SCIENTIFIC REPORTS

natureresearch

Check for updates

OPEN

Generation of Pseudo-CT using High-Degree Polynomial Regression on Dual-Contrast Pelvic MRI Data

Samuel C. Leu[1], Zhibin Huang[1,2] & Ziwei Lin[1]✉

Increasing interests in using magnetic resonance imaging only in radiation therapy require methods for predicting the computed tomography numbers from MRI data. Here we propose a simple voxel method to generate the pseudo-CT (pCT) image using dual-contrast pelvic MRI data. The method is first trained with the CT data and dual-contrast MRI data (two sets of MRI with different sequences) of multiple patients, where the anatomical structures in the images after deformable image registration are segmented into several regions, and after MRI intensity normalizations a regression analysis is used to determine a two-variable polynomial function for each region to relate a voxel's two MRI intensity values to its CT number. We first evaluate the accuracy via the Hounsfield unit (HU) difference between the pseudo-CT and reference-CT (rCT) images and obtain the average mean absolute error as 40.3 ± 2.9 HU from leave-one-out-cross-validation (LOOCV) across all six patients, which is better than most previous results and comparable to another study using the more complicated atlas-based method. We also perform a dosimetric evaluation of the treatment plans based on pCT and rCT images and find the average passing rate within 2% dose difference to be 95.4% in point-to-point dose comparisons. Therefore, our method shows encouraging results in predicting the CT numbers. This polynomial method needs less computer storage than the interpolation method and can be readily extended to the case of more than two MRI sequences.

The workflow of conventional radiation therapy (RT) uses computed tomography (CT) images to create the treatment plan and to position the patient at treatment. Magnetic resonance imaging (MRI) is also often used to provide precise delineation of RT target volumes due to its superior soft tissue contrast[1]. When these two modalities are both used, the workflow then necessitates an additional step of registering the images of the two modalities: MRI and CT[2]. Recently, the concept of MRI-Linac using magnetic resonance imaging only in radiation therapy, i.e., MRI-only RT, has become more popular, because MRI-only RT has the benefits of improving the workflow and removing systematic errors in registering MRI and CT images[3,4].

However, in MRI-only RT there is a problem in creating treatment plans with the MRI images because of the lack of CT images or the electron density information. To address this problem, various methods of generating CT images, called pseudo-CT (pCT), have been investigated[5–7]. Existing methods in deriving a pseudo-CT from MR images may be classified into different categories. They include the classification into segmentation-based, intensity-based, atlas-based, and hybrid methods[7], or the classification into segmentation-, atlas-, patch-, and learning-based methods[8], or the classification into bulk density assignment, atlas-based, and voxel-based methods[9]. These methods have produced mean absolute error (MAE) values ranging from 85 HU[10] to 137 HU[11] for the brain and from 36.5 HU[12] to 74.3 HU[13] for the prostate (pelvis region), for example.

Among these methods, atlas-based methods[12,14–17] align an MRI atlas, which has been derived from an MRI database pre-registered to the corresponding reference-CT (rCT) images, to the target patient's MRI images through registration. The atlas thus contains pre-determined correlations between the MRI voxels and the variables of interest such as the CT number or organ type. The same registration (with the translational, rotational, and deformable information) that maps the MRI atlas to the target patient's MRI images is then applied to the atlas CT images to create the target patient' pCT images. This approach is popular because of its potential in producing

[1]Department of Physics, C-209 Howell Science Complex, East Carolina University, Greenville, NC, 27858, USA. [2]Global Medical Consulting, LLC, Brentwood, TN, 37027, USA. ✉e-mail: linz@ecu.edu

1

reliable pCT images with conventional MRI images. However, it requires accurate deformable image registration between the atlas and the target patient's magnetic resonance (MR) images, which can be difficult, especially when large anatomical variations or pathological differences exist. This problem can be partially overcome by using multiple atlases or the hybrid method that combines the atlas method with other methods to increase the overall strength and reduce the overall weakness[18–20].

Learning-based methods employ model-fitting or statistical learning techniques to generate a mapping function that correlates the MRI information to the corresponding CT numbers[9,21–28]. Deep learning is a machine learning technique that is useful for processing low-level noisy data such as medical images. Currently, a main deep learning technique used in the field of radiation oncology is the deep convolutional neural network that is based on the U-net architecture[24,26,28]. An advantage of learning-based approaches is that they can take into account neighborhood voxels[23,26].

Voxel-based methods[9,11,13,23,29–35], such as the bulk-density method[36,37], mainly translate the voxel intensity information in the MRI images to CT numbers and pCT images. In general, a voxel-based method is specific to the MRI sequence(s) that the model is trained on but does not need the target patient anatomy to closely match the training patients. In contrast, atlas-based methods for generating the pCT do not depend on the MRI sequence(s) but depend on the patient anatomy, where a major difference between the target patient anatomy and the patient database population may result in inaccurate pCT images.

In this study, we investigate a voxel method that uses two MRI image sets acquired using different MRI sequences together with two-variable polynomial fitting functions to create pseudo-CT images. This method is validated by leave-one-out-cross-validation (LOOCV)[8,10,11,16,18,22], applied to a sample of six patients. We evaluate our method via MAE between the pCT and rCT images and also by comparing the dose distributions of simulated RT plans created on the pCT and rCT images.

Methods

Patient groups. Patients with carcinoma of the cervix were prospectively studied with serial MRI and CT scans during RT on an IRB-approved imaging protocol granted by the University & Medical Center Institutional Review Board (UMCIRB) of the East Carolina University (ECU). Informed consents were obtained from patients with all identifier information removed. Collection of data and methods were carried out in accordance with the guidelines and regulations of UMCIRB. We selected 6 female patient cervical image datasets. The 6 datasets selected for this study have none of the following: significant visible artifacts, considerable co-registration errors, or large rotational transformation for aligning the MRI and CT images. This study has no effect on the treatment of these patients.

Specifically, the six patients in this study were staged clinically with the International Federation of Gynecology and Obstetrics criteria[38], including physical examinations, chest radiograph, tumor biopsy, complete blood count, serum chemistries, intravenous pyelogram, and abdominal-pelvic computed tomography. There were 2 patients in Stage IB, 2 in Stage IIA, and 2 in Stage IVB (inguinal metastasis). Median age was 55 years (range 25–89). Two MRI scans were acquired on a Siemens Magneton Espree 1.5 T scanner. The T1-weighted MR sequence (MR_1) was a fast low angle shot with different repetition time TR, an echo time TE = 4.53 ms, a flip angle FA = 70°, a matrix size of 320 × 320, and a voxel size of 0.81 × 0.81 × 5.0 mm³. The T2-weighted MR sequence (MR_2) was a turbo spin echo with different repetition time TR, an echo time TE = 87 ms, a similar flip angle FA, a matrix size of 320 × 320, and a voxel size of 0.81 × 0.81 × 5.0 mm³. Both T1-weighted (T1w) and T2-weighted (T2w) MRI sets have more than 30 images per scan. In addition, a whole body CT scan using 120 kVp, various mAs values, a matrix size of 512 × 512 × 326, and a voxel size of 0.97 × 0.97 × 4.0 mm³ was acquired for each patient on a Siemens Biograph CT within one month of the acquisition of MRI. The detailed acquisition parameters for the MRI and CT scans of each patient are given in Supplementary Table S1.

