IPEM | Institute of Physics and Engineering in Medicine

PAPER • OPEN ACCESS

Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy

To cite this article: Matteo Maspero *et al* 2018 *Phys. Med. Biol.* **63** 185001

View the article online for updates and enhancements.

Recent citations

- Generation of abdominal synthetic CTs from 0.35T MR images using generative adversarial networks for MR-only liver radiotherapy
  Jie Fu *et al*

- Attenuation correction for human PET/MRI studies
  Ciprian Catana

- Comparison of the suitability of CBCT- and MR-based synthetic CTs for daily adaptive proton therapy in head and neck patients
  Adrian Thummerer *et al*
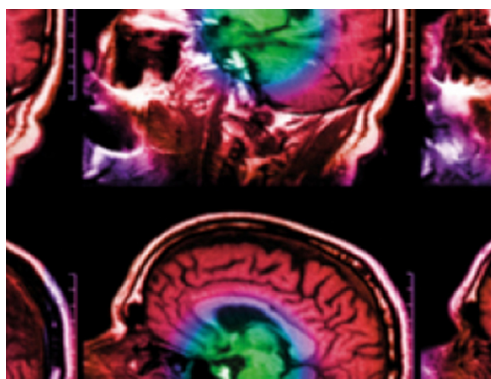
IPEM | IOP

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics, biomedical engineering and related subjects.

Start exploring the collection−download the first chapter of every title for free.

PAPER

Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy

Matteo Maspero[1,2,3,4], Mark H F Savenije[1,2,4], Anna M Dinkla[1,2], Peter R Seevinck[2,3], Martijn P W Intven[1], Ina M Jurgenliemk-Schulz[1], Linda G W Kerkmeijer[1] and Cornelis A T van den Berg[1,2]

[1] Department of Radiotherapy, University Medical Center Utrecht, Utrecht, Netherlands
[2] Center for Image Sciences, University Medical Center Utrecht, Utrecht, Netherlands
[3] Image Science Institute, University Medical Center Utrecht, Utrecht, Netherlands
[4] The authors equally contributed.

E-mail: m.maspero@umcutrecht.nl and matteo.maspero.it@gmail.com

Abstract

To enable magnetic resonance (MR)-only radiotherapy and facilitate modelling of radiation attenuation in humans, synthetic CT (sCT) images need to be generated. Considering the application of MR-guided radiotherapy and online adaptive replanning, sCT generation should occur within minutes. This work aims at assessing whether an existing deep learning network can rapidly generate sCT images for accurate MR-based dose calculations in the entire pelvis.

A study was conducted on data of 91 patients with prostate (59), rectal (18) and cervical (14) cancer who underwent external beam radiotherapy acquiring both CT and MRI for patients' simulation. Dixon reconstructed water, fat and in-phase images obtained from a conventional dual gradient-recalled echo sequence were used to generate sCT images. A conditional generative adversarial network (cGAN) was trained in a paired fashion on 2D transverse slices of 32 prostate cancer patients. The trained network was tested on the remaining patients to generate sCT images. For 30 patients in the test set, dose recalculations of the clinical plan were performed on sCT images. Dose distributions were evaluated comparing voxel-based dose differences, gamma and dose-volume histogram (DVH) analysis.

The sCT generation required 5.6 s and 21 s for a single patient volume on a GPU and CPU, respectively. On average, sCT images resulted in a higher dose to the target of maximum 0.3%. The average gamma pass rates using the 3%, 3 mm and 2%, 2 mm criteria were above 97 and 91%, respectively, for all volumes of interests considered. All DVH points calculated on sCT differed less than $\pm 2.5\%$ from the corresponding points on CT.

Results suggest that accurate MR-based dose calculation using sCT images generated with a cGAN trained on prostate cancer patients is feasible for the entire pelvis. The sCT generation was sufficiently fast for integration in an MR-guided radiotherapy workflow.

1. Introduction

'Magnetic resonance (MR)-only' radiotherapy refers to a radiotherapy workflow in which patient simulation and dose calculation are performed using only MR images. This workflow has been proposed to exploit the soft tissue contrast offered by magnetic resonance imaging (MRI) without recurring to inter-modality registration and thus reducing possible systematic errors in target definition (Nyholm *et al* 2009). Also, MR-only radiotherapy offers practical and logistical advantages reducing the overall treatment cost (Devic 2012), workload (Karlsson *et al* 2009) and patient exposure to ionising radiation (Schmidt and Payne 2015). In addition, MR-only is of interest considering the advent of MR-guided radiotherapy (MRgRT) systems

(Lagendijk *et al* 2014, Low 2017), where it may be exploited to perform online daily replanning based on the anatomy acquired before irradiation.

However, dose calculation cannot be performed directly on MR images since no correlation has been demonstrated between the nuclear magnetic properties, on which MRI depends, and the electron density, which is the property used to model radiation attenuation in humans (Brown *et al* 2014). Several groups have proposed methods to automatically generate the so-called 'synthetic' CT (sCT)[5] images to enable accurate MR-based dose calculation (Owrangi *et al* 2018).

When focusing on the pelvic area, many methods have been proposed and evaluated for radiotherapy of prostate cancer (Edmund and Nyholm 2017, Johnstone *et al* 2017), showing dose deviations below 2% with respect to CT-based dose calculations. Only three contributions investigated the accuracy of MR-based dose calculations for locations other than prostate within the pelvic area (Kemppainen *et al* 2017, Liu *et al* 2017, Wang *et al* 2018). Moreover, no attention has been specifically dedicated to the time required to generate sCT, which should be of the order of minutes to allow daily replanning during MRgRT as, for example, underlined by Raaymakers *et al* (2017).

Recently, deep learning-based sCT generation has been presented (Han 2017, Nie *et al* 2017, Wolterink *et al* 2017) enabling sCT generations of a full 3D volume in a minute. Han (2017) employed, for the first time, deep learning to generate sCT images using only a generative model, e.g. a U-net. Generative adversarial networks (GANs) have been used to synthesise medical images after training in a paired and an unpaired fashion by Nie *et al* (2017) and Wolterink *et al* (2017), respectively. They both obtained promising results, especially compared to networks in which only the generator was adopted (Han 2017). However, in none of these contributions, the sCT images were evaluated for dosimetric accuracy, which is the relevant metric for radiotherapy purposes. In this sense, we presented a dosimetric evaluation on deep learning-based sCT images for brain patient using a dilated convolutional network (Dinkla *et al* 2018). So far, no study has reported evaluation of dose calculation accuracy for deep-learning based sCT in the pelvis.

