Feasibility of synthetic computed tomography generated with an adversarial network for multi-sequence magnetic resonance-based brain radiotherapy

Yuhei Koike¹, Yuichi Akino²,*, Iori Sumida¹, Hiroya Shiomi¹,3, Hirokazu Mizuno⁴, Masashi Yagi⁵, Fumiaki Isohashi¹, Yuji Seo¹, Osamu Suzuki⁵, Kazuhiko Ogawa¹

¹Department of Radiation Oncology, Osaka University Graduate School of Medicine, Osaka, Japan
²Oncology Center, Osaka University Hospital, Osaka, Japan
³Miyakojima IGRTC Clinic, Osaka, Japan
⁴Department of Medical Physics and Engineering, Osaka University Graduate School of Medicine, Osaka, Japan
⁵Department of Carbon Ion Radiotherapy, Osaka University Graduate School of Medicine, Osaka, Japan

*Corresponding author. Oncology Center, Osaka University Hospital, 2-2 (D10) Yamadaoka, Suita, Osaka, 565-0871, Japan. Tel: +81-6-6879-3482, Fax: +81-6-6879-3489; Email: akino@radonc.med.osaka-u.ac.jp

(Received 30 May 2019; revised 15 July 2019; editorial decision 12 August 2019)

ABSTRACT

The aim of this work is to generate synthetic computed tomography (sCT) images from multi-sequence magnetic resonance (MR) images using an adversarial network and to assess the feasibility of sCT-based treatment planning for brain radiotherapy. Datasets for 15 patients with glioblastoma were selected and 580 pairs of CT and MR images were used. T1-weighted, T2-weighted and fluid-attenuated inversion recovery MR sequences were combined to create a three-channel image as input data. A conditional generative adversarial network (cGAN) was trained using image patches. The image quality was evaluated using voxel-wise mean absolute errors (MAEs) of the CT number. For the dosimetric evaluation, 3D conformal radiotherapy (3D-CRT) and volumetric modulated arc therapy (VMAT) plans were generated using the original CT set and recalculated using the sCT images. The isocenter dose and dose–volume parameters were compared for 3D-CRT plans. The differences in the dose received by 2% of the volume (D2%), D50% and D98% relative to the prescribed dose were <1.0%. The overall equivalent path length was shorter than that for real CT by 0.6 ± 1.9 mm. A treatment planning study using generated sCT detected only small, clinically negligible differences. These findings demonstrated the feasibility of generating sCT images for MR-only radiotherapy from multi-sequence MR images using cGAN.

Keywords: synthetic computed tomography; deep learning; generative adversarial network; dose calculation; treatment planning

INTRODUCTION

Magnetic resonance imaging (MRI) is used for the delineation of the organs in radiotherapy because its soft-tissue contrast is superior to that of computed tomography (CT). It has been recommended that MRI should be used for contouring in brain radiotherapy, especially for glioblastoma tumors [1–3]. However, the MR image does not correlate with electron density, and its geometrical accuracy is inferior to that of CT, so MR images are usually registered to reference CT images and the CT images are used for the dose calculation. However, the low contrast of soft tissues on CT images can result in considerable inter-observer variability in volume delineation [4–6], and image registration may lead to additional uncertainty with contouring [7,8]. These issues can...
be reduced in treatment planning by using MRI only. As additional benefits, the entire planning workload is reduced, there is no need for X-ray irradiation, and direct delineation on MR images allows the margin size to be minimized, thereby reducing toxicity to healthy organs.

The challenge for MR-based radiotherapy planning is to associate the CT number and electron density information with the MR image for dose calculation. To address this issue, several approaches to generate synthetic CT (sCT) images from MR images have been proposed. These include bulk density assignment [9–13], the atlas-based technique [14–17], and the voxel-based method [18–21].

An emerging approach to convert MR to CT images is the application of deep learning using a convolutional neural network (CNN) and a generative adversarial network (GAN) [22]. Han proposed the CNN-based method to generate sCT from T1-weighted (T1w) MRI sequences using U-Net architecture and reported that the CNN-based method performed better than the atlas-based method [23]. However, a limitation of CNN is that it may lead to blurry results due to misalignment between MR and CT [24]. GAN, which is trained by two competing networks, has been applied to MR-to-CT translation. Using adversarial loss, the GAN model can generate high-quality sCT images with less blurry results [24–26]. Emami et al. proposed a method of generating sCT from T1w MR using GAN and compared it with the CNN-based method using the same patient cohort [27]. They reported that the GAN-based method preserved anatomical details and improved the quality of sCT images as compared to the CNN approach. However, the primary aim of those studies was to generate sCT images using deep learning, although the use of sCT for radiotherapy has received little attention. Several recent studies [28–32] have investigated sCT-based treatment planning using the CNN or GAN models; however, most of them used single-sequence MR images of the pelvic area. For brain radiotherapy, Dinkila et al. generated sCT images from T1w MR images using a 2.5-dimensional convolutional network, which provided acceptable dosimetric results [33]. However, the feasibility of an adversarial network-based sCT generation from multi-sequence MR images with dosimetric evaluation for brain radiotherapy remains uncertain.