Pre-processing. We register the CT image set with the deformable setting in multi-modality registration on Velocity (Varian Medical Systems) using the MR_1 image set as the reference image; the same is done to register the MR_2 image set. The CT dataset is resampled to the same resolution as the MRI dataset: axial image dimension of 320 × 320 with the voxel size at ≈ 0.81 × 0.81 × 5.0 mm³. Because the MRI images of a given sequence have different image contrasts for different patients, MRI images were normalized[8,9,27,34,35,39,40]. Specifically, for the MRI datasets acquired with the same MRI sequence, e.g. MR_1, we first find the average intensity value for all patients and that for a given patient, then their ratio is used as the patient-specific correction factor to multiply the voxel MR_1 intensity values of this given patient, so that the average MR_1 intensity values for all patients become the same. In addition, a binary mask is created by using a voxel intensity threshold combined with the edge detection function in MATLAB to eliminate air voxels and cover only the imaged subject to reduce the computational burden and improve the pCT accuracy[11,40–42]. Air voxels are set to a value of −1000 HU.

Segmentation. The imaged anatomical structure of each patient is manually segmented with masks into three regions: bony, soft, and mixed regions. An example of the three regions superimposed on an MRI image is shown Fig. 1A. The bony region contains the cortical and cancellous bone tissues, while the soft region contains all soft tissues. For the bony region, a mask is defined at the edge of the cortical bone with approximately 1 mm margin left out. A margin of approximately 2–3 mm surrounding the bony tissue is not included within the soft region. Margins are generated automatically by expansion or reduction of the mask. The mixed region represents the region between the above two margins. Then voxels within the same region are grouped together for determining the prediction model for the region.

When necessary, an excluded region, as shown in Fig. 1B, is drawn to select misaligned anatomical structures due in large part to changes during the elapse time between MRI and CT acquisitions. To obtain the correct
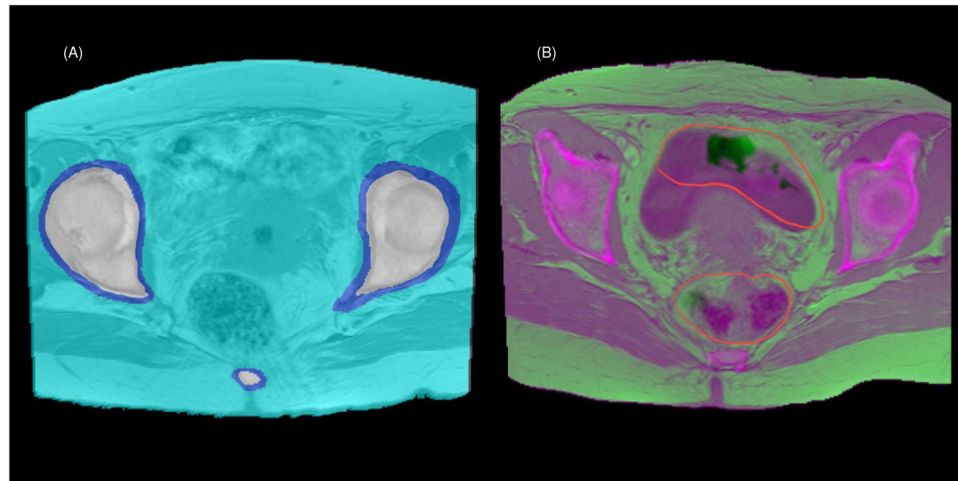
Figure 1. (**A**) Segmentation of the three regions superimposed on an MRI image: bony region (white), soft region (light blue), and mixed region (dark blue). (**B**) Example of a CT image (magenta) overlay on a corresponding MRI image (green) with the excluded region shown inside the red contours.

correlation between MRI intensity values and the CT number of each voxel, voxels from the excluded region are not included in the training data used to determine or evaluate our prediction model.

Modeling relation between two MRI intensity values and CT Number. The voxel method is based on the assumption that a relationship exists between a voxel's CT number and MRI intensity value on average, at least for voxels of the same tissue type. However, voxels with the same MRI intensity value, even of the same tissue type, may have very different CT numbers. Therefore, our strategy is to have a second set of MRI images acquired with different MRI sequences, so that voxels with the same MRI intensity value (in the first set) but very different CT numbers may have different MRI intensity values in the second set, which will allow us to distinguish those voxels and map them to their corresponding CT numbers.

For each region of the segmented anatomical structures (bony, soft and mixed regions), all the voxels in the same region are grouped together as the training data to determine the prediction model for the region. The data from each voxel consists of its CT-value, the normalized intensity value S_1 from the MR_1 image set, and the normalized intensity value S_2 from the MR_2 image set. The prediction model is trained to map a voxel's two MRI intensities to its CT number on the Hounsfield unit (HU) scale for each region. In this study, we used a two-variable nth-degree polynomial function that depends on a voxel's two MRI intensity values to predict the voxel's expected CT number:

$$
\begin{aligned}
pCT(S_1, S_2) &= \sum_{i_1,i_2=0}^{i_1+i_2 \le n} c_{i_1,i_2} S_1^{i_1} S_2^{i_2} \\
&= c_{0,0} + c_{1,0}S_1 + c_{0,1}S_2 + c_{1,1}S_1S_2 + \ldots + c_{n,0}S_1^n + c_{0,n}S_2^n.
\end{aligned}
\tag{1}
$$

In practice, the MRI data of each region are separated into different MR_1 and MR_2 intensity value bins to reduce the computational burden and noise. An equal number of bins (Nbin) for MR_1 and MR_2 is used, while the bin width is determined based on the maximum normalized MR_1 and MR_2 intensity values (1,420 and 37,100 respectively in this study). As a result, every voxel is associated with two bin indices, e.g., i for the MR_1 bin and j for the MR_2 bin, or (i, j). Inside a given (i, j) bin in these two dimensions, there can be multiple voxels with a range of CT numbers. We can define the average CT number of the bin as

$$
\overline{CT_{i,j}} = \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} CT_{i,j,k},
\tag{2}
$$

