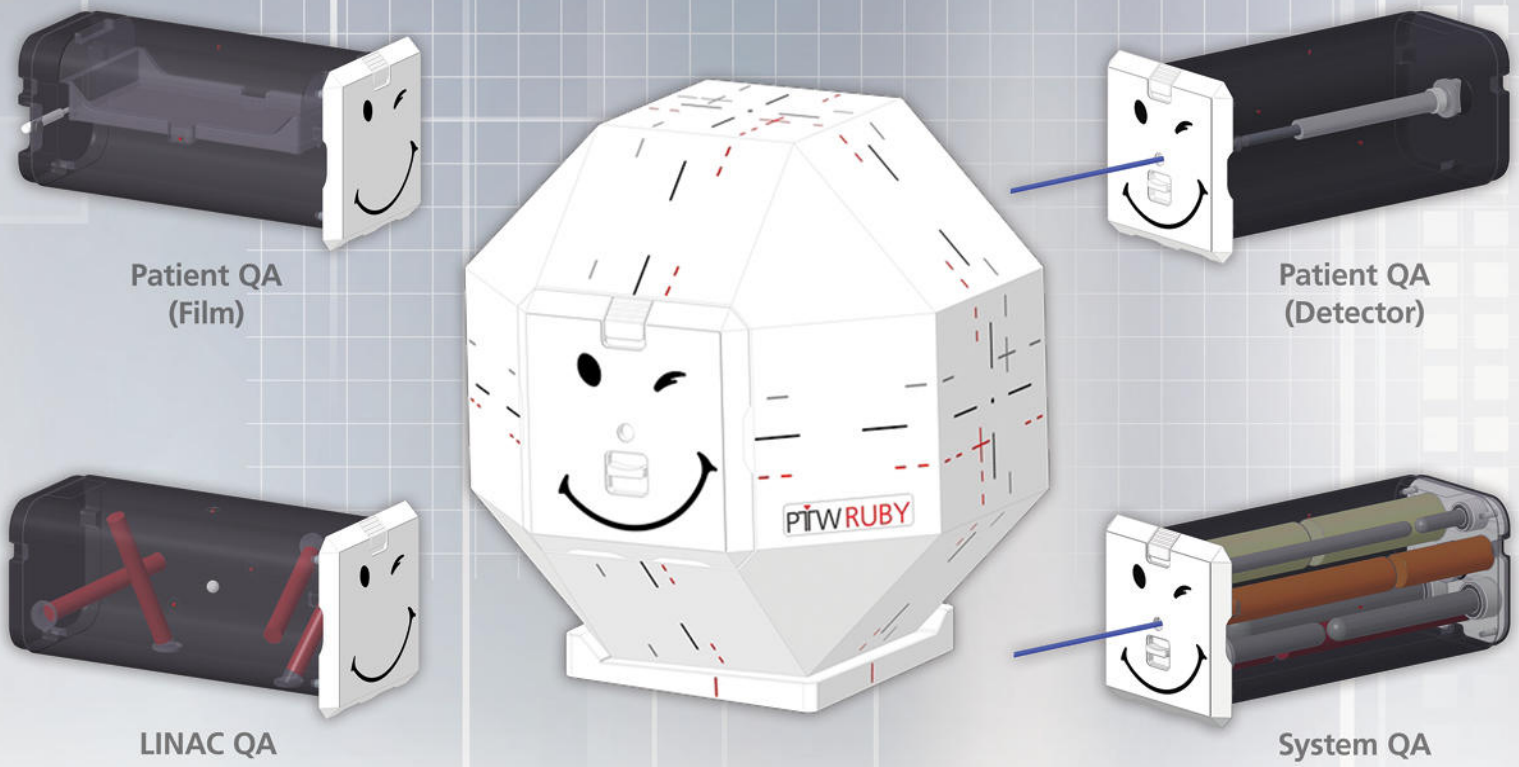


RUBY – The Modular QA Phantom

Multi-purpose, high-precision all-in-one phantom for System, LINAC and Patient-specific QA needs

The RUBY phantom delivers all the functionality you need to perform comprehensive end-to-end tests of the entire stereotactic treatment chain. Application-specific inserts enable you to perform system QA, machine QA or patient QA. And unlike other systems, RUBY enables you to **perform all tests with one phantom**. This saves time and eliminates the possibility of dangerous positional errors. With RUBY, you **get the flexibility and the precision** you need to ensure system accuracy. What else would you expect from the worldwide leader in dosimetry technology?

Contact PTW today for all your radiotherapy QA needs.

(516) 827-3181 • www.ptwdosimetry.com

PTW THE DOSIMETRY COMPANY

Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network

Anna M. Dinkla[a)]*
Department of Radiotherapy, Division of Imaging & Oncology, University Medical Center Utrecht, Utrecht, The Netherlands
Computational Imaging Group for MR Diagnostics & Therapy, Center for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

Mateusz C. Florkow
Centre for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

Matteo Maspero, and Mark H. F. Savenije
Department of Radiotherapy, Division of Imaging & Oncology, University Medical Center Utrecht, Utrecht, The Netherlands
Computational Imaging Group for MR Diagnostics & Therapy, Center for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

Frank Zijlstra
Centre for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

Patricia A. H. Doornaert
Department of Radiotherapy, Division of Imaging & Oncology, University Medical Center Utrecht, Utrecht, The Netherlands

Marijn van Stralen
Centre for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands
MRIguidance B.V, Utrecht, The Netherlands

Marielle E. P. Philippens
Department of Radiotherapy, Division of Imaging & Oncology, University Medical Center Utrecht, Utrecht, The Netherlands

Cornelis A. T. van den Berg
Department of Radiotherapy, Division of Imaging & Oncology, University Medical Center Utrecht, Utrecht, The Netherlands
Computational Imaging Group for MR Diagnostics & Therapy, Center for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

Peter R. Seevinck
Centre for Image Sciences, University Medical Center Utrecht, Utrecht, The Netherlands
MRIguidance B.V, Utrecht, The Netherlands

(Received 11 March 2019; revised 15 May 2019; accepted for publication 10 June 2019; published 9 July 2019)

Purpose: To develop and evaluate a patch-based convolutional neural network (CNN) to generate synthetic computed tomography (sCT) images for magnetic resonance (MR)-only workflow for radiotherapy of head and neck tumors. A patch-based deep learning method was chosen to improve robustness to abnormal anatomies caused by large tumors, surgical excisions, or dental artifacts. In this study, we evaluate whether the generated sCT images enable accurate MR-based dose calculations in the head and neck region.

Methods: We conducted a retrospective study on 34 patients with head and neck cancer who underwent both CT and MR imaging for radiotherapy treatment planning. To generate the sCTs, a large field-of-view T2-weighted Turbo Spin Echo MR sequence was used from the clinical protocol for multiple types of head and neck tumors. To align images as well as possible on a voxel-wise level, CT scans were nonrigidly registered to the MR (CT_{reg}). The CNN was based on a U-net architecture and consisted of 14 layers with $3 \times 3 \times 3$ filters. Patches of $48 \times 48 \times 48$ were randomly extracted and fed into the training. sCTs were created for all patients using threefold cross validation. For each patient, the clinical CT-based treatment plan was recalculated on sCT using Monaco TPS (Elekta). We evaluated mean absolute error (MAE) and mean error (ME) within the body contours and dice scores in air and bone mask. Also, dose differences and gamma pass rates between CT- and sCT-based plans inside the body contours were calculated.

