Comparison of the suitability of CBCT- and MR-based synthetic CTs for daily adaptive proton therapy in head and neck patients

Adrian Thummerer1,6,7, Bas A de Jong1,6, Paolo Zaffino2, Arturs Meijers1, Gabriel Guterres Marmitt1, Joao Seco3,4, Roel J H M Steenbakkers1, Johannes A Langendijk1, Stefan Both1, Maria F Spadea2,6 and Antje C Knopf1,5,6

1 Department of Radiation Oncology, University Medical Center Groningen, University of Groningen, The Netherlands
2 Department of Experimental and Clinical Medicine, Magna Graecia University, Catanzaro, Italy
3 Department of Biomedical Physics in Radiation Oncology, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany
4 Department of Physics and Astronomy, Heidelberg University, Heidelberg, Germany
5 Division for Medical Radiation Physics, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany
6 Both authors contributed equally to this work
7 Author to whom any correspondence should be addressed

E-mail: a.thummerer@umcg.nl

Keywords: adaptive proton therapy, NTCP-evaluation, CBCT-based synthetic CT, MR-based synthetic CT, neural network

Supplementary material for this article is available online

Abstract
Cone-beam computed tomography (CBCT)- and magnetic resonance (MR)-images allow a daily observation of patient anatomy but are not directly suited for accurate proton dose calculations. This can be overcome by creating synthetic CTs (sCT) using deep convolutional neural networks. In this study, we compared sCTs based on CBCTs and MRs for head and neck (H&N) cancer patients in terms of image quality and proton dose calculation accuracy.

A dataset of 27 H&N-patients, treated with proton therapy (PT), containing planning CTs (pCTs), repeat CTs, CBCTs and MRs were used to train two neural networks to convert either CBCTs or MRs into sCTs. Image quality was quantified by calculating mean absolute error (MAE), mean error (ME) and Dice similarity coefficient (DSC) for bones. The dose evaluation consisted of a systematic non-clinical analysis and a clinical recalculation of actually used proton treatment plans. Gamma analysis was performed for non-clinical and clinical treatment plans. For clinical treatment plans also dose to targets and organs at risk (OARs) and normal tissue complication probabilities (NTCP) were compared.

CBCT-based sCTs resulted in higher image quality with an average MAE of $40 \pm 4$ HU and a DSC of 0.95, while for MR-based sCTs a MAE of $65 \pm 4$ HU and a DSC of 0.89 was observed. Also in clinical proton dose calculations, sCT$_{CBCT}$ achieved higher average gamma pass ratios (2%/2 mm criteria) than sCT$_{MR}$ (96.1% vs. 93.3%). Dose-volume histograms for selected OARs and NTCP-values showed a very small difference between sCT$_{CBCT}$ and sCT$_{MR}$ and a high agreement with the reference pCT.

CBCT- and MR-based sCTs have the potential to enable accurate proton dose calculations valuable for daily adaptive PT. Significant image quality differences were observed but did not affect proton dose calculation accuracy in a similar manner. Especially the recalculation of clinical treatment plans showed high agreement with the pCT for both sCT$_{CBCT}$ and sCT$_{MR}$.

1. Introduction

Adaptive proton therapy (PT) attempts to spare healthy tissue and simultaneously increase the dose to tumor cells by reacting to interfractional anatomical changes with treatment plan adaptations (Lim-Reinders et al 2017, Albertini et al 2019). To monitor these anatomical changes and deploy adaptive workflows, repeated imaging throughout the treatment course is a necessity. In current clinical practice, it is feasible to acquire
conventional fan-beam computed tomography (CT) images on a weekly basis to observe the patient anatomy. However, these weekly CT acquisitions require a strong clinical motivation, since they come at the cost of additional imaging dose and increase the clinical workload. Recent literature suggests that PT plans should be adapted as soon as unusual anatomical variations occur (Hoffmann et al 2017, Nenoff et al 2019). In the future, online adaptive PT might be worth striving for (Albertini et al 2019). That would imply the necessity of daily or online repeated imaging.

As an alternative to using conventional fan-beam CTs for repeated image acquisition, cone-beam computed tomography (CBCT) or magnetic resonance (MR) imaging could be employed. In some PT centers, daily CBCT scans are already routinely acquired for accurate patient alignment (Hua et al 2017, Stock et al 2018). CBCT images provide the patient anatomy of the day and can be used as basis for daily adaptive workflows. With MR, volumetric images can be acquired without ionizing radiation and with superior soft tissue contrast. In current PT practice, MRs are acquired in the planning stage to aid delineations of target volumes and organs at risk (OAR, Karlsson et al 2009, Kupelian and Sonke 2014). Daily (online) in-room MR-image acquisition for PT is not yet clinically available. However, in sight of the rapid adoption of online MR-guided adaptive photon therapy, within few years, simple prototype systems for PT will likely exist, and in a decade, we could envisage coupled MR PT systems with integrated gantries (Oborn et al 2017, Hoffmann et al 2020).

The distinct advantages of CBCT and MR-systems make both favorable imaging modalities for daily or online adaptive treatment strategies. However, both imaging modalities are not directly suited for accurate proton dose calculations. Because of the imaging geometry, CBCT images suffer from severe scatter artifacts that impair the CT-number accuracy and as consequence the conversion into proton stopping power ratios (SPR), which are required for proton dose calculations. MR-image intensities correlate with magnetic relaxation properties of hydrogen atoms and thereby do not allow a derivation of electron densities and SPRs. The deficiencies of CBCTs and MRs can be overcome by creating so called synthetic CTs (sCTs), often referred to as virtual CTs or pseudo CTs. They act as a surrogate for CT images, containing accurate electron density information (HU-intensities) and enable proton dose calculations.

To enable dose calculations, various techniques to generate sCTs based on CBCT- and MR-images have been proposed in literature. Only a few have been assessed in the context of proton dose calculation accuracy. For CBCTs, projection- and deformable image registration (DIR)-based techniques have shown promising proton dose calculation accuracy. This includes anatomical locations such as lung (Veiga et al 2015, 2016), pelvis (Park et al 2015, Kurz et al 2016) and head and neck (Landry et al 2015, Kurz et al 2015). A downside of these methods is that they require a patient specific planning CT (pCT) to generate sCTs. For MR-to-sCT conversion, the investigated anatomical locations include brain (Koivula et al 2016, Pileggi et al 2018, Spadea et al 2019), prostate (Maspero et al 2018, Liu et al 2019), head and neck (Guerreiro et al 2017) and pediatric patients with abdominal tumours (Guerreiro et al 2019).