where $N_{i,j}$ is the number of voxels within the i-th MR_1 bin and j-th MR_2 bin, and $CT_{i,j,k}$ denotes the CT number of the $k-$th voxel within this bin. We use equation (1) to perform a regression analysis on the training data of each region to determine the coefficients c_{i_1,i_2} in the polynomial, then $pCT(S_1, S_2)$ best describes the average CT number $\overline{CT_{i,j}}$ as a function of the MRI intensity values S_1 and S_2. The regression analysis is performed by using the 'NonlinearModelFit' function in Mathematica (Wolfram, Champaign, United States). Note that a weighting factor based on the number of data points within a bin is used in the regression analysis. The 3-dimensional plots and contour plots of the polynomial function $pCT(S_1, S_2)$ for each region of Cycle1 are provided in Supplementary Figures S1–S3 and S4–S6, respectively. We have also put the Mathematica source code of the pCT regression analysis along with example input files and the output file for the calculated polynomial coefficients at http://myweb.ecu.edu/linz/pCT/.
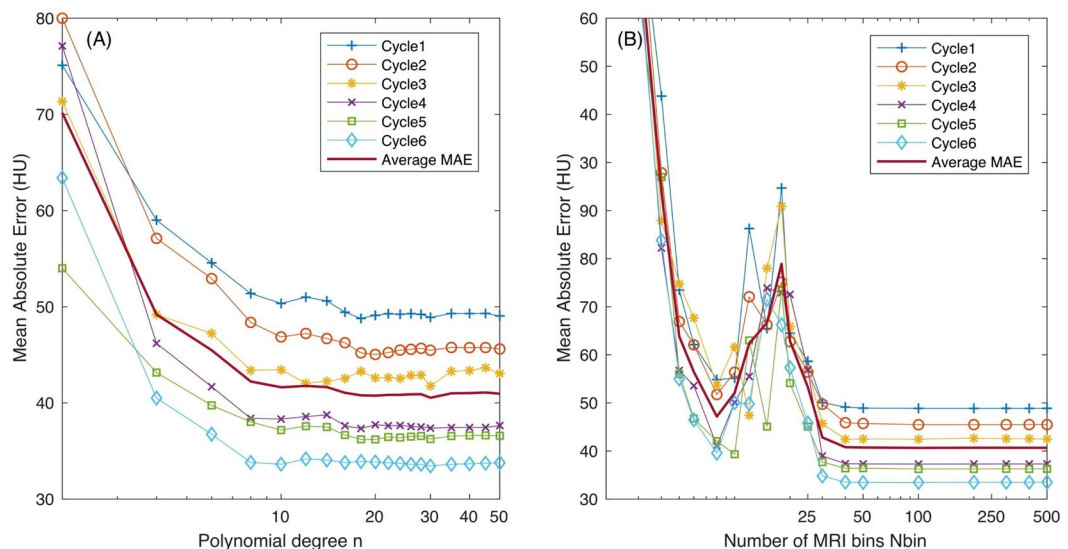
Figure 2. (**A**) Mean absolute error versus the polynomial degree n with Nbin = 200. (**B**) Mean absolute error versus the number of MRI bins (Nbin) with the polynomial degree $n = 30$.

Generation of pCT. To generate pseudo-CT images for a target patient, we first segment the anatomical structures on the MR_1 and MR_2 images of the target patient into the same types of regions: bony, soft, and mixed regions. For each region, the corresponding function from equation (1) is then applied to each voxel to map its S_1 and S_2 MRI intensity values to an expected CT number. Note that we restrict the predicted CT number to -1000 HU and 2000 HU. The pseudo-CT images are then obtained for the target patient after the CT number is generated for all the voxels. The total time to convert a whole pelvic MRI scan is within approximately 3 minutes using a Mac Pro (mid-2012) desktop with 2×2.4 GHz 6-Core Intel Xeon processor and 16 GB (8×2 GB) 1333 MHz DDR3 ECC memory.

To evaluate the accuracy of our method, we apply LOOCV for the six patients. There are six cycles, where five patients are used as the training data to determine the coefficients c_{i_1,i_2} in equation (1) for each region and then the prediction model is applied to the remaining (target) patient.

Mean absolute error. To evaluate the quality of the generated pCT images from our model, we use the mean absolute error as defined below:

$$MAE = \frac{1}{N}\sum_{k=1}^{N}\left|pCT_k - rCT_k\right|,$$

(3)

where N is the total number of body voxels (except voxels from the excluded region) of the target patient, pCT_k and rCT_k denote respectively the CT number from the generated pCT and the rCT for voxel k. The MAE thus measures the voxel-wise average error. When evaluated for each region, N in equation (3) then represents the total number of body voxels in that region of the target patient.

Optimization of parameters in the prediction model. We determine the optimal value of two key parameters in our prediction model: the polynomial degree n, and the number of MRI bins Nbin. From Fig. 2A, we see an overall decreasing trend in the average MAE value as the polynomial degree n increases (with Nbin = 200). In particular, the model's prediction improves as degree n increases to $n \sim 20$, and further increase of n changes the average MAE by no more than $\sim 1\%$. Therefore, we use $n = 30$ as the optimal value for the polynomial degree in equation (1). The effect of the number of MRI bins on the MAE values is shown in Fig. 2B. Although the MAE has a non-monotonous dependence on Nbin from ~10 to ~25, the average MAE values are the lowest for Nbin >40 and shows little further decrease with Nbin. Since the computing burden of using Nbin of 200 versus Nbin of 40 is similarly small, and a polynomial function with a higher Nbin should better correlate the MR intensities to the CT number when the number of patients gets bigger in future studies, we choose to use Nbin = 200.