Here, we describe a method to generate sCT images from T1w, T2-weighted (T2w) and fluid-attenuated inversion recovery (FLAIR) MR images using a conditional GAN (cGAN) model, and investigate its feasibility for dose calculations in brain radiotherapy.

**MATERIALS AND METHODS**

**Patient data and target volume**

Datasets that included CT and MR (T1w, T2w and FLAIR) images were retrieved from a freely available database provided by The Cancer Imaging Archive (TCIA) [34]. In total, 580 pairs of CT and MR images of 15 patients were collected. All CT images were acquired with GE Medical Systems scanners with a median resolution of 0.49 × 0.49 × 2.5 mm³ (range, 0.45–0.56 × 0.45–0.56 × 1.25–5.0 mm³). The matrix size was 512 × 512, with a median field-of-view (FOV) of 25 cm (range, 23–28 cm). For the MR images, the median voxel size was 0.45 × 0.45 × 5.0 mm³ (range, 0.43–0.94 × 0.43–0.94 × 0.90–5.0 mm³). The sequences included images acquired with Philips Medical Systems, GE Medical Systems and Siemens scanners. The median FOV was 23 cm for all sequences.

The gross tumor volumes (GTV) were contoured using registered MR images. GTV1 was identified from the T1w image and GTV2 from the T2w or FLAIR image by a single observer, which was then reviewed by an experienced radiation oncologist. The clinical target volumes (CTV) were defined as the GTVs plus a 1.5-cm margin within the brain. The planning target volumes (PTV) were defined as the CTVs plus a 0.5-cm margin. The prescribed doses to PTV1 and PTV2 were 60 and 40 Gy, respectively, delivered in 30 fractions.

**Image preprocessing**

The field inhomogeneities of MRI were corrected using the N4 bias field correction algorithm [35]. The coordinates were aligned between the CT and MR images by applying deformable image registration using the SmartAdapt application of the Eclipse version 15.1 treatment planning system (TPS) (Varian Medical Systems Inc., Palo Alto, CA). The pixel intensities of the MR images were normalized to values in the range 0–4095 using the method proposed by Nyul and Udupa [36]. To separate the patient body from the background, we used the binary mask by applying the Gaussian filter, the Otsu threshold method [37] and the morphology operation. The background voxels of the CT images were filled with −1000 and those in the MR images with 0. So that the MR image of all sequences could be input to the network as a single image, the T1w, T2w and FLAIR images of the same slice were combined into three channels in a color image.

**Network architecture**

A cGAN-based network [38,39] was used to generate sCT images from multi-sequence MR images. Fig. 1 illustrates the workflow of our proposed method for generating sCT images from multiple MR sequences. The architecture of the generator and discriminator is shown in Fig. 1a and b, respectively. For the generator, we used the U-Net architecture [40] proposed by Fu et al. [41], which was modified to handle color images and to fit to our datasets. The discriminator was based on the PatchGAN discriminator used in the pix2pix model.

**Training details**

Patch-based training was performed because of the limited number of medical images and the limited memory of the graphics processing unit (GPU). Patches of 128 × 128 were extracted from the MR and corresponding CT images for use as training data. The weights of the generator and discriminator were initialized by He initialization [42] and the biases were initialized to zero. As proposed by Isola et al. [39], the loss function by L1 norm of the generator was the mean absolute error (MAE) between the real CT (rCT) image and the sCT image as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |rCT_i - sCT_i|,$$

where n is the number of voxels, and rCT_i and sCT_i represent the CT numbers in the i-th voxels of the rCT and sCT images, respectively.

The patch size of the discriminator was 32 × 32. The network was evaluated with the 5-fold cross-validation method, where 15 datasets
were randomly divided into five groups. For each fold, the model was trained using four of the datasets and the sCT was predicted from the fifth dataset using the trained model. This process was repeated for the five groups. In the training phase, the training data were divided for training and validation at a ratio of 9:1. Data augmentation was performed using random horizontal flips, image shifts, zooms and rotations. Training was performed using an NVIDIA GTX 1080 Ti (11 GB) GPU with a mini-batch of 32.