By interpreting the generation of sCT images as an image-to-image problem, we aim at assessing whether an existing deep learning network can generate sCT images that enable accurate MR-based dose calculation using a conventional MR sequence in the pelvic area. We used a conditional generative adversarial network (cGAN) motivated by the results obtained by Isola *et al* (2016) who showed how this approach can promptly solve numerous image-to-image translation problems (Litjens *et al* 2017). More specifically, GANs are networks constituted by a generator and a discriminator network. Generator and discriminator are jointly trained, particularly aiming at generating realistically looking images by exploiting the capability of the discriminator network to discern between real and fake images (Goodfellow *et al* 2014, Isola *et al* 2016).

In this work, we trained an existing GAN (Isola *et al* 2016) with paired MRI-CT data to learn the generation of sCT images using multi-contrast Dixon reconstructed MR images from conventional multi-echo gradient echo. Training was performed on prostate cancer patients' images only. Finally, sCT images were evaluated for MR-based dose calculations for patients with prostate, rectal and cervical cancer.

2. Materials and methods

2.1. Patient data collection

This study was conducted on a total of 91 patients: 59 prostate, 18 rectal and 14 cervical cancer patients who had no hip implants and underwent external beam radiotherapy. Patients simulation was performed using both on CT and MRI images acquired between March 2016 and April 2017.

Fifty-nine patients were diagnosed with low, intermediate, high-risk prostate carcinoma stage T1c–T3b. Their mean age was 69.6 ± 5.1 years ($\pm 1\sigma$; range 59.8–82.9). Three intra-prostatic cylindrical gold fiducial markers were inserted in these patients for position verification purposes. Prostate cancer patients underwent 5-beam 10 MV intensity-modulated radiotherapy (IMRT) with a prescribed dose of 35×2.2 Gy to prostate and macroscopic tumour and 35×2.0 Gy to seminal vesicles.

Eighteen patients, of whom $5/18 = 27.8\%$ female, were diagnosed with intermediate and high-risk rectal cancer staged T2–T4. These patients were treated for neoadjuvant therapy with three fractionating regimes: short course treatment delivering 5×5 Gy (2), and long-course treatment 25×2.0 Gy without (14) and with (2) an integrated boost on extramesorectal pathological nodes of 25×2.4 Gy. All patients in this group were irradiated with volumetric modulated arc therapy (VMAT) consisting of two coplanar arcs of 10 MV between $50°$ and $310°$.

Fourteen patients were diagnosed with low, intermediate, high-risk cervical cancer staged T1–T4. Their mean age was 51.4 ± 15.1 years (range 29.1–83.0 years). These patients underwent external beam radiotherapy with 10 MV VMAT with a $360°$ irradiation arc and the following dose schemes: 25×1.8 Gy (2), 25×2.2 Gy with

---

[5] In the literature, the sCT images are also called 'pseudo-CT' or 'substitute CT'.

an integrated boost in the pelvic pathological nodes (3) and 25×2.3 Gy with an integrated boost in the common iliac and para-aortic region (9).

For all patients with prostate and rectal cancer, 3T MRI (Ingenia MR-RT, v 5.7.1, Philips Healthcare, The Netherlands) was acquired within 2.5 h the CT (Brilliance Big Bore, Philips Healthcare, Ohio, USA). For the cervical cancer patients, time between imaging protocols was up to one week. All patients were asked to drink between 200 and 300 ml of water one hour before the scans and after emptying the bladder (and rectum in the case of prostate cancer patients). Patients were positioned using a flat table and knee wedges. CT scans were performed with the following imaging parameters: 120 kV, 923 ms exposure time, 121–183 mA tube current, 512×512 in-plane matrix, and 3 mm slice thickness. In-plane resolution was variable depending on the field of view (FOV) used, with a typical pixel size of 1×1 mm^2 and maximum size of 1.2×1.2 mm^2. In the inferior-superior direction, the size of the FOV was variable ranging 33–77 cm.

To simulate treatment positioning, patients were marked with at least three skin tattoos at the CT scanner, which were then used to reposition the patient at the MR scanner with the aid of a laser system (Dorado3, LAP GmbH Laser Applikationen, Germany). MRI was acquired using anterior and posterior phased array coils (dS Torso and Posterior coils, 28 channels, Philips Healthcare, The Netherlands). To avoid compression of the patients, two in-house-built bridges supported the anterior coil.

For the generation of MR-based sCT images, a dual echo three-dimensional (3D) cartesian radio-frequency spoiled gradient-recalled echo sequence was acquired with the following parameters: 1.2/2.5 ms echo times, 3.9 ms repetition time, 10° flip angle, $552 \times 552 \times 300$ mm^3 FOV[6], anterior-posterior as the readout direction (frequency encoding), $284 \times 281 \times 120$ acquisition matrix, $1.05 \times 1.05 \times 2.5$ mm^3 reconstructed voxel, 1083 Hz/px bandwidth and 2 min 13 s acquisition time. A Dixon reconstruction (Dixon 1984, Eggers *et al* 2011) was performed obtaining in-phase, fat, and water images. This sequence was originally acquired to generate sCT for sole prostate patients with a proprietary method called MR for calculating attenuation (MRCAT, Philips Healthcare, The Netherlands) as presented in Tyagi *et al* (2016) and Maspero *et al* (2017a). In this work, MRCAT was used to automatically identify air regions based on the in-phase, fat and water images to avoid laborious manual segmentation during preparation of the training data. Identification of air regions was performed as specified in the following section to ensure consistency of air locations between CT and MR images during the training of the network.