In addition, not all (i, j) bins are filled, i.e., have voxels from the training data and thus have corresponding CT numbers. This often causes the polynomial function to rapidly deteriorate in the MRI region with empty bins. To improve the stability of the regression analysis, we fill the empty (i, j) bins with CT numbers, thus all MRI (i, j) bins have corresponding CT numbers before the regression analysis. In particular, we assign an empty bin with the CT number of the filled bin with the smallest distance. For example, for an empty bin with the MR_1 bin number i_e and MR_2 bin number j_e, its distance from a filled bin with the MR_1 bin number i and MR_2 bin number j is calculated as
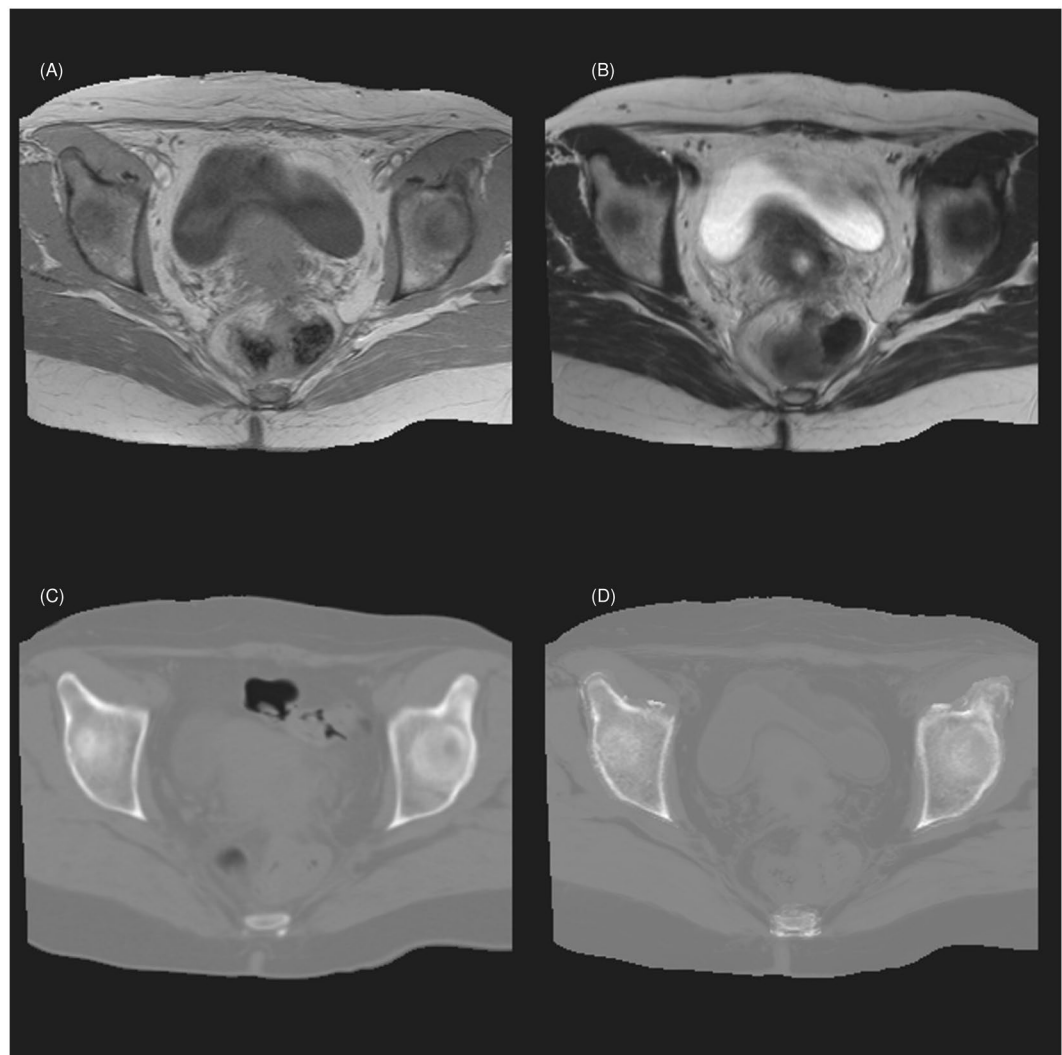
Figure 3. Axial views of the (**A**) MR$_1$, (**B**) MR$_2$, (**C**) reference-CT, and (**D**) generated pCT image at a given location of a target patient.

$$d = \sqrt{(i - i_e)^2 + (j - j_e)^2}. \qquad (4)$$

The HU-value of the filled bin with the smallest distance d is then assigned to the empty bin. If the smallest distance corresponds to multiple filled bins, then the average HU-value of those multiple bins is assigned to the empty bin. For our regression analysis, the weighting factor assigned to an empty bin is 10% of that of the nearest filled bin.

Results

Accuracy of the pCT image.   An example of the pCT image is shown in Fig. 3 with the axial view of the patient's MRI (MR$_1$ and MR$_2$) and CT images at the same slice. We see that the pCT image (Fig. 3D) closely matches the rCT (Fig. 3C). Noticeable structural differences are predominantly in the soft tissues, specifically the bladder and bowel. Similar differences are also observed between the rCT (Fig. 3C) and MRI images (Fig. 3A,B). Other minor differences are observed at the muscle-fat tissue and bony-tissue interfaces. These can be mostly attributed to the extended duration between the acquisition times of MRI and CT images. Therefore, the region with major structural differences (e.g. bladder and bowel due to daily changes) between the CT and MRI images are segmented as the excluded region and omitted in determining the prediction model or the MAE calculations.

Table 1 presents the MAE values for the overall patient volume and for each segmented (bony, soft, mixed) region of each LOOCV cycle as well as the average value for all six cycles. We obtained an average MAE of 40.3 ± 2.9 HU for the overall volume in the images, while the raw MAE values for the bony, soft, and mixed regions are 102.7 ± 9.9 HU, 24.6 ± 1.0 HU, and 143.5 ± 9.2 HU, respectively. The weighted MAE value for a region is determined by multiplying the raw MAE value by the percent of voxels in that region, thus the sum of the weighted MAE values of the three regions equals the overall MAE value. We find that the bone-tissue interface has the most

	Mean Absolute Error (HU)						
		Bony Region		Soft Region		Mixed Region	
	Overall	Raw	(Weighted)	Raw	(Weighted)	Raw	(Weighted)
Cycle1	44.4	118.5	(12.1)	25.8	(21.1)	138.3	(11.2)
Cycle2	42.7	100.2	(8.1)	25.6	(20.9)	134.3	(13.8)
Cycle3	40.9	104.0	(9.6)	24.8	(20.8)	151.8	(10.4)
Cycle4	38.7	105.6	(9.5)	23.0	(19.5)	157.0	(9.6)
Cycle5	38.1	99.7	(8.8)	24.6	(20.9)	135.8	(8.5)
Cycle6	36.7	88.0	(8.0)	23.9	(20.4)	143.6	(8.3)
Average	40.3 ± 2.9	102.7 ± 9.9	(9.4 ± 1.5)	24.6 ± 1.0	(20.6 ± 0.6)	143.5 ± 9.2	(10.3 ± 2.0)

Table 1. MAE values between pseudo-CT and reference-CT for the overall patient volume and for each segmented (bony, soft, and mixed) region of each LOOCV cycle and their average values for all six cycles. The weighted MAE value of a region is the raw MAE value multiplied by the percent of voxels in that region among all voxels.

	Cycle1	Cycle2	Cycle3	Cycle4	Cycle5	Cycle6	Average
Passing rate	95.7%	94.8%	95.8%	96.2%	95.7%	94.2%	$(95.4 \pm 0.7)\%$

Table 2. Passing rates (i.e., with a dose difference less than 2%) from the point-to-point dose comparisons between the pseudo-CT and reference-CT treatment plans.

noticeable differences, because there the deviations between MRI and reference-CT images are the highest due to registration errors. The high raw MAE values for the mixed region in Table 1 support this observation, since the mixed region consists mainly of the bone-tissue interface. Also, because the bony region and the mixed region account for only a small percentage (9% and 7% respectively) of all the voxels in the imaged volume, their contributions to the overall MAE (i.e., their weighted MAE values) are lower than that from the soft region.

Comparison of dosimetric calculations using pCT and rCT.

