Deep learning-enabled MRI-only photon and proton therapy treatment planning for paediatric abdominal tumours

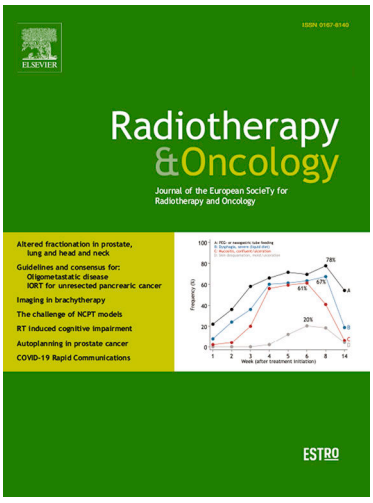
Florkow, Mateusz C; Guerreiro, Filipa; Zijlstra, Frank; Seravalli, Enrica; Janssens, Geert O; Maduro, John H; Knopf, Antje C; Castelein, René M; van Stralen, Marijn; Raaymakers, Bas W

Published in:
Radiotherapy and Oncology

DOI:
10.1016/j.radonc.2020.09.056

Publication date:
2020

Journal Pre-proofs



Original Article

Deep learning-enabled MRI-only photon and proton therapy treatment planning for paediatric abdominal tumours

Mateusz C. Florkow, Filipa Guerreiro, Frank Zijlstra, Enrica Seravalli, Geert O. Janssens, John H. Maduro, Antje C. Knopf, René M. Castelein, Marijn van Stralen, Bas W. Raaymakers, Peter R. Seevinck

PII: S0167-8140(20)30831-8
DOI: https://doi.org/10.1016/j.radonc.2020.09.056
Reference: RADION 8561

To appear in: *Radiotherapy and Oncology*

Received Date: 18 May 2020
Revised Date: 24 September 2020
Accepted Date: 28 September 2020

Please cite this article as: Florkow, M.C., Guerreiro, F., Zijlstra, F., Seravalli, E., Janssens, G.O., Maduro, J.H., Knopf, A.C., Castelein, R.M., van Stralen, M., Raaymakers, B.W., Seevinck, P.R., Deep learning-enabled MRI-only photon and proton therapy treatment planning for paediatric abdominal tumours, *Radiotherapy and Oncology* (2020), doi: https://doi.org/10.1016/j.radonc.2020.09.056

Deep learning-enabled MRI-only photon and proton therapy treatment planning for paediatric abdominal tumours

Mateusz C. Florkow[a,1,*], Filipa Guerreiro[b,1,*], Frank Zijlstra[a], Enrica Seravalli[b], Geert O. Janssens[c,d], John H. Maduro[e], Antje C. Knopf[e], René M. Castelein[f], Marijn van Stralen[a,g], Bas W. Raaymakers[b], Peter R. Seevinck[a,g]

[a]Image Sciences Institute, University Medical Centre Utrecht, Utrecht, The Netherlands

[b]Department of Radiotherapy, University Medical Centre Utrecht, Utrecht, The Netherlands

[c]Department of Radiation Oncology, University Medical Centre Utrecht, Utrecht, The Netherlands

[d]Princess Máxima Centre for Paediatric Oncology, Utrecht, The Netherlands

[e]Department of Radiation Oncology, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands

[f]Department of Orthopaedics, University Medical Centre Utrecht, Utrecht, The Netherlands

[g]MRIguidance B.V., Utrecht, The Netherlands

[*]Corresponding authors at: University Medical Center Utrecht 3508 GA Utrecht, The Netherlands

[1]authors contributed equally to this work

Graphical abstract

ga1

Highlights

- Satisfactory synthetic CT images were derived from planning T1w and T2w MR images.
- Deep learning-based MRI-only radiotherapy is feasible in pediatric abdominal tumors.
- CT-sCT dose differences were clinically acceptable (<2%) for photon & proton plans.
- Larger differences were caused by existing interscan differences (eg bowel filling)

Abstract

Purpose

To assess the feasibility of magnetic resonance imaging (MRI)-only treatment planning for photon and proton radiotherapy in children with abdominal tumours.

Materials and methods

The study was conducted on 66 paediatric patients with Wilms' tumour or neuroblastoma (age 4±2 years) who underwent MR and computed tomography (CT) acquisition on the same day as part of the clinical protocol. MRI intensities were converted to CT Hounsfield units (HU) by means of a UNet-like neural network trained to generate synthetic CT (sCT) from T1- and T2-weighted MR images. The CT-to-sCT image similarity was evaluated by computing the mean error (ME), mean absolute error (MAE), peak signal-to-noise ratio (PSNR) and Dice score coefficient (DSC). Synthetic CT dosimetric accuracy was verified against CT-based dose distributions for volumetric-modulated arc therapy (VMAT) and intensity-modulated pencil-beam scanning (PBS). Relative dose differences (D_{diff}) in the internal target volume and organs-at-risk were computed and a three-dimensional gamma analysis (2mm, 2%) was performed.

Results

The average ± standard deviation ME was -5±12 HU, MAE was 57±12 HU, PSNR was 30.3±1.6 dB and DSC was 76±8% for bones and 92±9% for lungs. Average D_{diff} were <0.5% for both VMAT (range [-2.5;2.4]%) and PBS (range [-2.7;3.7]%) dose distributions. The average gamma pass-rates were >99% (range [85;100]%) for VMAT and >96% (range [87;100]%) for PBS.

Conclusion

The deep learning-based model generated accurate sCT from planning T1w- and T2w-MR images. Most dosimetric differences were within clinically acceptable criteria for photon and proton radiotherapy, demonstrating the feasibility of an MRI-only workflow for paediatric patients with abdominal tumours.

Keywords

Synthetic CT

MRI only

Paediatric

Abdomen

Deep learning

Proton therapy

Photon therapy

Planning

MRI

Wilms Tumour

Neuroblastoma

Introduction

Wilms' tumour (WT) and neuroblastoma (NBL) are two of the most common abdominal solid tumours diagnosed in children. Depending on the risk stratification, the treatment involves a combination of surgery, chemotherapy and radiotherapy [1,2]. During radiotherapy treatment planning (RTP), magnetic resonance imaging (MRI) scans are commonly acquired to define the target volume thanks to their superior soft-tissue contrast [3]. In addition, computed tomography (CT) images are used for the computation of patient-specific dose deposition maps [4] using the conversion of the CT Hounsfield units (HU) to relative electron density for photon therapy or to stopping power ratio relative to water for proton therapy.

