Article

Synthetic CT in Carbon Ion Radiotherapy of the Abdominal Site

Giovanni Parrella [1,*], **Alessandro Vai** [2], **Anestis Nakas** [1], **Noemi Garau** [1], **Giorgia Meschini** [1], **Francesca Camagni** [1], **Silvia Molinelli** [2], **Amelia Barcellini** [3,4], **Andrea Pella** [5], **Mario Ciocca** [2], **Viviana Vitolo** [3], **Ester Orlandi** [6], **Chiara Paganelli** [1] and **Guido Baroni** [1]

[1] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
[2] Medical Physics Unit, National Center of Oncological Hadrontherapy (CNAO), Strada Campeggi, 53, 27100 Pavia, Italy
[3] Radiotherapy Unit, National Center of Oncological Hadrontherapy (CNAO), Strada Campeggi, 53, 27100 Pavia, Italy
[4] Department of Internal Medicine and Medical Therapy, University of Pavia, 27100 Pavia, Italy
[5] Bioengineering Unit, National Center of Oncological Hadrontherapy (CNAO), Strada Campeggi, 53, 27100 Pavia, Italy
[6] Clinical Unit, National Center of Oncological Hadrontherapy (CNAO), Strada Campeggi, 53, 27100 Pavia, Italy
[*] Correspondence: giovanni.parrella@polimi.it; Tel.: +39-02-2399-18-9022

Abstract: The generation of synthetic CT for carbon ion radiotherapy (CIRT) applications is challenging, since high accuracy is required in treatment planning and delivery, especially in an anatomical site as complex as the abdomen. Thirty-nine abdominal MRI-CT volume pairs were collected and a three-channel cGAN (accounting for air, bones, soft tissues) was used to generate sCTs. The network was tested on five held-out MRI volumes for two scenarios: (i) a CT-based segmentation of the MRI channels, to assess the quality of sCTs and (ii) an MRI manual segmentation, to simulate an MRI-only treatment scenario. The sCTs were evaluated by means of similarity metrics (e.g., mean absolute error, MAE) and geometrical criteria (e.g., dice coefficient). Recalculated CIRT plans were evaluated through dose volume histogram, gamma analysis and range shift analysis. The CT-based test set presented optimal MAE on bones (86.03 ± 10.76 HU), soft tissues (55.39 ± 3.41 HU) and air (54.42 ± 11.48 HU). Higher values were obtained from the MRI-only test set ($MAE_{BONE} = 154.87 \pm 22.90$ HU). The global gamma pass rate reached $94.88 \pm 4.9\%$ with 3%/3 mm, while the range shift reached a median (IQR) of 0.98 (3.64) mm. The three-channel cGAN can generate acceptable abdominal sCTs and allow for CIRT dose recalculations comparable to the clinical plans.

Keywords: synthetic CT; MRI guidance; MRI-only; image-guided radiotherapy; carbon ion radiotherapy; particle therapy; deep learning

1. Introduction

In the treatment of abdominal tumors such as those of the liver and pancreas, carbon ion radiotherapy (CIRT) is considered a promising therapeutic option, thanks to its excellent geometrical selectivity and radiobiological effectiveness [1,2]. However, tumors subject to respiratory motion may suffer from inter- and intra-fraction motion that needs to be properly accounted for during planning and delivery. So far, the repeated acquisition of computed tomography (CT), in the form of respiratory-correlated 4DCT, is the clinical routine for motion management, but concurrently exposes the patient to additional non-therapeutic radiation [3].

The growing interest in magnetic resonance imaging (MRI) in recent years has fostered the development of MRI-only workflows that would guarantee the absence of ionizing radiation while exploiting the superior soft tissue contrasts of MRI scans. In this regard,

online MR-guided radiotherapy (MRgRT) is a clinical reality in conventional radiotherapy [4–6], while no integrated system is clinically available for online MR guidance in particle therapy (PT), with only a few feasibility studies addressing protons but not carbon ions [7,8]. Nonetheless, the off-line use of MRI in CIRT may support treatment planning and adaptation through the generation of synthetic CT (sCT), used as alternatives to verification CTs, and avoid non-therapeutic radiation along with CT-MRI registration errors that are typically relevant in the abdominal site [9,10]. Different methods (i.e., bulk density override, atlas based, voxel based) have been studied to generate sCTs starting from MRI scans [8,11]. As an alternative, deep learning (DL) methods have been largely investigated [11–17], relying on their capability of autonomously learning hidden relationships among data [18]. In particular, deep convolutional neural networks (DCNN) showed promising results for the generation of sCTs of the head and neck with respect to conventional methods, as they are able to catch complex nonlinear relationships among MRI and CT [12,15]. Generative adversarial networks (GAN), especially in the conditional form (conditional GAN—cGAN) have been instead considered for more complex anatomical districts such as the abdomen, thanks to the adversarial learning process, which provides great efficacy in image-to-image translation applications [19–21]. The main limitation of these networks is the need for paired MRI-CT samples to perform the training, which is a critical aspect when dealing with multi-modal volumes. In recent years, new architectures have been proposed to overcome the need for paired CT-MRI training datasets, with cycleGAN and fully bidirectional networks showing great results in end-to-end synthetic image generation tasks [22]. Clearly, the use of such complex and multi-network architectures requires high dimensional training datasets, which still remain an evident concern in most clinical realities. The use of multiple MRI sequences (e.g., Dixon in-phase (IP), opposed-phase (OP), fat, and water) has been also investigated to increase the quality of critical anatomical structures, but no relevant improvements were shown [23].