To further assess the accuracy of the pCT image volume, we compare dosimetric calculations of treatment plans based on the pCT and the rCT. External beam pelvic radiotherapy with cisplatin and brachytherapy is the standard of care for patients with advanced cervical malignancy. The typical geometrical setup for cervical cancer is 4-field box technique, namely anterior-posterior (AP), posterior-Anterior (PA), right-lateral (RLAT) and left-lateral (LLAT) fields. A modified setup is a 3-field technique including RLAT, LLAT and AP, which can be used to reduce dose to rectum. In this study we use the 3-field treatment plan created with the Eclipse treatment planning system (Varian Medical System, Palo Alto, CA, USA), where the planning target volume (PTV) is the cervix. We then draw contours for the left and right femur heads, bladder, bowel, and rectum on the treatment planning system. For the excluded regions (mainly the bladder and bowel) where the CT and MR images differ significantly, the CT numbers for the bladder and bowel are overridden and assigned as 5 HU and 45 HU, respectively (for both the pCT and rCT). The center of the PTV is prescribed a dose of 56 Gy and for simplicity all three beams are equally weighted. The treatment plan is first generated on the rCT images and then mapped directly to the pCT images. Dose distributions of rCT and pCT treatment plans are calculated on the treatment planning system.

We compare the dose value of each voxel from the pCT and rCT dose distributions. A voxel that has a dose difference within the tolerance of $|2\%|$ is considered as passing the criterion. Passing voxels within the imaged volume are tabulated (not including the air voxels), and a passing rate is calculated. Table 2 shows the passing rate results, where the average passing rate is 95.4% in the point-to-point dose comparisons between the pCT and rCT treatment plans. When the pCT image is close to rCT, the dose distributions in a small region of interest would be very close to each other, which difference can be measured through the gamma index. Note that our point-to-point dose comparison method is similar to the gamma dose distribution comparison method except for the distance to agreement (DTA) value. In fact, our comparison method corresponds to using DTA = 0 and is thus stricter than the usual gamma index.

Sources of the pCT uncertainties.

Here we investigate in more detail the sources that contribute to the finite MAE between the generated pCT and the reference-CT images. The rCT image, pCT image, and the image of their absolute difference ($|rCT-pCT|$) at four different slices for patient 1 are provided in Supplementary Figures S7–S10. In general, the MAE has two main contributions: one from the imperfect fitting of the average CT number $\overline{CT}_{i,j}$ as a function of the MRI bin numbers (i, j) with equation (1), the other from the intrinsic fluctuation or spread of the CT numbers at given MRI bin numbers (i, j) around the mean value $\overline{CT}_{i,j}$.

Instead of MAE, it is simpler for this purpose to examine the root-mean-squared (rms) difference between pseudo-CT and reference-CT:

	σ_{pCT} (HU)			
	Overall	Bony Region	Soft Region	Mixed Region
Cycle1	86.1	160.0	48.4	188.1
Cycle2	80.7	133.7	46.0	180.6
Cycle3	79.4	142.5	41.1	209.0
Cycle4	80.4	147.8	43.6	216.5
Cycle5	69.0	128.1	39.3	179.3
Cycle6	70.0	118.2	41.7	192.7
Average	77.6 ± 6.7	138.4 ± 14.9	43.3 ± 3.4	194.4 ± 15.3

Table 3. σ_{pCT} values between pseudo-CT and reference-CT for the overall patient volume and for each segmented region of each LOOCV cycle and their average values for all six cycles.

	MAE_{fit} (HU)			σ_{fit} (HU)		
	Bony Region	Soft Region	Mixed Region	Bony Region	Soft Region	Mixed Region
Cycle1	10.4	2.0	16.6	21.5	5.3	31.6
Cycle2	11.0	2.1	17.3	22.6	4.5	33.3
Cycle3	10.4	2.1	16.0	20.0	5.4	31.3
Cycle4	10.5	2.1	15.8	21.4	5.5	29.8
Cycle5	11.3	2.1	16.5	23.2	5.7	32.4
Cycle6	10.9	2.1	15.9	23.3	5.4	31.7
Average	10.8 ± 0.4	2.1 ± 0.0	16.3 ± 0.6	22.0 ± 1.3	5.3 ± 0.4	31.7 ± 1.2

Table 4. MAE_{fit} and σ_{fit} values for fitting the mean reference-CT numbers with equation (1) for the overall patient volume and for each segmented region of each LOOCV cycle and their average values for all six cycles.

$$\sigma_{pCT} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(pCT_k - rCT_k)^2} \ ,$$

(5)

where N is the total number of voxels in the imaged volume (not including the voxels in the excluded region), and k is voxel index. When evaluated for each region, N in the above equation then represents the total number of body voxels in that region in the imaged volume. The σ_{pCT} values are presented in Table 3. Note that as expected σ_{pCT} is always higher than (or equal to) the corresponding raw MAE value, which can be verified by comparing Tables 1 and 3. Also, the largest values are observed for bony and mixed regions, consistent with the fact that their raw MAE values shown in Table 1 are higher than that of the soft region.

To evaluate the goodness of fit of the polynomial equation (1), we can calculate the mean absolute error and the rms difference between the predicted $pCT(S_1, S_2)$ from equation (1) and the averaged CT number $\overline{CT}_{i,j}$ in the reference-CT training data for each segmented region:

$$MAE_{fit} = \frac{1}{N}\sum_{i,j=1}\left|\overline{CT}_{i,j} - pCT_{i,j}\right|N_{i,j}, \qquad \sigma_{fit} = \sqrt{\frac{1}{N}\sum_{i,j=1}\left(\overline{CT}_{i,j} - pCT_{i,j}\right)^2 N_{i,j}},$$

(6)

where $N = \sum_{i,j=1}^{Nbin}N_{i,j}$ gives the total number of voxels in the imaged volume for that region. In the above equation, $pCT_{i,j} \equiv pCT(S_1, S_2)$ is the predicted CT number for the MRI bin (i, j), where S_1 and S_2 here represents the central intensity value of the i-th MR_1 bin and the j-th MR_2 bin, respectively. Table 4 shows the MAE_{fit} and σ_{fit} values for each cycle and their average values. We can see that the fit is better for the soft region than for the other two regions.

Another contribution to the MAE of the generated pCT comes from the fact that voxels at the same MRI bin numbers (i, j) do not have exactly the same CT number, although they average to the mean value $\overline{CT}_{i,j}$. For this, we can calculate for each region the mean absolute error and the rms difference between individual voxel's reference-CT number and the averaged value $\overline{CT}_{i,j}$ at the same MRI bin (i, j) as

$$MAE_{rCT} = \frac{1}{N}\sum_{i,j=1}^{Nbin}\sum_{k=1}^{N_{i,j}}\left|rCT_k - \overline{CT}_{i,j}\right|, \qquad \sigma_{rCT} = \sqrt{\frac{1}{N}\sum_{i,j=1}^{Nbin}\sum_{k=1}^{N_{i,j}}\left(rCT_k - \overline{CT}_{i,j}\right)^2},$$

(7)

where $N = \sum_{i,j=1}^{Nbin}N_{i,j}$ and k is voxel index within the MRI bin (i, j). These values are shown in Table 5. We see that the values for the bony region and the mixed region are much higher than that of the soft region, similar to Table 3.