In the last decade, MRI-only workflows have been developed using MRI-based synthetic CT (sCT) images to calculate the dose deposition [5–7]. Such workflows are especially useful for anatomies in which MRI-CT registration for contour propagation is difficult [8]. By omitting the CT acquisition, systematic uncertainties arising from the MRI-CT registration are eliminated and logistics are simplified in the treatment preparation phase.

In the last few years, the use of sCT has been reported for RTP of the brain [9], head and neck [10], lungs [11], prostate [12,13] and abdomen [14,15] in adult patients. Specifically in the abdomen, sCT images have been investigated for the treatment of liver tumours using photon [14] and proton [15] therapies. For the paediatric population, studies reporting sCT generation methods are scarce [16,17], with only one approach generating sCT in children with abdominal tumours [16]. Paediatric patients are certainly challenging because of their substantial inter-patient morphological variability. Nevertheless, they would potentially benefit the

most from the simplified MRI-only workflows, with a reduced time under anaesthesia during treatment preparation.

In this study, we investigated the feasibility of an MRI-only treatment planning workflow for children with abdominal tumours using an end-to-end deep learning approach to generate sCT from planning T1- and T2-weighted MR images. Dose calculations on the sCT were compared against CT-based dose distributions for volumetric-modulated arc therapy (VMAT) and intensity-modulated pencil beam scanning (PBS).

Materials and methods

This retrospective study was performed in accordance with the institutional review board (WAG/mb/17/008865).

Data collection

Images of 66 children, treated for WT (n=24) or NBL (n=42) at the Radiotherapy Department of the University Medical Centre Utrecht (UMCU) between April 2015 and July 2020, were collected for this study. Patients had a mean age (±standard deviation (SD)) of 4±2 years old (range: [1; 9] years) at the time of scanning and the female-to-male ratio was 28/38.

For treatment preparation, patients underwent four-dimensional (4D)-CT and MRI scanning less than an hour apart. During the image acquisition, 54/66 patients were under general anaesthesia and all patients were fixated in supine position in an individualized vacuum mattress (Bluebag, Elekta, Stockholm, Sweden) with the arms wide along the body.

- **Acquisition parameters**

The 4D-CTs were acquired by means of a respiratory belt (Philips Bellow System, Philips Medical System, Best, The Netherlands) from a 16-, 40-, or 64-channel detector scanner (Brillance, Philips Medical Systems, Best, The Netherlands) with

an in-plane isotropic resolution ranging from 0.8 to 1.2 mm and a slice spacing between 2 and 3 mm. The tube voltage varied between 90 and 120 kV and the tube current between 30 and 250 mA. The 4D-CT was acquired in ten phases which were averaged to obtain the planning-CT as per our institution clinical practice.

The MRI acquisition, performed on a 1.5T scanner (Achieva, Philips Medical Systems, Best, The Netherlands), included a T1-weighted (T1w) 3D gradient-echo sequence and a T2-weighted (T2w) 3D turbo spin echo sequence. T1w-MR images were obtained in the axial plane at a resolution of 0.7x0.7x1.5 mm^3. The echo/repetition times were 2.7 ms/5.4 ms with a flip angle of 10°. T2w-MR images were acquired in the coronal plane at a resolution of 0.8x0.8x1.1 mm^3. The echo/repetition times were 90 ms/454 ms with a flip angle of 90°. Respiratory motion artefacts were reduced during MRI using the phase encoding artefact reduction (PEAR) method. Both MRI acquisitions were part of the clinical protocol and were not optimized for sCT generation.

Data processing

- **Pre-processing**

MR images were independently and non-rigidly registered to the planning-CT using the Elastix toolbox [18] in order to get the best voxelwise match between modalities. The registration was a composition of a translation, an Euler transform and a cubic B-spline, optimized using mutual information and regularized with a rigidity penalty [19]. During registration, MR images were resampled to the planning-CT resolution using a cubic B-spline interpolation.

For the sCT generation only, MR and planning-CT images were normalized independently and per volume. MR intensities were clipped beyond the 95th percentile and the resulting intensities were linearly mapped to [-1; 1]. CT intensities were linearly mapped from [-1024; 3071] to [-1; 1].

- **Synthetic CT generation**

The proposed end-to-end deep learning sCT generation approach is an extension of a previously reported model [20], which used a neural network derived from a U-Net [21]. The network used 3D patches from the registered T1w- and T2w-MR images as inputs to generate sCT patches of the same resolution as the planning-CT images.

The network was trained simultaneously on both T1w- and T2w- MR images since it proved superior to models trained only on T1w- or T2w- images (Supplementary material Table 1 tbl1). Thus, the model presented in [20] was slightly modified to first process the T1w- and T2w-MR images separately with two convolution layers. The information extracted from the two images was concatenated before starting the standard encoding path of the U-Net, as described in Figure 1 fig1.

During the training phase, the generated sCT images were compared to the planning-CT using a L1 loss (absolute voxelwise difference). The resulting error was minimized using a Nadam [22] optimizer, given a learning rate of 10^{-4}. The training was stopped after 100,000 iterations. The remaining parameters were the same as in [20]. Synthetic CT images were generated in two phases. First, a 3-fold cross-validation (CV) was performed on 54 patients with 18 patients set aside for testing in each fold. Then a model was trained on the 54 patients of the CV set and evaluated on a fully independent test set consisting of 12 patients.

Treatment planning

For both WT and NBL patients, the clinical target and organs-at-risk (OARs) delineated by a radiation oncologist were used in this study. The gross tumour volume (GTV) consisted of the pre-operative tumour extension, including pathologic lymph nodes and residual disease if applicable. The clinical target volume was created by expanding the GTV by 5 mm for NBL patients and by 10

mm for WT patients. To account for the respiratory motion, an internal target volume (ITV) was delineated for each patient by using the minimum and maximum 4D-CT phases and surgical clips as motion surrogates [23]. OARs included the kidneys (divided into ipsilateral and contralateral kidneys), the liver and the spleen.