Results: sCT generation took 4 min per patient. The MAE over the patient population of the sCT within the intersection of body contours was 75 ± 9 Hounsfield Units (HU) (± 1 SD), and the ME was 9 ± 11 HU. Dice scores of the air and bone masks (CT_{reg} vs sCT) were 0.79 ± 0.08 and 0.70 ± 0.07, respectively. Dosimetric analysis showed mean deviations of $-0.03\% \pm 0.05\%$ for dose within the body contours and $-0.07\% \pm 0.22\%$ inside the >90% dose volume. Dental artifacts obscuring the CT could be circumvented in the sCT by the CNN-based approach in combination with

Turbo Spin Echo (TSE) magnetic resonance imaging (MRI) sequence that typically is less prone to susceptibility artifacts.

Conclusions: The presented CNN generated sCTs from conventional MR images without adding scan time to the acquisition. Dosimetric evaluation suggests that dose calculations performed on the sCTs are accurate, and can therefore be used for MR-only radiotherapy treatment planning of the head and neck. [https://doi.org/10.1002/mp.13663]

Key words: deep learning, head and neck cancer, MR-guided therapy, MR-only radiotherapy, synthetic CT

1. INTRODUCTION

The use of magnetic resonance imaging (MRI) in radiotherapy is growing rapidly, due to its superior soft tissue contrast as compared to computed tomography (CT).[1] In radiation treatment of head and neck cancers, MRI is often used in addition to CT to reduce interobserver variability of target delineation.[2,3] The excellent soft-tissue contrast of anatomical MRI makes this modality ideally suited for staging and follow-up in the head and neck region.[4] In addition, reduced dental artifacts compared to CT may improve delineation of target and organs at risk (OARs) for treatment planning.[5,6] However, dose calculation and patient positioning still rely on CT, which provides electron density information and high geometrical accuracy.[7] Therefore, a multimodality imaging workflow is currently used: MR images are registered to CT images allowing the use of both scans for treatment preparation. To facilitate rigid registration between CT and MR images, the MRI is acquired in treatment position with dedicated fixation devices used both during CT and MRI examination.[8–11] Nevertheless, the head and neck region is a challenging region for image registration and can lead to systematic errors of 2–3 mm.[12–15] An attractive approach making the registration step in the treatment simulation obsolete is an MR-only workflow. Using only MR images during simulation is beneficial since it simplifies the simulation workflow and exploits the benefits of MRI for high-precision treatment planning. Moreover, the emergence of combined MRI and treatment delivery machines[16] drives the development of fast online MR-guided treatment planning and MR-based dose calculation.

Heterogeneity correction is required for accurate dose calculations, since bulk assignment of the whole body to water density can lead to large dosimetric errors for the head and neck region.[17] To overcome the lack of electron densities on MRI, a so-called pseudo-CT or synthetic CT needs to be generated.[7,13,18–20] Several methods have been proposed in literature to produce a synthetic computed tomography (sCT). The two main approaches are voxel-based conversion[21] and atlas-based methods.[22–24] With voxel-based methods, often nonconventional sequences like ultrashort or zero echo time sequences are acquired to enhance bone visualization and aid the differentiation between air and bone on MRI.[25–27] In addition, these methods usually require multiple sequences to be added to the exam, including the aforementioned specialized sequences increasing total scan time. Most current sCT methods focus on the skull or pelvic region, and less work has been done on the entire head and neck region, possibly due to the large field of view (FOV) required for the MRI/sCT. Despite acquiring CT and MR imaging on the same day and using the same fixation devices, image registration between head and neck scans can be difficult. Nevertheless, atlas-based methods for sCT generation are most common in the head and neck region, but nonstandard patient anatomies and dental artifacts remain challenging. Farjam et al. presented a multi-atlas method with mean absolute error (MAE) of 124 ± 21 HU and good dosimetric results.[23] However, their method was quite time consuming (±20 min). Also, they noted that one of the most important challenges in using atlas-based synthetic CT generation is to deform the atlas to patients with large anatomical abnormalities such as surgical resection and large tumors. Guerreiro et al. presented another multi-atlas approach with lower MAE (90 + 12 HU) but for limited FOV neck scans, which contained only the planning target volume (PTV) region.[22] They also reported that patient-specific abnormalities not represented in the atlas are an exclusion criterion.

A relatively new approach in sCT generation is to train a convolutional neural network (CNN) to convert the MR images to CT.[18,28–30] A CNN is a fully trainable model that can learn the mapping between input MR and output CT images. The network consists of many convolutional filters with tunable weights whose values are optimized during training by minimizing an error function for paired input and output data. Once the network is trained, it can convert a newly presented MR to CT. Han trained a U-Net architecture[31] using sixfold cross validation with 19 patients to successfully produce sCT images of the head in two-dimensional (2D). To address issues with discontinuities across slices with a 2D approach, Nie et al. applied a three-dimensional (3D) patch-based approach to estimate the CTs for the prostate.[32] Training a 3D fully convolutional network requires more GPU memory and more data; moreover, they are generally difficult to train due to the large number of weights in the network.[28,32] Choosing a patch-based approach can exploit the 3D information that is available in medical image dataset, while limiting the number of parameters. In this work, we investigated the use of a 3D U-NET.[28,31,32] Another reason for the patch-based approach is that it is expected to be

relatively insensitive to variations in patient anatomy that may occur frequently in head and neck patients.

In this work, we aimed to generate sCTs of the head and neck with the proposed network to facilitate an MR-only radiotherapy workflow. We want to investigate the feasibility of generating sCTs without having to add an additional dedicated sequence but instead use a multipurpose sequence. The sequence that we selected was optimized for delineation of OARs. Priority was therefore given to the contrast of the OARs and obtaining sufficient geometrical fidelity. Moreover, since Turbo Spin Echo (TSE) MRI is typically less prone to susceptibility artifacts, the choice of these images as input for sCT generation for head and neck cancers has as particular advantage that it is less affected by dental artifacts compared to commonly used gradient echo sequences. We hypothesize that with a deep learning-based method, we can generate accurate sCTs from these MR images. We evaluated the dosimetric accuracy of dose calculations based on the obtained sCT images with respect to the gold standard CT.

2. MATERIALS AND METHODS

2.A. Patient data selection

We collected image data from 34 head and neck cases treated at our department between December 2017 and June 2018. Data were collected according to local ethical standards. Tumors were classified according to the TNM staging protocol between stage 1 and 4, with lymph node involvement between 0 and 3. For all oropharyngeal and nasopharyngeal cancer patients who underwent 3 T MRI as part of the treatment preparation, a large FOV T2-weighted MRI with Dixon reconstruction was available that could be used for sCT generation. This sequence was optimized for OAR delineation, having sufficient contrast and a large FOV (Fig. 1). The majority was treated for oropharyngeal cancer (24 patients, in base of mouth, base of tongue, or tonsil). Seven patients were treated in the nasopharynx and two had unilateral recurrent disease in the neck region. For most patients, the CT and the MRI were acquired on the same day (median spacing was 1 day); however, in some cases, there was much more time in between, up to a maximum of 28 days. These patients of whom the CT and MR were spaced further apart were not excluded, since we adopt deformable image registration in this study. To facilitate training, patients with severe dental streaking artifacts on the CT were excluded from the training data as well as insufficient quality of the MRI (due to motion/aliasing artifacts, or signal voids due to implants). Patients with limited streaking on the CT due to dental implants were not excluded.