The dosimetric evaluation of DL-based sCTs has been investigated for X-ray and proton radiotherapy in most anatomical districts [20–28], while a similar analysis for CIRT applications is still missing in the literature.

In the treatment of abdominal targets, the presence of inter- and intra-acquisition motions, which mainly translate into the different dispositions of the air fillings in the organs, highly affects the dose distribution for carbon ions plans [29]. An MRI-only workflow for CIRT would thus require more stringent constraints and tolerances in terms of accuracy in the definition of HU values, if compared to conventional radiotherapy, since a precise knowledge of particle stopping power, estimated from HU values within the patient anatomy, is essential for accurate treatment planning, to limit any range shift and damage to healthy tissues [30,31]. Up to now, to the authors' best knowledge, the application of sCT to CIRT plans has been investigated only by Knäusl et al. [32] for head and neck imaging, while no study has been performed on CIRT plans for abdominal tumors.

In this study, we focused on the generation of abdominal sCTs through a cGAN, with the final aim of simulating an MRI-only workflow for CIRT, based on recalculation of the clinical CIRT treatment plans for generated abdominal sCT volumes. This study, therefore, was conducted to evaluate, for the first time, the feasibility of sCT in CIRT for abdominal sites.

2. Materials and Methods

2.1. Patient Cohort

Image datasets were collected from 24 patients affected by liver or pancreatic cancer and treated with CIRT at the National Centre for Oncological Hadrontherapy (CNAO, Pavia, Italy) between 2017 and 2021. Standard clinical workflow comprised the acquisition of a 4DCT followed by a 3D MRI on the same day, for contouring and planning preparation. The same immobilization setup, consisting of customized pillows (MOLDCARE Cushion, QFix, Avondale, PA, USA) and non-perforated body thermoplastic masks (Klarity Medical Products, Heath, OH, USA), was used both for CT and MRI acquisitions and treatment

delivery. Acquisition in two different scanners along with re-positioning of the patient with the thermoplastic mask caused inter-acquisition motion. For 15 patients, re-evaluative images (both CT and MRI) were acquired during the treatment course with the same immobilizations setup and were considered independent from the first acquisition, leading to a total of 39 CT-MRI volume pairs collected. The study was approved by the local ethical committee, and all patients signed the informed consent (CNAO 37-2019 4D-MRI).

The 4DCTs were acquired during patient free breathing on a Siemens SOMATOM Sensation Open CT scanner (resolution $0.98 \times 0.98 \times 2$ mm^3). Clinical plans were optimized at end-exhale for gated treatments; as such, only this phase was used in this work to derive sCTs. CT acquisitions had a variable number of slices, resulting in a volume size of $512 \times 512 \times [96 - 145]$ voxels. MRI acquisitions were performed with a Siemens Magnetom Verio 3T scanner. Three-dimensional breath-hold T1-weighted volumetric interpolated breath-hold examination (VIBE) sequences were acquired at end-exhale with $1.06 \times 1.06 \times 3$ mm^3 resolution (repetition time TR = 3.87 ms, echo time TE = 1.92 ms). For two patients, MRI acquisitions had a voxel size of $1.25 \times 1.25 \times 3$ mm^3 and $1.125 \times 1.125 \times 3$ mm^3. Most of the MRI acquisitions had $320 \times 260 \times 64$ voxels, except for one having 88 transversal slices. Two CT-MRI volume pairs were discarded from the study because of the low quality of the acquired images. Therefore, 37 volume pairs were used: 32 pairs were exploited for cross-validation (CV) and training, while five pairs were randomly selected and held out for testing (Table 1) for more details on the treatment plans. All treatment plans were optimized with the RayStation (Raysearch Laboratories, Stockholm, Sweden—version 10.B) Treatment Planning System (TPS) on the end-exhale reconstructed CT phase and clinically approved. Corresponding organs at risk (OARs), gross tumor volume (GTV) and clinical target volume (CTV) were segmented by radiation oncologists. The relevant OARs included were kidneys, aorta, colon, duodenum, stomach and spinal cord.

Table 1. CIRT plan details for the patients used in the test set.

Patient	N of Beams	Prescribed Dose [Gy(RBE)]	Fractions	Position	Tumor Location
P17	1	43	10	Prone	Pancreas
P20	1	38.4	8	Prone	Pancreas
P21	2	57.6	12	Supine	Pancreas
P27	1	38.4	8	Prone	Pancreas
P31	1	48	10	Prone	Pancreas

2.2. Pre-Processing

Although CT and MRI scans were acquired the same day and with an immobilization setup, inter-acquisition motion (i.e., anatomical changes between CT and MRI scans) was present. A manual rigid registration was firstly performed to align CT and MRI scans. The application of deformable image registration (DIR) was investigated but did not show relevant improvements on the quality of sCTs; therefore, it was not applied (Supplementary Material S1). Indeed, in most cases, the multi-modal DIR could hardly compensate for air filling mismatches and caused large bone deformations and artefacts that reduced the size of the training dataset, making it a less effective approach. Nonetheless, given the lack of a real ground truth (i.e., a CT with MRI anatomical configuration), DIR was applied for the generation of pseudo ground truths (see Section 2.4).

CT slices were resampled to the corresponding MRI spacing to guarantee the voxel consistency among the two volumes and clipped to $[-1000, +1047]$ HU as previously performed by Maspero et al. [14] to reduce the discretization step, while background values were set to -1000 HU. For MRI volumes, the pre-processing consisted of: (i) bias field correction to reduce low frequency noise due to magnetic field inhomogeneities [33]; (ii) reduction of Gaussian noise through a non-linear bilateral filter; (iii) contrast enhancement through histogram clipping to 99th percentile of intensity [14,21,24]: (iv) setting of background

values to zero; and (v) histogram matching to similarly distribute grey levels across all MRI volumes [34].

