT, cCBCT): The cCBCT method available in RayStation 12A-R (v13.0.100.0, RaySearch Laboratories, Stockholm, Sweden) was applied in a fully automated workflow. The method is based on an iterative process applying a voxel-wise conversion between the CT numbers on the CBCT and the registered pCT followed by a low frequency artifact correction. Neither of these steps alter the anatomy [24].

DIR-based method (adapted CT, aCT): The aCTs were manually generated by an experienced medical physicist according to our clinical practice. A DIR between pCT and CBCT was iteratively established in Velocity™ v.4.1 (Varian) to generate a deformed pCT. Anatomically mismatching areas that could not be properly deformed (e.g. tongue, patient outline) were manually corrected by bulk override of CT

numbers in Eclipse. Areas representing air were overridden with a CT number of -960 HU while soft tissue areas were overridden with a CT number of 40 HU, leaving small homogenous override regions [25].

Hybrid method (virtual CT, vCT): The vCT approach available in RayStation was applied in a fully automated workflow. This method is a hybrid between a CBCT correction and a DIR-based method, which deforms the pCT to the anatomy of the CBCT, similar to the aCT method, but this method includes an automatic CT number override in mismatched low-density areas with the corresponding voxels from the cCBCT [26].

Deep learning method (deep-learning sCT, DLsCT): The DLsCTs were generated using an in-house, three-dimensional (3D) Cycle-consistent Contrastive Unpaired Translation (CycleCUT) network, described in detail elsewhere [21].

All four sCT generation methods applied stitching of the pCT to obtain image information outside the CBCT. The stitching techniques varied between the sCT methods (see [Supplementary material S3](#)).

2.3. Post-processing and structure propagation

The cCBCTs and vCTs were initially defined in the frame-of-reference of the CBCT, while the DLsCTs and aCTs were defined in the frame-of-reference of the pCT. To align all sCTs, the cCBCTs and vCTs were resampled to the frame-of-reference of the pCT (NiftyReg [27]) using the clinical rigid registration between the CBCT and pCT.

The clinical structure set on the pCT was deformed to the reCTs and edited if needed as part of the clinical workflow for monitoring and replanning. The structure set on the reCT was rigidly propagated to all sCTs without further editing, ensuring identical structures were used across the reCT and sCTs. Here the underlying rigid registration between the reCT and pCT was used, exploiting that the pCT and the sCTs had the same frame-of-reference.

2.4. Evaluations

All sCTs were imported into Eclipse for evaluation. The sCTs were evaluated based on i) anatomical preservation and image quality, ii) CT number accuracy, iii) proton range calculation accuracy, and iv) proton dose calculation accuracy.

Anatomical preservation and image quality was qualitatively assessed by visual inspection. The individual comparisons of the sCTs with the respective CBCT were performed using the available tools in Eclipse, including image/color blending, moving window and image difference. The observations were collected and reviewed by an experienced medical physicist.

CT number accuracy was assessed by comparing the mean and standard deviation of CT numbers in spherical regions-of-interest (ROIs; 1 cm in diameter) on the reCT and sCTs. The ROIs were placed on the reCTs (in regions where the reCTs and CBCTs anatomies were similar) and rigidly transferred to the sCTs. A total of twelve ROIs were placed: four in muscle tissue, four in fat tissue, and four in bone tissue.

Proton range calculation accuracy on the sCTs was evaluated through three single-spot proton plans, created using a 100 MeV proton beam with 100 MU at a gantry angle of 90 degrees and a 1 mm dose-grid resolution. The three spots were placed 2 cm apart in the cranial-caudal direction, centered around the clinical iso-center, with individual spot positions adjusted as described in [Supplementary material S4](#). The integral depth-dose curve was fitted with a B-spline and the range was defined as the depth of the dose falling to 80% (R80) of its maximum [28]. The differences in R80 ($\Delta R80$) between the sCT and reCT, as well as among the different sCTs, were evaluated.

Proton dose calculation accuracy was evaluated by recalculating the clinical plan for the nominal and range uncertainty scenarios ($\pm 3.5\%$) on the reCT and each sCT, using identical dose calculation settings. The dose-volume-histogram (DVH) parameters for the targets and the clinically relevant organs-at-risk were extracted from the reCT and sCTs and

compared. The following DVH parameters were evaluated for the nominal scenario: near-max dose ($D_{0.03cc}$) in the body and mandible; D_{mean} for the oral cavity, parotids (left and right), pharyngeal constrictor muscles (PCM; lower/middle/upper), and the total submandibular gland (i.e. the left and right combined). For the robustly evaluated plan, the worst-case-scenario was extracted for the $V_{95\%}$ for the three targets (CTV1-3, with different prescribed doses) and $D_{0.03cc}$ for the spinal cord.

3. Results

3.1. Image quality and anatomical evaluation

The image quality of the sCTs varied between the sCT generation methods (Fig. 1). The cCBCTs exhibited higher levels of image noise than the other sCT types. The streaking artifacts seen on the CBCTs around the cartilage, hyoid bone, and throat persisted on the cCBCTs, and were also partially replicated on the DLsCTs. Additionally, the DLsCTs appeared to have slightly less fine structure details than the vCTs and aCTs. Stitching discontinuities were occasionally observed on the DLsCTs and aCTs.