	MAE_{rCT} (HU)			σ_{rCT} (HU)		
	Bony Region	Soft Region	Mixed Region	Bony Region	Soft Region	Mixed Region
Cycle1	95.3	23.4	138.0	128.4	40.2	187.5
Cycle2	97.6	23.7	141.3	132.5	40.3	190.6
Cycle3	97.6	23.7	133.3	131.9	41.4	183.6
Cycle4	96.8	24.1	133.3	129.9	41.0	182.7
Cycle5	98.8	23.6	138.1	134.4	42.5	189.3
Cycle6	100.3	23.9	135.1	135.3	41.2	186.8
Average	97.7 ± 1.7	23.7 ± 0.2	136.5 ± 3.2	132.1 ± 2.6	41.1 ± 0.8	186.7 ± 3.1

Table 5. MAE_{rCT} and σ_{rCT} values for the spread of CT numbers in the same MRI bin for the overall patient volume and for each segmented region of each LOOCV cycle and their average values for all six cycles.

	Mean Absolute Error (HU)						
	Current Method	Only MR_1	Only MR_2	Interpolation	Without Segmentation	Without Using Excluded Region	LOOCV 3 + 1
Cycle1	44.4	48.2	55.6	44.5	66.0	45.3	48.5
Cycle2	42.7	47.3	64.4	42.9	65.3	47.4	
Cycle3	40.9	46.1	81.5	41.1	61.9	43.1	40.7
Cycle4	38.7	44.1	64.6	38.8	60.8	47.4	39.7
Cycle5	38.1	41.7	53.2	38.3	64.7	40.8	
Cycle6	36.7	40.8	66.4	36.9	59.5	41.0	37.2
Average	40.3 ± 2.9	44.7 ± 3.0	64.3 ± 10.0	40.4 ± 2.9	63.0 ± 2.7	44.1 ± 2.9	41.5 ± 4.9
p-value		<0.001	0.0026	<0.001	<0.001	0.017	0.25

Table 6. MAE values from our method using two MRI sets with segmentation, in comparison with those using only one MRI set, using the interpolation method, using two MRI sets but without segmentation, without using the excluded region, and using only 4-patient datasets. The p-value is the two-sided value determined by performing a paired t-test between a given method and the 'Current Method' for sample size 6 (except that the sample size is 4 for the t-test of 'LOOCV 3 + 1').

We also find that the MAE_{rCT} or σ_{rCT} values are much higher than the corresponding MAE_{fit} or σ_{fit} values for the same region. Furthermore, if the uncertainties from these two sources were independent, one would expect the following relation for each region:

$$\sigma_{pCT}^2 = \sigma_{rCT}^2 + \sigma_{fit}^2. \tag{8}$$

From Tables 3–5, we indeed see that the average σ_{pCT} value of each region is higher than the corresponding average value of σ_{rCT} or σ_{fit}, and the above relation is approximately satisfied even quantitatively. Similar features can also be observed for the MAE values from Tables 1, 4 and 5. Specifically, the average raw MAE value of each region is higher than the corresponding average value of MAE_{rCT} or MAE_{fit}, with $MAE_{rCT} \gg MAE_{fit}$, and the relation $MAE^2 \approx MAE_{rCT}^2 + MAE_{fit}^2$ holds approximately. Therefore, we conclude that the intrinsic fluctuation of the reference-CT numbers for voxels with similar MRI intensity values, not the fitting procedure from the regression analysis, is the dominant source to the uncertainty of the generated pCT images.

Discussions

We have presented a voxel method that uses two sets of MRI images acquired with different MRI sequences to generate the pCT images. This method is straightforward to implement and extend. For example, this method can be easily generalized to use more than two sets of MRI contrast data, where equation (1) would just include additional independent MRI variables. Our method differs from earlier pCT works[11,13,23,29–35], mainly in the use of high-degree polynomials with more than one MRI variables together with the use of deformable image registration and MRI intensity normalizations.

Using multiple MRI sets with different sequences allows for improved delineation and identification of voxels with different CT numbers and thus leads to better accuracy of the pCT. Table 6 compares our MAE values with those using only one set of MRI image (MR_1 or MR_2), including the corresponding two-sided p-values from the paired t-test to compare the 'Current Method' to the single-MRI methods. The small p-values (p < 0.001 and p = 0.0026) suggest that our method of using dual-contrast MRI data improves the pCT accuracy.

We note that the use of multiple MR sequence parameters has been explored earlier and the advantage of using more than one MRI sequence has been pointed out[8,30]. Aouadi et al.[16] used a patch-based method for the brain and also combined T1- and T2-weighted MRI intensities to have an enhanced description of tissue properties.

8

Burgos et al.[17] used an atlas-based method to generate pCT images, where each subject had a T1w MR image, T2w MR image, and a CT image for pelvis acquired on the same day. Note that a disadvantage of the atlas method is that it will be unable to extrapolate to a feasible pCT image without pre-existing templates for other anatomic areas or an atypical anatomy. Speier et al.[34] investigated the generation of pCT images from T1w and T2w MRI images for the brain including a voxel-based method that used a localized lookup table, where the lookup table algorithm relied on segmentation and regionalization steps in the data preprocessing. Pileggi et al.[35] used T1w and T2w MR for brain to generate a pCT for proton therapy treatment, where a voxel-based lookup table was generated by binning HU in matrixes of 10×10 MR intensity units together with rigid image registration. Koike et al.[9] described a method to generate pCT images from T1w, T2w and fluid-attenuated inversion recovery MR images using an adversarial network for the head region. Zhong et al.[8] used a patch-based approach and reported a MAE of 97.72 ± 15.78 HU for combined T1w and T2w training but 113.73 ± 16.86 HU for only T1w, showing an improvement in the pCT accuracy from using both T1w and T2w MR sets.

In addition to using a polynomial to calculate the average CT number as a continuous function of the two MRI variables, we have also used an interpolation method that is similar to using a lookup table. We first use the center points of all the MR_1-MR_2 bins to divide the MR_1-MR_2 plane into ~$Nbin^2$ number of cells. Then we calculate the average CT number of an arbitrary point in the MR_2-MR_2 plane by using a bilinear interpolation of the average CT numbers of the four neighboring center points. Table 6 shows that the MAE values from this interpolation method are only slightly higher than those from our current method. The p-value ($p < 0.001$) is small because the paired MAE differences between the interpolation method and our current method from Table 6 (0.1, 0.2, 0.2, 0.1, 0.2, and 0.2 HU, respectively) all have the same sign and are close in magnitude, even though the mean MAE difference is very small (~0.17 HU). We note that, compared to using a lookup table[34,35], using a continuous polynomial function to represent the complicated relation between the MR values and the corresponding CT value has an advantage in that the computer storage of such a polynomial function needs less space since only the coefficients are needed. This advantage in storage will be even greater if more than two MRI sequences or contrasts will be used. We also note that other methods, such as probability functions[19,30], the gaussian mixture regression model[11,22], and weighted summation[13,33], have also been used in other studies to calculate the CT number from the input MRI data.