CT and MRI preprocessed volumes were all resized to 256×256 to match the network input dimensions [19]. Then, the input MRI and target CT transversal slices were segmented in three channels by means of CT-based masks (i.e., air [−1000,−800] HU, bone [150,1047] HU, soft tissues [−800,150] HU) and then linearly scaled to [−1,1].

As a final step, eight transformations were applied to each CT/MRI channel triplet used for training, including horizontal or vertical flip, Gaussian noise adding, shear, rotation and cropping. Thus, the initial training dataset composed of 2014 CT-MRI slice pairs was enlarged to 16,112. Pre-processing was performed through Python scripts implemented using SimpleITK modules (version 2.0.2).

2.3. Neural Network

The neural network used in this work consisted of a cGAN derived from the open-source network "Pix2Pix" by Isola et al. [19], adapted to work on three channels (air, bone, soft tissues) to better account for the anatomical complexity of the abdomen [35] and to deal with the limited dataset of 39 volume pairs.

The net was trained on transversal slices and composed of a $256 \times 256 \times 3$ U-net generator (Figure 1a), used to generate sCTs, and a $70 \times 70 \times 3$ PatchGAN discriminator (Figure 1b), used to judge the quality of the output with high resolution during the training [19]. The U-net is composed of eight encoder blocks, each comprising convolution, batch normalization and leaky rectified linear unit (ReLU) layers, and seven decoder blocks, each composed of transposed convolution, batch normalization, dropout and ReLU layers. The PatchGAN architecture is made of four encoding blocks, such that each pixel of the 30×30 output classifies a 70×70 pixel patch of the input image. For the detailed architecture, refer to Figure 1a,b.

The loss function (Equation (3)) combined L_1 norm (Equation (1)) and cGAN adversarial cross-entropy loss (Equation (2)) to reduce blurring effects and artefacts [19]:

$$L_1 = E_{x,y,z}[\|x - G(y,z)\|_1] \tag{1}$$

$$L_{cGAN} = -\left[\mathrm{E}_{y,x}[\log D(y,x)] + E_{y,z}[\log(1 - D(y, G(y,z)))] \right] \tag{2}$$

$$L_{tot} = L_{cGAN}(\theta_g, \theta_d) + \lambda \cdot L_1(\theta_g) \tag{3}$$

where x is the target CT, y is the input MRI, z the noise, $G(y,z)$ the sCT, $D(y,x)$ the PatchGAN output, θg, θd are the trainable parameters of generator and discriminator, and $\lambda = 100$ is a weight for the L_1 norm, as from Isola et al. [19]. The role of this additional loss is to have the generator not only mislead the discriminator, but also generate synthetic images that mimic the target CT in an L_1 sense, reducing the blurring and improving the representation of structures. The cGAN loss was evaluated on the channel triplets, while the L_1 norm was evaluated after reassembling the sCT slice.

The MRI represents the conditional input of the net, with the noise z being provided in the form of dropout on several layers of the U-net generator during training and testing. The training was performed by alternating one gradient descent step on the discriminator and one on the generator, using ADAM optimizer with momentum parameters β1 = 0.5, β2 = 0.999. The training was stopped after 20 epochs, which were sufficient to achieve convergence thanks to data augmentation.

The cGAN was optimized through a six-fold cross-validation (CV), setting the batch size to 1 and discriminator learning rate to 2×10^{-7} (Supplementary Material S2). This confirmed that the instance normalization (i.e., the use of batch-normalization layers with batch = 1) is well-suited for image generation tasks [36]. The network was trained on the full dataset (i.e., 32 volumes) and tested on the remaining five volumes (Figure 1c). Before stacking all of the output slices to rebuild the synthetic volumes, the values of each channel were re-scaled to the corresponding HU range and re-assembled by means of the same