In recent years, technological development lead to significant progress in the field of artificial intelligence and deep learning. These developments have been translated to the field of medical physics and radiotherapy (Meyer et al 2018, Shen et al 2020). Deep learning techniques, such as deep convolutional neural networks (DCNNs) and generative adversarial networks (GANs), have shown their potential for sCT generation based on CBCTs and MRs (Han 2017, Kida et al 2018, Maspero et al 2018, Wang et al 2019, Liang et al 2019). DCNNs are trained with paired MR/CT or CBCT/CT images and learn a nonlinear mapping of intensities from the original imaging modality, CBCT or MR, to CT. sCTs, based on deep learning approaches, have recently been discussed for adaptive PT (Hansen et al 2018, Liu et al 2019, Landry et al 2019) and have shown promising performance when compared to previous techniques in various anatomical locations (Arabi et al 2018, Thummerer et al 2020).

Previous studies have only looked into either CBCT- or MR-based sCTs and only a limited number of studies investigated the suitability of the resulting sCTs for proton dose calculations. Our aim here is to perform a direct comparison of MR- and CBCT-based sCTs, generated using the same DCNN network architecture, for a comprehensive head & neck patient cohort. By simultaneously assessing the dosimetric suitability of CBCT- and MR-based sCTs for PT, we will identify differences relevant for their employment in daily or online adaptive workflows.

2. Materials and methods

2.1. Patient dataset
To evaluate the performance of CBCT- and MR-based sCTs, imaging data from 27 head and neck cancer patients who received a PT treatment at the University Medical Center Groningen (UMCG) were used. The included patients were aged between 27 and 79 years (mean age: 62) and 2/3 were of male sex. Out of the 27 patients, 26 received primary RT (14 patients chemoradiation, 8 conventional RT, 2 RT + cetuximab and 2...
accelerated RT) and one postoperative RT. For 24 patients the tumor was located in the pharynx, for two in the oral cavity and for one in the larynx. Tumors had varying extent (T-stage 1–4) and spread to regional lymph nodes (N stage 0–3). The datasets included pCTs, repeated CTs (rCT), CBCTs and MR-images. pCTs were acquired on a Siemens SOMATOM Definition AS Open scanner (Siemens Healthineers, Germany) and rCT scans on a Siemens SOMATOM Confidence scanner. Similar imaging protocols were used for pCT and rCT and besides being acquired on different scanners, pCT and rCT can be considered equal in image quality. CBCTs were acquired with the onboard imaging device of an IBA Proteus PLUS gantry (IBA, Belgium), using a 190°-arc with a rotation speed of 4.9°/s, a total projection number of 258 and an acquisition time of 39 s. Detailed imaging parameters for pCT, rCT and CBCT are presented in table 1. MR scans were performed on a 3 T Siemens MAGNETOM Skyra system after administration of a single dose of gadoterate meglumine (0.2 ml kg\(^{-1}\)) contrast agent. A 3D spoiled gradient recalled echo (SPGRE) sequence was used to generate MR-images. MR-imaging parameters were: echo time \(= 2.46\) ms, repetition time \(= 5.5\) ms, flip angle \(= 9\) degree, FOV = 229 mm × 237 mm × 229 mm, bandwidth = 455 hz px\(^{-1}\) and acquisition time \(= 42\) s. Additional parameters are provided in table 1. In the case of pCT and MR, which were usually acquired for planning purposes several weeks before the start of treatment, only one instance was available. rCT scans and CBCTs on the other hand were acquired periodically (rCTs weekly, CBCTs daily) during treatment progression. Radiotherapeutic immobilization devices were used during all image acquisitions to assure consistency of patient immobilization.

### Table 1. Summary of imaging parameters of CBCT, pCT, rCT and MR. For tube current, acquisition matrix and FOV, min-max ranges are reported for some modalities.

|              | CBCT         | pCT          | rCT          | MR          |
|--------------|--------------|--------------|--------------|-------------|
| **Scanner**  | IBA Proteus PLUS | Siemens SOMATOM | Siemens SOMATOM | Siemens MAGNETOM |
| **Gantry**   |              |              |              |             |
| **Voltage [kV]** | 100          | 120          | 120          |             |
| **Current [mA]** | 160          | 61–165 (min-max) | 61–165       |             |
| **Exposure [ms]** | 12.5         | 1106         | 1106         |             |
| **Acq. matrix** | 512 × 512 × 140 | 512 × 512 × (181–262) | 512 × 512 × (181–262) | 256 × 264 × 256 |
| **Resolution [mm]** | 0.5 × 0.5 × 2.5 | 1.0 × 1.0 × 2.0 | 1.0 × 1.0 × 2.0 | 0.9 × 0.9 × 0.9 |
| **FOV [mm]**  | RL:260 AP: 260 IS: 350 | 500 500 362–524 (min-max) | 50 500 362–524 | 229 237 229 |

2.2. Neural network training

Both, CBCT- and MR-based sCTs, were created utilizing the same DCNN architecture, initially described in the work of Spadea et al (2019). The DCNN consists of an encoding and decoding path to convert either CBCT-HU or MR intensities into CT numbers with an accuracy comparable to pCT scans. For training of the networks, mean absolute error (MAE) between sCT and ground truth ‘pCT’ (in case of MR) or ‘rCT’ (CBCT) was used as similarity metric in the loss function. Following the approach of Spadea et al (2019), three individual networks were trained with axial, coronal or sagittal slices exclusively. After training, images from each view were combined into the final sCT. A three-fold cross validation approach was chosen. Therefore, patients were randomly divided into three equal sets of nine patients. Two sets were then used for training and one for evaluation. This was repeated so each set was used for evaluation once. Additionally, two cases from each training set were withheld as validation cases during training. When no improvement in validation loss was observed within five consecutive epochs, the network training was stopped.

2.3. CBCT data preparation

The DCNN requires paired image sets of CBCTs and CTs to successfully learn a conversion of image intensities. For each patient, the first acquired rCT and a CBCT from the same day were selected to minimize anatomical differences. Plastimatch, an image processing toolbox (www.plastimatch.org, Zaffino et al 2016), was used to automatically segment the patient outline on CBCT and rCT. The resulting masks were manually edited to assure full patient coverage. Voxels outside these masks were set to a HU value of \(-1000\). An additional crop was performed to remove the area below the shoulders. This was necessary because a low dose imaging protocol was used for CBCT acquisition and led to scatter artifacts and very poor image quality in this area. Afterwards, a rigid registration utilizing Plastimatch and a DIR were performed. For DIR, a diffeomorphic morphons algorithm with a 4 level resolution pyramid was used. This algorithm is implemented in the open-source MATLAB toolbox openREGGUI (www.openreggui.org). The principal suitability of this algorithm for CBCT to CT image registration has been demonstrated previously (Landry et al 2015, Kurz et al 2016). As a last step, masks from CBCT and CT were combined to only include voxels.
present in both CBCT and rCT. The deformed and masked CBCT-rCT image pairs were then used to train the CBCT-network.