RTP was performed in RayStation (RaySearch Laboratories, Stockholm, Sweden) for both VMAT and PBS dose distributions. In accordance with the clinical protocol at the UMCU, VMAT plans consisted of a 6MV full-arc. PBS plans were optimized for the purpose of this study using 2-3 posterior-oblique irradiation fields to exploit the dorsal and superficial location of the target. The number and direction of the proton beams (range [120^{o}, 240^{o}]) were patient-specific. The prescribed dose ranged from 10.8 to 36 Gy and was delivered in 6 to 20 fractions, depending on the histology and the presence of residual disease.

3D plan optimization was performed using the planning-CT, a uniform 3 mm dose grid, a collapse cone engine for VMAT and a pencil-beam algorithm for intensity-modulated PBS dose distributions. To ensure a fair dosimetric comparison, both dose distributions were ITV-based robustly optimized and evaluated accounting for a 5 mm patient set-up uncertainty and a 3% range uncertainty (only for PBS). In addition, a relative biological effectiveness of 1.1 was included during planning for PBS.

VMAT and PBS plans were optimized using a minimax optimization method [24] whilst its robustness was evaluated using a minimum evaluation dose map (V_{wmin}) [25]. Plans were considered robust if, in the V_{wmin}, 98% of the ITV received at least 95% of the prescribed dose [25]. In addition, to reduce the risk of asymmetric skeletal growth, a homogeneous dose avoiding left-right dose gradients higher than

5 Gy was aimed for the primary ossification centres of the vertebra volume adjacent to the ITV [26].

Evaluation

- **Technical evaluation**

Synthetic CT images were compared against the planning-CT using distance and structural metrics. The distance metrics consisted of voxelwise differences between the HU of the planning-CT and of the sCT, within the body contour, the soft-tissue, the bone and the lungs. In particular, mean error (ME) was computed to estimate any systematic bias in the sCT generation and mean absolute error (MAE) to estimate the overall error.

Structural metrics included Dice score coefficient (DSC) [27] and peak-signal-to-noise ratio (PSNR). DSC measured the overlap of bone and lungs between the planning-CT and sCT images. PSNR measured the reconstruction quality and was computed as:

$$PSNR = 10\log_N\left(\frac{4095^2}{1/N\sum_{i=1}^{N}\left(I_{CT}(i)-I_sCT(i)\right)^2}\right)$$

All metrics were computed only in regions where T1w-, T2w-MR and CT information was available. Regions where HU > 200 were labelled as bone and regions where -200 < HU < 200 were labelled as soft tissue. Lungs were segmented independently on the planning-CT and sCT images using a dedicated in-house clinical tool [28].

- **Dosimetric evaluation**

To evaluate the dosimetric acceptability of sCT, CT-based VMAT and PBS dose distributions were re-calculated on the sCT images.

Planning-CT and sCT dose differences were compared using dose-volume histogram (DVH) metrics and 3D global gamma analysis [29]. DVH metrics

included $D_{98\%}$, $D_{50\%}$, $D_{2\%}$ and $V_{95\%}$ for the ITV as well as mean dose (D_{mean}) and $D_{2\%}$ for the OARs. CT-sCT DVH dose differences in the target and OARs smaller than 2% at a 95% confidence interval were considered clinically acceptable, as proposed by Korsholm et al. [30]. 3D global gamma pass-rates were computed using the CT dose as reference, multiple dose thresholds (10%, 50%, 90%), a 2 mm distance to agreement and a 2% dose difference (2mm, 2%). Gamma pass-rates were calculated within a body mask excluding the arms to neglect CT-MR mismatch in this region on the results (eg. registration errors, arms not in the field of view of one image). In addition, the correlation between the DVH and gamma results and the use of anaesthesia during image acquisition was verified using a two-tailed student's t-test at a 0.05 significance level.

Results

For an example patient with a MR matrix of size 429x429x149, sCT images were generated under 200s on a GeForce GTX 1080 Ti (NVIDIA, California, USA) graphics processing unit.

Table 1 summarizes the average ME, MAE, DSC and PSNR obtained separately for both CV and test set patients and across the entire population. Notably, no differences were found between the CV and test set in terms of technical performance with metrics within the same range in both sets. Most errors occurred in the bone with HU largely underestimated on the sCT images, as indicated by the high ME. Large errors were also found in the lungs, despite a high overlap of their bulk structure between CT and sCT.

Figure 2 fig2, which compares planning-CT and sCT images for two patients, shows that voxelwise differences were partially caused by inter-scan differences. This includes differences originating from patient positioning during image acquisition (i.e skin outline, clavicle position), breathing motion (i.e volume and internal

structures of the lungs and ribs) and physiological changes (i.e bowel filling) that were not corrected by the registration.

Relative dose differences of clinically relevant DVH parameters between the planning-CT and sCT dose distributions are shown in box-and-whisker plots in Figure 3 fig3(a) for the ITV and in Figure 3(b)(c) for the OARs. Similarly to the technical performance, no dosimetric differences were found between the CV and test set results with DVH metrics within the same range in both sets.

For the ITV, DVH parameters differed on average by less than 0.5% (range [-2.1;1.9]%) for VMAT and by less than 0.1% (range [-1.1;1.2]%) for PBS dose distributions for all patients (Figure 3(a)). For the OARs, average DVH differences were under 0.3% for both VMAT and PBS dose distributions (Figure 3(b)(c)). D_{mean} differences ranged from [-1.1;0.6]% for VMAT and [-1.4;1.3]% for PBS. Similarly, $D_{2\%}$ differences ranged from [-2.5;2.4]% for VMAT and [-2.7;3.7]% for PBS.

Figure 4 fig4 shows relative dose difference maps for three example patients. For the VMAT dose distributions, larger dosimetric differences were seen due to (1) changes in the position of air cavities between the body outline and the ITV (Figure 4(b)(c)) and (2) changes in the CT-sCT body outline (Figure 4(c)). For the PBS dose distributions, because posterior irradiation fields were used, CT-sCT dose differences were predominantly a result of different air cavities distributions between images in the vicinity of the ITV (Figure 4(b)(c)).

Gamma pass-rates (2%, 2mm) obtained between the planning-CT and sCT dose distributions are given in Table 2 tbl2. Average gamma pass-rates were above 99% for VMAT and above 96% for PBS dose distributions. Individual gamma pass-

rates were lower than 90% for 2 patients on the VMAT and for 3 patients on the PBS dose distributions (example patients in Figure 4(b)(c)).

In addition, no significant statistical differences were found in the dosimetric DVH and gamma results between patients imaged with and without anaesthesia ($p>0.05$).