The cCBCTs and DLsCTs showed the closest anatomical agreement with the CBCT. Inadequate deformation of the pCT caused larger disagreements for the aCTs and vCTs, which partially remained despite the use of overriding techniques. For instance, in two patients, a nasogastric tube inserted after the pCT acquisition was not restored on the aCT and was only partially recovered on the vCT with the use of cCBCT for overriding (Fig. 2A). Moreover, the homogenous CT number overrides in

aCTs lacked the subtle intensity variations of the surrounding tissue (due to image noise and tissue heterogeneity), making these override regions poorly integrated (Fig. 1). Furthermore, the manual overrides of the patient outline (to capture weight loss/gain) were occasionally discontinuous with the rest of the image. For the vCT, the incorporation of cCBCT data occasionally introduced areas with increased image noise and artifacts, including large intensity artifacts in the back of the patient (Fig. 2B).

3.2. CT number evaluation

The cCBCTs had the overall largest CT number deviations, with a median difference of -1 HU in muscle tissues and 40 HU in bone tissues (Fig. 3). All four sCT methods were found to have median absolute CT number differences close to zero, however, the DLsCTs and cCBCTs were found to slightly under- and over-estimate the CT numbers in bony tissues (median difference of -25 HU and 40 HU, respectively). Compared to the reCT, the cCBCTs were in general found to have higher standard deviations (median difference of 24 HU for bone), while the remaining three sCT methods had lower standard deviations in the bony tissues (median difference of -17 HU (aCT) to -24 HU (DLsCT)).

3.3. Proton range evaluation

The median absolute differences in R_{80} between the reCT and the sCTs were 1.0 mm for vCT and DLsCT and 1.1 mm for cCBCT and aCT, while the maximum ΔR_{80} was between 4.5 and 5.4 mm (Table 1). The