2.4. MR data preparation
In contrast to the CBCT-network, the MR-network was trained with pCT-MR image pairs. This was advantageous since both images were acquired either on the same day or with a maximum of one day in-between. Similarly, to CBCT and rCT, the patient outline was segmented on MR and pCT. For the pCT, voxels outside the patient were set to $-1000$ HU, while for MRs a value of 0 was assigned to this area. An initial rigid registration was performed using Plastimatch. For DIR of pCT and MR, the Elastix (Klein et al. 2010, https://elastix.lumc.nl/) registration toolbox with a three level resolution pyramid was used. Because of the multimodal imaging data, registration of pCT and MR was more challenging than the CBCT-rCT registration. Mutual information was used as similarity metric and a penalty, with a weighting of 600:1, was introduced to suppress un-anatomical deformations (Staring et al. 2007). Afterwards masks were combined, and the training of the MR-DCNN was performed.

2.5. Evaluation of image quality
Anatomical differences between pCT, CBCT and MR had to be minimized to allow a meaningful comparison of conversion characteristics. MR-images were already deformably registered to the pCT during data preparation. CBCTs, on the other hand, were registered to the first available repeat CT for training of the DCNN CBCT. Therefore, prior to the conversion into sCTs, CBCTs were also deformed to the pCT using openREGGUI. This further minimized anatomical differences and allowed to focus almost entirely on the conversion characteristics of the DCNNs and eliminated the influence of anatomical differences. The pCT was used as ‘ground truth’ for image quality and the dosimetric evaluation. MAE (equation (1)) and mean error (ME, equation (2)) were used to evaluate the similarity between the pCT and the sCT in terms of Hounsfield units.

\[
MAE = \frac{\sum_{i=1}^{n} |sCT_i - pCT_i|}{n}
\]

\[
ME = \frac{\sum_{i=1}^{n} (sCT_i - pCT_i)}{n}
\]

Furthermore, average MAE spectrums for CBCT and MR were calculated by binning voxels in HU intervals of 20 HU and calculating the MAE for each bin. Error bars were added to visualize the standard deviation within the dataset. To analyze the similarity in bones, the Dice similarity coefficient (DSC, equation (3)) was calculated for various threshold levels between 100 and 1000 HU. All image quality metrics were calculated within the union of patient contours of pCT, CBCT and MR to enable a meaningful comparison.

\[
DSC = \frac{2|sCT_{threshold} \cap pCT_{threshold}|}{sCT_{threshold} + pCT_{threshold}}
\]

2.6. Evaluation of proton dose calculation accuracy
To determine the proton dose calculation accuracy of CBCT- and MR-based sCTs, we performed two types of dosimetric analysis. Firstly, non-clinical single-beam proton treatment plans were used to systematically investigate proton dose accuracy and range errors introduced by sCT conversion. Secondly, clinically used treatment plans were recalculated on the sCTs to show the accuracy in clinical conditions.

For the systematic evaluation, an intracranial target was defined in the brainstem region using the treatment planning system RayStation (Raysearch, Sweden). This target was irradiated with a single field from a 45-degree gantry angle and a homogenous dose of 2 GyRBE (constant RBE of 1.1). The dose was calculated using the RayStation Monte Carlo dose engine on a 1 mm isotropic dose grid. For comparison between the sCT dose distributions and the reference dose, which was calculated on the pCT, we performed a gamma analysis with 2%/2 mm and 3%/3 mm passing criteria. Furthermore, range uncertainties introduced by the sCT conversion were investigated by calculating range shifts. Range shifts were determined by shifting depth-dose profiles to minimize the sum of squared differences between sCT and pCT profiles. Only profiles with a maximum dose of at least 80% of the planned dose were included for this range error assessment.

Clinically used proton treatment plans, based on the pCT, were recalculated on the CBCT- and MR-based sCTs. This allowed a comparison of the resulting dose distributions and thereby an assessment of the clinical suitability of the sCTs. Since CBCT and MR were deformally registered to the pCT, OARs and target volumes could be transferred from the pCT to both sCTs. Because of the cropping of CBCTs during data preparation,
sCTs were not always covering the entire low-risk CTVs of the clinical treatment plans. To still allow a clinical plan recalculation, using original target volumes, parts of the pCT (e.g. headrest, couch and shoulder area) were stitched to the sCT. A visualization of the cropping and stitching procedure is available in the supplementary materials (available online at stacks.iop.org/PMB/65/235036/mmedia). The clinical treatment plans consisted of two CTVs. The first one targeted the primary tumor and pathological lymph nodes and was irradiated with 70 Gy\textsubscript{RBE}. The second one was used to irradiate the elective lymph node areas with 54.25 Gy\textsubscript{RBE}. In most cases, the primary tumor was within the region covered by the sCT, while the elective area, extending towards the lower neck, also had substantial parts of its volume on the stitched pCT.

Similar to the systematic dose analysis using single-beam plans, we calculated gamma pass ratios for 2%/2 mm and 3%/3 mm criteria. To eliminate the influence of the pCT on the clinical gamma analysis, a mask, corresponding to the synthetic part of the image, was applied to the dose volume. We also compared the mean dose, calculated on the pCT, sCT\textsubscript{CBCT} and sCT\textsubscript{MR}, for target volumes (CTV) and selected OARs (brainstem, mandible, parotid glands, submandibular glands and inferior-, middle- and, superior-pharyngeal constrictor muscles). The selected OARs were almost always fully covered by the synthetic part of the stitched image. Exceptions included nine patients where a minor part of the CTV also extended towards the lower part of the neck and five patients where the inferior pharyngeal constrictor muscle (PCM) was not entirely covered by the sCTs.

In addition, we used normal tissue complication probability (NTCP) models for xerostomia (dry mouth) and dysphagia (swallowing difficulties) to investigate differences between pCT and sCTs. NTCP models establish a relation between the dose to certain OARs and the probability of radiation induced side effects. Clinically, NTCP models are used in the so called ‘model-based approach’ for treatment selection (e.g. photon vs. proton radiotherapy) (Langendijk et al 2013, Widder et al 2016). We made use of models for xerostomia ($\geq$ grade 2 and $\geq$ grade 3) and for dysphagia ($\geq$ grade 2 and $\geq$ grade 3) which have recently been defined in the Dutch nationwide indication protocol for PT (LIPPv2.2) of head and neck cancer patients. All included patients qualified for PT based on such NTCP models.