Discussion

In this study, the feasibility of a deep learning-based sCT generation method for photon and proton treatment planning in paediatric patients with abdominal tumours was evaluated. We presented an end-to-end method to generate sCT from planning T1w- and T2w-MR images. We found clinically acceptable differences between the planning–CT and sCT dose distributions for both photon and proton therapy on this morphologically heterogeneous dataset. To the authors' knowledge, the present study is the first evaluating the feasibility of performing dose calculations on deep-learning based sCT for children younger than 10 years old with abdominal tumours.

Facing the low number of patients treated per year and the ethical adversities within the children population to acquire imaging data, datasets including both CT and MR images in treatment position for over 60 patients in a single institute are rare. To thoroughly exploit the entire population, technical and dosimetric evaluation of the sCT images were performed on both a CV and a fully independent test sets. No differences were found between the CV and test sets in terms of technical and dosimetric performance confirming the generalization performance of the proposed network.

The technical accuracy of the sCT generated in this study, with average MAE_{body} of 57 HU, $MAE_{soft\ tisue}$ of 33 HU, MAE_{bone} of 158 HU and MAE_{lungs} of 105 HU, compared favourably with literature. A study in adult patients with hepatic cancer

[15] reported a MAE_{body} of 73±18 HU, a MAE_{bone} of 217±63 HU and a $MAE_{soft\ tissue}$ of 59±31 HU. A previous study, performed on a subset of this dataset [16], obtained a MAE_{bone} of 212±40 HU, a $MAE_{soft\ tissue}$ of 53±7 HU and a MAE_{lungs} of 125±24 HU using a combined atlas- and voxel-based method. Although performed on similar anatomies, these two studies had different MRI-to-CT registration techniques, using rigid [16] or non-rigid [15] registration, which make them difficult to compare. In the present study, we made the choice to non-rigidly register the planning-CT and MR images aiming for an optimal voxelwise match between modalities while minimizing distortions with a rigidity penalty. Therefore, dose differences related to registration errors were minimized to focus on the HU differences between CT and sCT. Note that for a prospective clinical use of the model, only a standard rigid registration between the T1w- and T2w-MR images would be required. Despite the non-rigid registration, registration errors were still observed, especially for the body outline and for the bones and lungs, which resulted in blurring [31].

The dosimetric accuracy of the generated sCT was evaluated for VMAT and PBS dose distributions optimized using the planning-CT. Clinically acceptable dose differences (≤2%) were obtained for all patients on the ITV and for 61/66 patients on the OARs for both VMAT and PBS. Differences larger than 2% were only obtained for the $D_{2\%}$ point only. Nevertheless, these $D_{2\%}$ dose differences are not clinically significant since the tolerance doses for these OARs are far from being reached using the clinical prescribed doses. For the ITV, larger differences between the planning-CT and sCT doses were detected for the VMAT dose distributions due to bowel filling variations and body outline differences between images (Figure 4). A potential reduction of these differences could be expected if a posterior partial arc and/or air density override during optimization of the VMAT

dose distributions were used. For PBS, results showed that when using posterior-oblique irradiation fields, the differences between planning-CT and sCT dose distributions on the ITV were minimized (Figure 4). For the OARs, larger dosimetric differences were observed for both VMAT and PBS dose distributions especially for the liver and spleen as a result of their location in the vicinity of the lungs. At soft tissue-to-air interfaces (i.e lungs, skin), higher HU variations can arise due to misalignments introduced by motion or registration errors between the planning-CT and MR images. Concerning the skin outline, previous studies [16,32,33] have applied the skin outline extracted from the planning-CT on the sCT to correct for these misalignments and lower the CT-to-sCT dose difference. In this study, the planning-CT outline information was not used for the dose re-calculation on the sCT images. The present CT-sCT skin outline misalignment might have led to an overestimation of the observed dosimetric differences.

In terms of gamma pass-rates, both VMAT and PBS showed clinically acceptable results with average gamma pass-rates over 96% for a (2%, 2mm) criterion. Photon gamma pass-rates with a (2%, 2mm) criterion and a 10% dose threshold obtained in this study (99.7±0.5%) compared favourably with the literature, reporting pass-rates of 97.7±2.2% for the same patient category [16]. Similarly, proton gamma pass-rates in this study (97.2±3.1%) were in agreement with values reported in literature with pass-rates of 97.0±2.9% for liver cancer treatment [15] and of 87.1±5.4% for the same patient category [16]. Nevertheless, for a reduced number of patients, gamma pass-rates were below 90% for both VMAT (n=2) and PBS (n=3) potentially due to (1) target size and location (i.e at the lungs interface) and (2) inter-scan differences (i.e body oultine, bowel filling). Because of these inter-scan differences, which would not be present in a clinical setting, the dosimetric

differences showed in this study might be interpreted as an overestimation of the real difference between the planning-CT and sCT.

Overall, the results of this study demonstrated that MRI-only photon and proton dose calculations on the generated sCT images were clinically feasible, even when using MR images not optimized for the sCT generation. The deep learning methodology used to generate sCT enabled the creation of a model which coped with the morphological variability of the paediatric population. With an application in the abdominal region, in which MRI-CT registration is particularly hard, this method is promising for radiotherapy treatment planning and would ease the clinical workflow for paediatric population whilst potentially improving its accuracy.

Funding

Support of NVIDIA Corporation with the donation of GPU for this research.
This work is part of the research program Applied and Engineering Sciences (TTW) with project number 15479 which is (partly) financed by the Netherlands Organization for Scientific Research (NWO).

Acknowledgements

The authors would like to acknowledge the support of NVIDIA Corporation with the donation of GPU for this research. This work is part of the research program Applied and Engineering Sciences (TTW) with project number 15479 which is (partly) financed by the Netherlands Organization for Scientific Research (NWO).

References

[1] Van Den Heuvel-Eibrink MM, Hol JA, Pritchard-Jones K, Tinteren H Van,

Furtwängler R, Verschuur AC, et al. Position Paper: Rationale for the treatment of Wilms tumour in the UMBRELLA SIOP-RTSG 2016 protocol. Nat Rev Urol 2017;14:743–52. https://doi.org/10.1038/nrurol.2017.163.