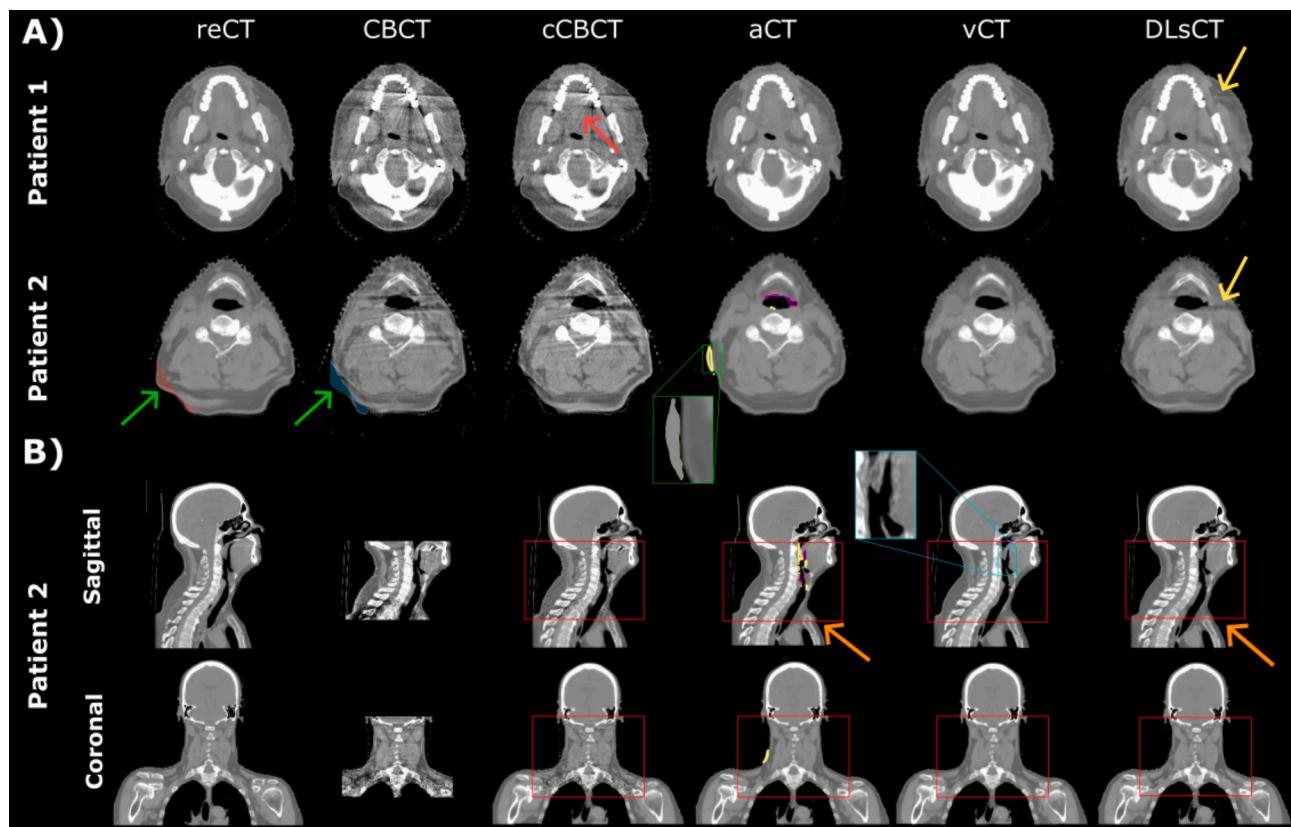


Fig. 1. A) Axial slices of the repeated CT (reCT), CBCT, and the four types of synthetic CTs (cCBCT, aCT, vCT, and DLsCT) are shown for two patients; one with good anatomical agreement between the reCT and CBCT (Patient 1) and one with poor agreement (Patient 2). The green arrows points to anatomical differences between the reCT and CBCT for Patient 2: Additional tissue is present on the reCT (red) compared to the CBCT (light blue). The red arrow points to streaking artifacts on the cCBCT also seen on the CBCT, and the yellow arrows point to residual artifacts on the DLsCT. For the aCT, the overriding with air and soft tissue is shown by magenta and yellow structures, respectively. The green insert shows a zoom of a region overridden with soft tissue on the aCT (CT number override of 40 HU), which is disconnected from the body outline. B) Sagittal and coronal views of Patient 2. The CBCT field-of-view and scan length is indicated on the synthetic CTs by a red box. The orange arrows point to a discontinuity resulting from the stitching technique used for the aCT and DLsCT. The cyan insert on the vCT shows a zoom of a region where voxels from the cCBCT have been used for overriding.

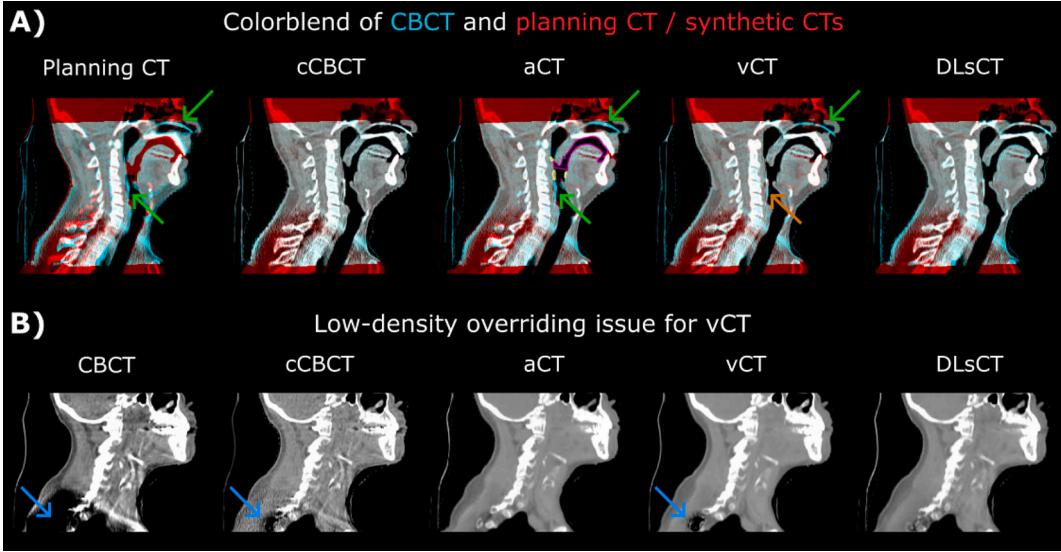


Fig. 2. A) Colorblend of the CBCT (blue) and planning CT or synthetic CTs (red) for a patient with nasogastric tube inserted after the planning CT acquisition. On the aCT, the CT number override with air and soft tissue is shown by magenta and yellow structures, respectively. The cCBCT and DLsCT correctly replicate the nasogastric tube. The aCT fails to restore the nasogastric tube (green arrows), while the vCT partially recovers the nasogastric tube through overriding with the cCBCT (orange arrow). B) Sagittal views of the CBCT, cCBCT, aCT, vCT, and DLsCT for a patient where the usage of cCBCT in seemingly mismatching low-density regions for the vCT leads to the introduction of large intensity artifacts (blue arrows).