2.B. Image acquisition

Patients were scanned head first in treatment position and with the head immobilized in a custom thermoplastic mask during both CT and MR imaging. Both at the CT and at the MRI, patients were positioned on an in-house developed dedicated base plate that accommodates individualized head support and the immobilization mask. For the MRI, two medium-sized flexible general-purpose surface receive coils (dStream Flex M, Philips Healthcare, the Netherlands) were placed on the immobilization mask on lateral sides of the head/neck opposing each other.[9] An additional torso coil/abdominal receiver array (dS Torso, Philips Healthcare, The Netherlands) was placed anteriorly, supported by two coil bridges to avoid compression (Fig. 2). MRI acquisition was performed on the same dedicated MR-RT simulator (3T Ingenia, Philips Healthcare, the Netherlands) and a multislice T2-weighted TSE with mDixon reconstruction was used for training the CNN to generate the sCT images. Transversal images were acquired with the following imaging parameters: 90° flip angle, TR/TE = 9023/100 ms, TSE factor = 27, and readout bandwidth of 876 Hz/vox in the right–left direction, corresponding to 0.5 pixel. In the phase encoding direction (anterior–posterior), an acceleration factor SENSE of 2 was applied resulting in a total acquisition time of 5 min 24 s. The FOV was $45 \times 45 \times 24.9$ cm^3, acquired voxel size $1.3 \times 1.18 \times 3$ mm^3, and reconstructed voxel size $0.9375 \times 0.9375 \times 3$ mm^3. 2D Geometry correction as provided by the vendor was applied and uniformity correction "CLEAR." Finally, a Dixon reconstruction was performed to obtain in-phase, water, and fat images.[33] All planning CT scans were acquired helically on the same scanner (Philips Brilliance Big Bore, Philips Healthcare, the Netherlands), with fixed tube potential of 120 kV, while current (mAs) varied with FOV. Slices were 2 mm thick with in-plane resolution of around 1 mm.

2.C. Image registration

Despite the uniformity of scanning position and patient immobilization, head and neck images are prone to misalignment due to differences in neck flexion,[11] especially with a large FOV containing the shoulder region. Since paired learning of a neural network may be sensitive to pixel-wise misalignments between input training data,[18,32,34,35] we aligned the CTs to the MRI with a nonrigid registration. First, the CT was rigidly registered to the water and in-phase images simultaneously using Euler registration based on mutual information with Elastix 4.7,[36] which was subsequently used as initial transformation for the nonrigid registration of the CT to the in-phase MR images only. A B-Spline transform was then performed based on two metrics: normalized mutual information (MI) and a transform rigidity penalty (TRP) on the CT inside the bone mask (>200 HU), resulting in a cost function of $0.2 \times \text{MI} + 0.8 \times \text{TRP}$. After registration, the CT and MR images were resampled to an isotropic resolution of $1 \times 1 \times 1$ mm^3. To evaluate the performance of the registration between CT and MRI/sCT, we calculated the Dice similarity coefficients (DSC) of the body contours. The found transformations were applied to the body contours, to obtain a body mask in the frame of reference of the MR and CT_{reg}. For training, we applied a dilation of 50 voxels to these masks to include a region outside the body

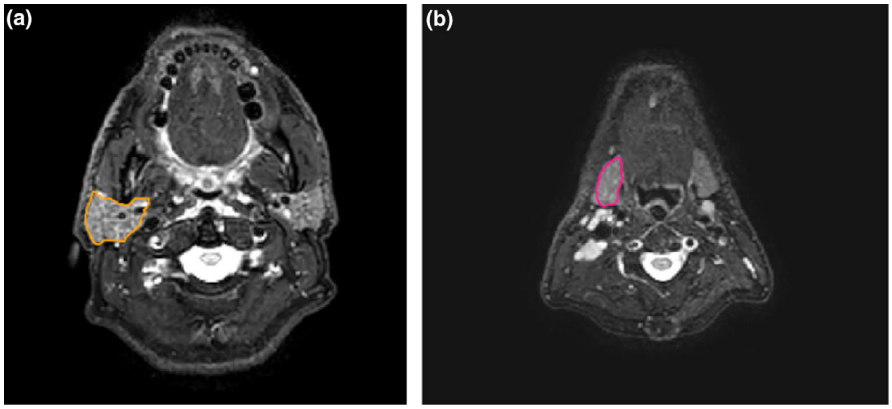


FIG. 1. Two (a&b) transversal slices of a T2 Turbo Spin Echo (TSE) magnetic resonance imaging (MRI) (Dixon water reconstruction), showing sufficient contrast for OAR delineation, since the T2-weighting results in good visibility of the parotid and submandibular glands containing a relatively large amount of free water.(a) The right parotid is delineated in orange. (b) The right submandibular gland is delineated in pink (same patient). [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 2. Patient setup during magnetic resonance (MR) acquisition. The patient is immobilized in a thermoplastic mask attached to an in-house developed base plate. The Flex-M coils are placed on both sides of the neck and fixated using sand bags and the headphones. The anterior body coil is supported by two coil bridges. [Color figure can be viewed at wileyonlinelibrary.com]

contour to also include (MR) image background in the training. The background of the CT images was masked by setting all voxels outside the body contour to −1000 HU. Subsequently, CT Hounsfield Units (HU) were normalized using constant scaling parameters to ensure that most values lie between −1 and 1 by: $2 * \frac{CT - CT_{Min}}{CT_{Max} - CT_{Min}} - 1$, where CT_{Min} and CT_{Max} were set to −1000 and 3000, respectively. The registered and masked CT images will be referred to as CT_{reg}. The in-phase MR images were selected for sCT generation, whose intensities were also normalized from the minimum to maximum intensities to −1 and 1.

2.D. sCT generation

A patch-based deep neural network was used to generate sCTs from MR inputs. The architecture of the network was derived from the U-Net[31] and composed of 14 layers. A schematic representation of the method including our network is given in Fig. 3. One layer consists of multiple convolution filters with size 3 × 3 × 3 voxels. In addition, every layer includes instance normalization[37] and a rectified linear unit activation function. Specifically for this image generation task, nearest-neighbor interpolation was implemented as an upsampling method to avoid the checkerboard artifact reported by Odena et al.[38] The network was implemented using Keras 2.1.3 with a Tensorflow 1.7 backend.

The network was trained for approximately 11 h on mini-batches of eight 3D patches of size 48 × 48 × 48 voxels. Corresponding patches from the in-phase MRI and CT_{reg} were randomly extracted from within body contour, balancing the presence of soft tissue and bone.

A L_1 loss function measured the error between the ground truth CT and the generated sCT. Optimization was done using Nadam[39] with a learning rate of 10^{-3} and 4 million parameters to learn. After 100 epochs with 1000 iterations per epochs, convergence was reached using a single Nvidia GeForce GTX 1080 GPU.

We trained the network using 22 datasets and kept the remaining patients for validation. By performing a threefold cross validation, we obtained synthetic CT for all 34 patients. The sCT reconstruction took approximately 4 min on the aforementioned GPU.

2.E. sCT evaluation

Synthesized CT images were compared to CT_{reg} on a voxel-wise basis. Only voxels within the intersection of the body contours were evaluated to compare HU. An additional body contour was created based on the sCT for this purpose. For each patient, the accuracy of the sCT was evaluated by calculating the MAE and Mean Error (ME) in the body region (Matlab release 2015a, Mathworks, Natick, MA, USA).