Similar to other studies[10,29,31,43], setting masks to segment tissues is an important step in improving the pCT accuracy. The importance of segmentation can be seen from the 'Without Segmentation' column of Table 6, where MAE values obtained without segmentation (still with the same excluded region and using two MRI sets) are significantly higher. The small p-value ($p < 0.001$) from the paired t-test suggests that the inclusion of segmentation in the approach of deriving pCT improves its accuracy. Furthermore, we believe the implementation of an automated segmentation process[33] to our method will not only further improve the efficiency but also further improve the accuracy and robustness of our method by ensuring better agreements in the region delineation between patients.

One issue of this study was that the MRI and CT patient data could not be acquired on the same day. This created uncertainties during the image registration process due to the rectal/bladder filling inconsistency between sessions and other setup inconsistencies. To address this issue, we have used a deformable image registration and the excluded region. The excluded region as shown in Fig. 1B contains mainly of organs, bowel and bladder that could have significant internal motion and anatomical differences on a daily basis[44–46]. When we include voxels in the excluded region by applying the corresponding fitting functions, we get the MAE values in the 'Without Using Excluded Region' column of Table 6. The small p-value ($p = 0.017$) from the paired t-test suggests that using an excluded region improves the pCT accuracy. The implementation of better alignment and a shorter duration between the acquisition of CT and MRI datasets into our method could help in reducing the size of the excluded region and thus improve the pCT accuracy. Note that the excluded region is not needed when applying the prediction model to the actual MRI data to generate the pCT images of the target patient.

Since our datasets of 6 patients are rather limited, we investigate the effect of sample size by applying LOOCV for 4 out of the 6 datasets (i.e., assuming that we only have data for patient number 1, 3, 4, and 6). The 'LOOCV $3 + 1$' column of Table 6 shows the corresponding MAE values of using three training datasets and one dataset as the target patient in generating the pCT, which are close to the MAE values of our current 'LOOCV $5 + 1$' method. We also conduct the paired t-test to compare the MAE values from 'LOOCV $3 + 1$' with the corresponding values from our 'Current Method' (i.e., the MAE values under 'Current Method' for Cycle1, 3, 4, and 6). The rather large p-value ($p = 0.25$) indicates that there is insufficient evidence to conclude there is a significant difference between the MAE of 'LOOCV $5 + 1$' and 'LOOCV $3 + 1$'. However, to bring our proposed method to implementation in the clinical routine, we would like to have more training patients (than five used in this study) and also have the MRI and CT data of a given patient acquired on the same day.

In the pre-processing, a deformable image registration[12,19,47] is used to ensure that the anatomical position of each voxel between the MRI and the CT images matches and anatomical structures between the image modalities are aligned. Matching of anatomical positions is necessary in training our model to establish the relationship between the MRI voxel intensity values and the CT number while suppressing the noise from mismatched voxels. For generating the pCT for the target patient, however, a deformable registration will not be necessary, thus the geometric integrity of the target patient images is retained.

We could further improve this model by determining the optimal MRI sequences for acquiring the dual-contrast MRI data. We also recommend using the same MRI sequence parameters to acquire each set of the dual-contrast MRI data (for the training patients as well as the target patient); then the MRI intensities of different patients would be more consistent, which would further improve the pCT accuracy. Note that different anatomical locations (e.g., lung or head and neck) from the pelvic region used in this study could possibly require different optimal MRI sequences for the MRI acquisition, therefore extending the method into different anatomical sites is warranted.

In conclusion, we have developed a voxel-based method that uses two different MR sequence sets (dual-contrast MRI) to create pseudo-CT images, where after deformable image registration and MRI intensity normalizations a regression analysis is used to determine the two-variable high-degree polynomial function for each segmented region. Using pelvic data from six patients, the HU values and dose distributions from the pseudo-CT are in close agreements with those from the reference-CT. Therefore, the pseudo-CT generation using a multi-variable polynomial prediction model with deformable image registration, anatomical segmentation and MRI intensity normalizations shows promising results for MRI-only radiation treatment planning. Our proposed polynomial method is easy to extend to more MRI sequences and saves computer storage. In addition, its accuracy can be further improved in the future by optimizing sequence parameters of the dual-contrast MRI or by using MRI data with more than two different sequences.