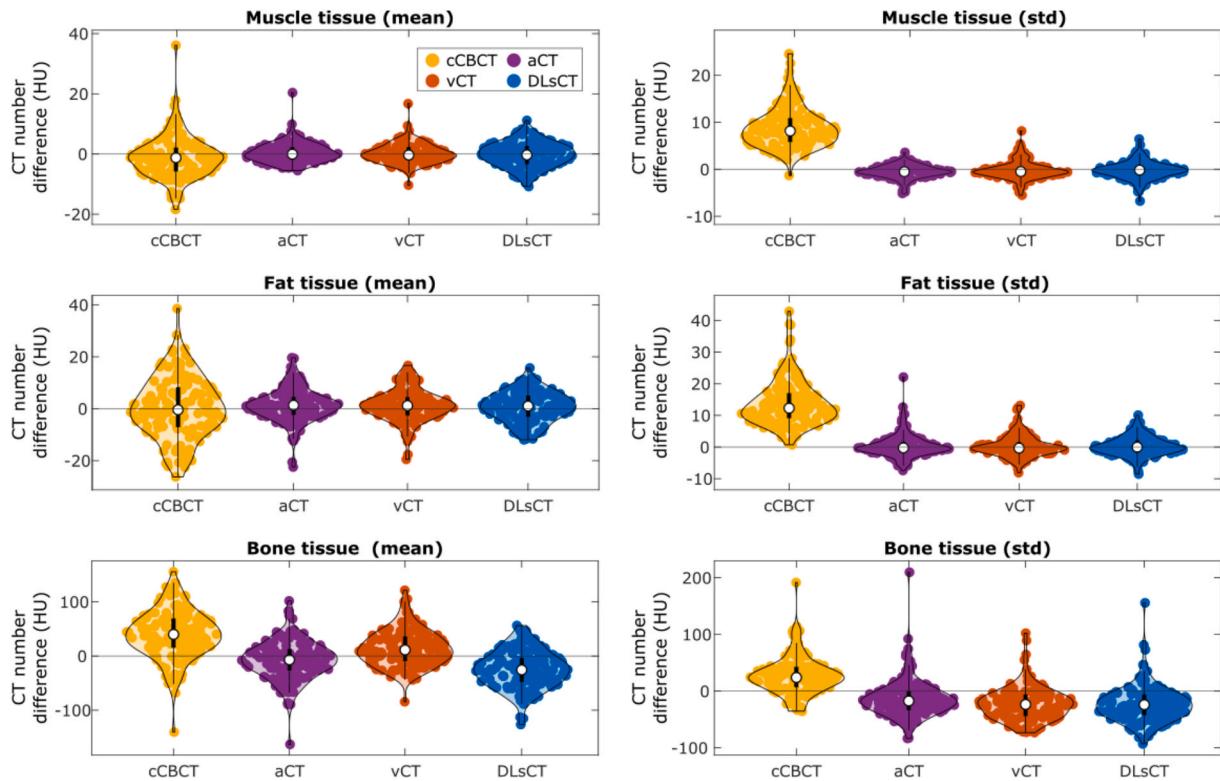


Fig. 3. Violin plots of the difference in mean CT number (left) and standard deviation (std; right) for the regions-of-interest placed in muscle (top row), fat (middle row), and bone (bottom row) tissues on the repeated CT and synthetic CTs. The white dot on each violin plot indicates the median value.

percentage of $\Delta R80$ values larger than 1 mm varied between 46% (DLsCT) and 55% (cCBCT).

The $\Delta R80$ between the individual sCT methods were smaller than those between the reCT and sCTs, with median absolute $\Delta R80$ ranging from 0.3 to 0.5 mm. Moreover, there were notably fewer $\Delta R80$ values exceeding 1 mm in the comparison between the individual sCT methods, ranging from 6% (DLsCT-cCBCT) to 17% (DLsCT-vCT). A few outliers

were though still seen, with a maximum $\Delta R80$ of 3.0 mm (DLsCT-aCT).

3.4. Dose evaluation

For all four sCT methods, the DVH differences relative to the reCT had median values close to zero, however, large outliers were found for all four methods for most of the evaluated DVH parameters (Fig. 4). The

Table 1

Median of absolute range differences, $\Delta R80$, maximum $\Delta R80$, and percentage of occurrences of $\Delta R80 > 1$ mm for the different image comparison pairs across the 3 single spot plans on each of the 35 images.

	cCBCT	aCT	vCT	DLSCT
	Median $\Delta R80$ (mm)			
reCT	1.1	1.1	1.0	1.0
cCBCT		0.4	0.5	0.3
aCT			0.4	0.5
vCT				0.5
	Maximum $\Delta R80$ (mm)			
reCT	4.5	5.3	5.4	4.6
cCBCT		2.5	1.9	1.4
aCT			2.5	3.0
vCT				1.9
	Number of occurrences of $\Delta R80 > 1$ mm (%)			
reCT	55	54	49	46
cCBCT		15	10	6
aCT			10	15
vCT				17

mean absolute values of the medians across the V95% ranged from 0.02 percentage points (pp; vCT) to 0.04 pp (cCBCT), while for the $D_{0.03cc}/D_{\text{mean}}$ it ranged from 0.6 Gy (aCT) to 0.7 Gy (cCBCT). The overall largest median difference was found for the lower PCM, ranging from -0.1 Gy (cCBCT) to -0.8 Gy (vCT). However, a large outlier was found for the CTV1 V95% for three out of four sCT methods, caused by missing tissue at the entry of one of the posterior oblique fields in the CBCT compared to the reCT (see Patient 2 in Fig. 1). This outlier was less pronounced for the cCBCT, likely due to its reduced image quality and inaccurate CT numbers.

4. Discussion

In this study, we compared sCTs generated using four different approaches (cCBCT, aCT, vCT, and DLSCT) to same-day reCTs to evaluate the feasibility of using the sCTs for dose re-calculation in proton therapy of head-and-neck cancer patients. The dose calculation accuracies were similar across all four sCT generation methods. Most outliers were present across all methods and originated from discrepancies between the CBCT and the same-day reCT (despite efforts to select reCTs with matching anatomy). This is consistent with the range analysis, which