Synthetic CT generation
sCT images were generated from full-sized MR images (512 × 512) using the trained generator with the GPU calculation, even though training was performed using patch images (128 × 128) as shown in Fig. 1c. The datasets were not created for use in radiotherapy, so the full MR volume corresponding to the registered CT was not obtained. As a result, the generated sCT image lacked the top or bottom slices, as compared to the rCT volume. This issue was compensated by filling the missing slices with the corresponding slices of the rCT and replacing the CT numbers with values calculated from the correlation between the predicted and original voxel values. In particular, we replaced the original CT number in the missing regions with the median of all the predicted hounsfield unit (HU) values, which were converted from all voxels with the corresponding CT number. For the purposes of comparison, sCT images were also generated from only a single T1w sequence. Thus, sCT$_{\text{single}}$ was defined as sCT generated from a single T1w MR image and sCT$_{\text{multi}}$ from multiple MR sequences (T1w, T2w and FLAIR). The networks for sCT$_{\text{single}}$ and sCT$_{\text{multi}}$ were trained separately.

Evaluation of image quality
Image quality was evaluated by calculating the voxel-wise MAE between the sCT and rCT images with respect to the CT number. The MAE of the CT number was evaluated for the entire body region, for soft tissue and for bone regions. The soft tissue and bone region were extracted by setting CT number thresholds of −100 to 150 HU and ≥250 HU, respectively.

Treatment planning
The sCT-based dose calculation was evaluated by applying 3D conformal radiotherapy (3D-CRT) and volumetric modulated arc therapy (VMAT) techniques. The treatment plans were calculated with the Eclipse TPS. In the sCT plans, the body was delineated using the sCT images and the other structures were copied from those of the rCT.

For 3D-CRT evaluation, boost plans up to 40–60 Gy were compared to clarify the impact of using the sCT for the dose calculation. A dose of 20 Gy in 10 fractions was prescribed to the isocenter to cover the PTV1 at a dose of 95% using the anisotropic analytical algorithm (version 13.7.14) of the Eclipse TPS.

The simultaneous integrated boost technique was used for VMAT evaluation. Single-arc VMAT plans were created with a collimator angle of 30°. The VMAT plans were calculated using the Acuros XB algorithm (version 15.1.51) with a 6-MV photon beam and were optimized by applying the constraints reported by the Radiation Therapy Oncology Group (RTOG 0825) [43]. The VMAT plans were normalized to 95% of the prescribed dose to PTV1. An sCT-based plan was created by recalculating the rCT plan with the sCT images, with the same multileaf collimator sequence and monitor unit as calculated in the rCT. A calculation grid size of 2.5 mm was used in both algorithms.

Dose differences were evaluated with respect to the isocenter dose for the 3D-CRT and dose–volume parameters for the VMAT. In particular, we compared the dose received by 2% of the volume (D$_{2\%}$), D$_{95\%}$, and D$_{98\%}$, the volume receiving at least 95% of the prescribed dose (V$_{95\%}$) for PTV1, the mean dose for the body, and the maximum dose (D$_{\text{max}}$) for the organs at risk.
Synthetic CT for MR-only radiotherapy

The impact of sCT on the equivalent path length (EPL) was also evaluated. Assuming a single-arc VMAT, EPLs were compared between the rCT and sCT for 180 beams at angles from 181°–179° at intervals of 2°.

The dose distribution agreement with the original plan was compared by the 2D gamma evaluation [44] with the criteria of 3%/3 mm, 2%/2 mm and 1%/1 mm for the dose difference and the distance.

**Statistical analysis**

All statistical analyses were performed using JMP Software (SAS Institute Inc., Cary, NC). Depending on the normality of the data distributions, dose differences between the rCT and sCT plans were compared using either the paired t-test or the Wilcoxon signed-rank test. Statistical significance was defined as \( P < 0.05 \).

**RESULTS**

**Image quality evaluation**

Each training session of 500 epochs took \( \sim 2.5 \) days. For testing, sCT images were predicted using the trained network in 17.4 ± 5.4 s per patient.