masks used for the channels' segmentation. The stacked volumes were finally resized to the original MRI [$320 \times 260 \times N_{slices}$] size.
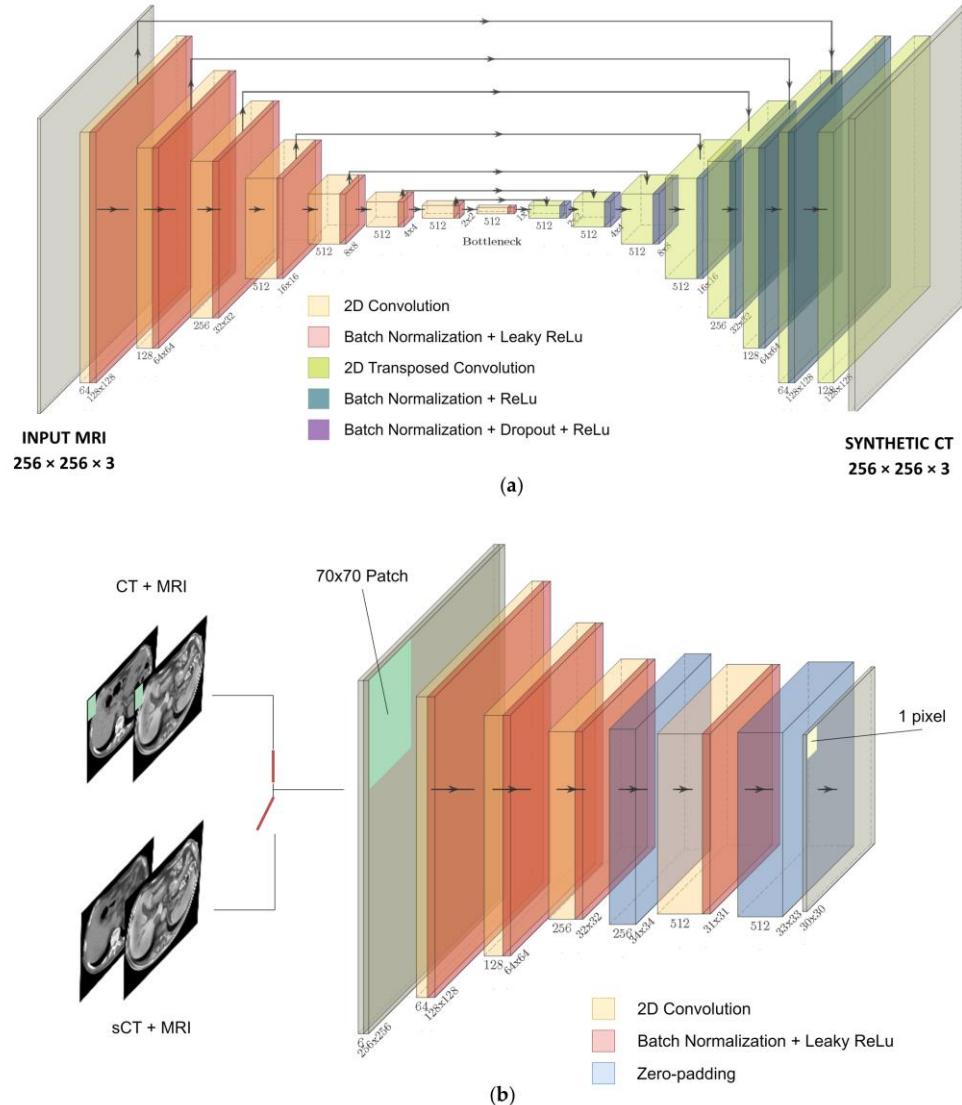

Figure 1. (**a**) U-net generator. (**b**) PatchGAN discriminator.

2.4. Experiments

The results were evaluated by means of similarity metrics such as mean absolute error (MAE), root mean squared error (RMSE), normalized cross correlation (NCC), structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) between sCT and target CT (Supplementary Material S3), with the exclusion of the background, both within CV and testing scenarios. All metrics were evaluated on the basis of reassembled volumes, applying the corresponding tissue mask. Specifically, the five held-out patients were used (i) to create the test set, where CT-based masks were used for the analysis, and (ii) to build an MRI-only simulation set, where the use of CT-based masks was replaced by manual segmentation of the three channels directly on MRI (Figure 1c).

Within the MRI-only workflow, the evaluation of similarity metrics was performed between sCTs and reference volumes obtained by applying DIR between the target CT and MRI (i.e., pseudo ground truths, CT_{PGT}) to compensate for MRI-CT inter-acquisition motion (Figure 2a).
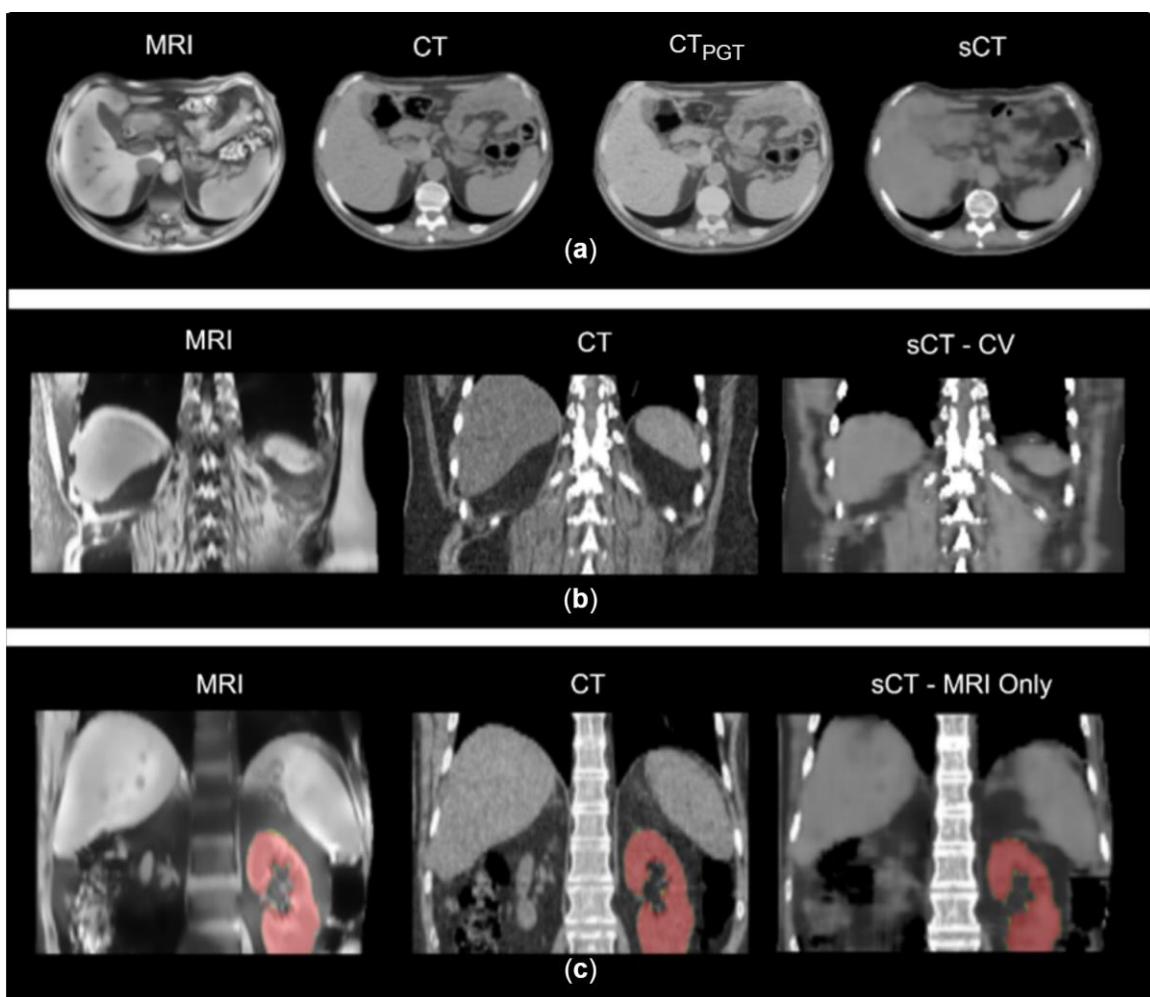
Figure 2. (**a**) Example of MRI, CT, sCT and pseudo ground truth (CT$_{PGT}$) on axial plane. CT$_{PGT}$ still shows visible discrepancies with MRI anatomical condition. (**b**) Example of inter-acquisition motion in a CT-MRI pair and the resulted sCT in CV. (**c**) Example of MRI, planning CT and synthetic CT from MRI-only scenario. In red, the segmentation of kidneys used for the geometrical analysis.