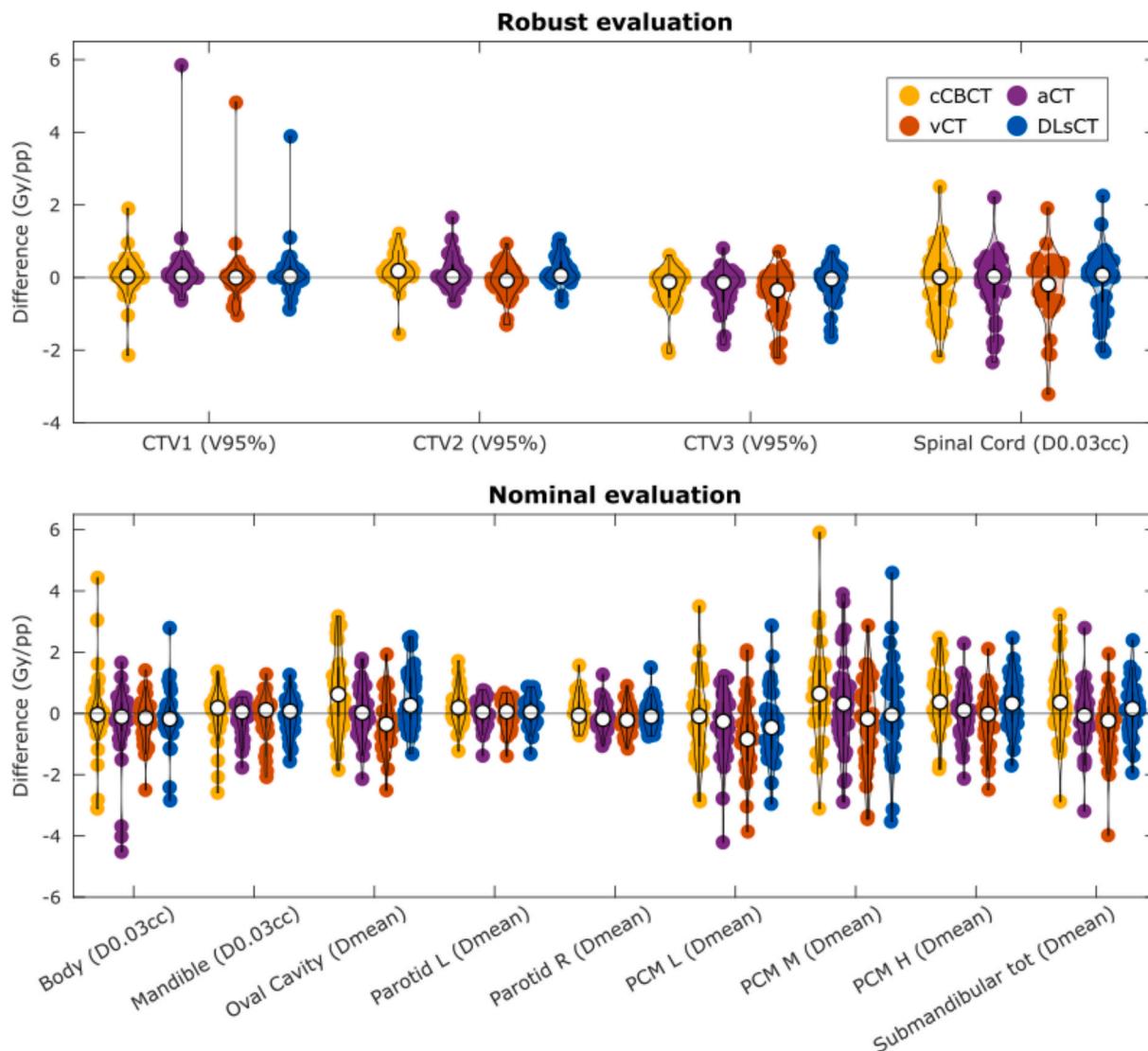


Fig. 4. Violin plots of the difference in dose-volume-histogram (DVH) parameters (in Gray (Gy) or percent point (pp)) for the worst-case scenario of the robust evaluation (0% and $\pm 3.5\%$ range uncertainty; top) and nominal scenario (bottom) re-calculated on the repeated CT and synthetic CTs. The white dot on each violin plot indicates the median value.

showed smaller differences between the individual sCTs than between any of the sCTs and the reCT. These findings highlight a key limitation of using reCT for dose monitoring, as anatomical and/or positional differences between treatment and reCT acquisition can impact the accuracy.

The four investigated sCT generation methods were found to have different advantages and limitations that potentially influence their use in a clinical workflow. The cCBCT method preserved anatomy and was fully automated (processing time < 1 minute), making it suitable for unsupervised daily dose monitoring. However, the high level of image noise reduced the CT number accuracy, impacting the dose calculation accuracy, and decreased the tissue contrast, complicating delineation. The aCT method on the other hand provided high image quality, with homogenous CT number patches resulting from manual CT number overrides to address larger anatomical deviations not captured by the DIR. However, the aCT method was highly user-dependent and time-consuming, with processing times ranging from 15 minutes to several hours. This limits its feasibility for daily dose monitoring. The vCT method offered high image quality, though small local regions of reduced quality were observed where the cCBCT data was automatically incorporated, e.g. in low-density regions (Fig. 2). Although the generation process was fully automated and fast (< 1 minute), anatomical inaccuracies observed in some cases suggest that manual evaluation and potential fine-tuning of the DIR may be necessary for clinical implementation [29–31]. Finally, the DLsCT method was fully automated (processing time < 2 minutes). It maintained good image quality and preserved anatomical features well, including cases with rare anatomical variations (e.g. nasogastric tube insertion; Fig. 2). However, minor residual artifacts could persist on the DLsCT (Patient 2, Fig. 1). As with all DL methods, performance depends on the quality and diversity of training data [32], but no failures were observed in this study.