Data availability

All data generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Barillot, I. & Reynaud-Bougnoux, A. The use of MRI in planning radiotherapy for gynaecological tumours. *Cancer Imaging* **6**, 100 (2006).
2. Greer, P. B. *et al*. A magnetic resonance imaging-based workflow for planning radiation therapy for prostate cancer. *Medical Journal of Australia* **194**, S24 (2011).
3. Karlsson, M., Karlsson, M. G., Nyholm, T., Amies, C. & Zackrisson, B. Dedicated magnetic resonance imaging in the radiotherapy clinic. *International Journal of Radiation Oncology* Biology* Physics* **74**, 644–651 (2009).
4. Nyholm, T., Nyberg, M., Karlsson, M. G. & Karlsson, M. Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments. *Radiation Oncology* **4**, 54 (2009).
5. Edmund, J. M. & Nyholm, T. A review of substitute CT generation for MRI-only radiation therapy. *Radiation Oncology* **12**, 28 (2017).
6. Johnstone, E. *et al*. Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging–only radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* **100**, 199–217 (2018).
7. Wafa, B. & Moussaoui, A. A review on methods to estimate a CT from MRI data in the context of MRI-alone RT. *Medical Technologies Journal* **2**, 150–178 (2018).
8. Zhong, L. *et al*. Flexible prediction of CT images from MRI data through improved neighborhood anchored regression for PET attenuation correction. *IEEE J. Biomed. Health Inform*. (2019).
9. Koike, Y. *et al*. Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy. *Journal of Radiation Research* **61**, 92–103 (2020).
10. Andreasen, D., Van Leemput, K., Hansen, R. H., Andersen, J. A. & Edmund, J. M. Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain. *Medical Physics* **42**, 1596–1605 (2015).
11. Johansson, A., Karlsson, M. & Nyholm, T. CT substitute derived from MRI sequences with ultrashort echo time. *Medical physics* **38**, 2708–2714 (2011).
12. Siversson, C. *et al*. MRI only prostate radiotherapy planning using the statistical decomposition algorithm. *Medical Physics* **42**, 6090–6097 (2015).
13. Kim, J. *et al*. Implementation of a novel algorithm for generating synthetic CT images from magnetic resonance imaging data sets for prostate cancer radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* **91**, 39–47 (2015).
14. Dowling, J. A. *et al*. An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* **83**, e5–e11 (2012).
15. Sjölund, J., Forsberg, D., Andersson, M. & Knutsson, H. Generating patient specific pseudo-CT of the head from MR using atlas-based regression. *Physics in Medicine & Biology* **60**, 825 (2015).
16. Aouadi, S. *et al*. Generation of synthetic CT using multi-scale and dual-contrast patches for brain MRI-only external beam radiotherapy. *Physica Medica* **42**, 174–184 (2017).
17. Burgos, N. *et al*. Iterative framework for the joint segmentation and CT synthesis of MR images: application to MRI-only radiotherapy treatment planning. *Physics in Medicine & Biology* **62**, 4237 (2017).
18. Hofmann, M. *et al*. MRI-based attenuation correction for whole-body PET/MRI: quantitative evaluation of segmentation-and atlas-based methods. *Journal of Nuclear Medicine* **52**, 1392–1399 (2011).
19. Gudur, M. S. R. *et al*. A unifying probabilistic bayesian approach to derive electron density from MRI for radiation therapy treatment planning. *Physics in Medicine & Biology* **59**, 6595 (2014).
20. Lee, J., Carass, A., Jog, A., Zhao, C. & Prince, J. L. Multi-atlas-based CT synthesis from conventional MRI with patch-based refinement for MRI-based radiotherapy planning. In *Medical Imaging 2017: Image Processing*, vol. 10133, 101331I (International Society for Optics and Photonics, 2017).
21. Navalpakkam, B. K., Braun, H., Kuwert, T. & Quick, H. H. Magnetic resonance–based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps. *Investigative Radiology* **48**, 323–332 (2013).
22. Johansson, A., Garpebring, A., Karlsson, M., Asklund, T. & Nyholm, T. Improved quality of computed tomography substitute derived from magnetic resonance (MR) data by incorporation of spatial information–potential application for MR-only radiotherapy and attenuation correction in positron emission tomography. *Acta Oncologica* **52**, 1369–1373 (2013).
23. Edmund, J. M. *et al*. A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo times. *Physics in Medicine & Biology* **59**, 7501 (2014).
24. Han, X. MR-based synthetic CT generation using a deep convolutional neural network method. *Medical physics* **44**, 1408–1419 (2017).
25. Lei, Y. *et al*. Pseudo CT estimation using patch-based joint dictionary learning. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 5150–5153 (IEEE, 2018).
26. Leynes, A. P. *et al*. 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. *Journal of Nuclear Medicine* **59**, 852–858 (2018).
27. Maspero, M. *et al*. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Physics in Medicine & Biology* **63**, 185001 (2018).

28. Torrado-Carvajal, A. *et al*. Dixon-VIBE deep learning (DIVIDE) pseudo-CT synthesis for pelvis PET/MR attenuation correction. *Journal of Nuclear Medicine* **60**, 429–435 (2019).
29. Kapanen, M. & Tenhunen, M. T1/T2*-weighted MRI provides clinically relevant pseudo-CT density data for the pelvic bones in MRI-only based radiotherapy treatment planning. *Acta Oncologica* **52**, 612–618 (2013).
30. Hsu, S.-H., Cao, Y., Huang, K., Feng, M. & Balter, J. M. Investigation of a method for generating synthetic CT models from MRI scans of the head and neck for radiation therapy. *Physics in Medicine & Biology* **58**, 8419 (2013).
31. Korhonen, J., Kapanen, M., Keyriläinen, J., Seppälä, T. & Tenhunen, M. A dual model HU conversion from MRI intensity values within and outside of bone segment for MRI-based radiotherapy treatment planning of prostate cancer. *Medical Physics* **41** (2014).
32. Su, K.-H. *et al*. Generation of brain pseudo-CTs using an undersampled, single-acquisition UTE-mDixon pulse sequence and unsupervised clustering. *Medical physics* **42**, 4974–4986 (2015).
33. Zheng, W. *et al*. Magnetic resonance–based automatic air segmentation for generation of synthetic computed tomography scans in the head region. *International Journal of Radiation Oncology* Biology* Physics* **93**, 497–506 (2015).
34. Speier, C. *et al*. Advanced multimodal methods for cranial pseudo-CT generation validated by IMRT and VMAT radiation therapy plans. *International Journal of Radiation Oncology* Biology* Physics* **102**, 792–800 (2018).
35. Pileggi, G. *et al*. Proton range shift analysis on brain pseudo-CT generated from T1 and T2 MR. *Acta Oncologica* **57**, 1521–1531 (2018).
36. Lee, Y. K. *et al*. Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. *Radiotherapy and Oncology* **66**, 203–216 (2003).
37. Jonsson, J. H., Karlsson, M. G., Karlsson, M. & Nyholm, T. Treatment planning using MRI data: an analysis of the dose calculation accuracy for different treatment regions. *Radiation Oncology* **5**, 62 (2010).
38. Pecorelli, S., Benedet, J. L., Creasman, W. T. & Shepherd, J. H. FIGO staging of gynecologic cancer. *International Journal of Gynecology & Obstetrics* **65**, 243–249 (1999).
39. Nyúl, L. G. & Udupa, J. K. On standardizing the mr image intensity scale. *Magnetic Resonance in Medicine* **42**, 1072–1081 (1999).
40. Rank, C. M. *et al*. MRI-based simulation of treatment plans for ion radiotherapy in the brain region. *Radiotherapy and Oncology* **109**, 414–418 (2013).
41. Keereman, V. *et al*. MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences. *Journal of Nuclear Medicine* **51**, 812 (2010).
42. Rank, C. M. *et al*. MRI-based treatment plan simulation and adaptation for ion radiotherapy using a classification-based approach. *Radiation Oncology* **8**, 51 (2013).
43. Berker, Y. *et al*. MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/Dixon MRI sequence. *Journal of Nuclear Medicine* **53**, 796 (2012).
44. van Herk, M. *et al*. Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. *International Journal of Radiation Oncology* Biology* Physics* **33**, 1311–1320 (1995).
45. Jhingran, A., Salehpour, M., Sam, M., Levy, L. & Eifel, P. J. Vaginal motion and bladder and rectal volumes during pelvic intensity-modulated radiation therapy after hysterectomy. *International Journal of Radiation Oncology* Biology* Physics* **82**, 256–262 (2012).
46. Jadon, R. *et al*. A systematic review of organ motion and image-guided strategies in external beam radiotherapy for cervical cancer. *Clinical Oncology* **26**, 185–196 (2014).
47. Akbarzadeh, A. *et al*. Evaluation of whole-body MR to CT deformable image registration. *Journal of Applied Clinical Medical Physics* **14**, 238–253 (2013).

Author contributions

Competing interests

Additional information