Fig. 2 shows difference maps between the sCT and rCT images for representative patient datasets. The MAE values of sCT\(_\text{single}\) and sCT\(_\text{multi}\) for the whole body are shown in Fig. 3. The MAE values of sCT\(_\text{multi}\) were lower than those of sCT\(_\text{single}\) in all patients. The mean ± standard deviation (SD) MAE values for the CT numbers for the whole body, soft tissue and bone regions are summarized in Table 1. The overall MAE values of sCT\(_\text{multi}\) were significantly smaller than sCT\(_\text{single}\). The proposed approach to use multiple MR sequences was found to generate better quality sCT images, similar to rCT rather than sCT\(_\text{single}\).

**Dosimetric evaluation**

In the 3D-CRT plans, the mean ± SD isocenter dose for the sCT\(_\text{multi}\) plans was 2003.3 ± 8.0 cGy, which was not significantly different from the reference dose of 2000.0 cGy in the rCT plans (\( P = 0.14 \)). Representative dose distributions of sCT\(_\text{multi}\) for the 3D-CRT and VMAT plans are shown in Figs. 5 and 6.

Table 2 compares the dose–volume parameters between rCT and sCT\(_\text{multi}\) for the VMAT plans. The dose parameters tended to be higher in the sCT\(_\text{multi}\) plans. The target coverage (\( V_{95\%} \)) of the sCT\(_\text{multi}\) plan was preserved, with no significant difference. \( D_{2\%} \), \( D_{50\%} \) and \( D_{98\%} \) for PTV1, and \( D_{\text{max}} \) for the brainstem, optic chiasm and left optic nerve were significantly higher with the sCT\(_\text{multi}\) than the rCT. Dose differences relative to the prescribed dose were within 1.0% for all parameters.

**Equivalent path length**

The mean ± SD EPL error for the sCT\(_\text{multi}\) was −0.6 ± 1.9 mm (range, −12.0 to 8.0 mm) compared with that for the rCT. An example of the difference in EPL is illustrated in Fig. 7a and b. Fig. 7c shows the variation in the difference of EPL (sCT—rCT) for each patient.
Fig. 3. A comparison of MAE within the body contour between sCT single and sCT multi of all 15 patients.

Table 1. The MAE of CT number between the sCT single and sCT multi images

| Region                        | Mean ± SD (HU) | P-value |
|-------------------------------|----------------|---------|
| Body                          | 120.1 ± 20.4   | 108.1 ± 24.0 | <0.001 |
| Soft tissue region (-100 HU < | 46.3 ± 9.3     | 38.9 ± 10.7 | <0.001 |
| (−150 HU) < 150 HU)           |                |         |
| Bone region (≥250 HU)         | 399.4 ± 51.8   | 366.2 ± 62.0 | <0.001 |

Fig. 4. The difference in CT numbers between the synthetic and real CT images (sCT and rCT, respectively). The median, 25th and 75th percentiles of absolute differences (sCT—rCT) are shown with a histogram of the distribution of the voxels inside the patient body contour (bin size, 20). The horizontal axis indicates the rCT CT numbers and the upper vertical axis indicates the absolute difference in CT number. The inset expands the region between −100 and 100 HU, which included more than 70% of all the voxels. In this region, the median differences were within ±50 HU.
Fig. 5. Dose distributions of sCT\textsubscript{multi} of representative 3D-CRT plans and associated difference maps showing the dose difference relative to the prescribed dose of 20 Gy.

EPLs with errors between −5.0 and 5.0 mm accounted for >98% of all beams for whole datasets. A large error was found in the metal placement (Fig. 7a).

**Dosimetric comparison of sCT\textsubscript{single} and sCT\textsubscript{multi}**

The dose difference maps of sCT\textsubscript{single} and sCT\textsubscript{multi} of a representative patient are shown in Fig. 8. Dose errors around the left temporal bone were improved for the sCT\textsubscript{multi}, as compared with sCT\textsubscript{single}.

The gamma evaluation results are summarized in Table 3. Although the pass rates for both sCT\textsubscript{single} and sCT\textsubscript{multi} were >94% even with the 1%/1 mm criterion, the pass rates of sCT\textsubscript{multi} were significantly higher than those of sCT\textsubscript{single} with all criteria. For EPL evaluation, the mean EPLs of sCT\textsubscript{single} and sCT\textsubscript{multi} were −0.5 ± 2.1 mm (range, −16.0 to 10.0 mm) and −0.6 ± 1.9 mm (range, −12.0 to 8.0 mm), respectively. Although the effects of the MR sequences on the dose-volume histogram (DVH) parameters of the target and organs at risk were <1% (data not shown), multi-sequence MR improved the accuracy of sCT generations in terms of CT values and dose calculations, particularly in the bone region.