Delineations of the target volumes and organs at risk (OARs) were performed by radiation oncologists.
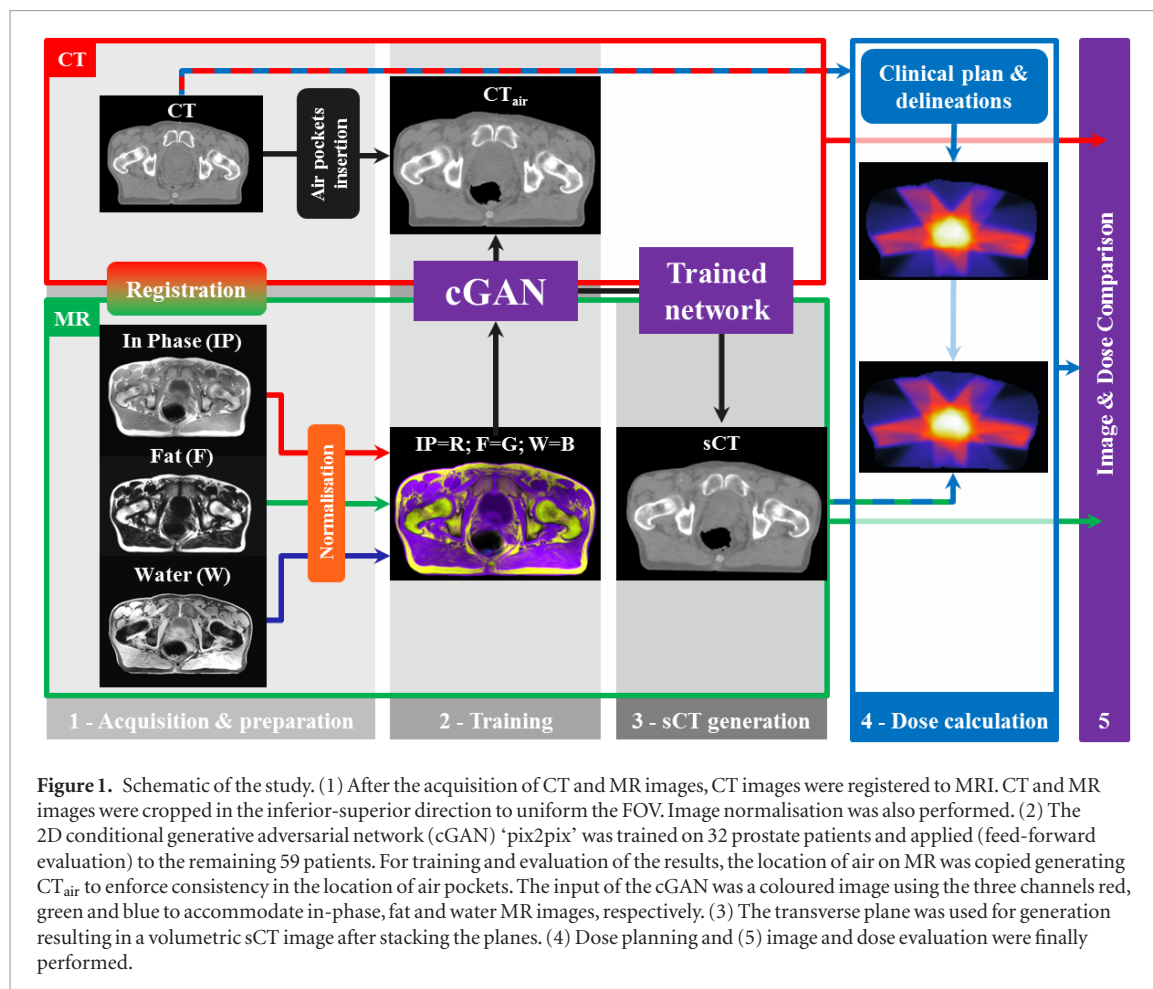
2.2. The network

A cGAN called 'pix2pix' consisting of a 256×256 U-net generator network and a 70×70 PatchGAN discriminator architecture was employed as provided in the PyTorch implementation by Isola *et al* (2016). As a proof-of-concept, considering that our main goal is the dosimetric evaluation of sCT images generated with a generative adversarial network, we kept the network implementation as similar as possible to what was originally presented by Isola and co-workers. Optimisation was performed as in Goodfellow *et al* (2014) alternating between one gradient descendent step on the discriminator network and one step on the generator network. A structured loss function cGAN$+\lambda\cdot$L1 with $\lambda = 100$ was adopted. As already investigated by Mathieu *et al* (2015) and Isola *et al* (2016), the use of a loss function constituted by L1 alone leads to reasonable but blurred results; on the other hand, cGAN alone will lead to sharp results but introducing artefacts in the images. Isola *et al* (2016) showed that training in an adversarial setting together with an L1 norm generates sharp images with a low amount of artefacts. The weights of the network were randomly initialised from a Gaussian distribution with mean 0 and standard deviation 0.02. The implementation of pix2pix can be applied to 8-bit grey-scale (1 channel) or coloured (3 channels) two-dimensional (2D) images. All the patient data underwent pre-processing to normalise the MR images and prepare a paired experiment by registering input (MRI) and target data (CT). In this work, we hypothesised that maximising the number of input images enriches contextual information per subject and facilitates learning of the image-to-image relation between MRI and CT. Therefore, we used all the available multi-contrast MR images to generate sCT images: in-phase, fat and water images were coded as colours of the images (red, green and blue) and input of the network. A schematic representation of the study is presented in figure 1.

2.3. sCT generation

2.3.1. Image pre-processing

First, all the CT images were rigidly registered and resampled to MR with Elastix v4.7 (Klein *et al* 2010). To keep the FOV consistent between CT and MRI, images were cropped in the superior-inferior direction to the smallest FOV ensuring that training may be conducted in a paired fashion. In the following, we use the term CT$_{reg}$, IP, F and W to refer the cropped images of registered CT, in-phase, water and fat images, respectively.

---

[6] Expressed in terms of anterior-posterior, right-left and superior-inferior directions.

Figure 1. Schematic of the study. (1) After the acquisition of CT and MR images, CT images were registered to MRI. CT and MR images were cropped in the inferior-superior direction to uniform the FOV. Image normalisation was also performed. (2) The 2D conditional generative adversarial network (cGAN) 'pix2pix' was trained on 32 prostate patients and applied (feed-forward evaluation) to the remaining 59 patients. For training and evaluation of the results, the location of air on MR was copied generating CT_{air} to enforce consistency in the location of air pockets. The input of the cGAN was a coloured image using the three channels red, green and blue to accommodate in-phase, fat and water MR images, respectively. (3) The transverse plane was used for generation resulting in a volumetric sCT image after stacking the planes. (4) Dose planning and (5) image and dose evaluation were finally performed.

Before feeding the images to the cGAN, the voxel intensity of CT_{reg} was clipped within the interval $[-1000;1047]$ HU to avoid a too large discretisation step after conversion to 8-bits. Also, MR images were normalised to their 95% intensity interval over the whole patient. Finally, all the images were converted to 8-bits to conform the pix2pix implementation.

Before training, we tried to remove the impact of mismatch of air pockets location, e.g. in the rectum and bowel loops and enforce that air pockets were consistently located between CT_{reg} and MRI. For this purpose, air cavities were filled in CT_{reg} and bulk-assigned (-1000 HU) as located in MR images using an automatic method previously described by Maspero *et al* (2017b). Without this pre-processing step, the generated sCT images may be characterised by inconsistent depiction of air between MR and sCT images, as reported in the supplementary material (stacks.iop.org/PMB/63/185001/mmedia). Manual inspection was performed to verify correct assignment of the air on CT. In the following, we use the term CT_{air} to refer to these datasets. The generation of CT_{air} has been introduced after noticing that the location of air when training directly on CT_{reg} was not consistent with MR images.