2.F. Dosimetric evaluation

To assess dose calculation accuracy, the clinically optimized dose plan was recalculated on the sCT and CT_{reg} for

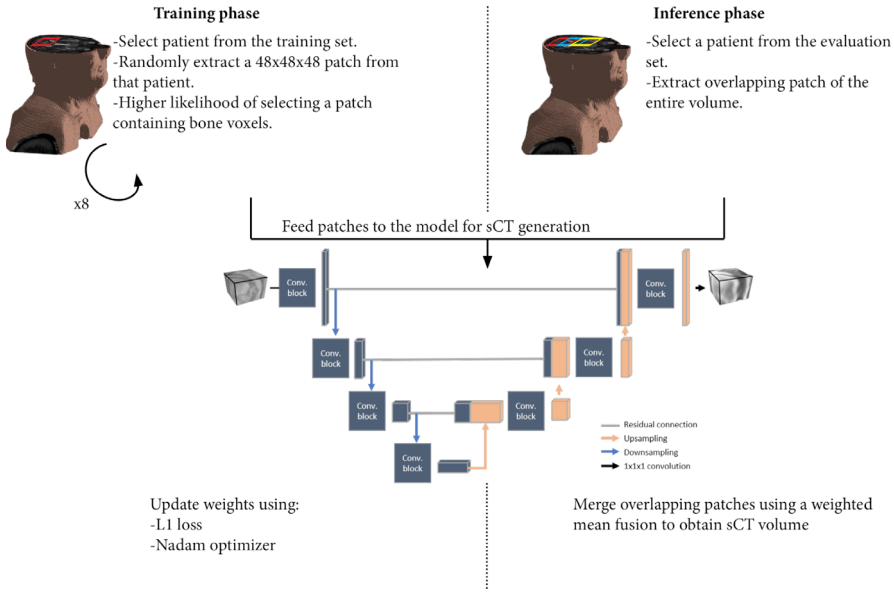


FIG. 3. A schematic representation of our method, including network architecture. [Color figure can be viewed at wileyonlinelibrary.com]

32 patients, since two patients did not have a treatment plan. Dose calculation was performed with the Monte Carlo dose algorithm with a 3-mm^3 grid, with 3% statistical uncertainty using the Monaco treatment planning system (TPS) (v 5.11.02, Elekta AB, Sweden). Dose cubes were exported and analyzed in Matlab. Dosimetric performance of the sCTs was evaluated by calculating the mean error and 3D gamma pass rates of the dose distributions (CT_{reg}-sCT) calculated using two different criteria: 2 mm/2% and 3 mm/3%, with a 10% dose threshold.

3. RESULTS

3.A. sCT evaluation

The body masks of the CT_{reg} and sCTs agreed well with a mean dice score of 0.98 + 0.01 (±1 SD, range: 0.96–0.99). An example of an sCT is given in Fig. 4. Despite some CT scans having streaking artifacts from dental implants, the CNN can reconstruct the sCT without artifacts, as can be appreciated in Fig. 5. Quantitatively, the average of the MAEs over all 34 generated sCTs was 75 ± 9 HU. Average ME was 9 ± 11 HU. Values of the MAE and ME of bone, soft tissue, and air regions are given in Table I. Dice scores of the air and bone masks (CT_{reg} vs sCT) were 0.79 ± 0.08 (±1 SD, range: 0.63–0.91) and 0.70 ± 0.07 (±1 SD, range: 0.52–0.84), respectively.

3.B. Dosimetric evaluation

Mean voxel-wise differences of the dose distribution calculated on the sCT were within 1% (Table II). An example of the dose distributions calculated on CT_{reg} and sCT is given in Fig. 6. The dose inside the target volume is represented by calculating voxel-wise dose differences in the high-dose region only, for voxels receiving dose above 90% of the prescribed dose. The mean difference in this volume was limited to −0.07% ± 0.22%, with a maximum mean deviation is −0.52% over all patients (Fig. 7). Mean gamma pass rates for $\gamma_{2mm2\%}$ and $\gamma_{3mm3\%}$ were 95.6% ± 2.9% (±1 SD, range: 87.5%–98.6%) and 98.7% ± 1.4% (±1 SD, range: 93.0%–99.7%), respectively. With the patch-based approach, also nonstandard patient anatomies were reconstructed correctly, as shown in Fig. 8. For this postoperative case, the mean dose difference was 0.0% ± 0.6% within the body volume and −0.36% ± 2.3% in the high-dose volume (>90%).

4. DISCUSSION

In this study, we evaluated the feasibility of using a 3D patch-based deep learning approach for generating sCT images of the head and neck to be used for dose calculation. We aimed at finding a generic solution to enable fast and accurate MR-based dose calculations in the head and neck region, without the need to acquire additional dedicated MR sequences for sCT generation such as ultrashort or zero echo time MR to discriminate bone from air. This is advantageous for clinical implementation of an MR-only workflow, since time on the scanner is generally limited. Such dedicated 3D scans are relatively time consuming, especially for the large FOVs that are required for this patient group, who find it particularly difficult to endure the entire exam. Furthermore, specialized sequences are difficult to tune, may not be available at all scanners, and would need to be registered to the other image. The presented method utilized just one MRI sequence for both OAR delineation and sCT generation, mitigating the need for additional image registration. With this method, locations of the air and bone are learned in the network, resulting in a high mean dose calculation accuracy of within 1% when compared to CT. Therefore, the benefit of

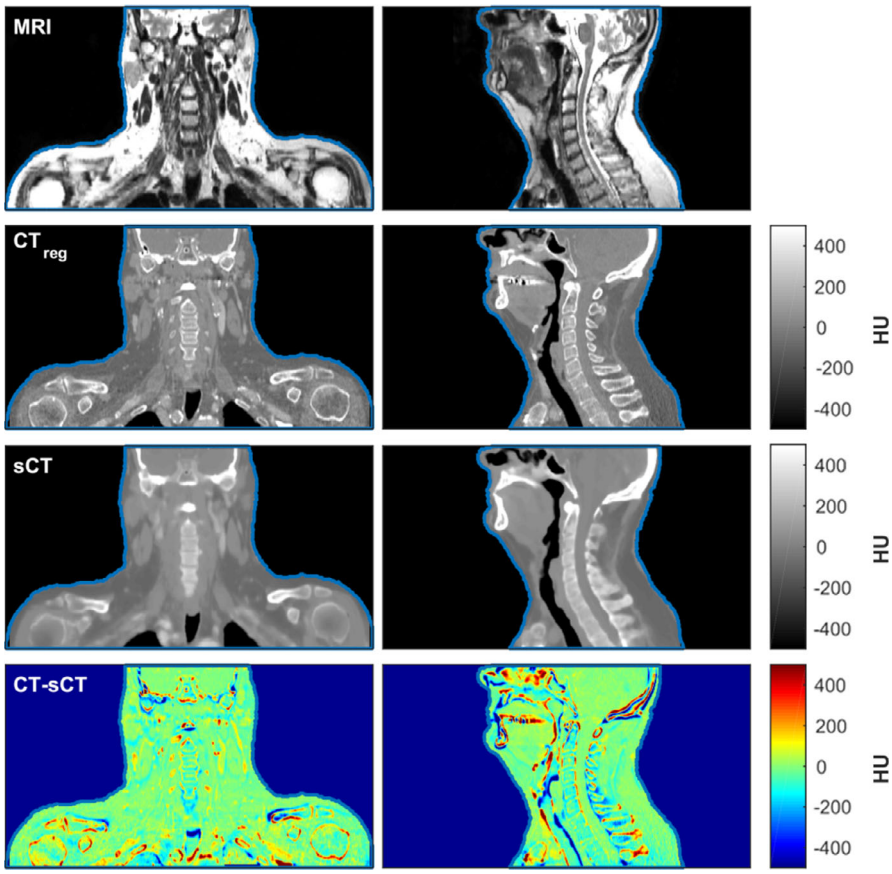


FIG. 4. Example (patient 1) of a coronal (left) and sagittal (right) slice of a magnetic resonance imaging (MRI), CT_{reg}, synthetic computed tomography (sCT), and the difference for an oropharyngeal cancer case. [Color figure can be viewed at wileyonlinelibrary.com]