[1] M.M. van den Heuvel-Eibrink J.A. Hol K. Pritchard-Jones H. van Tinteren R. Furtwängler A.C. Verschuur G.M. Vujanic I. Leuschner J. Brok C. Rübe A.M. Smets G.O. Janssens J. Godzinski G.L. Ramírez-Villar B. de Camargo H. Segers P. Collini M. Gessler C. Bergeron F. Spreafico N. Graf Rationale for the treatment of Wilms tumour in the UMBRELLA SIOP–RTSG 2016 protocol Nat Rev Urol 14 12 2017 743 752 http://www.nature.com/articles/nrurol.2017.163

[2] Ladenstein R, Pötschger U, Pearson ADJ, Brock P, Luksch R, Castel V, et al. Busulfan and melphalan versus carboplatin, etoposide, and melphalan as high-dose chemotherapy for high-risk neuroblastoma (HR-NBL1/SIOPEN): an international, randomised, multi-arm, open-label, phase 3 trial. Lancet Oncol 2017;18:500–14. https://doi.org/10.1016/S1470-2045(17)30070-0.

[2] R. Ladenstein U. Pötschger A.D.J. Pearson P. Brock R. Luksch V. Castel I. Yaniv V. Papadakis G. Laureys J. Malis W. Balwierz E. Ruud P. Kogner H. Schroeder A.F. de Lacerda M. Beck-Popovic P. Bician M. Garami T. Trahair A. Canete P.F. Ambros K. Holmes M. Gaze G. Schreier

A. Garaventa G. Vassal J. Michon D. Valteau-Couanet Busulfan and melphalan versus carboplatin, etoposide, and melphalan as high-dose chemotherapy for high-risk neuroblastoma (HR-NBL1/SIOPEN): an international, randomised, multi-arm, open-label, phase 3 trial The Lancet Oncology 18 4 2017 500 514

https://linkinghub.elsevier.com/retrieve/pii/S1470204517300700

[3] Lagendijk JJW, Raaymakers BW, Van Den Berg CAT, Moerland MA, Philippens ME, Van Vulpen M. MR guidance in radiotherapy. Phys Med Biol 2014;59:R349–69. https://doi.org/10.1088/0031-9155/59/21/R349.

[3] J.J.W. Lagendijk B.W. Raaymakers C.A.T. Van den Berg M.A. Moerland M.E. Philippens M. van Vulpen MR guidance in radiotherapy Phys. Med. Biol. 59 21 2014 R349 R369

https://iopscience:iop.org/article/10.1088/0031-9155/59/21/R349

[4] Chernak ES, Rodriguez-Antunez A, Jelden GL, Dhaliwal RS, Lavik PS. The use of computed tomography for radiation therapy treatment planning. Radiology 1975;117:613–4. https://doi.org/10.1148/117.3.613.

[4] E.S. Chernak A. Rodriguez-Antunez G.L. Jelden R.S. Dhaliwal P.S. Lavik The Use of Computed Tomography for Radiation Therapy Treatment Planning Radiology 117 3 1975 613 614

[5] Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. Radiat Oncol 2017;12:28. https://doi.org/10.1186/s13014-016-0747-y.

[5] J.M. Edmund T. Nyholm A review of substitute CT generation for MRI-only radiation therapy Radiat Oncol 12 1 2017 10.1186/s13014-016-0747-y http://ro-journal.biomedcentral.com/articles/10.1186/s13014-016-0747-y

[6] Tenhunen M, Korhonen J, Kapanen M, Seppälä T, Koivula L, Collan J, et al. MRI-only based radiation therapy of prostate cancer: workflow and early clinical experience. Acta Oncol (Madr) 2018;57:902–7. https://doi.org/10.1080/0284186X.2018.1445284.

[6] M. Tenhunen J. Korhonen M. Kapanen T. Seppälä L. Koivula J. Collan K. Saarilahti H. Visapää MRI-only based radiation therapy of prostate cancer: workflow and early clinical experience Acta Oncologica 57 7 2018 902 907

[7] Hoffmann A, Oborn B, Moteabbed M, Yan S, Bortfeld T, Knopf A, et al. MR-guided proton therapy: A review and a preview. Radiat Oncol 2020;15:129.

https://doi.org/10.1186/s13014-020-01571-x.

[7] A. Hoffmann B. Oborn M. Moteabbed S. Yan T. Bortfeld A. Knopf H. Fuchs D. Georg J. Seco M.F. Spadea O. Jäkel C. Kurz K. Parodi MR-guided proton therapy: a review and a preview Radiat Oncol 15 1 2020 10.1186/s13014-020-01571-x https://ro-journal.biomedcentral.com/articles/10.1186/s13014-020-01571-x

[8] Jonsson J, Nyholm T, Söderkvist K. The rationale for MR-only treatment planning for external radiotherapy. Clin Transl Radiat Oncol 2019;18:60–5. https://doi.org/10.1016/j.ctro.2019.03.005.

[8] J. Jonsson T. Nyholm K. Söderkvist The rationale for MR-only treatment planning for external radiotherapy Clinical and Translational Radiation Oncology 18 2019 60 65 https://linkinghub.elsevier.com/retrieve/pii/S2405630819300552

[9] Kazemifar S, Mcguire S, Timmerman R, Wardak Z, Nguyen D, Park Y, et al. MRI-only brain radiotherapy: Assessing the dosimetric accuracy of synthetic CT images generated using a deep learning approach. Radiother Oncol 2019;136:56–63. https://doi.org/10.1016/j.radonc.2019.03.026.

[9] S. Kazemifar S. McGuire R. Timmerman Z. Wardak D. Nguyen Y. Park S. Jiang A. Owrangi MRI-only brain radiotherapy: Assessing the dosimetric

accuracy of synthetic CT images generated using a deep learning approach Radiotherapy and Oncology 136 2019 56 63 https://linkinghub.elsevier.com/retrieve/pii/S0167814019301483

[10] Dinkla AM, Florkow MC, Maspero M, Savenije MHF, Zijlstra F, Doornaert PAH, et al. Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network. Med Phys 2019. https://doi.org/10.1002/mp.13663.

[10] A.M. Dinkla M.C. Florkow M. Maspero M.H.F. Savenije F. Zijlstra P.A.H. Doornaert M. Stralen M.E.P. Philippens C.A.T. Berg P.R. Seevinck Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network Med. Phys. 46 9 2019 4095 4104 https://onlinelibrary.wiley.com/toc/24734209/46/9

[11] Hsu SH, Peng Q, Tomé WA. On the generation of synthetic CT for a MRI-only radiation therapy workflow for the abdomen. J. Phys. Conf. Ser., vol. 1154, Institute of Physics Publishing; 2019. https://doi.org/10.1088/1742-6596/1154/1/012011.