### 3. Results

#### 3.1. DCNN training

Neural network training was stopped if the validation loss did not decrease within five consecutive epochs. This condition was reached after 9–26 epochs. Detailed numbers for each fold and anatomical view can be found in the supplementary materials. The neural network was implemented using the python framework Theano (the Theano Development Team et al 2016). A Nvidia GeForce 1080TI was used for training and validation purposes. With this configuration, the training duration of a single epoch was approximately two hours for axial trainings and four hours for sagittal/coronal trainings. This variation is caused by the difference in slices available for each view. After training, the conversion of an entire CBCT or MR-image took approximately three minutes (axial, sagittal and coronal view combined).

#### 3.2. Evaluation of image quality

Central slices of axial, sagittal and coronal views of CBCT, MR, sCT\textsubscript{CBCT}, sCT\textsubscript{MR} and the reference pCT are presented for patient 20 in figure 1. A Hounsfield-unit window of 1250/250 was applied to all images (except CBCT). In figure 2(a) slices from sCT\textsubscript{CBCT} and sCT\textsubscript{MR} have been subtracted from the corresponding pCT slices to create difference images. This reveals that MR-based sCTs have higher errors in bone tissue and at tissue boundaries. This can be a result of geometric distortions of the MR-images and the more difficult image registration between MRI and CT compared to CBCT and CT. In soft tissue, sCT\textsubscript{MR} and sCT\textsubscript{CBCT} show a comparable error magnitude. Figure 2(b) shows selected details of pCT, sCT\textsubscript{CBCT} and sCT\textsubscript{MR} to highlight the differences in bone structures. The loss of bone-details in sCT\textsubscript{MR} is clearly visible.

The mismatch is quantified in figure 3, which shows MAE of sCT\textsubscript{CBCT} and sCT\textsubscript{MR} for all patients. On average CBCT-based sCTs resulted in a MAE of 40.2 ± 3.9 HU and a ME of −1.7 ± 7.4 HU. For sCT\textsubscript{MR} a significantly higher MAE of 65.4 ± 3.6 HU and a comparable ME of 2.9 ± 9.4 HU was observed. These results confirm the visual impression of figures 1 and 2. Additional image metrics (PSNR and SSIM) are presented in the supplementary materials (Section C).

Figure 4 depicts the average DSC for various thresholds between 100 and 1000 HU. The highest DSC, with a value of 0.95 for sCT\textsubscript{CBCT} and 0.89 for sCT\textsubscript{MR}, was observed for a threshold of 200 HU. With increasing threshold values, which corresponds to increasing bone density, the DSC decreases down to 0.91 for sCT\textsubscript{CBCT} and 0.81 for sCT\textsubscript{MR} at a threshold of 1000 HU.

An average MAE spectrum for sCT\textsubscript{CBCT} and sCT\textsubscript{MR} is reported in figure 5. The standard deviation among all patients is indicated by the shaded areas. CBCT- and MR-based sCTs follow a similar trend although, as expected from findings presented in the previous figures, the CBCT spectrum shows lower MAE over the
Figure 1. Overview showing axial, sagittal and coronal view of CBCT, MR, sCT \textsubscript{CBCT}, sCT \textsubscript{MR} and the reference pCT. A Hounsfield-unit window of 1250/250 was used (except CBCT).

Figure 2. (a) Difference image sCT \textsubscript{CBCT}-pCT and sCT \textsubscript{MR}-pCT. (b) Image details that show the difference in bone structures of pCT, sCT \textsubscript{CBCT} and sCT \textsubscript{MR}.

entire HU range (~1000 HU to 1500 HU). The grey area indicates the HU-range where partial volume artifacts are partially responsible for increased MAE. Overlapping with the MAE-spectrum, an average image histogram is presented. This shows that the overall MAE is mainly determined by soft tissue and that the MAE for bone structures but also for air cavities is higher.

3.3. Evaluation of proton dose calculation accuracy
3.3.1. Gamma analysis
Performing gamma analysis of dose distributions based on the single-beam plan and using 2%/2 mm criteria resulted in mean pass ratios of 99.31% with a standard deviation (SD) of 0.80% for sCT \textsubscript{CBCT} and 98.22% with a SD of 1.88% for sCT \textsubscript{MR}. Average passing ratios for the 3%/3 mm acceptance criteria were 99.97% (sCT \textsubscript{CBCT}, SD: 0.08%) and 99.42% (sCT \textsubscript{MR}, SD: 1.25%). For the clinically used treatment plans lower average
pass ratios were observed. The 2%/2 mm passing criteria resulted in mean pass ratios of 96.57% for $\text{sCT}_{\text{CBCT}}$ (SD: 3.26%) and 93.45% for $\text{sCT}_{\text{MR}}$ (SD: 3.42%). The less strict 3%/3 mm criteria lead to mean pass ratios of 98.77% (SD: 1.17%) and 97.04% (SD: 1.75%) for $\text{sCT}_{\text{CBCT}}$ and $\text{sCT}_{\text{MR}}$, respectively. Figure 6(a) presents pass ratios of the single beam plan for the stricter 2%/2 mm acceptance criteria for the entire dataset. Figure 6(b) shows a similar plot for the clinically used treatment plans.

### 3.3.2. Range error

The single beam plan was used to assess the range error between pCT and sCTs and results are shown in figure 7. For $\text{sCT}_{\text{CBCT}}$ the median range error is always within ± 2% and only for a few patients whiskers, indicating maximum/minimum values, are above or below ± 2%, indicating good agreement between $\text{sCT}_{\text{CBCT}}$ and pCT. For $\text{sCT}_{\text{MR}}$ larger range errors were observed. Although median range errors and also 25th to 75th percentile range are comparable to $\text{sCT}_{\text{CBCT}}$, $\text{sCT}_{\text{MR}}$ shows significantly larger maximum range

![Figure 3. MAE for $\text{sCT}_{\text{CBCT}}$ and $\text{sCT}_{\text{MR}}$ for all patients individually.](image3)

![Figure 4. DSC for $\text{sCT}_{\text{CBCT}}$ and $\text{sCT}_{\text{MR}}$ for bone thresholds between 100 and 1000 HU.](image4)
deviations (indicated by the whiskers). This might be caused by the higher reconstruction errors in small bone structures on sCTMR.