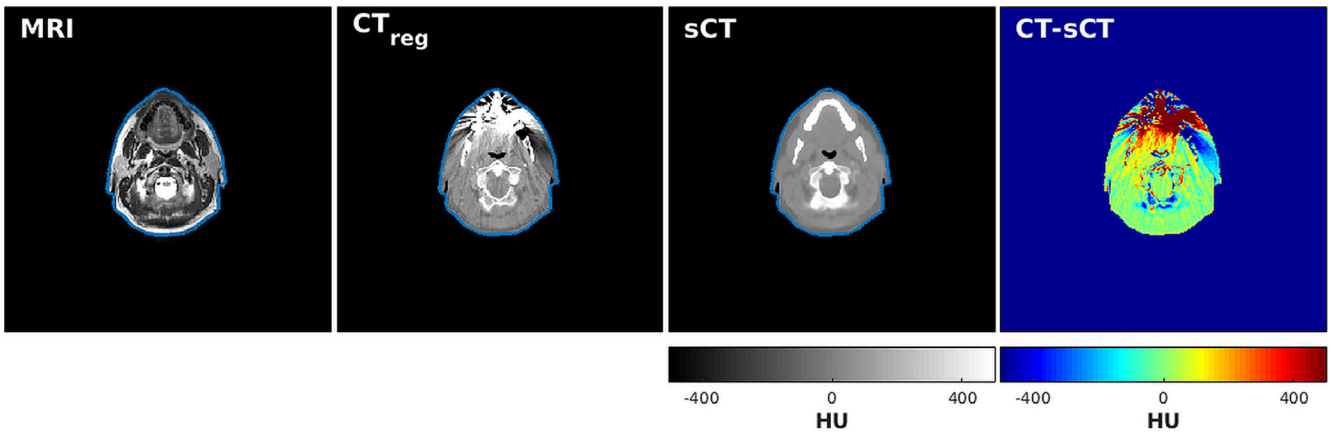


FIG. 5. Example of a patient with streaking artifacts on the computed tomography (CT) from dental implants, which are not manifested on the synthetic computed tomography (sCT). This patient was not excluded in the MAE and ME evaluation, since we did not observe any impact of a few slices with artifacts on the total mean absolute error (MAE) and mean error (ME). [Color figure can be viewed at wileyonlinelibrary.com]

adding a dedicated sequence to aid in bone/air discrimination will likely be limited.

The sCTs obtained in this study had a MAE of 75 ± 9 HU and a ME of 9 ± 11 HU, which was comparable to what is found in literature.[18,22,23,28] In our previous work, we obtained a MAE of 67 ± 11 HU and ME of 13 ± 9 HU, despite using a different network architecture and using head images for brain radiotherapy only.[18] These values are comparable to the first synthetic CTs obtained by training a CNN, reported in a study in which a 2D U-Net was trained on head images, resulting in a MAE of 84.8 ± 17.3.[28] Head and neck synthetic CT is somewhat more challenging due to a wider variety in patient anatomies and often lower resolution images. Nevertheless, Farjam et al. obtained a MAE of

TABLE I. Mean (±1 SD) and range of the mean absolute error (MAE) and mean error (ME) [in Hounsfield Units (HU)] for the whole body, bone (>250 HU), soft tissue, and air (<−200 HU) regions.

	Mean (±1 SD)	Range
MAE		
Body	75 (9)	[62; 94]
Bone	214 (26)	[167; 266]
Soft tissue	35 (3)	[27; 44]
Air	130 (24)	[100; 193]
ME		
Body	9 (11)	[−17; 32]
Bone	106 (51)	[8; 208]
Soft tissue	−2 (5)	[−10; 7]
Air	−57 (41)	[−134; 28]

TABLE II. Mean (±1 SD) and range of voxel-wise dose differences (in percentage of prescribed dose).

	Mean error dose (%)	
Parameter	Mean (±1 SD)	Range
Body	−0.03 (0.05)	[−0.15; 0.06]
Dose > 10%	0.32 (0.25)	[−0.03; 0.85]
Dose > 50%	0.12 (0.28)	[−0.36; 0.75]
Dose > 90%	−0.07 (0.22)	[−0.52; 0.45]

64 ± 10 HU of head and neck synthetic CT by using a multi-atlas approach with bone suppression during deformable registration and a generalized registration error metric for weighting of the Hounsfield Unit.[23] Also using an atlas-based approach for (prostate and) head and neck images, Guerreiro et al obtained a MAE of 90.7 ± 12.1 HU.[22] MAEs are generally sensitive to differences in image resolution, field of view, and image resampling or other image processing, making these difficult to compare. Moreover, head and neck images are more challenging to properly align due to positioning and other interscan differences, hampering the evaluation of differences between CT and sCT by calculating MAEs in the head and neck region.

As is more often the case for synthetically generated CTs, the average HU values of bone were lower in the sCTs than in the CTs.[18] We hypothesize that this is mainly caused by registration issues, as can be observed in the difference image in Fig. 4, where the borders of bony structures show misalignment between the CT and the MRI (and thus sCT). This slight misalignment may result in blurring of the cortical bone and therefore a lowering of the attenuation value on the sCT. A second explanation could be that there were less training patches containing the shoulder and ribs and the estimation of the sCT in these regions suffered from border effects. Patch-based methods are prone to border effects, meaning that voxels in the center of a patch can be predicted with a higher certainty than voxels at the border of an image, who are never at the center of a patch. Lastly, streaking artifacts present on the CT but absent on the sCT may also have contributed slightly to a higher ME of the bones.

In terms of dose, this method showed minor dosimetric differences between the dose calculated on CT and the dose calculated on sCT. Mean difference in dose to the high dose region of the target volume (dose > 90%) was within 1% (close to 0.5%) for all patients. On average, this difference was close to zero, so no bias could be observed. For larger regions, containing lower doses, these differences were even smaller. These results are comparable to what was reported by Farjam et al., who compared several DVH parameters like D_{mean}, D_{max}, and $D_{95\%}$ of the PTVs. Largest differences were observed for the $D_{95\%}$ of the PTV50 (1.12% ± 2.1%) and D_{mean} to the (left) parotid glands (1.58% ± 2.09%). It should be noted, however, that comparing DVH parameters based on a structure could result in different numbers than performing a voxel-wise comparison of doses. DVH parameters are more sensitive to how a structure is defined on the two image sets and how the DVHs are calculated. Similarly, Guerreiro et al. reported a difference in PTV D_{mean} of −0.09% ± 0.33% [range: −0.60–0.23) $D_{98\%}$ of −0.67% ± 0.62% [range: 0.19–1.78].

We chose a CNN-based approach as an alternative to atlas-based sCT generation which is based on deformable registration(s) that may be time consuming and not robust. It should be noted that despite circumventing the need for deformable registration in the training and testing phase, deformable image registration was required to optimally prepare our training data. Nevertheless, after data preparation and training, creating a synthetic CT is fast and requires minimal preprocessing. In addition, by training our network on available data from the clinical scan protocol, it is not necessary to add a new sequence to the protocol. Not having to add scan time to the protocol is advantageous since MR scan time is generally limited and it is difficult for patients to endure long scan times while wearing the immobilization mask. The MR sequence that was used in this study is used for delineation of the OARs, which means that these contours will be present on the sCT as well, avoiding any rigid or deformable registration to propagate these contours. CT scan quality is often degraded by the presence of streaking artifacts from dental implants. In this study, patients with severe dental artifacts were excluded in the training. This eliminates streaking artifacts for patients of whom the CT would present with these artifacts. In addition, since TSE MRI is less prone to susceptibility artifacts at tissue–air interfaces, it is easier to obtain artifact-free MR images which will be reconstructed to an sCT in which no streaking artifacts are present. Lastly, with our 3D patch-based approach, we exploit the 3D nature of these images while being robust to anatomical outliers or nonstandard patients.