[11] S.-H. Hsu Q. Peng W.A. Tomé On the generation of synthetic CT for a MRI-only radiation therapy workflow for the abdomen J. Phys.: Conf. Ser.

1154 2019 012011 10.1088/1742-6596/1154/1/012011

https://iopscience:iop.org/article/10.1088/1742-6596/1154/1/012011

[12] Depauw N, Keyriläinen J, Suilamo S, Warner L, Bzdusek K, Olsen C, et al. MRI-based IMPT planning for prostate cancer. Radiother Oncol 2020;144:79–85. https://doi.org/10.1016/j.radonc.2019.10.010.

[12] N. Depauw J. Keyriläinen S. Suilamo L. Warner K. Bzdusek C. Olsen H. Kooy MRI-based IMPT planning for prostate cancer Radiotherapy and Oncology 144 2020 79 85

https://linkinghub.elsevier.com/retrieve/pii/S0167814019331469

[13] Maspero M, Savenije MHF, Dinkla AM, Seevinck PR, Intven MPW, Jurgenliemk-Schulz IM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. Phys Med Biol 2018;63:185001. https://doi.org/10.1088/1361-6560/aada6d.

[13] M. Maspero M.H.F. Savenije A.M. Dinkla P.R. Seevinck M.P.W. Intven I.M. Jurgenliemk-Schulz L.G.W. Kerkmeijer C.A.T. van den Berg Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy Phys. Med. Biol. 63 18 185001 10.1088/1361-6560/aada6d

https://iopscience:iop.org/article/10.1088/1361-6560/aada6d.

[14] Bredfeldt JS, Liu L, Feng M, Cao Y, Balter JM. Synthetic CT for MRI-based liver stereotactic body radiotherapy treatment planning. Phys Med Biol 2017;62:2922–34. https://doi.org/10.1088/1361-6560/aa5059.

[14] J.S. Bredfeldt L. Liu M. Feng Y. Cao J.M. Balter Synthetic CT for MRI-based liver stereotactic body radiotherapy treatment planning Phys. Med. Biol. 62 8 2017 2922 2934 https://iopscience:iop.org/article/10.1088/1361-6560/aa5059

[15] Liu Y, Lei Y, Wang Y, Wang T, Ren L, Lin L, et al. MRI-based treatment planning for proton radiotherapy: Dosimetric validation of a deep learning-based liver synthetic CT generation method. Phys Med Biol 2019;64. https://doi.org/10.1088/1361-6560/ab25bc.

[15] Y. Liu Y. Lei Y. Wang T. Wang L. Ren L. Lin M. McDonald W.J. Curran T. Liu J. Zhou X. Yang MRI-based treatment planning for proton radiotherapy: dosimetric validation of a deep learning-based liver synthetic CT generation method Phys. Med. Biol. 64 14 145015 10.1088/1361-6560/ab25bc https://iopscience:iop.org/article/10.1088/1361-6560/ab25bc.

[16] Guerreiro F, Koivula L, Seravalli E, Janssens GO, Maduro JH, Brouwer CL, et al. Feasibility of MRI-only photon and proton dose calculations for pediatric patients with abdominal tumors. Phys Med Biol 2019;64. https://doi.org/10.1088/1361-6560/ab0095.

[16] F. Guerreiro L. Koivula E. Seravalli G.O. Janssens J.H. Maduro C.L. Brouwer E.W. Korevaar A.C. Knopf J. Korhonen B.W. Raaymakers Feasibility of MRI-only photon and proton dose calculations for pediatric patients with abdominal tumors Phys. Med. Biol. 64 5 055010 10.1088/1361-6560/ab0095 https://iopscience:iop.org/article/10.1088/1361-6560/ab0095.

[17] Ladefoged CN, Marner L, Hindsholm A, Law I, Højgaard L, Andersen FL. Deep learning based attenuation correction of PET/MRI in pediatric brain tumor patients: Evaluation in a clinical setting. Front Neurosci 2019;13. https://doi.org/10.3389/fnins.2018.01005.

[17] C.N. Ladefoged L. Marner A. Hindsholm I. Law L. Højgaard F.L. Andersen Deep Learning Based Attenuation Correction of PET/MRI in Pediatric Brain Tumor Patients: Evaluation in a Clinical Setting Front. Neurosci. 12 10.3389/fnins.2018.01005 10.3389/fnins.2018.01005.s001 10.3389/fnins.2018.01005.s002 10.3389/fnins.2018.01005.s003 https://www.frontiersin.org/article/10.3389/fnins.2018.01005/full https://www.frontiersin.org/articles/10.3389/fnins.2018.01005/supplementary-material/10.3389/fnins.2018.01005.s001 https://www.frontiersin.org/articles/10.3389/fnins.2018.01005/supplementar

y-material/10.3389/fnins.2018.01005.s002

https://www.frontiersin.org/articles/10.3389/fnins.2018.01005/supplementary-material/10.3389/fnins.2018.01005.s003.

[18] Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. Elastix: A toolbox for intensity-based medical image registration. IEEE Trans Med Imaging 2010;29:196–205. https://doi.org/10.1109/TMI.2009.2035616.

[18] S. Klein M. Staring K. Murphy M.A. Viergever J. Pluim elastix: A Toolbox for Intensity-Based Medical Image Registration IEEE Trans. Med. Imaging 29 1 2010 196 205 http://ieeexplore.ieee.org/document/5338015/

[19] Staring M, Klein S, Pluim JPW. A rigidity penalty term for nonrigid registration. Med Phys 2007;34:4098–108. https://doi.org/10.1118/1.2776236.

[19] M. Staring S. Klein J.P.W. Pluim A rigidity penalty term for nonrigid registration: A rigidity penalty term for nonrigid registration Med. Phys. 34 11 2007 4098 4108

[20] Florkow MC, Zijlstra F, Willemsen K, Maspero M, Berg CAT, Kerkmeijer LGW, et al. Deep learning–based MR-to-CT synthesis: The influence of varying gradient echo–based MR images as input channels. Magn Reson Med 2019. https://doi.org/10.1002/mrm.28008.