### 2.3.2. Training of the network

Training of the cGAN in a paired fashion was performed in the transverse plane randomly selecting 32 prostate cancer patients (training set). The network was trained for 200 epochs on a Tesla P100 (NVIDIA, California, USA) graphical processing unit (GPU) with batch size of one. Data augmentation was applied during training by flipping the images left and right and randomly cropping input and corresponding output images.

To verify the impact of enforcing the consistency of the air location between CT and MRI during training, we repeated training using CT_{reg} as a target of the network.

2.3.3. Image generation

The sCT generation was performed by applying the trained generator model and stacking all generated 2D transverse planes for each patient not used during training. This patient group is considered as the 'test' set. The 3D volumes were further post-processed using DCMTK (http://dicom.offis.de/dcmtk.php.en) to create files that conform to the DICOM standard, usable in a treatment planning system.
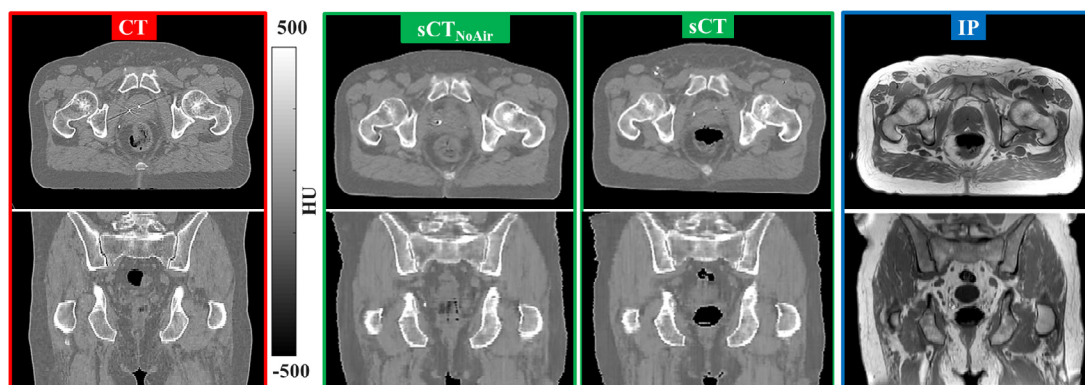
Figure 2. CT (left), sCT$_{NoAir}$ (middle, left) sCT (middle, right) and IP (right) for an exemplary patient treated for prostate cancer (top and central rows). The images on the top and bottom rows present a corresponding transverse plane, while in the bottom a corresponding coronal plane.

Also, sCT generation for the prostate cancer patients in the test set (27 patients) was performed applying the generator trained using CT$_{reg}$ to evaluate the impact of enforcing the consistency of the air location during training. In the following, we use the term sCT$_{NoAir}$ to refer to this dataset.

2.4. Evaluation

The performance of the network was evaluated reporting the time needed to train the cGAN and to infer the generator on a GPU and a central processing unit (CPU) framework (quad-core Intel Xeon 3.4 GHz).

2.4.1. Image comparison

Image evaluation was performed in the test set by calculating the mean absolute error (MAE) and mean error (ME) on sCT with respect to CT$_{reg}$ (CT$_{reg}$ minus sCT) in the intersection of the body contours. The body contours were automatically calculated by thresholding CT$_{reg}$ and sCT images at -500 HU.

In addition, MAE and ME were also calculated for the prostate patients in the test set (27) between sCT$_{NoAir}$ and CT$_{reg}$.

2.4.2. Dose comparison

Thirty patients (ten for each tumour type) were randomly selected from the test set to undergo dose comparison. For these patients, clinical plans were recalculated (QA modality) on sCT images in Monaco (v 5.11.02, Elekta AB, Sweden) using the Montecarlo photon algorithm on a grid of $3 \times 3 \times 3$ mm^3 with 3% statistical uncertainty for VMAT plans (rectum and cervix) and 1% for IMRT plans (prostate).

Before planning, the sCT images were rigidly registered, resampled and linearly interpolated to the planning CT using the inverse transformation of the registration that was previously found between planning CT and MR images. The delineations used on the planning CT were also adopted for sCT except for the body contour, which was recalculated.

In almost all cervical cancer patients, except the two without integrated boost, the FOV acquired during MRI was insufficient to include all OARs and calculate the plan on the patient. This was expected considering that MRI in the clinical settings is used for delineation of the sole primary tumour, while the plans may also include a boost to nodes in the common iliac and para-aortic regions. We, therefore, performed dose recalculation for the two patients that had comparable FOV in the inferior-superior direction, and for the remaining eight patients a new plan on the sCT was calculated prescribing 25×1.8 Gy to the PTV, excluding pathological nodes. The obtained plan was then recalculated on CT images.

Dose distributions were analysed through dose differences ($\frac{CT - sCT}{D_{Presc}}$) and gamma analysis at 3%, 3 mm and 2%, 2 mm (Low 2010) within dose threshold regions of 90%, 50%, and 10% prescription dose and in the body contour intersection after a 15 mm cropping to exclude dose build-up. Analysis of dose-volume histogram (DVH) points was performed to verify target (CTV, PTV) dose coverage and adherence to OARs constraints considering the differences between dose points (D_{98}, D_{50}, D_2, V_{95}, V_{75}) on the CT and sCT plans for CTV, PTV and rectum for prostate patients or bladder for rectal and cervical cancer patients. Note that DVH on sCT images were calculated on structures that were propagated after rigid registration; therefore, the DVH analysis does not take into account inter-scan differences of the structures.

Figure 3. Doses calculated on CT$_{reg}$ (left) and sCT (middle) in the transverse plane corresponding to the isocentre for one prostate cancer patient. (Right) The dose difference (CT$_{reg}$-sCT) is presented as the percentage of the prescribed dose (77 Gy) for the corresponding plane.

Table 1. Statistics of the image comparison between patients in the test set with prostate, rectal and cervical cancer in terms of mean absolute error (MAE) and mean error (ME) between CT$_{reg}$ or CT$_{air}$ and sCT (and sCT$_{NoAir}$) over the intersection of the body contours. The values are reported in terms of average ($\pm 1\sigma$) and range [min; max] and expressed in Hounsfield Units [HU].