Training a CNN for sCT generation also comes with some limitations. The training is generally performed on a carefully selected set of images that are scanned within a fixed scan protocol, having exactly the same scan acquisition parameters. If a scan protocol is adapted over the course of time, the

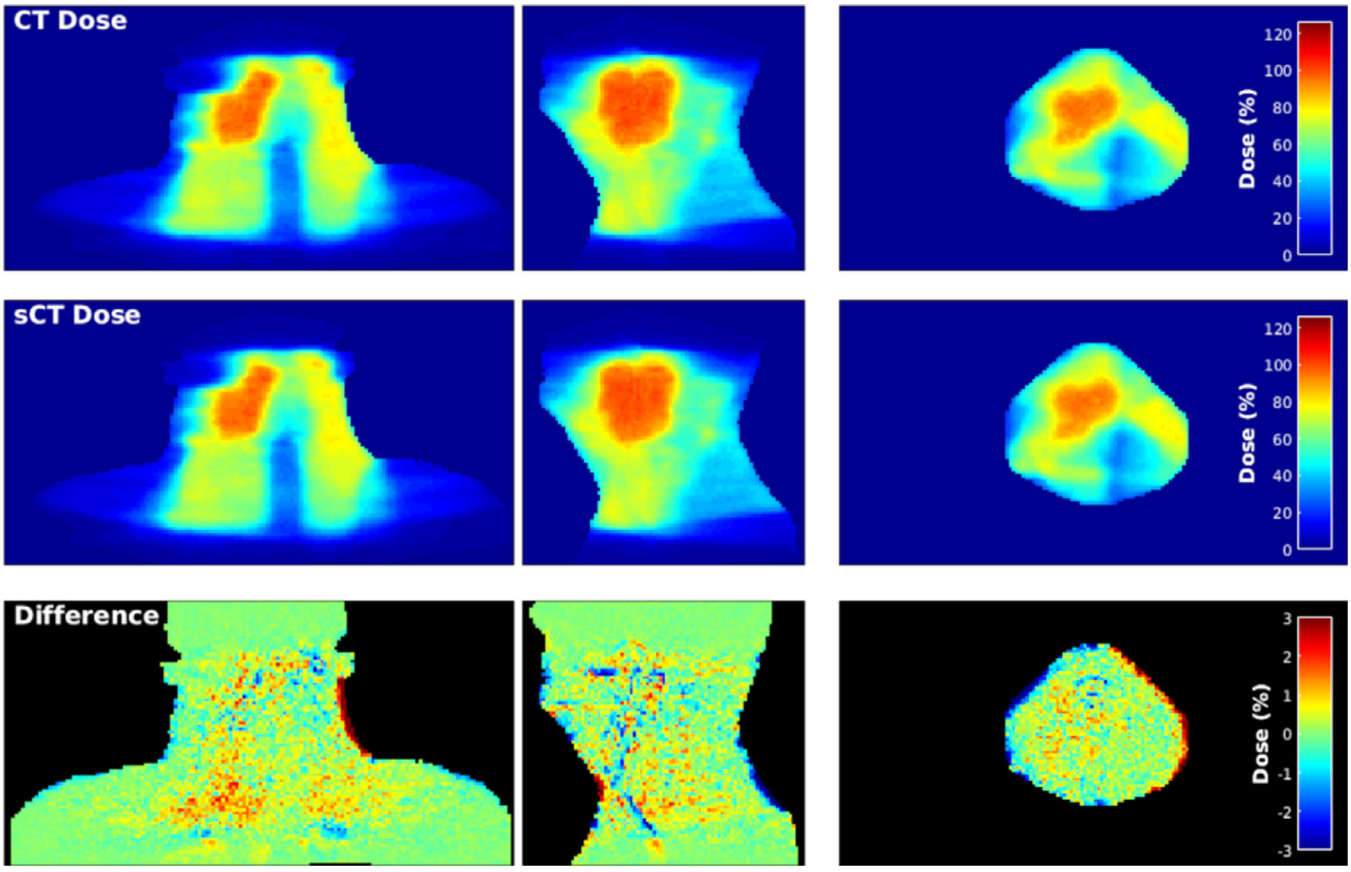


FIG. 6. Coronal, sagittal, and transversal slice through the dose calculated on the CT_{reg}, synthetic computed tomography (sCT), and difference of a 30 × 2.3 Gy VMAT plan for an oropharyngeal cancer patient with a tumor in the right tonsil (patient 1). [Color figure can be viewed at wileyonlinelibrary.com]

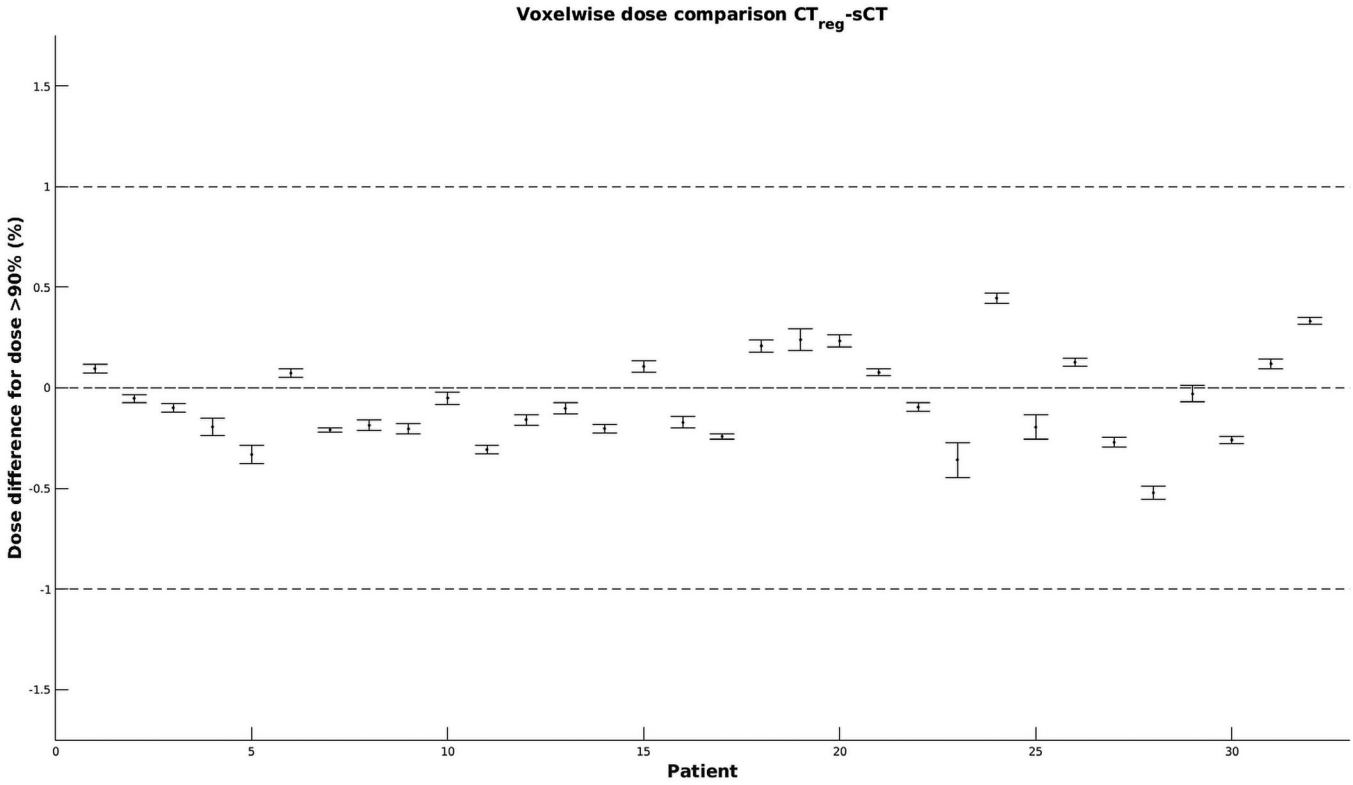


FIG. 7. Mean dose difference inside high-dose region (>90% of PD) for all patients. Error bars represent 95% confidence intervals.