3.3.3. Dose-volume parameters

Figures 8(a) and (b) compare the absolute and relative difference in mean dose of selected OARs between pCT and sCTCBCT/CTMRI. Highest absolute and relative dose differences were observed for the inferior PCMs for sCTMRI. Together with superior and middle PCM and the oral cavity, this structure is relatively close to the upper airways and is influenced by the inconsistent positioning caused by swallowing and breathing motions between and during image acquisitions of CBCT, MR and pCT. Therefore, these larger errors are not solely caused by conversion errors of the sCTs but also influenced by anatomical differences. A difference between sCTCBCT and sCTMR is mainly present in the PCMs, for other OARs sCTCBCT and sCTMR show similar absolute and relative dose differences. In a similar manner, relative and absolute differences in mean dose to CTV-targets of sCTCBCT and sCTMR were compared to the pCT (figures 8(c) and (d)). For sCTCBCT absolute dose differences for CTVs were within ±0.1 Gy. Also, sCTMR resulted in low dose errors for CTVs with all values between −0.2 and +0.1 Gy.

In figure 9, DVHs for OARs and targets are presented for ‘worst’ and ‘best’ case scenarios. These scenarios were defined based on the gamma analysis of clinical treatment plans. With 98.5%, patient 11 resulted in the highest 2%/2 mm pass ratio (sCTMR) and was selected for the ‘best-case’ scenario. Patient 24 showed the lowest pass ratio (87.7%) on sCTMR and was therefore used to illustrate the worst-case scenario. Excellent agreement of DVH-curves between pCT, sCTCBCT and sCTMR was observed for the ‘best case’. The ‘worst-case’ scenario reveals some deviations in OARs, especially in the PCMinf and the oral cavity. One must consider that these OARs are close to moving structures which can have a significant influence on the dose distribution. The worst-case scenario shows that even if there is a significant difference in the global dose distribution, indicated by the low gamma pass ratio, the dose to the target volumes and OAR is not disturbed in a similar manner. The worst-case scenario does not contain the worst-case for each OAR. As seen in figure 8(b), relative mean dose differences of up to 8% were observed for some OARs in some patients.

3.3.4. NTCP evaluation

Figure 10 compares the NTCP for dysphagia (figure 10(a)) and xerostomia (figure 10(b)) of grade two or higher, calculated on pCT, sCTMR and sCTCBCT. The data in figure 10 shows that there is a very good agreement between NTCP calculated on the reference pCT and both sCTs. For dysphagia, grade 2 or higher, the maximum ΔNTCP, defined as NTCPsCTCBCT/CTMRI − NTCPpCT, was 2.0% for sCTMR (patient 8) and 1.4% for sCTCBCT (patient 18). The mean ΔNTCP value for the entire patient cohort was −0.1 ± 0.7% for sCTMR and −0.1 ± 0.5% for sCTCBCT. For xerostomia (grade 2 or higher) maximum ΔNTCP values were 0.5% for sCTMR (patient 18) and −0.68% for sCTCBCT (patient 12). The mean ΔNTCP value for xerostomia was 0.0 ± 0.2% for both sCTMR and sCTCBCT. For dysphagia and xerostomia grade 3 or higher similar results
were observed. Figures for grade 3 are presented in the supplementary materials. Due to the low ∆NTCP values, all investigated patients would have also qualified for PT if the planning comparison would have been performed on sCT\textsubscript{CBCT} or sCT\textsubscript{MR}.

4. Discussion

The necessity of accurate volumetric images for daily or online adaptive PT is unquestioned. Various image modalities are potentially suited to provide an up to date representation of the patient anatomy. In a daily adaptive workflow both CBCTs or MRs could be deployed, but MRs might be more suited due to the absence of additional imaging dose. However, it is not clear which imaging modality results in sCTs with the best image quality and subsequently the most accurate proton dose calculations. For both CBCTs and MRs, various methods to generate sCTs have been proposed. This work aimed at comparing CBCT- and MR-based sCTs for a common set of patients. sCTs were generated using a DCNN and evaluated in terms of image quality and proton dose calculation accuracy. Thereby we could identify characteristics relevant for daily adaptive PT.
Visual comparison of sCT\textsubscript{CBCT}, sCT\textsubscript{MR} and the ground-truth pCT images revealed higher image fidelity for sCT\textsubscript{CBCT} than for sCT\textsubscript{MR}. Especially in areas with fine bone structures, sCT\textsubscript{CBCT} showed more details than sCT\textsubscript{MR}. This was confirmed by quantitative image similarity metrics, such as MAE (sCT\textsubscript{CBCT}: 40.2 HU vs. sCT\textsubscript{MR}: 65.4 HU), ME (sCT\textsubscript{CBCT}: −1.7 HU vs. sCT\textsubscript{MR}: 2.9 HU) and the DSC of bony anatomy (sCT\textsubscript{CBCT}: 0.95 vs. sCT\textsubscript{MR}: 0.89). This quite clear image quality difference can be explained by two main reasons. Firstly, CBCT and the reference pCT are both based on the same physical principle to generate a volumetric image, the interaction of x-rays with tissues of different electron density. For MR-imaging, the underlying physical mechanism is fundamentally different. Image intensities do not correlate with electron density and show a different contrast than CT images. As a consequence, the sCT generation based on MR-images is more challenging for the DCNN than a conversion based on CBCTs. Secondly, and this is also connected to the image intensities, image registration between MR and CT images is more challenging than a registration between CBCTs and CTs. This has a direct influence on the training of the DCNN, which depends on paired CBCT-CT and MR-CT image sets. Furthermore, we assume that CBCT or MR and the reference pCT are perfectly aligned when we calculate image similarity metrics on a voxel by voxel wise manner. This means that a slight registration error can lead to increased (MAE, ME)/decreased (DSC) similarity metrics during image evaluation. However, the slight misalignment of MR and CT should only have minor influence on our results since we visually observed clear differences between the images and we carefully optimized the registration between MR and CT.

For sCT\textsubscript{CBCT} the obtained MAE is comparable to Maspero \textit{et al} (2020) who achieved a MAE of 51 ± 12 HU for head and neck patients using a cycle-consistent GAN. Chen \textit{et al} (2020) achieved a significantly lower MAE of 19 HU for head and neck cancer patients but also used a registered pCT, together with the CBCT, as input images for a U-net neural network. For sCT\textsubscript{MR}, good agreement with the patch based 3D-convolutional network of Dinkla \textit{et al} (2019) was achieved. For head and neck patients they reported a MAE of 75 ± 9 HU. For brain tumor patients Spadea \textit{et al} (2019) reported a slightly lower MAE of 54 ± 7 HU using the same DCNN-architecture as in this work.