Tumour location	Number of patients	Data	MAE [HU]	ME [HU]
Prostate	27	CT$_{reg}$−sCT	65 ± 10 [50;97]	1 ± 6 [−12;15]
		CT$_{air}$−sCT	60 ± 6 [48;71]	-3 ± 5 [−18;8]
		CT$_{reg}$−sCT$_{NoAir}$	65 ± 9 [51;95]	-2 ± 10 [−21;13]
Rectum	18	CT$_{reg}$−sCT	56 ± 5 [48;67]	2 ± 9 [−16;23]
Cervix	14	CT$_{reg}$−sCT	59 ± 6 [50;69]	4 ± 10 [−16;22]

3. Results

3.1. Performance of the network

In total, 3495 transverse planes were used for training, which required about 11 h on the GPU (200 epochs). After cropping the FOV on CT and MR images, each patient was trained using a volume of about 109 transverse planes on average over the training set. After training, inferring the generator network to obtain sCT images was performed in 5.6 s per patient on GPU and four times longer (21 s) on CPU. Figure 2 presents the transverse (top) and coronal (center) planes of CT$_{reg}$ (left), sCT (middle) and IP (right) for an exemplary prostate patient. It can be noticed that the different rectal filling between CT$_{reg}$ and IP is consistent to filling in the sCT image. In the supplementary material, corresponding figures for exemplary rectal and cervical cancer patients are reported.

3.2. Image comparison

Table 1 reports the statistics of the image comparison in terms of MAE and ME in the patient test set between CT$_{reg}$/CT$_{air}$ and sCT and sCT$_{NoAir}$. Over the entire test set (59), the MAE and ME were, on average, 61 ± 9 and 2 ± 8 HU. It can be noticed that the MAE and ME are comparable among patients with different tumour location and the MAE decreases when comparing sCT to CT$_{air}$. This demonstrates that CT$_{air}$ is more similar to sCT than CT$_{reg}$, which justifies its use during the training of the network in a paired fashion.

When considering MEA and ME of the network trained with and withour enforcing air consistency, we observe that the metrics are comparable in the two scenarios, but the ME slightly decreases without enforced air location consistency. This result may be explained by the fact that the size of air pockets is much smaller than the entire body contour. In this sense, voxelwise differences to the air location are not expected to greatly impact the reported MAE and ME. Nevertheless, if we consider figure 2, we can observe that air is filled with soft tissue in case training was performed without enforcing air consistency.

3.3. Dose comparison

An example of dose calculated on CT$_{reg}$ and sCT along with their difference is presented in figure 3 for the same prostate patient shown in figure 2. On average (table 2), it was observed that sCT images result in a higher dose to the target of about 0.1–0.3%. In the worst case, the mean dose difference was 1.6%. The average gamma pass rates using the 3%, 3 mm and 2%, 2 mm criteria were >97 and 91%, respectively, for all volumes of interests considered.

As part of the supplementary material, the dose difference of each individual patient in the high dose region (D> 90%) is presented for all thirty patients included in the dose comparison.

Table 2. Statistics of the 10/27 prostate, 10/18 rectal and 10/14 cervical cancer patients among the test set. Mean dose difference relative to the prescribed dose and gamma pass rate among the average dose difference calculated on a threshold of 10%, 50%, and 90% of the prescribed dose and the intersection of the body contour between CT and sCT images (Body). The values are reported in terms of average ($\pm 1\sigma$) and range [min; max].

Tumour location	Volume of interest	Dose difference $\frac{CT-sCT}{D_{Prescr}}$ [%]	Pass rate $\gamma_{3\%,3\,mm}$ [%]	Pass rate $\gamma_{2\%,2\,mm}$ [%]
Prostate (10 patients)	D> 10%	-0.1 ± 0.1 [$-0.3;0.2$]	98.1 ± 1.2 [96.3;99.4]	95.0 ± 2.3 [91.4;97.8]
	D> 50%	-0.1 ± 0.2 [$-0.4;0.5$]	99.4 ± 0.6 [98.1;100]	97.4 ± 1.6 [93.8;99.7]
	D> 90%	-0.3 ± 0.4 [$-1.1;0.4$]	99.7 ± 0.2 [99.3;100]	97.6 ± 2.3 [91.8;99.9]
	Body	0.0 ± 0.1 [$-0.2;0.1$]	98.8 ± 0.7 [97.5;99.6]	96.8 ± 2.1 [93.2;98.9]
Rectum (10 patients)	D> 10%	-0.2 ± 0.2	97.1 ± 1.1	91.6 ± 3.3
	D> 50%	-0.3 ± 0.3 [$-0.8;0.0$]	98.5 ± 1.1 [96.0;99.7]	93.2 ± 3.6 [86.5;97.7]
	D> 90%	-0.3 ± 0.5 [$-1.0;0.6$]	98.5 ± 2.1 [93.2;99.9]	92.0 ± 6.6 [77.9;98.2]
	Body	-0.2 ± 0.1 [$-0.4;-0.1$]	97.6 ± 1.2 [95.7;98.8]	94.0 ± 3.0 [88.7;96.8]
Cervix (10 patients)	D> 10%	-0.1 ± 0.3 [$-0.6;0.2$]	97.1 ± 1.7 [93.8;98.7]	92.9 ± 3.7 [84.3;96.1]
	D> 50%	-0.2 ± 0.5 [$-1.5;0.4$]	99.6 ± 1.9 [94.0;100]	94.5 ± 4.6 [83.1;98.6]
	D> 90%	-0.1 ± 0.7 [$-1.6;1.0$]	98.5 ± 3.1 [89.9;99.9]	90.6 ± 6.8 [72.3;96.3]
	Body	-0.1 ± 0.3 [$-0.9;0.2$]	97.7 ± 1.7 [94.3;99.4]	93.6 ± 4.0 [86.6;98.1]

Figure 4 presents the boxplot of the DVH point difference (CT–sCT) for targets (PTV, CTV) and OARs showing that all the DVH points on sCT-derived plans were within $\pm 2.5\%$ with respect to the corresponding points on CT-derived plans.

4. Discussion

Here, we demonstrated that deep learning enabled fast generation of sCT images facilitating accurate MR-based dose calculation for irradiation of patients with cancer in the pelvic area. In particular, we showed that training a cGAN on prostate cancer patients results in MR-based dose calculation within an average dose difference of 0.5% compared to CT-based calculations (CT–sCT). Although the network was trained on prostate cancer patients, sCT images generation in rectal and cervical cancer patients resulted in accurate dose calculations (see the supplementary material). This is of particular interest, considering that the network seems to accurately solve image-to-image translation problems also for female patients that were not included in the training set. However, for the cervical patients, the FOV coverage is insufficient to recalculate clinical plans. This means that for this patient group the MR sequence should be revised (i.e. extension of the FOV) and newly evaluated before clinical use. Also, by enforcing consistency of air location during training, we demonstrated that the network was able to depict internal air in location consistent to MRI. Please note that a 'correct' depiction of internal air is not crucial in the scope of standard radiotherapy, given its limited size and the fact that it may be differently located at each fraction. However, it may become of larger interests when considering MRgRT, given the additional dose at tissue interfaces (Raaijmakers *et al* 2008) and considering the scenario of hypofractionated treatments (Benjamin *et al* 2017).