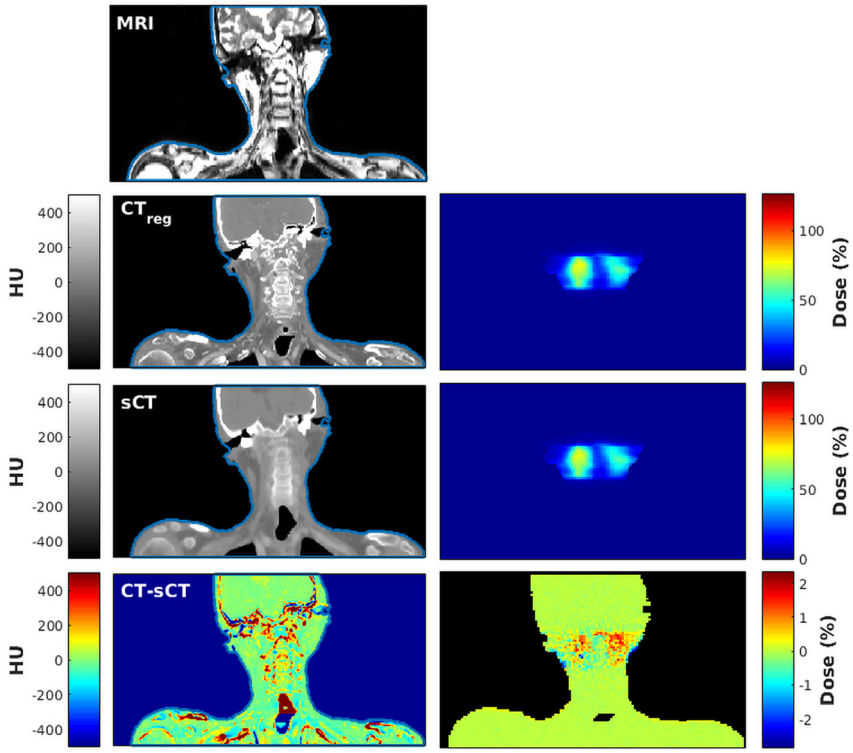


FIG. 8. Coronal slice of the MRI, CT_{reg}, sCT, and difference image (left) and the dose calculated on the CT_{reg}, synthetic computed tomography (sCT), and difference (right) for a postoperative case with nonstandard anatomy. This oropharyngeal cancer patient was treated with 35 × 2 Gy for a tumor in the base of tongue. [Color figure can be viewed at wileyonlinelibrary.com]

network may need to be retrained and evaluated. If retraining is required, a new set of training data (including CT) will need to be acquired first. The same issue holds for changing between systems with different field strengths. The generalizability of this method will need to be further investigated. In this study, we used the in-phase images of a Dixon reconstruction from T2 TSE MRI scans, as this already existing sequence in our protocol has a large scan volume and was used to delineate OARs. The disadvantage of this sequence is that Dixon reconstruction may not be available on all scanners and Dixon reconstruction may sometimes fail.

Besides Dixon reconstruction-related issues, other image quality issues may be a limiting factor in sCT generation. Image artifacts may hamper the quality of the sCT and may result in incorrect dose calculations. The head and the neck are a challenging body site for MRI, due to the large FOV, (breathing) motion, bone–air interfaces, and coil setup. The sCT solution proposed here is most likely not robust to all image artifacts, such as breathing motion. As insufficient image quality may be a limiting factor in the quality of the sCT, an optimized standardized imaging protocol has been used to minimize artifacts and maximize image quality.[9] Nevertheless, QA of the MRI and of the sCT is required and of vital importance to ensure an sCT solution that can safely be introduced clinically.

The use of MRI in radiotherapy is growing rapidly. Clinics are adopting MR imaging for delineation of OARs and target volumes to prepare (conventional) radiation treatment. By eliminating the CT in this preparation phase, we improve patient comfort and avoid image registration. For MR-only treatments, all delineations and treatment planning are based on MR imaging. Specifically for MR-guided treatment treatments on an MR-linac, it is of particular importance to obtain HU in a fast and reliable way, to limit treatment time, and to avoid the need for a CT-based treatment simulation and image registration. Especially if this scan is performed first, the 4 min that our method currently requires to synthesize the sCT could be used to verify the delineations, after which plan optimization and dose calculations can start.

5. CONCLUSIONS

Synthetic CTs were generated from T2-weighted 2D MRI of 34 head and neck cancer patients with deep learning using a 3D patch-based U-net model. This method is relatively fast and robust to anatomical variation in the head and neck region. Mean differences between the dose calculated on the reference CT and sCT were below 1% of the prescribed dose, showing that it is feasible to generate large field-of-view sCTs for dose calculation purposes from a generic MRI scan from the clinical protocol, without adding scan time.

ACKNOWLEDGMENTS

This research is funded by ZonMw IMDI Programme, project number: 1040030. The research is also partly sponsored by Philips Healthcare.

CONFLICT OF INTEREST

Peter R. Seevinck and Marijn van Stralen declare to be minority shareholder of MRIguidance B.V. Cornelis A.T. van den Berg declares to be a minority shareholder of MRCode B.V.

*Present address: Department of Radiotherapy, Amsterdam University Medical Center, Amsterdam, The Netherlands
[a)]Author to whom correspondence should be addressed. Electronic mail: a.dinkla@amsterdamumc.nl.