Results from proton dose calculations using single-beam plans confirm the findings from the image quality analysis. On average, sCT\textsubscript{CBCT} resulted in a 2%/2 mm gamma pass ratio of 99.3% (SD: 0.8%) which is slightly higher than the mean pass ratios of 98.2% (SD: 1.9%) observed for sCT\textsubscript{MR}. The dosimetric
Figure 8. (a) Absolute dose difference of selected organs at risk (OARs) for sCT_{CBCT} and sCT_{MR}, (b) relative dose difference for OARs, (c) absolute dose difference for target volumes and (d) relative dose difference for target volumes. Whiskers extend to 1.5*IQR and outliers are marked by red and blue dots.

differences between sCT_{CBCT} and sCT_{MR} for the single beam proton plans seem to be not as pronounced as the image quality differences. The recalculation of clinical treatment plans lead to overall lower pass ratios of 96.6% (SD: 3.3%) for sCT_{CBCT} and 93.5% (SD: 3.4%) for sCT_{MR} (2%/2 mm criteria). The used clinical treatment plans for head and neck cancer patients usually consisted of four beam angles and covered a much larger area than the single beam plans with the artificially created target volume. The target area of clinical plans involved the entire neck while the artificial target was positioned intracranial. Thus, the clinical target area can be considered more challenging than the intracranial target, since the neck is more susceptible to anatomical changes and positioning variations.

The analysis of the mean dose to selected OARs and target volumes and the dose-volume histograms revealed that the lower image quality of sCT_{MR} seems to have a measurable effect on the global dose distribution (gamma pass ratio) but a very low effect on the actual dose to anatomical structures (targets...
volumes and OARs) relevant for treatment planning and dose calculations. For target volumes (PTV and CTV) a maximum absolute dose deviation of 0.2 Gy was observed. The clinical relevancy of sCTs was further confirmed by the very high agreement of NTCP-values calculated on pCT and sCTs. Since pCT and sCTs were deformably registered during data preparation, target and OAR delineations could be transferred from the pCT to the sCTs. However, especially in soft tissues DIR is challenging and can lead to errors. This could be overcome by experts delineating OARs and targets individually on each image. Although, for the extent of the dataset we used, this was not feasible. In the future, deep learning auto segmentation might also enable delineation on the sCTs.

Training the networks with image pairs acquired on the same day (rCT-CBCT and pCT-MR) insured equal learning conditions for MR- and CBCT-based networks. For evaluation and comparison of sCTs however, CBCTs had to be registered to the pCT as well. Since the time between pCT and first CBCT is around three weeks, this could have introduced a small bias towards MR-based sCTs, which were acquired on the same day as the reference pCT. DIR between CBCT and pCT was used to minimize this effect.

In contrast to CBCTs, various acquisition sequences and techniques, that alter the appearance and tissue contrast, exist for MR. In literature, a variety of sequences has been used for sCT generation based on neural networks. Since we used retrospective clinical data, we had no influence on the acquired MR-sequences. The used sequences are routinely acquired for head and neck cancer patients and therefore have clinical relevancy.
and represent data that is already available. We chose the in-phase image of a 3D SPGRE sequence since it resulted in images with the highest resolution and visual quality. The chosen sequence was also used for creation of sCTs in previous works reported in literature (Maspero et al 2018, Florkow et al 2020). Contrast agents were used for MR-image acquisition and could in principle interfere with the neural networks ability to learn a correct translation of MR to CT intensities. In our resulting sCTs we did not observe any visual impairment caused by the contrast agent, but a general influence on the training performance cannot be ruled out.

The used network architecture, which is derived from a U-net, has shown its potential for MR- and CBCT-based sCT conversion in previous works (Spadea et al 2019, Thummerer et al 2020). Our results have confirmed these findings. Recently also GANs have been applied to radiotherapy related image synthesis tasks (Liang et al 2019, Liu et al 2019). GANs have the advantage that they cannot only be trained with paired image data (Jin et al 2019) but also with unpaired imaging data (Wolterink et al 2017, Maspero et al 2020). This eliminates the need of DIR during data preparation and therefore speeds up the training process and removes a possible source of error, introduced by using DIR in the first place. Another approach to improve

---

**Figure 10.** Normal tissue complication probability for (a) dysphagia grade 2 or higher and (b) xerostomia grade 2 or higher calculated on pCT, sCT_{CBCT} and sCT_{MR}.
MR-based sCTs in terms of network architecture might be the use of multiple MR-sequences for network training. This can support the neural network training to better distinguish between different tissues and thereby lead to improved sCTs (Florkow et al 2020).

A limitation of this study is the reduced axial field of view due to the severity of the CBCT scatter artifacts below the shoulders. In order to allow a meaningful comparison, this FOV reduction was also performed on the MR-images. The removed part below the shoulders was still necessary for the calculation of clinical treatment plans and therefore parts of the pCT were stitched to the sCTs for clinical dose calculations. The influence of these stitched image parts on the results is minor. Only very limited parts of the target volumes and some OARs, located in the lower neck, were not covered by the cropped field of view of CBCT and MR. The image stitching was also used to add patient couch and head support to the image. These structures are required for dose calculations since beams from certain angles traverse them and influence the dose distributions. Our dataset was almost completely limited to patients who received primary RT. Only a single patient received postoperative RT. Postoperative cases are likely to contain surgery-related features (e.g. surgical clips, staples, flaps) that can cause image artifacts and interfere with the image synthesis. This potential influence warrants investigation in future work.

We performed the comparison of sCT\textsubscript{CBCT} and sCT\textsubscript{MR} solely for head and neck cancer patients. CBCT- and MR-based deep learning techniques have been reported for many other anatomical locations, including brain (Spadea et al 2019, Han 2017, Kazemifar et al 2019, Koike et al 2019), breast (Maspero et al 2020), lung (Maspero et al 2020) and pelvis (Liu et al 2019, Maspero et al 2018, Harms et al 2019, Kurz et al 2019). Further work is necessary to also perform a comparison of sCT\textsubscript{CBCT} and sCT\textsubscript{MR} in these anatomical locations. For CBCT imaging, the head and upper neck are advantageous sites, since the patient diameter is quite limited and with increasing diameter also scatter artifacts increase. Therefore, it can be expected that the CBCT image quality in anatomical locations such as lung or pelvis is lower than for head and neck cancer patients. This might lead to a reduced image quality of sCT\textsubscript{CBCT}. MR-based sCTs do not suffer from these scatter artifacts and therefore the image quality difference might be smaller in other anatomical locations. In the thorax, breathing motion can lead to image artifacts on CBCT and MR. These image artifacts might impair the image quality and thereby also the accuracy of sCTs. MR-images were acquired on a diagnostic MR-scanner while for CBCTS an on-board imaging device was used. Gantry-mounted MR scanners are not yet clinically available but the image quality would likely be lower on such a device. This would probably also influence the image quality of MR-based sCTs and future investigations are required to study this impact.