The dosimetric evaluation performed in this work has been restricted to the pelvic area. The sCT generation method here adopted, however, could be applied also to different anatomical locations after proper training of the network. This makes the approach generic.

Within this work, for the first time, sCT generated with a deep learning technique underwent dosimetric evaluation in the pelvic region. In general, the dose differences obtained in this study are in line with previously published studies on prostate cancer patients, which reported dose differences within 1% (Korhonen *et al* 2014, Dowling *et al* 2015, Kim *et al* 2015, Siversson *et al* 2015, Prior *et al* 2016, Tyagi *et al* 2016, Maspero *et al* 2017a,
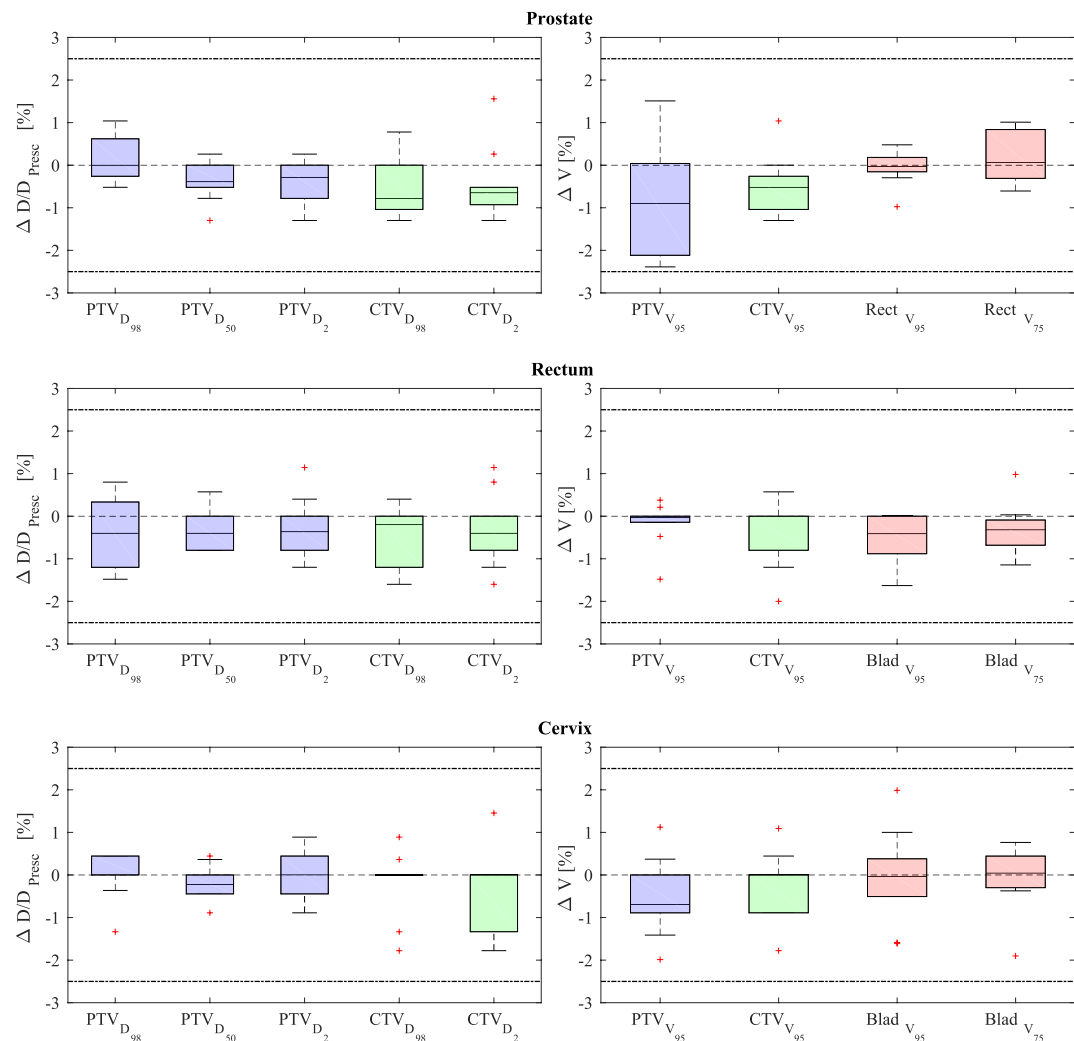
Figure 4. Boxplots of targets (CTV, PTV) and OARs DVH parameter differences between dose on CT and sCT (CT–sCT) for the prostate (top), rectal (middle) and cervical (bottom) cancer patients. The values refer to the whole course of fractionated treatment and are rescaled to the prescribed dose (left) or the total volume of the specific structure (right).

Persson *et al* 2017), and to other studies in the pelvic area (Kemppainen *et al* 2017, Liu *et al* 2017, Wang *et al* 2018). Dose deviations should be interpreted in the context of the clinically acceptable uncertainty in radiation therapy. When considering the complete radiotherapy pathway, including uncertainties in beam calibration, relative dosimetry, dose calculations, and dose delivery, the International Commission on Radiation Protection estimated an uncertainty of 5% in a clinical set-up (ICRP 2000, Thwaites 2013). The dosimetric deviation of an MR-based dose calculation (assuming CT to be the ground truth) of less than 0.5% only makes up for a small fraction of the total uncertainty (Persson *et al* 2017).

In this study, a conventional multi-gradient echo sequence was employed as already used by Tyagi *et al* (2016), Maspero *et al* (2017a) and Kemppainen *et al* (2017). This is the first time that multi-contrast MR images acquired with a single sequence have been used as input of a deep learning network for sCT generation. Previous work showed that multi-contrast images from different sequences, e.g. Dixon and ultra-short echo time could be used for generating images for attenuation correction for MR-PET (Leynes *et al* 2018). It is still unclear whether the use of multi-contrast images effectively increase the quality sCT generation or facilitate the training of the network. Future investigations may clarify this aspect, however, based on our findings, we believe that the use of deep learning makes specialised MR sequences, e.g. ultra-short echo time for direct bone visualisation, obsolete. This may lower the requirements for MR sequences used for sCT generation as well as the acquisition time. Of course, a high geometric fidelity, e.g. by means of high bandwidth, of the sequence is still required as adopted in this work. In particular, for the sCT generation method here proposed, evaluation of geometric accuracy, especially in the case these sCT images may be used for position verification purposes is still required.