[20] M.C. Florkow F. Zijlstra K. Willemsen M. Maspero C.A.T. Berg L.G.W. Kerkmeijer R.M. Castelein H. Weinans M.A. Viergever M. Stralen P.R. Seevinck Deep learning–based MR-to-CT synthesis: The influence of varying gradient echo–based MR images as input channels Magn Reson Med 83 4 2020 1429 1441

https://onlinelibrary.wiley.com/toc/15222594/83/4

[21] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, and Hornegger J, and Wells WM, and Frangi AF, editors. Lect. Notes Comput. Sci., vol. 9351, Cham: Springer International Publishing; 2015, p. 234–41. https://doi.org/10.1007/978-3-319-24574-4_28.

[21] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, and Hornegger J, and Wells WM, and Frangi AF, editors. Lect. Notes Comput. Sci., vol. 9351, Cham: Springer International Publishing; 2015, p. 234–41. https://doi.org/10.1007/978-3-319-24574-4_28.

[22] Dozat T. Incorporating Nesterov Momentum into Adam. ICLR Work 2016:2013–6.

[22] T. Dozat Incorporating Nesterov Momentum into Adam ICLR Work 2016 2013 2016

[23] Guerreiro F, Seravalli E, Janssens GO, van de Ven CP, van den Heuvel-Eibrink MM, Raaymakers BW. Intra- and inter-fraction uncertainties during IGRT for Wilms' tumor. Acta Oncol (Madr) 2018;57:941–9. https://doi.org/10.1080/0284186X.2018.1438655.

[23] F. Guerreiro E. Seravalli G.O. Janssens C.P. van de Ven M.M. van den Heuvel-Eibrink B.W. Raaymakers Intra- and inter-fraction uncertainties during IGRT for Wilms' tumor Acta Oncologica 57 7 2018 941 949

[24] Fredriksson A, Forsgren A, Hårdemark B. Minimax optimization for handling range and setup uncertainties in proton therapy. Med Phys 2011;38:1672–84. https://doi.org/10.1118/1.3556559.

[24] A. Fredriksson A. Forsgren B. Hårdemark Minimax optimization for handling range and setup uncertainties in proton therapy: Minimax optimization for handling uncertainties in proton therapy Med. Phys. 38 3 2011 1672 1684

[25] Korevaar EW, Habraken SJM, Scandurra D, Kierkels RGJ, Unipan M, Eenink MGC, et al. Practical robustness evaluation in radiotherapy – A photon and proton-proof alternative to PTV-based plan evaluation. Radiother Oncol 2019;141:267–74. https://doi.org/10.1016/j.radonc.2019.08.005.

[25] E.W. Korevaar S.J.M. Habraken D. Scandurra R.G.J. Kierkels M. Unipan M.G.C. Eenink R.J.H.M. Steenbakkers S.G. Peeters J.D. Zindler M. Hoogeman J.A. Langendijk Practical robustness evaluation in radiotherapy – A photon and proton-proof alternative to PTV-based plan evaluation Radiotherapy and Oncology 141 2019 267 274 https://linkinghub.elsevier.com/retrieve/pii/S0167814019330440

[26] Hoeben BA, Carrie C, Timmermann B, Mandeville HC, Gandola L, Dieckmann K, et al. Management of vertebral radiotherapy dose in paediatric patients with cancer: consensus recommendations from the SIOPE radiotherapy working group. Lancet Oncol 2019;20:e155–66. https://doi.org/10.1016/S1470-2045(19)30034-8.

[26] B.A. Hoeben C. Carrie B. Timmermann H.C. Mandeville L. Gandola K. Dieckmann M. Ramos Albiac H. Magelssen Y. Lassen-Ramshad B. Ondrová T. Ajithkumar C. Alapetite B.V. Balgobind S. Bolle A.L. Cameron R. Davila Fajardo S. Dietzsch D. Dumont Lecomte M.M. van den Heuvel-Eibrink R.D. Kortmann A. Laprie P. Melchior L. Padovani B. Rombi G. Scarzello R. Schwarz K. Seiersen E. Seravalli N. Thorp G.A. Whitfield T. Boterberg G.O. Janssens Management of vertebral radiotherapy dose in paediatric patients with cancer: consensus recommendations from the

SIOPE radiotherapy working group The Lancet Oncology 20 3 2019 e155 e166 https://linkinghub.elsevier.com/retrieve/pii/S1470204519300348

[27] Dice LR. Measures of the Amount of Ecologic Association Between Species. Ecology 1945;26:297–302. https://doi.org/10.2307/1932409.

[27] L.R. Dice Measures of the Amount of Ecologic Association Between Species 26 3 1945 297 302.

[28] Bol GH, Kotte ANTJ, van der Heide UA, Lagendijk JJW. Simultaneous multi-modality ROI delineation in clinical practice. Comput Methods Programs Biomed 2009;96:133–40. https://doi.org/10.1016/j.cmpb.2009.04.008.

[28] G.H. Bol A.N.T.J. Kotte U.A. van der Heide J.J.W. Lagendijk Simultaneous multi-modality ROI delineation in clinical practice Computer Methods and Programs in Biomedicine 96 2 2009 133 140 https://linkinghub.elsevier.com/retrieve/pii/S0169260709001412

[29] Low DA. Gamma dose distribution evaluation tool. J Phys Conf Ser 2010;250:349–59. https://doi.org/10.1088/1742-6596/250/1/012071.

[29] D.A. Low Ph.D. Gamma Dose Distribution Evaluation Tool J. Phys.: Conf. Ser. 250 2010 012071 10.1088/1742-6596/250/1/012071 https://iopscience:iop.org/article/10.1088/1742-6596/250/1/012071

[30] Korsholm ME, Waring LW, Edmund JM. A criterion for the reliable use of

MRI-only radiotherapy. Radiat Oncol 2014;9:16. https://doi.org/10.1186/1748-717X-9-16.

[30] M.E. Korsholm L.W. Waring J.M. Edmund A criterion for the reliable use of MRI-only radiotherapy Radiat Oncol 9 1 2014 10.1186/1748-717X-9-16 https://ro-journal.biomedcentral.com/articles/10.1186/1748-717X-9-16

[31] Florkow MC, Zijlstra F, Kerkmeijer LGW, Maspero M, van den Berg CAT, van Stralen M, et al. The impact of MRI-CT registration errors on deep learning-based synthetic CT generation. In: Angelini ED, Landman BA, editors. Med. Imaging 2019 Image Process., vol. 10949, SPIE; 2019, p. 116. https://doi.org/10.1117/12.2512747.