This study was performed with a relatively limited dataset of 27 patients. Consistent results were observed across the patient cohort and no major outliers were detected. A larger patient cohort would be desirable since it is more likely to include rare edge cases that might lead to sCT-conversion and dose calculation errors. Stringent quality assurance procedures (van Harten et al 2020) are required to detect these errors and establish trust into the accuracy of sCTs. Especially for MR-systems, which are known to be susceptible to geometric distortions, further evaluation and QA-mechanisms for sCTs have to be introduced. Only then MR-based sCTs can also provide reliable position verification, useful in future MR-only scenarios.

### 5. Conclusion

In this work, we presented a comparison of CBCT- and MR-based sCTs generated with DCNNs, using the same set of patients. CBCT-based sCTs showed a higher image similarity when compared to pCT images than MR-based sCTs. As a consequence, the dosimetric evaluation using gamma analysis showed higher agreement for sCT\textsubscript{CBCT} than for sCT\textsubscript{MR}. A recalculation of clinical treatment plans however revealed that the influence of the lower image quality is insignificant for dose-volume parameters of target volumes and selected OARs. From a dosimetric point of view, sCT\textsubscript{CBCT} and sCT\textsubscript{MR} for head and neck patients seem to be equally suited for daily adaptive PT.

### Acknowledgments

Support by the European Association for Cancer Research in form of a travel grant for Paolo Zaffino is gratefully acknowledged. The authors would also like to thank the developer teams of openREGGUI, Plastimatch and Elastix for the provision of their open source software tools. This work was financially supported by a grant of the Dutch Cancer Society (KWF research project 11518).

### ORCID iDs

Adrian Thummerer @ https://orcid.org/0000-0002-1874-5030
Bas A de Jong @ https://orcid.org/0000-0002-6535-0840
References