The MRI-only scenario was evaluated also through geometrical criteria: dice coefficient (DSC), 95th percentile Hausdorff distance (HD) and the center of mass distance (CoMD) were calculated on kidney segmentations of MRI, CT and sCT, to assess the quality of the sCT in terms of correct reproduction of soft tissues with respect to the MRI anatomy. HD, DSC and CoMD were assessed for each couple of segmented volumes (CT-MRI, sCT-CT, sCT-MRI). sCTs were compared to MRIs since we expected the sCT to be representative of the MRI anatomical condition; CT was compared to MRI to determine the initial mismatch, while sCTs were compared to target CTs to confirm that sCT was far from matching CT structures.

Due to the lack of a real ground truth, the net was also validated on a CT-MRI volume pairs obtained through a computational phantom (i.e., XCAT [37] for CT and correspondent ComBAT [38] for MRI) that guaranteed an improved match of anatomical structures between the two volumes, avoiding any inter-acquisition motion (Supplementary Material S4).

The clinical CIRT plans were recalculated on the MRI-only sCT for each patient through the TPS and evaluated on the basis of DVH-based metrics (GTV D$_{95\%}$, CTV D$_{95\%}$, D$_{2\%}$ on OARs) and dose difference maps with respect to the original plan. A two one-sided test of equivalence for paired samples (TOST-P) was used to compare the DVH metrics, considering a confidence interval of 95% and an equivalence interval of $\pm 0.5\%$. The global gamma analysis was also performed with 1 mm/1%, 2 mm/2%, 3 mm/3% as

tolerance criteria, and three dose thresholds: 10%, 50%, and 90% on the prescription dose. A range shift (*RS*) analysis was performed to take into account possible dose shifts, which are generally averaged in gamma analysis [15,39]. In this regard, both range shift (*RS*) and relative range shift (*RRS*) were evaluated on a beam-by-beam basis, considering an acceptability threshold set to 5 mm, in accordance with clinical margins used at CNAO [40]. RS and RRS are defined as:

$$RS = (R_{sCT80} - R_{CT80}) \tag{4}$$

$$RRS = \left(RS / R_{CT80} \right) \tag{5}$$

where R_{sCT80} and R_{CT80} are the beam ranges computed at 80% of the dose peak. This evaluation was performed on the dose profile along the central axis of each beam, considering 10 transversal slices, for a total of 60 RSs.

All computational steps were performed on a Precision 5820 Tower DELL workstation equipped with a 16 GB RAM Nvidia GPU (QUADRO P5000). A full training procedure took around 12 h, while the generation of a synthetic volume took ~5 s.

3. Results

The similarity metrics from CV, testing and MRI-only simulation are presented in Table 2. Comparable results were found between CV and test performance, with MAE on soft tissues and air channels being lower than that on bone. Notwithstanding inter-acquisition discrepancies (Figure 2b), the air channel showed an error of 54.42 ± 11.48 HU, the soft tissues 55.39 ± 3.41 HU, and the bone structures 86.03 ± 10.76 HU.

Table 2. Average results for CV, test and MRI-Only procedures, compared to the literature. Average (St. Dev.). * Soft tissue. ** Lungs. *** Vertebral bodies. **** Bidirectional network.