[31] Florkow MC, Zijlstra F, Kerkmeijer LGW, Maspero M, van den Berg CAT, van Stralen M, et al. The impact of MRI-CT registration errors on deep learning-based synthetic CT generation. In: Angelini ED, Landman BA, editors. Med. Imaging 2019 Image Process., vol. 10949, SPIE; 2019, p. 116. https://doi.org/10.1117/12.2512747.

[32] Maspero M, van den Berg CAT, Landry G, Belka C, Parodi K, Seevinck PR, et al. Feasibility of MR-only proton dose calculations for prostate cancer radiotherapy using a commercial pseudo-CT generation method. Phys Med Biol 2017;62:9159–76. https://doi.org/10.1088/1361-6560/aa9677.

[32] M. Maspero C.A.T. van den Berg G. Landry C. Belka K. Parodi P.R. Seevinck B.W. Raaymakers C. Kurz Feasibility of MR-only proton dose

calculations for prostate cancer radiotherapy using a commercial pseudo-CT generation method Phys. Med. Biol. 62 24 9159 9176 10.1088/1361-6560/aa9677 https://iopscience:iop.org/article/10.1088/1361-6560/aa9677.

[33] Koivula L, Wee L, Korhonen J. Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: Dose calculation accuracy in substitute CT images. Med Phys 2016;43:4634–42. https://doi.org/10.1118/1.4958677.

[33] L. Koivula L. Wee J. Korhonen Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: Dose calculation accuracy in substitute CT images: Feasibility of MRI-only treatment planning for proton therapy Med. Phys. 43 8Part1 2016 4634 4642

Figure 1. Architecture of the U-Net. T1-weighted (T1w) and T2-weighted (T2w) MR patches of size 24x24x24 were input to the neural network separately. After two convolutions, patches were concatenated to undergo the standard encoding-decoding process of the U-Net. Encoding was performed using max pooling and decoding with nearest neighbour upsampling. The resulting sCT image has the same resolution as the planning-CT. The spatial dimension of each level is given on the left hand side.

Figure 2. T1-weighted MR, T2-weighted MR, planning-CT and sCT images in the sagittal and coronal planes for two example patients from the cross-validation set (a) and from the independent test set (b). The error map (Δ [CT-sCT]) on the right shows the difference in Hounsfield units (HU) between the planning-CT and sCT images.

Figure 3. Box-and-whisker plots of the dose-volue histogram (DVH) parameters relative differences between the planning-CT, used as reference, and the sCT dose distributions for VMAT (left) and PBS (right) for patients from the cross-validation (CV) (in white)

and test (in grey) sets. (a) ITV DVH parameters differences (b) Mean dose (D_{mean}) differences for the OARs (c) $D_{2\%}$ differences for the OARs. Crosses indicate differences per patient. Values are presented as a percentage of the respective prescribed dose. Abbreviations: $Kidney_I$= ipsilateral kidney; $Kidney_C$= contralateral kidney.

Figure 4. Dose maps optimized on the planning-CT and re-calculated on the sCT images for VMAT and PBS dose distributions for three example patients from the cross validation (a,b) and independent test (c) sets. The last column displays the dose difference maps (CT-sCT) overlaid on the planning-CT image. The scale is expressed as a percentage of the prescribed dose (21.6 Gy for (a,b) and 10.8Gy for (c)). Patients in (b) and (c) had gamma pass-rates below 90% as a result of (1) CT-sCT different air cavities distributions and (2) CT-sCT body outline differences.

Table 1. Average ± standard deviation (SD) of the mean error (ME), mean absolute error (MAE), Dice score coefficient (DSC), and peak signal-to-noise ratio (PSNR) computed between the planning-CT and sCT across the cross-validation (CV) set, the test set and the entire population.

Metric	**Body site**	**Average (± SD [Range])**				
		CV (n=54)		**Test (n=12)**		**Overall (n=66)**
ME (HU)	**Body**	-5 ± 11	[-30; 42]	-7 ± 13	[-22; 15]	-5 ± 12
	Soft tissue	0 ± 8	[-17; 27]	1 ± 12	[-10; 35]	0 ± 9
	Bone	114 ± 59	[-12; 307]	132 ± 35	[66; 178]	117 ± 55
	Lungs	-9 ± 61	[-155; 146]	-5 ± 77	[-200; 90]	-9 ± 67
MAE (HU)	**Body**	56 ± 11	[35; 86]	62 ± 13	[39; 86]	57 ± 12
	Soft tissue	33 ± 6	[22; 55]	36 ± 8	[26; 53]	33 ± 7
	Bone	156 ± 43	[79; 322]	167 ± 22	[118; 198]	158 ± 40
	Lungs	105 ± 33	[50; 212]	104 ± 38	[66; 215]	105 ± 34
DSC (%)	**Bone**	76 ± 7	[50; 87]	76 ± 6	[65; 83]	76 ± 8
	Lungs	92 ± 9	[45; 99]	88 ± 7	[71; 94]	92 ± 9
PSNR (dB)	**Body**	30.4 ± 1.6	[26.7; 33.7]	30.0 ± 1.8	[27.7; 34.3]	30.3 ± 1.6

Table 2. Average ± standard deviation [range] gamma pass-rates using a 2% and 2mm criterion with multiple dose thresholds (10%, 50%, 90%) for VMAT and PBS dose distributions optimized on the planning-CT and re-calculated on the sCT across the cross-validation (CV) set, the test set and the entire population.

Threshold	VMAT (%)			PBS (%)		
	CV (n=54)	Test (n=12)	Overall	CV (n=54)	Test (n=12)	Overall
10%	99.7±0.5 [96.2;100]	99.7±0.3 [99.2;100]	99.7±0.5	97.4±2.9 [87.5;100]	96.2±4.0 [86.8;100]	97.2±3.1
50%	99.7±1.0 [94.1;100]	99.8±0.5 [98.4;100]	99.7±0.9	99.4±1.1 [94.8;100]	98.6±2.0 [94.4;100]	99.2±1.3
90%	99.5±2.4 [85.4;100]	99.5±1.5 [94.9;100]	99.5±2.3	99.6±1.1 [93.6;100]	99.2±1.1 [97.2;100]	99.5±1.1