**DISCUSSION**

In this study, sCT images were generated from multi-channel MR images using a cGAN model. The input was multi-sequence T1w,
Fig. 6. Dose distributions of sCT\textsubscript{multi} of representative VMAT plans and associated difference maps showing the dose difference relative to the prescribed dose of 60 Gy. Each row shows the same slices as in Fig. 5.

T2w and FLAIR MR images combined into a single-input image with three channels. The images were used to simulate sCT-based treatment planning and the resulting dose distributions were compared with the ground truth rCT doses. The results showed that our proposed approach was able to generate sCT images that provided dose distributions similar to those of the reference images. To the best of our knowledge, this is the first report to describe the use of multi-sequence brain MR images to create sCT images and then to investigate the impact of using the sCT images for the dose calculation. Maspero et al. generated sCT images using three MRI sequences, including in-phase, fat and water MR images, of the pelvic region [32]. Their sampling size for sCT generation was 256 × 256 of 8-bit CT images, whereas our method used 512 × 512 samplings of 16-bit images, providing better image quality.

sCT for the head region has been reported to have higher MAEs (80–200 HU) when compared with those for the pelvic region, in most cases because of the complex structures and the high ratios of air and bone [28, 45]. Our MAE results were superior or comparable to those of the atlas-based [17] and classification-based results [18, 20, 46]. Although Andreasen et al. achieved a MAE of 85 HU using the patch-based method, it took 15 h to generate sCTs from the T1w images [47]. Uh et al. used a multi-atlas approach to generate sCTs from MRI, with
Table 2. Comparison of dose–volume parameters between the sCT_{multi} and rCT plans. The data show the mean dose–volume parameters for each structure and the differences relative to the prescribed dose of 60 Gy. $D_{x\%}$ indicates the dose that covered $x\%$ of the volume. $V_{x\%}$ indicates the volume that received $x\%$ of the prescribed dose.

| Structures       | Parameters | Mean ± SD | Difference (%) | P-value |
|------------------|------------|-----------|----------------|---------|
|                  | rCT        | sCT_{multi} |                |         |
| PTV1             | $D_{2\%}$ (Gy) | 64.0 ± 0.5 | 64.4 ± 0.6     | 0.68    | 0.008 |
|                  | $D_{50\%}$ (Gy) | 61.9 ± 0.3 | 62.1 ± 0.4     | 0.34    | 0.011 |
|                  | $D_{98\%}$ (Gy) | 59.4 ± 0.2 | 59.6 ± 0.3     | 0.30    | 0.010 |
|                  | $V_{95\%}$ (%) | 100.0 ± 0.1 | 100.0 ± 0.1    | 0.01    | 0.064 |
| Body             | $D_{max}$ (Gy) | 19.7 ± 6.2 | 19.8 ± 6.3     | 0.09    | 0.14  |
| Brainstem        | $D_{max}$ (Gy) | 46.4 ± 16.6 | 46.6 ± 16.7    | 0.35    | 0.018 |
| Optic chiasm     | $D_{max}$ (Gy) | 33.0 ± 21.8 | 33.1 ± 21.8    | 0.20    | 0.004 |
| Optic nerve (left) | $D_{max}$ (Gy) | 22.0 ± 18.8 | 22.1 ± 18.9    | 0.23    | 0.027 |
| Optic nerve (right) | $D_{max}$ (Gy) | 23.1 ± 20.0 | 23.2 ± 19.9    | 0.04    | 0.59  |
| Eye (left)       | $D_{max}$ (Gy) | 20.0 ± 15.0 | 20.0 ± 15.1    | 0.10    | 0.42  |
| Eye (right)      | $D_{max}$ (Gy) | 21.2 ± 15.8 | 21.3 ± 16.1    | 0.27    | 0.068 |
| Lens (left)      | $D_{max}$ (Gy) | 6.1 ± 4.1  | 6.1 ± 4.1      | 0.00    | 1     |
| Lens (right)     | $D_{max}$ (Gy) | 6.0 ± 4.5  | 6.0 ± 4.5      | 0.00    | 1     |