		MAE_Body [HU]	RMSE [HU]	SSIM	PSNR [dB]	NCC	MAE_Air [HU]	MAE_Bone [HU]	MAE_Soft [HU]
Our work	CV	56.52 (8.31)	97.24 (17.56)	0.651 (0.043)	27.73 (1.23)	0.857 (0.054)	46.19 (6.30)	90.76 (7.86)	54.79 (8.98)
	TEST	57.08 (2.79)	99.69 (4.90)	0.67 (0.06)	27.64 (0.68)	0.92 (0.02)	54.42 (11.48)	86.03 (10.76)	55.39 (3.41)
	MRI-ONLY	88.22 (9.88)	181.10 (11.84)	0.59 (0.08)	20.99 (1.49)	0.76 (0.10)	279.01 (142.46)	154.87 (22.90)	75.00 (8.12)
Literature	[20]	78.71 (18.46)	-	-	-	-	-	152.71 (30.14)	53.89 (10.7)
	[24]	62 (13)	-	-	30.0 (1.8)	-	104 (38) **	167 (22)	36 (8) *
	[25]	72.48 (18.16)	-	-	22.65 (3.63)	0.92 (0.04)	108.06 (49.45)	216.81 (63.0)	58.62 (30.61)
	[26]	55.56 (2.27)	106.43 (11.45)	-	-	0.87 (0.03)		-	-
	[27]	-	-	-	-	-	-	-	90 (29)
	[28]	-	-	-	-	-	-	110.09 (29.23) ***	-
	[21]	89.8 (18.7)	-	-	27.4 (1.6)	-	-	-	-
	[23]	60.42 (2.27)	-	-	-	0.88 (0.03)	-	-	-
	[22]	6.30 (0.56) ****	-	0.90 (0.42)	-	-	-	-	-

The results relative to the MRI-only evaluation were characterized by larger values on all metrics. In particular, the errors grew to 279.01 ± 142.46 HU on the air channel, 154.87 ± 22.90 HU in the case of bone structures, and 88.22 ± 9.88 HU within the body.

The geometrical analysis of the MRI-only workflow (Figure 2c) showed the lowest discrepancies between CT and MRI segmentation, confirming good accuracy in replicating MRI anatomy. As an example, the average CoMD in sCT-MRI was 5.85 ± 4.87 mm versus 7.75 ± 5.92 mm in CT-MRI and 13.30 ± 10.42 mm for sCT-CT. Detailed results are reported in Supplementary Material S5.

The validation performed on the phantom showed an MAE of 73.3 HU on the whole volume, while the MAEs on the three channels were of 66.11 HU, 77.93 HU and 167.36 HU for soft tissues, air, and bones, respectively (see Supplementary Material S4).

As for dose accuracy, Figure 3 shows the DVH comparison for patient and dose re-calculations on P17 and P27 (complete results reported in Supplementary Materials S6–S8). The GTV and CTV $D_{95\%}$ as well as the $D_{2\%}$ on the organs at risk (OARs) are displayed in Figure 4, expressed in terms of dose difference Δ(sCT-CT), MAE and relative error (E[%]) with respect to the prescribed dose. For all five patients, good reproducibility was shown relative to the dose to the GTV, with patients P20 and P21 being characterized by errors of -2.04 Gy[RBE] and -1.88 Gy[RBE] respectively, corresponding to -5.3% and -3.3% of the prescribed dose. The MAE on the GTV $D_{95\%}$ was 0.86 ± 0.90 Gy[RBE].
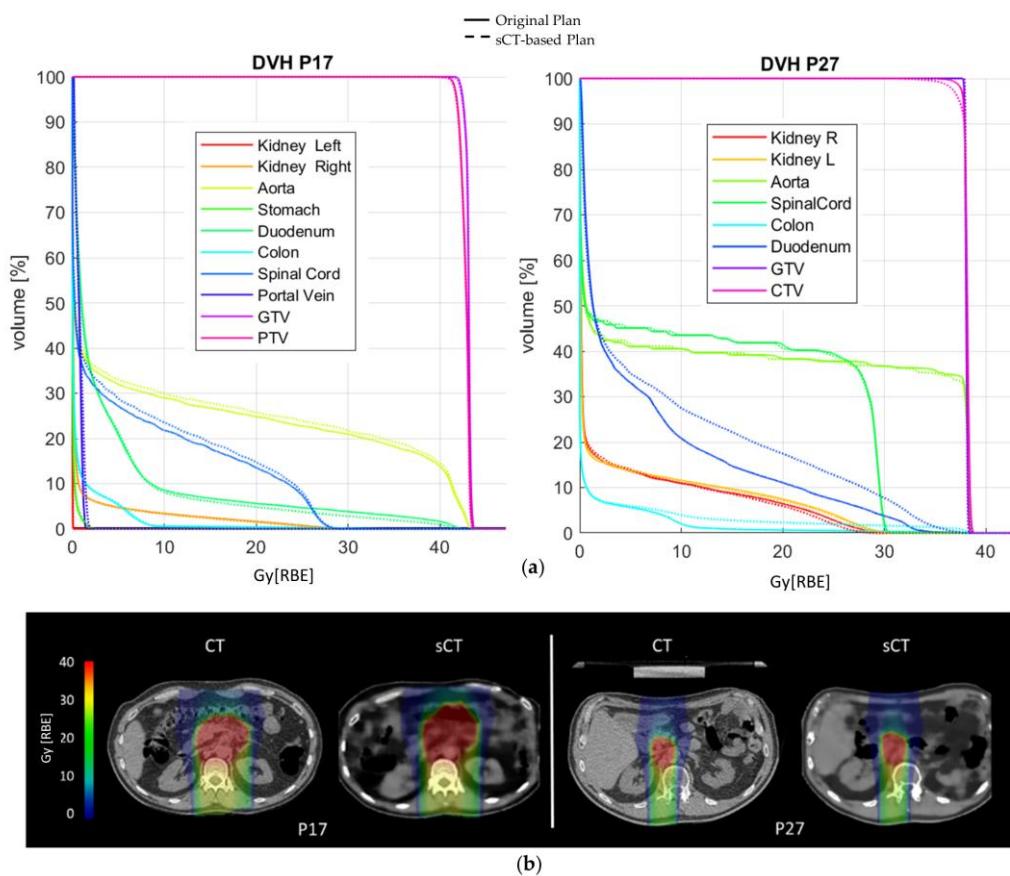