Albertini F, Matter M, Nenoff P, Zhang Y and Lomax A 2019 Online daily adaptive proton therapy Br. J. Radiol. 93 20190594
Arabi H, Dowling J A, Burgos N, Han X, Greer P B, Koutsouvelis N and Zaidi H 2018 Comparison of synthetic CT generation algorithms for MRI-only radiation planning in the pelvic region 2018 IEEE Nuclear Science Symp. and Medical Imaging Conf. Proc. (NSS/MIC) (Piscataway, NJ: IEEE) pp 1–3
Chen L, Liang X, Shen C, Jiang S and Wang J 2020 Synthetic CT generation from CBCT images via deep learning Med. Phys. 47 1115–25
Dinkla A M, Florkow M C, Maspero M, Savenije M H F, Zijlstra E, Doornenbal P A H and Stralen M 2019 Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network Med. Phys. 46 4095–104
Florkow M C et al 2020 Deep learning–based MR-to-CT synthesis: the influence of varying gradient echo–based MR images as input channels Magn. Reson. Med. 83 4123–41
Guerreiro F, Burgos N, Dunlop A, Wong K, Petkar I, Nutting C and Harrington K 2017 Evaluation of a multi-atlas CT synthesis approach for MRI-only radiotherapy treatment planning Med. Physica 55 7–17
Guerreiro F, Koivula I, Seravalli E, Janssens G O, Maduro J H, Brouwer C L and Koevevaa E W 2019 Feasibility of MRI-only photon and proton dose calculations for pediatric patients with abdominal tumors Phys. Med. Biol. 64 5
Han X 2017 MR-based synthetic CT generation using a deep convolutional neural network method Med. Phys. 44 1408–19
Hansen D C, Landry G, Kamp F, Li M, Belka C, Parodi K and Kurz C 2018 ScatterNet: a convolutional neural network for cone-beam CT intensity correction Med. Phys. 45 4916–26
Harms J, Lei Y, Wang T, Zhang R, Zhou J, Tang X and Currar W J 2019 Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography Med. Phys. 46 3998–4009
Hoffmann A, Oborn B, Moteabbed M, Yan S, Bortfeld T, Knopf A and Fuchs H 2020 MR-guided proton therapy: a review and a preview Radiother. Oncol. 151 1–13
Hoffmann L, Alber M, Jensen M F, Holt M I and Møller D S 2017 Adaptation is mandatory for intensity modulated proton therapy of advanced lung cancer to ensure target coverage Radiother. Oncol. 122 400–5
Hua C, Yao W, Kidani T, Tomida K, Ozawa S, Nishimura T and Fujisawa T E 2017 A robotic C-arm cone beam CT system for image-guided proton therapy: design and performance Br. J. Radiol. 90 1079
Jin C-B, Kim H, Liu M, Jung W, Joo S, Park E and Cui X 2019 Deep CT to MR synthesis using paired and unpaired data Sensors 19 23611
Karlsson M, Karlsson M G, Nyholt T, Amy C and Zackrisson E 2009 Magnetic resonance imaging in the radiotherapy clinic Int. J. Radiat. Oncol. Biol. Phys. 74 644–51
Kazemifar S, McGuire S, Timmerman R, Wardak Z, Nguyen D, Park Y, Jiang S and Owrangi A 2019 MRI-only brain radiotherapy: assessing the dosimetric accuracy of synthetic CT images generated using a deep learning approach Radiother. Oncol. 136 56–63
Kida S, Nakamoto T, Nakano M, Nawa K, Haga A, Koteiku J and Nakagawa K 2018 Cone beam computed tomography image quality improvement using a deep convolutional neural network Careus 10 4
Klein S, Staring M, Murphy K, Viergever M A and Pluim J P W 2010 Elastic: a toolbox for intensity-based medical image registration IEEE Trans. Med. Imaging 29 196–205
Koike Y, Akino Y, Sumida I, Shiomi H, Mizuno H, Yagi M and Isohashi F 2019 Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy J. Radiat. Res. 61 92–103
Koivula I, Wee L and Korhonen J 2016 Feasibility of MRI-only treatment planning for proton therapy in brain and prostate cancers: dose calculation accuracy in substitute CT images Med. Phys. 43 1634–42
Kupelian P and Sonke -J-J 2014 Magnetic resonance-guided adaptive radiotherapy: a solution to the future Semin. Radiat. Oncol. 24 227–32
Kurz C, Dedes G, Resch A, Reiner M, Ganswindt U, Nijhuis R and Thieke C 2015 Comparing cone-beam CT intensity correction methods for dose recalculation in adaptive intensity-modulated proton and photon therapy for head and neck cancer Acta Oncol. (Maalr) 54 1651–7
Kurz C, Dedes G, Resch A, Reiner M, Ganswindt U, Nijhuis R and Thieke C 2015 Comparing cone-beam CT intensity correction methods for dose recalculation in adaptive intensity-modulated photon and proton therapy for head and neck cancer Acta Oncol. (Maalr) 54 1651–7
Kurz C, Kamp F, Park Y-K, Zöllner C, Rit S, Hansen D and Podesta M 2016 Investigating deformable image registration and scatter correction for CBCT-based dose calculation in adaptive IMPT Med. Phys. 43 3655–46
Kurz C, Maspéro M, Savenije M H F, Landry G, Kamp F, Pinto M and Li M 2019 CBCT correction using a cycle-consistent generative adversarial network and unpaired training to enable photon and proton dose calculation Phys. Med. Biol. 64 22
Landry G, Hansen D, Kamp F, Hoyle B, Weller J, Parodi K and Kurz C 2019 Comparing Unet training with three different datasets to correct CBCT images for prostate radiotherapy dose calculations Phys. Med. Biol. 64 035011
Landry G, Nijhuis R, Dedes G, Handrack J, Thieke C, Janssens G and Orban de Xivry J 2015 Investigating CT to CBCT image registration for head and neck proton therapy as a tool for daily dose recalculation Med. Phys. 42 3
Langendijk J A, Lambin P, De Ruyscher D, Widder J, Bos M and Verheij M 2013 Selection of patients for radiotherapy with protons aiming at reduction of side effects: the model-based approach Radiother. Oncol. 107 267–73
Liang X, Chen L, Nguyen D, Zhou Z, Gu X, Yang M and Wang J 2019 Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy Phys. Med. Biol. 64 12
Lim-Reinders S, Keller B M, Al-Ward S, Sahgal A and Kim A 2017 Online adaptive radiation therapy Int. J. Radiat. Oncol. Biol. Phys. 99 994–1003
Liu Y, Lei Y, Wang Y, Shafai-Erfani G, Wang T, Tian S and Yang X 2019 Evaluation of a deep learning-based pelvic synthetic CT generation network for MRI-based prostate proton treatment planning Med. Phys. Biol. 64 20
Maspéro M, Houssein A C, Savenije M H F, van Heijst T C F, Verhoef J J C, Kotte A N T J and van den Berg C A T 2020 A single neural network for cone-beam computed tomography-based radiotherapy of head-and-neck, lung and breast cancer Phys. Imaging Radiother. Oncol. 14 24–31
Maspéro M, Savenije M, Dinkla A M, Seevinck P R, Intven M, Jurgentiem-Schulz J M, Kerkmeijer L and van den Berg C 2018 Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy Phys. Med. Biol. 63 185001
Meyer P, Noblet V, Mazzara C and Lallemand A 2018 Survey on deep learning for radiotherapy Comput. Biol. Med. 98 126–46
Nenoff L, Matter M, Hedlund Lindmar J, Weber D C, Lomax A J and Albertini F 2019 Daily adaptive proton therapy— the key to
innovative planning approaches for paranasal cancer treatments Acta Oncol. (Madr) 58 1423–8
Oborn B M, Dowdell S, Metcalfe P E, Crozier S, Mohan R and Keall P J 2017 Future of medical physics: real-time MRI-guided proton
therapy: real-time Med. Phys. 44 e77–e90
Park Y-K, Sharp G C, Phillips J and Winey B A 2015 Proton dose calculation on scatter-corrected CBCT image: feasibility study for
adaptive proton therapy Med. Phys. 42 4449–59
Pileggi G, Speier C, Sharp G C, Izquierdo D, Catana C, Pursley J and Amato M F 2018 Proton range shift analysis on brain pseudo-CT
generated from T1 and T2 MR Acta Oncol. (Madr) 57 1521–31
Shen C, Nguyen D, Zhou Z, Jiang S B, Dong B and Jia X 2020 An introduction to deep learning in medical physics: advantages, potential,
and challenges Phys. Med. Biol. 65 08TR01
Spadea M F, Pileggi G, Zaffino P, Salome P, Catana C, Izquierdo-Garcia D and Amato F 2019 Deep Convolution Neural Network
(DCNN) multiplane approach to synthetic CT generation from MR images—application in brain proton therapy Int. J. Radiat.
Oncol. Biol. Phys. 105 495–503
Staring M, Klein S and Pluim J P W 2007 A rigidity penalty term for nonrigid registration Med. Phys. 34 4098–108
Stock M, Georg D, Ableitinger A, Zechner A, Utz A, Mumot M and Kragl G 2018 The technological basis for adaptive ion beam therapy
at MedAustron: status and outlook. Z. Med. Phys. 28 196–210
Theano Development Team et al 2016 A Python framework for fast computation of mathematical expressions (arXiv:1605.02688)
Thummerer A, Zaffino P, Meijers A, Marmitt G G, Seco J, Steenbakkers R J H M and Knop A-C 2020 Comparison of CBCT based
synthetic CT methods suitable for proton dose calculations in adaptive proton therapy Phys. Med. Biol. 65 095002
van Harten L, Wolterink J M, Verhoeven J J C and Išgum I 2020 Automatic online quality control of synthetic CTs Proc. SPIE
11313 113131M
Veiga C, Alshaikhi J, Arnos R, Lourenço A M, Modat M, Ourselin S and Royle G 2015 Cone-beam computed tomography and
deformable registration-based “Dose of the Day” calculations for adaptive proton therapy Int. J. Part. Ther. 2 404–14
Veiga C, Janssens G, Teng C-L, Baudier T, Hotoiu L, Mcclelland J R and Royle G 2016 First clinical investigation of cone beam computed
tomography and deformable registration for adaptive proton therapy for lung cancer Int. J. Radiat. Oncol. Biol. Phys. 93 549–59
Wang T, Manohar N, Lei Y, Dhabaan A, Shu H-K, Liu T and Curran W J 2019 MRI-based treatment planning for brain stereotactic
radiosurgery: dosimetric validation of a learning-based pseudo-CT generation method Med. Dosim. 44 199–204
Widder J, Van Der Schaaf A, Lambin P, Marijnen C A M, Pignol J-P, Rasch C R and Slotman B J 2016 The quest for evidence for proton
therapy: model-based approach and precision medicine Int. J. Radiat. Oncol. Biol. Phys. 95 30–36
Wolterink J M, Dinkla A M, Savenije M H F, Seevinck P R, van den Berg C A T and Išgum I 2017 Deep MR to CT synthesis using
unpaired data Simulation and Synthesis in Medical Imaging. SASHIMI 2017. Lecture Notes in Computer Science vol 10557,
ed S Tsaftaris et al (Berlin: Springer) pp 14–23
Zaffino P, Raudaschl P, Fritscher K, Sharp G C and Spadea M F 2016 Plastimatch MABS, an open source tool for automatic image
segmentation Med. Phys. 43 3153–60