REFERENCES

1. Schmidt MA, Payne GS. Radiotherapy planning using MRI. *Phys Med Biol*. 2015;60:R323–R361.
2. Rasch C, Keus R, Pameijer FA, et al. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. *Int J Radiat Oncol*. 1997;39:841–848.
3. Jager EA, Ligtenberg H, Caldas-Magalhaes J, et al. Validated guidelines for tumor delineation on magnetic resonance imaging for laryngeal and hypopharyngeal cancer. *Acta Oncol*. 2016;55:1305–1312.
4. Dai YL, King AD. State of the art MRI in head and neck cancer. *Clin Radiol*. 2018;73:45–59.
5. Metcalfe P, Liney GP, Holloway L, et al. The potential for an enhanced role for MRI in radiation-therapy treatment planning. *Technol Cancer Res Treat*. 2013;12:429–446.
6. Khoo VS. MRI–"magic radiotherapy imaging" for treatment planning? *Br J Radiol*. 2000;73:229–233.
7. Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. *Radiat Oncol*. 2017;12:28.
8. Brunt JNH. Computed tomography–magnetic resonance image registration in radiotherapy treatment planning. *Clin Oncol*. 2010;22:688–697.
9. Verduijn GM, Bartels LW, Raaijmakers CPJ, Terhaard CHJ, Pameijer FA, van den Berg CAT. Magnetic resonance imaging protocol optimization for delineation of gross tumor volume in hypopharyngeal and laryngeal tumors. *Int J Radiat Oncol*. 2009;74:630–636.
10. Hanvey S, McJury M, Tho LM, et al. The influence of MRI scan position on patients with oropharyngeal cancer undergoing radical radiotherapy. *Radiat Oncol*. 2013;8:129.
11. Chuter R, Prestwich R, Bird D, et al. The use of deformable image registration to integrate diagnostic MRI into the radiotherapy planning pathway for head and neck cancer. *Radiother Oncol*. 2017;122:229–235.
12. Burgos N, Guerreiro F, McClelland J, et al. Iterative framework for the joint segmentation and CT synthesis of MR images: application to MRI-only radiotherapy treatment planning. *Phys Med Biol*. 2017;62:4237–4253.
13. Nyholm T, Jonsson J. Counterpoint: opportunities and challenges of a magnetic resonance imaging–only radiotherapy work flow. *Semin Radiat Oncol*. 2014;24:175–180.
14. Ulin K, Urie MM, Cherlow JM. Results of a multi-institutional benchmark test for cranial CT/MR image registration. *Int J Radiat Oncol*. 2010;77:1584–1589.
15. Fortunati V, Verhaart RF, Angeloni F, et al. Feasibility of multimodal deformable registration for head and neck tumor treatment planning. *Int J Radiat Oncol*. 2014;90:85–93.
16. Lagendijk JJW, Raaymakers BW, van Vulpen M. The magnetic resonance imaging–linac system. *Semin Radiat Oncol*. 2014;24:207–209.
17. Chin AL, Lin A, Anamalayil S, Teo B-KK. Feasibility and limitations of bulk density assignment in MRI for head and neck IMRT treatment planning. *J Appl Clin Med Phys*. 2014;15:100–111.
18. Dinkla AM, Wolterink JM, Maspero M, et al. MR-only brain radiation therapy: dosimetric evaluation of synthetic CTs generated by a dilated convolutional neural network. *Int J Radiat Oncol Biol Phys*. 2018;102:801–812.
19. Owrangi AM, Greer PB, Glide-Hurst CK. MRI-only treatment planning: benefits and challenges. *Phys Med Biol*. 2018;63:05TR01.
20. Johnstone E, Wyatt JJ, Henry AM, et al. A systematic review of synthetic CT generation methodologies for use in MRI-only radiotherapy. *Int. J Radiat Oncol*. 2017;100:199–217.
21. Hsu S, Cao Y, Huang K, Feng M, Balter JM. Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy. *Phys Med Biol*. 2013;58:8419–8435.
22. Guerreiro F, Burgos N, Dunlop A, et al. Evaluation of a multi-atlas CT synthesis approach for MRI-only radiotherapy treatment planning. *Phys Medica*. 2017;35:7–17.
23. Farjam R, Tyagi N, Veeraraghavan H, et al. Multiatlas approach with local registration goodness weighting for MRI-based electron density mapping of head and neck anatomy. *Med Phys*. 2017;44:3706–3717.
24. Burgos N, Cardoso MJ, Guerreiro F, et al.Robust CT synthesis for radiotherapy planning: application to the head and neck region. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015*. Lecture Notes in Computer Science. Vol. 9350. Cham: Springer; 2015:476–484.
25. Zheng W, Kim JP, Kadbi M, Movsas B, Chetty IJ, Glide-Hurst CK. Magnetic resonance–based automatic air segmentation for generation of synthetic computed tomography scans in the head region. *Int J Radiat Oncol*. 2015;93:497–506.
26. Hsu S, Cao Y, Lawrence TS, et al. Quantitative characterizations of ultrashort echo (UTE) images for supporting air–bone separation in the head. *Phys Med Biol*. 2015;60:2869–2880.
27. Wiesinger F, Bylund M, Yang J, et al. Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning. *Magn Reson Med*. 2018;80:1440–1451.
28. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys*. 2017;44:1408–1419.
29. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging–based attenuation correction for PET/MR imaging. *Radiology*. 2018;286:676–684.
30. Maspero M, Savenije MHF, Dinkla AM, 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.
31. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation.In:Navab N, Hornegger J, Wells W, Frangi A, eds. *MICCAI*. Cham: Springer; 2015:234–241.
32. Nie D, Cao X, Gao Y, Wang L, Shen D.Estimating CT Image from MRI Data Using 3D Fully Convolutional Networks. In: *Deep Learning and Data Labeling for Medical Applications : First International Workshop, LABELS 2016, and Second International Workshop, DLMIA 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 21, 2016, Proceedings. LA...* Vol. 2016; 2016:170–178. https://doi.org/10.1007/978-3-319-46976-8_18
33. Eggers H, Brendel B, Duijndam A, Herigault G. Dual-echo Dixon imaging with flexible choice of echo times. *Magn Reson Med*. 2011;65:96–107.
34. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks;2017.
35. Wolterink JM, Dinkla AM, Savenije MHF, Seevinck PR, van den Berg CAT, Išgum I. Deep MR to CT synthesis using unpaired data. Vol 10557. LNCS;2017. https://doi.org/10.1007/978-3-319-68127-6_2
36. Klein S, Staring M, Murphy K, Viergever MA, Pluim J. Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29:196–205.
37. Ulyanov D, Vedaldi A, Lempitsky V. Instance Normalization: The Missing Ingredient for Fast Stylization. July 2016. http://arxiv.org/abs/1607.08022. Accessed September 17, 2018.
38. Odena A, Dumoulin V, Olah C. Deconvolution and checkerboard artifacts. *Distill*. 2016;1:e3.
39. Dozat T.Incorporating Nesterov Momentum into Adam. February 2016. https://openreview.net/forum?xml:id=OM0jvwB8jIp57ZJjtNEZ. Accessed September 17, 2018.

Proven Tissue Equivalent **Phantom Solutions**

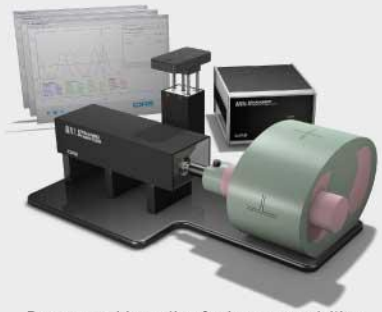
High fidelity simulation for your SRS program

CIRS is recognized as a leader in the manufacturer of tissue equivalent phantoms for medical imaging, radiation therapy and procedural training.

CIRS is one of a few companies worldwide with proprietary technology that permits the manufacture of phantoms for every commercial imaging modality.

Visit Us at RSNA, Booth 2700

Our most realistic maxillofacial phantom for dental cone beam CT & panoramic X-Ray

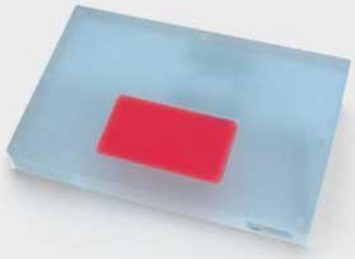


Programmable motion for image acquisition, treatment planning & dose delivery

Measure known kPa's with shear wave systems

MR, CT & Ultrasound compatible

Required for MQSA Program



Evaluate FFDM system performance

Most complete solution for performance and QA testing

CIRS

900 Asbury Ave. • Norfolk, Virginia 23513 • USA • Tel: 757-855-2765 • **WWW.CIRSINC.COM**