Figure 3. (**a**) DVH comparison on patients P17 and P27; (**b**) original CIRT plan (RBE) and sCT-based recalculation for patients P17 and P27.
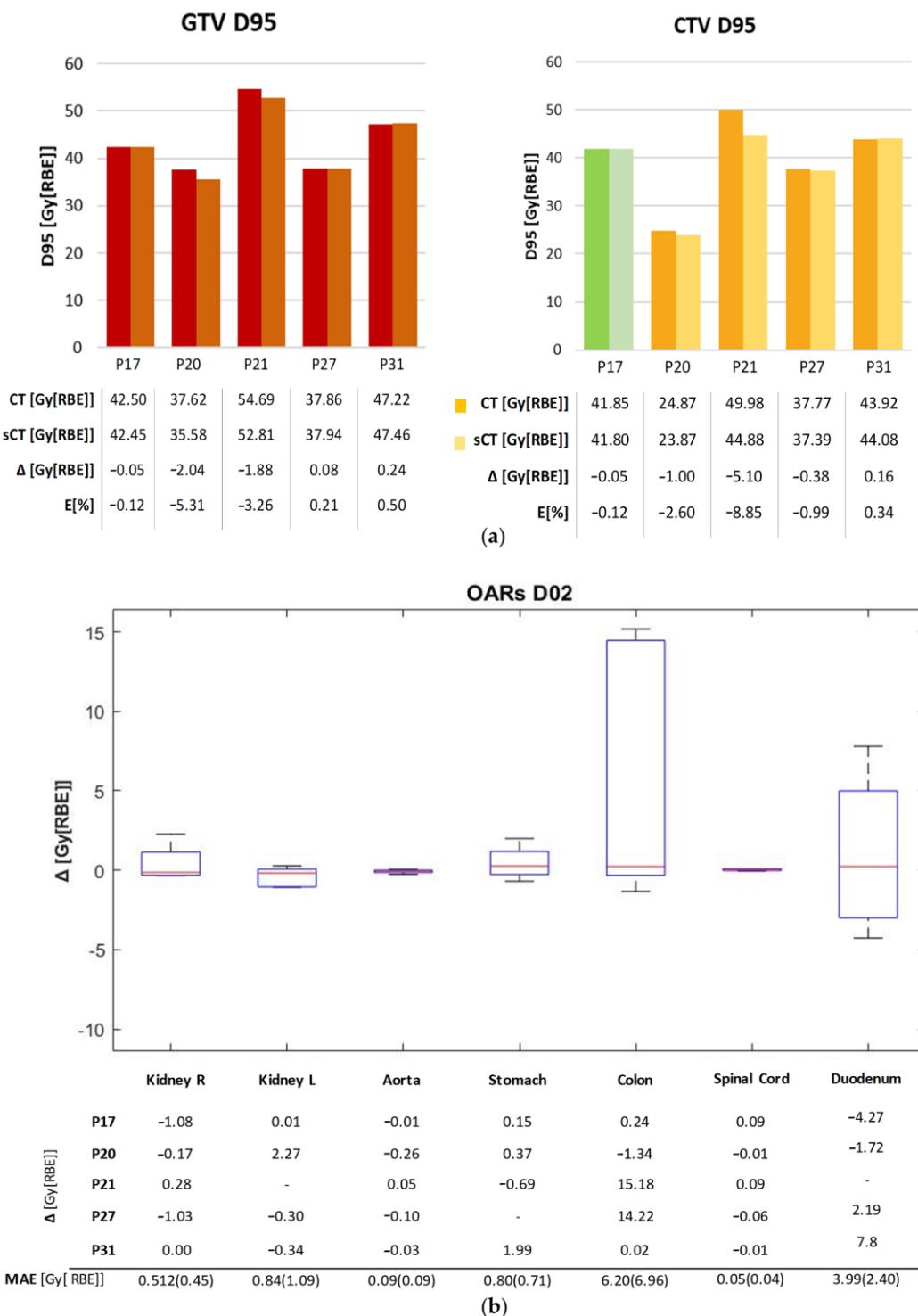
GTV D95	P17	P20	P21	P27	P31
CT [Gy[RBE]]	42.50	37.62	54.69	37.86	47.22
sCT [Gy[RBE]]	42.45	35.58	52.81	37.94	47.46
Δ [Gy[RBE]]	−0.05	−2.04	−1.88	0.08	0.24
E[%]	−0.12	−5.31	−3.26	0.21	0.50

CTV D95	P17	P20	P21	P27	P31
CT [Gy[RBE]]	41.85	24.87	49.98	37.77	43.92
sCT [Gy[RBE]]	41.80	23.87	44.88	37.39	44.08
Δ [Gy[RBE]]	−0.05	−1.00	−5.10	−0.38	0.16
E[%]	−0.12	−2.60	−8.85	−0.99	0.34

(**a**)

	Kidney R	Kidney L	Aorta	Stomach	Colon	Spinal Cord	Duodenum
P17	−1.08	0.01	−0.01	0.15	0.24	0.09	−4.27
P20	−0.17	2.27	−0.26	0.37	−1.34	−0.01	−1.72
P21	0.28	-	0.05	−0.69	15.18	0.09	-
P27	−1.03	−0.30	−0.10	-	14.22	−0.06	2.19
P31	0.00	−0.34	−0.03	1.99	0.02	−0.01	7.8
MAE [Gy[RBE]]	0.512(0.45)	0.84(1.09)	0.09(0.09)	0.80(0.71)	6.20(6.96)	0.05(0.04)	3.99(2.40)

(**b**)

Figure 4. (**a**) D95% values for GTV and CTV in the original plan (CT) and the recalculated one (sCT). For P17, PTV was considered in this comparison (shown in Green). The table contains the dose values and the dose difference Δ[Gy[RBE]], as well as the error relative to the prescribed dose E [%]. (**b**) D2% difference (sCT-CT) for the main OARs on each patient, and the D2% MAE over each OAR. The red mark indicates the median, and edges of the box show the 25th and 75th percentiles.