Fig. 7. Comparison of the EPLs between the real and synthetic CT (rCT and sCT, respectively) plans, where the latter was generated from multiple MR sequences plans. (a) A representative example of the EPL (for Patient #1). The dashed line indicates the physical depth of rCT (in cm) and the solid lines indicate the EPL (in cm) for rCT and sCT_{multi}. The arrow indicates the implantation of metal. (b) Differences in EPL between the sCT_{multi} and rCT plans (sCT_{multi}—rCT). (c) Boxplot of the EPL differences for all 15 patients. The diamonds denote the mean values and the gray area indicates the region with differences between −0.5 and 0.5 cm.
Table 3. Gamma pass rates (%) of the sCT<sub>single</sub> and sCT<sub>multi</sub>

| Criterion | Mean ± SD (range) | P-value |
|-----------|-------------------|---------|
|           | sCT<sub>single</sub> | sCT<sub>multi</sub> |         |
| 3%/3 mm   | 99.7 ± 0.5 (98.3–100.0) | 99.8 ± 0.3 (99.1–100.0) | 0.045   |
| 2%/2 mm   | 98.9 ± 1.2 (96.0–99.8) | 99.2 ± 1.0 (97.1–100.0) | 0.004   |
| 1%/1 mm   | 94.2 ± 4.9 (83.4–98.6) | 95.3 ± 4.7 (85.0–99.0) | 0.002   |

In contrast, the mean error for the EPLs was $-0.6 \pm 1.9$ mm for sCT<sub>multi</sub>. The impact on dose calculations of such a small difference in EPLs would be negligible. The comparison of 3D-CRT plans showed that the difference in isocenter dose between the rCT and sCT<sub>multi</sub> images was within 0.2%, which might be attributed to the small proportion of bone relative to the whole beam path for a brain site in spite of a large MAE for the bone regions. Dinkla et al. also reported small differences in the EPLs and dose calculations for
the brain, although the input MR sequence and network structure in their study differed from those used in our study [33]. Paradis et al. investigated sCT-based brain VMAT treatment planning using the voxel-based method with fuzzy c-means clustering, achieving mean differences of $D_{5\%,5\%}$ and $D_{95\%,1\%}$ for PTV of <1.0%, with the largest difference of 0.6 Gy [21]. In our proposed method, the percentage error relative to the prescribed dose was <1.0% and the largest error in $D_{95\%}$ for PTV1 was an absolute difference of 0.4 Gy; these differences may be acceptable from a clinical viewpoint. However, deviations of CT numbers in bone may lead to large dosimetric errors in regions where the proportion of bony anatomy is higher than that of soft tissue.

The MAE values of the $sCT_{\text{multi}}$ images were significantly lower than those of the $sCT_{\text{angle}}$ images, resulting in larger dose errors around the bone regions (Fig. 8). Consequently, the pass rates for $sCT_{\text{multi}}$ were improved compared with those for $sCT_{\text{angle}}$. These results indicate that the multiple MR sequences generated potentially similar dose distributions as those of the original CT image. Although our proposed method using multi-sequence MR will require additional time and cost in clinical application, the use of multiple sequences provided better results than those provided by a single T1w sequence in the same patient population.

There were some limitations to this study. First, the patient data used in this study were derived from The Cancer Imaging Archive (TCIA) online database. The patients were not actually treated with the treatment protocol of this study. Analysis of the data of patients who received radiotherapy will provide more realistic results. Second, the sCT used for treatment planning included slices that were not predicted using the trained network, in particular at the top and/or bottom of the volume. As described in the Materials and Methods section, the datasets used in this study were not for radiotherapy, so the full MR volume corresponding to CT was not available. The compensation process may have affected the dose–volume histogram data, although there was no effect on the isocenter dose or EPL results. Third, there were additional uncertainties associated with the deformable image registration to match the alignment between the CT and MR images, which was applied three times for each sequence. As reported previously [49, 50], errors between the three sequences may therefore have affected the MAE results. These errors will have less of an impact by applying rigid image registration, rather than deformable image registration, for patients scanned at one time for all sequences.

In conclusion, we developed a technique to generate sCT images from multi-sequence brain MR images using an adversarial network. The performance of the model was evaluated by comparing the image quality and the treatment planning with those of the original CT images. The use of multiple MR sequences for sCT generation using cGAN provided better image quality and dose distribution results compared with those from only a single T1w sequence. The CT number of the generated sCT images showed good agreement with the original images, but not in the bone regions. Impacts on the dose calculations were within 1%. These findings demonstrate the feasibility and utility of sCT-based treatment planning and support the use of deep learning for MR-only radiotherapy.

ACKNOWLEDGMENTS

The results shown here are in whole or in part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov/. Data used in this publication were generated by the National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC). This work was presented at the American Association of Physicists in Medicine (AAPM) 61st Annual Meeting & Exhibition in San Antonio, TX (2019).