Overall, our study highlights that while each method has its trade-offs, automated approaches hold promise for routine daily dose monitoring, provided their limitations are recognized and validated in a clinical setting. With continued advances in model robustness and generalizability, DL-based methods are likely to play the central role in future clinical workflows for daily CBCT-based dose monitoring.

As this study used real patient data, no ground truth images and contours were available, complicating the dose accuracy assessment. While some studies have used deformed CTs as ground truth [19,21], we avoided this to reduce uncertainties resulting from DIR and to compare to the standard clinical practices for dose monitoring, as in the study of Allen *et al.* [15]. As in their study, consistent structure sets were used across all images, to avoid variability from deformable propagation. While this limits anatomical accuracy and may have contributed to outliers (Fig. 4), it was intended to relate DVH deviations among the different sCT methods primarily to image differences rather than to structure variations [21]. Likewise, the CT number analysis was based on ROIs rigidly propagated from the reCT, which may have led to misalignment with the underlying anatomy in the sCTs and thereby contributed to some of the observed CT number differences.

Several studies have explored different methods for sCT generation in proton therapy. Pang *et al.* [33] applied different GAN-based DL approaches for proton dose monitoring in the head-and-neck region. Hu *et al.* [34] and Rusanov *et al.* [35] improved on the widely employed CycleGAN by incorporating vision transformers to tackle the difficulties of CycleGANs in capturing global features [36]. Recent studies have also explored alternatives to the GAN-based approaches, in the form of diffusion models, with promising results [37,38]. Yeap *et al.* [19] and Chang *et al.* [39] explored the cCBCT and vCT methods in RayStation (as also used in this study). Both studies found the cCBCT to work best in the head-and-neck region but reported limitations for proton dose recalculations in other anatomical sites. Thummerer *et al.* [20] compared sCT generation using DL (u-net), DIR, and an earlier version of cCBCT in RayStation (v.7.99) in the head-and-neck region. They found similar proton dose calculation accuracy between their DL- and DIR-based

approaches, while slightly worse performance was found for the earlier version of the cCBCT.

Commercial sCT generation methods are currently available in treatment planning systems [40]. Moreover, numerous in-house DL methods have been proposed and clinically validated for dose monitoring, and workflows for integrating such methods in treatment planning systems are available [41]. However, accurate contouring of targets and organs-at-risk on the sCTs is needed for their use in daily dose monitoring. While fast contouring approaches have been developed for CT data [42–45], their performance on sCTs generated by different methods requires further investigation.

In conclusion, all four investigated sCT generation methods were found to have distinct advantages and disadvantages regarding image quality and anatomy preservation. However, when compared to the reCTs, all four sCT methods demonstrated similar accuracy in proton range estimation and dose calculations, making them viable for proton dose monitoring.

Declaration of competing interest

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

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Marie Skłodowska-Curie Actions under Grant Agreement No. 955956 (NV), the Department of Clinical Medicine, Aarhus, Denmark (CDV), and from the Danish Cancer Society, Copenhagen, Denmark (LPM; grant no.: KBVU-BK R325-A19098).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.phro.2025.100784>.

References

- [1] Durante M, Loeffler JS. Charged particles in radiation oncology. Nat Rev Clin Oncol 2010;7:37–43. <https://doi.org/10.1038/nrclinonc.2009.183>.
- [2] Terasawa T, Dvorak T, Ip S, Raman G, Lau J, Trikalinos TA. Systematic review: charged-particle radiation therapy for cancer. Ann Intern Med 2009;151:556–65. <https://doi.org/10.7326/0003-4819-151-8-200910200-00145>.
- [3] Teoh S, Fiorini F, George B, Vallis KA, Van den Heuvel F. Proton vs photon: A model-based approach to patient selection for reduction of cardiac toxicity in locally advanced lung cancer. Radiother Oncol 2020;152:151–62. <https://doi.org/10.1016/j.radonc.2019.06.032>.
- [4] Szeto YZ, Witte MG, van Kranen SR, Sonke J-J, Belderbos J, van Herk M. Effects of anatomical changes on pencil beam scanning proton plans in locally advanced NSCLC patients. Radiother Oncol 2016;120:286–92. <https://doi.org/10.1016/j.radonc.2016.04.002>.
- [5] Nenoff L. Daily adaptive proton therapy: challenges and clinical potential. Doctoral Thesis. ETH Zurich, 2020. Doi: 10.3929/ethz-b-000462595.
- [6] Caciceda J, Casquero F, Martinez-Indart L, del Hoyo O, Gomez de Iturriaga A, Navarro A, et al. A prospective analysis of factors that influence weight loss in patients undergoing radiotherapy. Chin. J Cancer 2014;33:204–10. <https://doi.org/10.5732/cjc.013.10009>.
- [7] van Beek S, Jonker M, Hamming-Vrieze O, Al-Mamgani A, Navran A, Remeijer P, et al. Protocolised way to cope with anatomical changes in head & neck cancer during the course of radiotherapy. Tech Innov Patient Support Radiat Oncol 2019; 12:34–40. <https://doi.org/10.1016/j.tipsro.2019.11.001>.
- [8] Chao M, Xie Y, Moros EG, Le Q-T, Xing L. Image-based modeling of tumor shrinkage in head and neck radiation therapy. Med Phys 2010;37:2351–8. <https://doi.org/10.1118/1.3399872>.
- [9] Miracle AC, Mukherji SK. Conebeam CT of the head and neck, part 1: physical principles. AJNR Am J Neuroradiol 2009;30:1088–95. <https://doi.org/10.3174/ajnr.A1653>.
- [10] Fotina I, Hopfgartner J, Stock M, Steininger T, Lütgendörff-Caucig C, Georg D. Feasibility of CBCT-based dose calculation: comparative analysis of HU adjustment techniques. Radiother Oncol 2012;104:249–56. <https://doi.org/10.1016/j.radonc.2012.06.007>.