In general, lowering requirements for the quality of MR sequence used for sCT generation may be of particular interest for MRgRT. For example, investigating the use of accelerated MR imaging techniques (Brix *et al* 2014, Feng *et al* 2014, Hollingsworth 2015) to speed up the acquisition time may facilitate an MR-based dose

calculation also for locations affected by high tissue mobility, which is a currently unmet need of MR-only radiotherapy (Owrangi *et al* 2018).

In this study, we enforced consistency of patient anatomy during training by performing rigid registrations and assigning air locations from MRI to CT. As an alternative approach, non-rigid registrations could have been adopted. However, we decided to avoid this approach since it would have introduced additional geometrical uncertainties, due to registration errors (Thor *et al* 2011, 2013), that we preferred not to introduce in this study. Also, non-rigid registration could have masked possible image deformation that are inherent in MR images (Fransson *et al* 2001, Wang *et al* 2004, Walker *et al* 2014). Future studies are advocated to clarify this aspect. Also, it is of interest to investigate the use of GANs in an unpaired fashion (Zhu *et al* 2017) since it may eliminate the need of a perfectly aligned dataset in the training phase as shown by Wolterink *et al* (2017) for brain cancer patients. Training in an unpaired fashion may be of particular interest since it may avoid the need of minimising inter-scan differences by copying air pockets from sCT to CT. Moreover, a 2.5D or 3D network could also be investigated to solve the discontinuities observed in the inferior-superior direction after stacking the transverse planes.

The sCT generation was performed in less than 6 s for a single patient volume. The computation time can be affected by many factors, e.g. the type of GPU adopted, the matrix and FOV size. When compared to existing methods, the sCT images presented in our work are generated faster even when compared to other deep-learning based methods (Han 2017, Wolterink *et al* 2017, Dinkla *et al* 2018). This can facilitate daily sCT generation for application where time constraints are crucial, e.g. in MRgRT (Lagendijk *et al* 2014). A limitation of the current study is that MR-based dose calculations were assessed in the absence of magnetic field. For MRgRT, dose calculations require particular attention due to the presence of a magnetic field affecting the dose distribution, especially near air cavities (Raaijmakers *et al* 2008). Given the promising results, a future study will investigate whether MR-based dose calculation on sCT obtained from the same network can be considered accurate also in the presence of magnetic fields. Also, before clinical usage in a complete MR-only workflow, the sCT generation method still needs to be thoroughly tested for accuracy in position verification. Moreover, a safe clinical implementation may also require designing quality assurance methods to validate the sCT images in the absence of the 'gold standard' offered by CT.

5. Conclusion

To conclude, this study shows, for the first time, that sCT images generated with a deep learning approach employing a cGAN and multi-contrast MR images acquired with a single acquisition facilitated accurate dose calculations in prostate cancer patients. It was further shown that without retraining the network, the cGAN could generate sCT images in the pelvic region for accurate dose calculations for rectal and cervical cancer patients. A particularly attractive feature of our method is its speed as it allows sCT generation within 6 s on a GPU and within 21 s on a CPU. This could be of particular benefit for MRgRT applications.

Acknowledgments

ORCID iDs

Matteo Maspero ⬤ https://orcid.org/0000-0003-0347-3375

References

Benjamin L C, Tree A C and Dearnaley D P 2017 The role of hypofractionated radiotherapy in prostate cancer *Curr. Oncol. Rep.* **19** 30

Brix L, Sørensen T S, Berber Y, Ries M, Stausbøl-Grøn B and Ringgaard S 2014 Feasibility of interactive magnetic resonance imaging of moving anatomy for clinical practice *Clin. Physiol. Funct. Imaging* **34** 32–8

Brown R W, Cheng Y C N, Haacke E M, Thompson M R and Venkatesan R 2014 *Magnetic Resonance Imaging: Physical Properties and Sequence Design* (New York: Wiley)

Devic S 2012 MRI simulation for radiotherapy treatment planning *Med. Phys.* **39** 6701

Dinkla A M, Wolterink J, Maspero M, Savenije M, Verhoeff J, Seravalli E, Isgum I, Seevinck P and van den Berg C 2018 MR-only brain radiotherapy: dosimetric evaluation of synthetic CTs generated by a dilated convolutional neural network *Int. J. Radiat. Oncol. Biol. Phys.* accepted

Dixon W T 1984 Simple proton spectroscopic imaging *Radiology* **153** 189–94

Dowling J A *et al* 2015 Automatic substitute computed tomography generation and contouring for magnetic resonance imaging (MRI)-alone external beam radiation therapy from standard MRI sequences *Int. J. Radiat. Oncol. Biol. Phys.* **93** 1144–53

Edmund J M and Nyholm T 2017 A review of substitute CT generation for MRI-only radiation therapy *Radiat. Oncol.* **12** 28

Eggers H, Brendel B, Duijndam A and Herigault G 2011 Dual-echo Dixon imaging with flexible choice of echo times *Magn. Reson. Med.* **65** 96–107

Feng L, Grimm R, Block K T, Chandarana H, Kim S, Xu J, Axel L, Sodickson D K and Otazo R 2014 Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI *Magn. Res. Med.* **72** 707–17

Fransson A, Andreo P and Pötter R 2001 Aspects of MR image distortions in radiotherapy treatment planning *Strahlentherapie Onkologie* **177** 59–73

Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial networks *Advances in Neural Information Processing Systems* (Cambridge, MA: MIT Press) pp 2672–80

Han X 2017 MR-based synthetic CT generation using a deep convolutional neural network method *Med. Phys.* **44** 1408–19

Hollingsworth K G 2015 Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction *Phys. Med. Biol.* **60** R297–322

ICRP 2000 *International Commission on Radiation Protection: Prevention of Accidents to Patients Undergoing Radiation Therapy* vol 86 (Amsterdam: Elsevier)

Isola P, Zhu J Y, Zhou T and Efros A A 2016 Image-to-image translation with conditional adversarial networks *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (July 2017)* (http://doi.org/10.1109/cvpr.2017.632)

Johnstone E, Wyatt J J, Henry A M, Short S C, Sebag-Montefiore D, Murray L, Kelly C G, McCallum H M and Speight R 2017 A systematic review of synthetic CT generation methodologies for use in MRI-only radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **100** 199–217

Karlsson M, Karlsson M G, Nyholm T, Amies C and Zackrisson B 2009 Dedicated magnetic resonance imaging in the radiotherapy clinic *Int. J. Radiat. Oncol. Biol. Phys.* **74** 644–51