FUNDING

This work was supported by JSPS KAKENHI Grant number 17 K15802.

CONFLICT OF INTEREST

None declared.

REFERENCES

1. Fiorentino A, Caivano R, Pedicini P et al. Clinical target volume definition for glioblastoma radiotherapy planning: Magnetic resonance imaging and computed tomography. Clin Transl Oncol. 2013;15:754–8.
2. Niyazi M, Brada M, Chalmers AJ et al. ESTRO-ACROP guideline ‘target delineation of glioblastomas’. Radiother Oncol. 2016;118:35–42.
3. Zhao F, Li M, Kong L et al. Delineation of radiation therapy target volumes for patients with postoperative glioblastoma: A review. Onco Targets Ther. 2016;9:3197–204.
4. Weltens C, Menten J, Feron M et al. Interobserver variations in gross tumor volume delineation of brain tumors on computed tomography and impact of magnetic resonance imaging. Radiother Oncol. 2001;60:49–59.
5. Coles CE, Hoole ACF, Harden SV et al. Quantitative assessment of inter-clinician variability of target volume delineation for medulloblastoma: Quality assurance for the SIOP PNET 4 trial protocol. Radiother Oncol. 2003;69:189–94.
6. Cattaneo GM, Reni M, Rizzo G et al. Target delineation in postoperative radiotherapy of brain gliomas: Interobserver variability and impact of image registration of MR (pre-operative) images on treatment planning CT scans. Radiother Oncol. 2005;75:217–23.
7. Ulin K, Urie MM, Cherlow JM. Results of a multi-institutional benchmark test for cranial CT/MR image registration. Int J Radiat Oncol Biol Phys. 2010;77:1584–9.
8. Nyholm T, Nyberg M, Karlsson MG et al. Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments. Radiat Oncol. 2009;4:54.
9. Lee YK, Bollet M, Charles-Edwards G et al. Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. Radiother Oncol. 2003;66:203–16.
10. Jonsson JH, Karlsson MG, Karlsson M et al. Treatment planning using MRI data: An analysis of the dose calculation
accuracy for different treatment regions. Radiat Oncol. 2010; 5:62.

11. Lambert J, Greer PB, Menk F et al. MRI-guided prostate radiation therapy planning: Investigation of dosimetric accuracy of MRI-based dose planning. Radiother Oncol. 2011;98: 330–4.

12. Kim J, Garbarino K, Schultz L et al. Dosimetric evaluation of synthetic CT relative to bulk density assignment-based magnetic resonance-only approaches for prostate radiotherapy. Radiat Oncol. 2015;10:239.

13. Prior P, Chen X, Gore E et al. Technical note: Is bulk electron density assignment appropriate for MRI-only based treatment planning for lung cancer? Med Phys. 2017;44:3437–43.

14. Dowling JA, Lambert J, Parker J et al. An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy. Int J Radiat Oncol Biol Phys. 2012;83: e5–11.

15. Dowling JA, Sun J, Pichler P et al. 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. 2015;93:1144–53.

16. Guerreiro F, Burgos N, Dunlop A et al. Evaluation of a multi-atlas CT synthesis approach for MRI-only radiotherapy treatment planning. Phys Med. 2017;35:7–17.

17. Sjolund J, Forsberg D, Andersson M et al. Generating patient specific pseudo-CT of the head from MR using atlas-based regression. Phys Med Biol. 2015;60:825–39.

18. Johansson A, Karlsson M, Nyholm T. CT substitute derived from MRI sequences with ultrashort echo time. Med Phys. 2011;38:2708–14.

19. Navalpakkam BK, Braun H, Kuwert T et al. Magnetic resonance-based attenuation correction for PET/MR hybrid imaging using continuous valued attenuation maps. Investigative radiology. 2015;48:323–32.

20. Edmund JM, Kjer HM, Van Leemput K et al. A voxel-based investigation for MRI-only radiotherapy of the brain using ultra short echo time. Phys Med Biol. 2014;59:7501–19.

21. Paradis E, Cao Y, Lawrence TS et al. Assessing the Dosimetric accuracy of magnetic resonance-generated synthetic CT images for focal brain VMAT radiation therapy. Int J Radiat Oncol Biol Phys. 2015;93:1154–61.

22. Goodfellow IJ, Pouget-Abadie J, Mirza M et al. Generative adversarial nets. In: Advances in Neural Information Processing Systems, Montreal, QC Canada 2014;2672–80.