- [11] 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>.
- [12] Kidar HS, Azizi H. Enhancement of Hounsfield unit distribution in cone-beam CT images for adaptive radiation therapy: Evaluation of a hybrid correction approach. *Phys Med* 2020;69:269–74. <https://doi.org/10.1016/j.ejmp.2020.01.002>.
- [13] Marchant TE, Moore CJ, Rowbottom CG, MacKay RI, Williams PC. Shading correction algorithm for improvement of cone-beam CT images in radiotherapy. *Phys Med Biol* 2008;53:5719. <https://doi.org/10.1088/0031-9155/53/20/010>.
- [14] Elström UV, Wysocka BA, Muren LP, Petersen JBB, Grau C. Daily kV cone-beam CT and deformable image registration as a method for studying dosimetric consequences of anatomic changes in adaptive IMRT of head and neck cancer. *Acta Oncol* 2010;49:1101–8. <https://doi.org/10.3109/0284186X.2010.500304>.
- [15] Allen C, Yeo AU, Hardcastle N, Franich RD. Evaluating synthetic computed tomography images for adaptive radiotherapy decision making in head and neck cancer. *Phys Imaging Radiat Oncol* 2023;27:100478. <https://doi.org/10.1016/j.phro.2023.100478>.
- [16] Kurz C, Kamp F, Park Y-K, Zöllner C, Rit S, Hansen D, et al. Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive IMPT. *Med Phys* 2016;43:5635. <https://doi.org/10.1118/1.4962933>.
- [17] Landry G, Nijhuis R, Dedes G, Handrack J, Thieke C, Janssens G, et al. Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation. *Med Phys* 2015;42:1354–66. <https://doi.org/10.1118/1.4908223>.
- [18] Veiga C, Janssens G, Teng C-L, Baudier T, Hotoiu L, McClelland JR, et al. First Clinical Investigation of Cone Beam Computed Tomography and Deformable Registration for Adaptive Proton Therapy for Lung Cancer. *Int J Radiat Oncol Biol Phys* 2016;95:549–59. <https://doi.org/10.1016/j.ijrobp.2016.01.055>.
- [19] Yeap PL, Wong YM, Lee KH, Koh CWY, Lew KS, Chua CGA, et al. A treatment-site-specific evaluation of commercial synthetic computed tomography solutions for proton therapy. *Phys Imaging Radiat Oncol* 2024;31:100639. <https://doi.org/10.1016/j.phro.2024.100639>.
- [20] Thummerer A, Zaffino P, Meijers A, Marmitt GG, Seco J, Steenbakkers RJHM, et al. Comparison of CBCT based synthetic CT methods suitable for proton dose calculations in adaptive proton therapy. *Phys Med Biol* 2020;65:095002. <https://doi.org/10.1088/1361-6560/ab7d54>.
- [21] Vestergaard CD, Elstrom UV, Muren LP, Ren J, Nørrevang O, Jensen K, et al. Proton dose calculation on cone-beam computed tomography using unsupervised 3D deep learning networks. *Phys Imaging Radiat Oncol* 2024;32:100658. <https://doi.org/10.1016/j.phro.2024.100658>.
- [22] Kurz C, Maspero M, Savenije MHF, Landry G, Kamp F, Pinto M, et al. CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation. *Phys Med Biol* 2019;64:225004. <https://doi.org/10.1088/1361-6560/ab4d8c>.
- [23] Jensen K, Friborg J, Hansen C, Samsoe E, Johansen J, Andersen M, et al. The Danish Head and Neck Cancer Group (DAHANCA) 2020 radiotherapy guidelines. *Radiother Oncol* 2020;151:149–51. <https://doi.org/10.1016/j.radonc.2020.07.037>.
- [24] Hamming VC, Andersson S, Maduro JH, Langendijk JA, Both S, Sijtsema NM. Daily dose evaluation based on corrected CBCTs for breast cancer patients: accuracy of dose and complication risk assessment. *Radiat Oncol* 2022;17:205. <https://doi.org/10.1186/s13014-022-02174-4>.
- [25] Kadoya N, Fujita Y, Katsuta Y, Dobashi S, Takeda K, Kishi K, et al. Evaluation of various deformable image registration algorithms for thoracic images. *J Radiat Res* 2014;55:175–82. <https://doi.org/10.1093/jrr/rrt093>.
- [26] Thing RS, Nilsson R, Andersson S, Berg M, Lund MD. Evaluation of CBCT based dose calculation in the thorax and pelvis using two generic algorithms. *Phys Med* 2022;103:157–65. <https://doi.org/10.1016/j.ejmp.2022.10.012>.
- [27] Modat M, Cash DM, Daga P, Winston GP, Duncan JS, Ourselin S. Global image registration using a symmetric block-matching approach. *J Med Imaging* 2014;1:024003. <https://doi.org/10.1117/1.JMI.1.2.024003>.