Kemppainen R, Suilamo S, Tuokkola T, Lindholm P, Deppe M H and Keyriläinen J 2017 Magnetic resonance-only simulation and dose calculation in external beam radiation therapy: a feasibility study for pelvic cancers *Acta Oncol.* **56** 792–8

Kim J, Glide-Hurst C, Doemer A, Wen N, Movsas B and Chetty I J 2015 Implementation of a novel algorithm for generating synthetic CT images from magnetic resonance imaging data sets for prostate cancer radiation therapy *Int. J. Radiat. Oncol. Biol. Phys.* **91** 39–47

Klein S, Staring M, Murphy K, Viergever M A and Pluim J P 2010 Elastix: a toolbox for intensity-based medical image registration *IEEE Trans. Med. Imaging* **29** 196–205

Korhonen J, Kapanen M, Keyriläinen J, Seppälä T and Tenhunen M 2014 A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer *Med. Phys.* **41** 011704

Lagendijk J J W, Raaymakers B W, Van den Berg C A T, Moerland M A, Philippens M E and van Vulpen M 2014 MR guidance in radiotherapy *Phys. Med. Biol.* **59** R349–69

Leynes A P, Yang J, Wiesinger F, Kaushik S S, Shanbhag D D, Seo Y, Hope T A and Larson P E 2018 Zero-echo-time and dixon deep pseudo-ct (zedd ct): direct generation of pseudo-ct images for pelvic pet/mri attenuation correction using deep convolutional neural networks with multiparametric mri *J. Nucl. Med.* **59** 852–8

Litjens G, Kooi T, Bejnordi B E, Setio A A A, Ciompi F, Ghafoorian M, van der Laak J A W M, van Ginneken B and Sánchez C I 2017 A survey on deep learning in medical image analysis **42** 60–88

Liu L, Jolly S, Cao Y, Vineberg K, Fessler J A and Balter J M 2017 Female pelvic synthetic CT generation based on joint intensity and shape analysis *Phys. Med. Biol.* **62** 2935–49

Low D A 2010 Gamma dose distribution evaluation tool *J. Phys.: Conf. Ser.* **250** 012071

Low D A 2017 MRI guided radiotherapy *Advances in Radiation Oncology* (*Cancer Treatment and Research*) (New York: Springer) pp 41–67

Maspero M, Seevinck P R, Schubert G, Hoesl M A U, van Asselen B, Viergever M A, Lagendijk J J W, Meijer G J and van den Berg C A T 2017a Quantification of confounding factors in MRI-based dose calculations as applied to prostate IMRT *Phys. Med. Biol.* **62** 948–65

Maspero M, van den Berg C A T, Landry G, Belka C, Parodi K, Seevinck P R, Raaymakers B W and Kurz C 2017b Feasibility of MR-only proton dose calculations for prostate cancer radiotherapy using a commercial pseudo-CT generation method *Phys. Med. Biol.* **62** 9159–76

Mathieu M, Couprie C and LeCun Y 2015 Deep multi-scale video prediction beyond mean square error *CoRR* (arXiv:abs/1511.05440)

Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q and Shen D 2017 Medical image synthesis with context-aware generative adversarial networks *Medical Image Computing and Computer-Assisted Intervention* (Cham: Springer) pp 417–25

Nyholm T, Nyberg M, Karlsson M G and Karlsson M 2009 Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments *Radiat. Oncol.* **4** 1–9

Owrangi A M, Greer P B and Glide-Hurst C K 2018 MRI-only treatment planning: benefits and challenges *Phys. Med. Biol.* **63** 05TR01

Persson E *et al* 2017 Mr-opera-a multi-center/multi-vendor validation of mri-only prostate treatment planning using synthetic ct images *Int. J. Radiat. Oncol. Biol. Phys.* **99** 692–700

Prior P, Chen X, Botros M, Paulson E S, Lawton C, Erickson B and Li X A 2016 MRI-based IMRT planning for MR-linac: comparison between CT- and MRI-based plans for pancreatic and prostate cancers *Phys. Med. Biol.* **61** 3819–42

Raaijmakers A J E, Raaymakers B W and Lagendijk J J W 2008 Magnetic-field-induced dose effects in MR-guided radiotherapy systems: dependence on the magnetic field strength *Phys. Med. Biol.* **53** 909–23

Raaymakers B W *et al* 2017 First patients treated with a 1.5 T MRI-Linac: clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment *Phys. Med. Biol.* **62** L41–50

Schmidt M A and Payne G S 2015 Radiotherapy planning using MRI *Phys. Med. Biol.* **60** R323–61

Siversson C, Nordström F, Nilsson T, Nyholm T, Jonsson J, Gunnlaugsson A and Olsson L E 2015 Technical note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm *Med. Phys.* **42** 6090–97

Thor M, Bentzen L, Elstrøm U V, Petersen J B B and Muren L P 2013 Dose/volume-based evaluation of the accuracy of deformable image registration for the rectum and bladder *Acta Oncol.* **52** 1411–6

Thor M, Petersen J B B, Bentzen L, Høyer M and Muren L P 2011 Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer *Acta Oncol.* **50** 918–25

Thwaites D 2013 Accuracy required and achievable in radiotherapy dosimetry: have modern technology and techniques changed our views? *J. Phys.: Conf. Ser.* **444** 012006

Tyagi N, Fontenla S, Zhang J, Cloutier M, Kadbi M, Mechalakos J, Zelefsky M, Deasy J and Hunt M 2016 Dosimetric and workflow evaluation of first commercial synthetic CT software for clinical use in pelvis *Phys. Med. Biol.* **62** 2961–75

Walker A, Liney G, Metcalfe P and Holloway L 2014 MRI distortion: considerations for MRI based radiotherapy treatment planning *Australas. Phys. Eng. Sci. Med.* **37** 103–13

Wang D, Doddrell D M and Cowin G 2004 A novel phantom and method for comprehensive 3-dimensional measurement and correction of geometric distortion in magnetic resonance imaging *Magn. Reson. Imaging* **22** 529

Wang H, Du K, Qu J, Chandarana H and Das I J 2018 Dosimetric evaluation of magnetic resonance-generated synthetic CT for radiation treatment of rectal cancer *Plos One* **13** e0190883

Wolterink J M, Dinkla A M, Savenije M H F, Seevinck P R, van den Berg C A T and Isgum I 2017 Deep MR to CT synthesis using unpaired data *Simulation and Synthesis in Medical Imaging* ed S A Tsaftaris *et al* (Cham: Springer) pp 14–23

Zhu J Y, Park T, Isola P and Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks *IEEE Int. Conf. on Computer Vision (October 2017)* (IEEE) (http://doi.org/10.1109/iccv.2017.244)