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

24. Wolterink JM, Dinkla AM, Savenije MHF et al. Deep MR to CT synthesis using unpaired data. arXiv preprint 2017; arXiv:1708.01155.

25. Lei Y, Harms J, Wang T, et al. MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. Med Phys. 2019.

26. Nie D, Trullo R, Lian J, et al. Medical image synthesis with context-aware generative adversarial networks. In: Medical Image Computing and Computer Assisted Intervention (MICCAI), Cham, 2017, 417–25. Springer.

27. Emami H, Dong M, Nejad-Davarani SP et al. Generating synthetic CTs from magnetic resonance images using generative adversarial networks. Med Phys. 2018;45:3627–36.

28. Andreassen D, Van Leemput K, Edmund JM. A patch-based pseudo-CT approach for MRI-only radiotherapy in the pelvis. Med Phys. 2016;43:4742–52.

29. Arabi H, Dowling JA, Burgos N et al. Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI-guided radiation planning in the pelvic region. Med Phys. 2018;45:5218–33.

30. Chen S, Qin A, Zhou D et al. Technical note: U-net-generated synthetic CT images for magnetic resonance imaging-only prostate intensity-modulated radiotherapy treatment planning. Med Phys. 2018;45:5659–65.

31. Hansen DC, Landry G, Kamp F et al. ScatterNet: A convolutional neural network for cone-beam CT intensity correction. Med Phys. 2018;45:4916–26.

32. Maspero M, Savenije MHF, Dinkla AM et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. Phys Med Biol. 2018;63:185001.

33. Dinkla AM, Wolterink JM, Maspero M et al. MR-only brain radiation therapy: Dosimetric evaluation of synthetic CTs generated by a dilated convolutional neural network. Int J Radiat Oncol Biol Phys. 2018;102:801–12.

34. Clark K, Vendt B, Smith K et al. The cancer imaging archive (TCIA): Maintaining and operating a public information repository. J Digit Imaging. 2013;26:1045–57.

35. Tustison NJ, Avants BB, Cook PA et al. N4ITK: Improved N3 bias correction. IEEE Trans Med Imaging. 2010;29:1310–20.

36. Nyul LG, Udupa JK. On standardizing the MR image intensity scale. Magn Reson Med. 1999;42:1072–81.

37. Otsu N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics. 1979;9: 62–6.

38. Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint 2014; arXiv:1411.1784.

39. Isola P, Zhu J-Y, Zhou T et al. Image-to-image translation with conditional adversarial networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017; 5967–76.

40. Ronneberger, O, Fischer, P, Brox, T. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer Assisted Intervention (MICCAI), 2015. Springer.

41. Fu, J, Yang, Y, Singhrao, K, et al. Male pelvic synthetic CT generation from T1-weighted MRI using 2D and 3D convolutional neural networks. arXiv preprint 2018; arXiv:1803.00131.

42. He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: IEEE International Conference on Computer Vision (ICCV), 2015,1026–34.

43. Gilbert MR, Dignam J, Won M et al. RTOG 0825: Phase III double-blind placebo-controlled trial evaluating bevacizumab
(Bev) in patients (Pts) with newly diagnosed glioblastoma (GBM). *Journal of Clinical Oncology* 2013;31:1.

44. Low DA, Harms WB, Mutic S et al. A technique for the quantitative evaluation of dose distributions. *Med Phys.* 1998;25:656–61.

45. Johnstone E, Wyatt JJ, Henry AM et al. Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging-only radiation therapy. *Int J Radiat Oncol Biol Phys.* 2018;100:199–217.

46. Zheng W, Kim JP, Kadbi M et al. Magnetic resonance-based automatic air segmentation for generation of synthetic computed tomography scans in the head region. *Int J Radiat Oncol Biol Phys.* 2015;93:497–506.

47. Andreasen D, Van Leemput K, Hansen RH et al. Patch-based generation of a pseudo CT from conventional MRI sequences for MRI-only radiotherapy of the brain. *Med Phys.* 2015;42:1596–605.

48. Uh J, Merchant TE, Li Y et al. MRI-based treatment planning with pseudo CT generated through atlas registration. *Med Phys.* 2014;41:051711.

49. Brock KK. Deformable registration accuracy C. results of a multi-institution deformable registration accuracy study (MIDRAS). *Int J Radiat Oncol Biol Phys.* 2010;76:583–96.

50. Murphy MJ, Salguero FJ, Siebers JV et al. A method to estimate the effect of deformable image registration uncertainties on daily dose mapping. *Med Phys.* 2012;39:573–80.