- [28] Schuemann J, Dowdell S, Grassberger C, Min CH, Paganetti H. Site-specific range uncertainties caused by dose calculation algorithms for proton therapy. *Phys Med Biol* 2014;59:4007–31. <https://doi.org/10.1088/0031-9155/59/15/4007>.
- [29] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. *Med Phys* 2017;44:e43–76. <https://doi.org/10.1002/mp.12256>.
- [30] Hussein M, Akintonde A, McClelland J, Speight R, Clark CH. Clinical use, challenges, and barriers to implementation of deformable image registration in radiotherapy – the need for guidance and QA tools. *Br J Radiol* 2021;94:20210001. <https://doi.org/10.1259/bjr.20210001>.
- [31] Nenoff L, Amstutz F, Murr M, Archibald-Heeren B, Fusella M, Hussein M, et al. Review and recommendations on deformable image registration uncertainties for radiotherapy applications. *Phys Med Biol* 2023;68:24TR01. <https://doi.org/10.1088/1361-6560/ad0d8a>.
- [32] Munappy AR, Bosch J, Olsson HH, Arpteg A, Brinne B. Data management for production quality deep learning models: Challenges and solutions. *J Syst Softw* 2022;191:111359. <https://doi.org/10.1016/j.jss.2022.111359>.
- [33] Pang B, Si H, Liu M, Fu W, Zeng Y, Liu H, et al. Comparison and evaluation of different deep learning models of synthetic CT generation from CBCT for nasopharynx cancer adaptive proton therapy. *Med Phys* 2023;50:6920–30. <https://doi.org/10.1002/mp.16777>.
- [34] Hu Y, Zhou H, Cao N, Li C, Hu C. Synthetic CT generation based on CBCT using improved vision transformer CycleGAN. *Sci Rep* 2024;14:11455. <https://doi.org/10.1038/s41598-024-61492-7>.
- [35] 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;69:035014. <https://doi.org/10.1088/1361-6560/ad1fcf>.
- [36] Li X, Jiang Y, Rodriguez-Andina JJ, Luo H, Yin S, Kaynak O. When medical images meet generative adversarial network: recent development and research opportunities. *Discov Artif Intell* 2021;1:5. <https://doi.org/10.1007/s44163-021-00006-0>.
- [37] Chen X, Qiu RLJ, Peng J, Shelton JW, Chang C-W, Yang X, et al. CBCT-based synthetic CT image generation using a diffusion model for CBCT-guided lung radiotherapy. *Med Phys* 2024;51:8168–78. <https://doi.org/10.1002/mp.17328>.
- [38] Peng J, Qiu RLJ, Wynne JF, Chang C-W, Pan S, Wang T, et al. CBCT-Based synthetic CT image generation using conditional denoising diffusion probabilistic model. *Med Phys* 2024;51:1847–59. <https://doi.org/10.1002/mp.16704>.
- [39] Chang C-W, Nilsson R, Andersson S, Bohannon D, Patel SA, Patel PR, et al. An optimized framework for cone-beam computed tomography-based online evaluation for proton therapy. *Med Phys* 2023;50:5375–86. <https://doi.org/10.1002/mp.16625>.
- [40] Janson M, Glimelius L, Fredriksson A, Traneus E, Engwall E. Treatment planning of scanned proton beams in RayStation. *Med Dosim* 2024;49:2–12. <https://doi.org/10.1016/j.meddos.2023.10.009>.
- [41] Rasmussen ME, Vestergaard CD, Kallehauge JF, Ren J, Guldberg MH, Nørrevang O, et al. RadDeploy: A framework for integrating in-house developed software and artificial intelligence models seamlessly into radiotherapy workflows. *Phys Imaging Radiat Oncol* 2024;31:100607. <https://doi.org/10.1016/j.phro.2024.100607>.
- [42] Barber J, Yuen J, Jameson M, Schmidt L, Sykes J, Gray A, et al. Deforming to Best Practice: Key considerations for deformable image registration in radiotherapy. *J Med Radiat Sci* 2020;67:318–32. <https://doi.org/10.1002/jmrs.417>.
- [43] Ren J, Hochreiter K, Rasmussen ME, Kallehauge JF, Korreman SS. Gradient Map-Assisted Head and Neck Tumor Segmentation: A Pre-RT to Mid-RT Approach in MRI-Guided Radiotherapy 2024. <https://doi.org/10.48550/arXiv.2410.12941>.
- [44] Ren J, Teuwen J, Nijkamp J, Rasmussen ME, Gouw Z, Eriksen JG, et al. Enhancing the reliability of deep learning-based head and neck tumour segmentation using uncertainty estimation with multi-modal images. *Phys Med Biol* 2024;69:165018. <https://doi.org/10.1088/1361-6560/ad682d>.
- [45] Luan S, Ding Y, Shao J, Zou B, Yu X, Qin N, et al. Deep learning for head and neck semi-supervised semantic segmentation. *Phys Med Biol* 2024;69:055008. <https://doi.org/10.1088/1361-6560/ad25c2>.