Similarly, the values relative to the CTV showed the two dose distributions to have comparable results, with a maximum error of −5.10 Gy[RBE] for P21 (−8.9% of the prescribed dose). The MAE on the $D_{95\%}$ CTV was 1.34 ± 1.33 Gy[RBE]. The $D_{2\%}$ relative errors on OARs lay in an interquartile range (IQR) of [−0.24,0.22]%, although the colon reached a

peak relative error of 37.03% (14.22 Gy[RBE]), and the duodenum of 16.25% (7.8 Gy[RBE]) (Supplementary Material S6). The voxel-wise dose difference maps are shown in Figure S5 in Supplementary Material S7, highlighting the worst-case slice for each test patient. Regions with a high dose difference can be mainly seen with correspondence of air pocket mismatches, while the overall distribution of dose to the body is comparable. Indeed, as in Table S6 in Supplementary Material S7, all test patients showed a median dose difference close to zero, with the widest IQR being 0.158 Gy[RBE] on patient P21. The maximum errors were in the range [22.64,42.36] Gy[RBE]. In this regard, an incomplete reproduction of the kidney affected the dose distribution in patient P27 (Figure S6), while the limited field of view of MRI introduced dose artefacts on recalculation for patient P21 (Figure S7).

A peak gamma pass rate of 94.88% was obtained in the 3%/3 mm analysis (Supplementary Material S9).

The range shift analysis showed a median (IQR) RS of 0.98 (3.64) mm and RRS of 0.61 (2.14)% (Figure 5). Considering each beam individually, as shown in Table S9 in Supplementary Material S10, patient P27 was the one showing the highest errors, with a median RS of 5.69 (6.97) mm.

(a)

(b)

Figure 5. Representative range shift analysis on patient P27. (**a**) The graph shows both dose and HU profiles in CT (light blue) and sCT (red), evaluated along the yellow line shown in (**b**). The 80% reference is marked by the horizontal line. (**b**) Corresponding CT and sCT sagittal views are compared. The yellow line is one of the 10 considered for each beam, on different transversal slices.

4. Discussion

In this work, we investigated for the first time the feasibility of a cGAN in generating sCTs of the abdominal site for applications in CIRT. The cGAN, trained with transversal CT-MRI slice pairs, was optimized to work on three channels corresponding to air, bone and soft tissues to better account for the anatomical complexity of the abdomen.

The performance of the network was firstly evaluated on the basis of similarity metrics of the test set built with CT-based segmentation. Despite MRI to CT inter-acquisition motion, the MAE on the body (57.08 ± 2.79 HU) was comparable to results obtained by other works on the abdominal site that used multiple Dixon sequences and more complex architectures (55.56 ± 2.27 HU [26] and 60.42 ± 2.27 HU [23]), as well as U-nets trained on T1w–T2w MRI acquisitions (62 ± 13 HU) [24]. In addition, this work favorably compares to other studies exploiting cGAN or cycleGAN in terms of MAE_{BODY} [20,21,25], even if the bidirectional network from Xu et al. achieved outstanding results, although working on a much wider unpaired dataset [22]. The NCC (0.92 ± 0.02) was shown to be consistent with those of other works [23,25,26], while the PSNR (27.64 ± 0.68 dB) was comparable to the work by Fu et al. [21]. Optimal MAE metrics were obtained in the generation of bone structures (86.03 ± 10.76 HU) and soft tissues (55.39 ± 3.4 HU), outperforming other works in the literature [20,24,25,28].

The MAE on the air channel and bones was also low with respect to other approaches with cycleGAN [25]. This could be due to the use of CT-derived masks on the test set, which may have aided the replication of CT air pockets and bone structures on the sCTs.

In order to cope with this and to simulate a real-case scenario, the network was then tested on an MRI-only simulation set, where the MRI segmentation was completely independent from the use of CT. Given the limited performance of multi-modal DIR on the considered dataset, manual segmentation of the three channels on MRI was considered the most accurate approach, despite the time-consuming task for clinical purposes. Nonetheless, DIR was applied between the target CT and MRI to overcome the lack of a ground truth CT representative of MRI anatomical condition, notwithstanding the minor contribution of such registration. Indeed, this process did not fully compensate for different air cavities and inter-acquisition motion; therefore, the volumes (i.e., CT_{PGT}) used as reference were still showing visible discrepancies with respect to the MRI, biasing the evaluation of the performance of the network (Figure 2a). A similar consideration applied by Florkow et al. highlighted errors in HU intensities caused by inter-acquisition variations that were not compensated by the deformable registration [24]. Moreover, the manual segmentation of MRI represented a complex step, especially for bones, since they are not clearly visible on VIBE volumes. Nonetheless, the results were shown to still be acceptable when compared to the literature, with MAE on soft tissues (75.00 ± 8.12 HU) being aligned to results obtained with GAN (90 ± 29 HU) or cycleGAN (58.62 ± 30.61 HU) [25,27].

The geometrical analysis, performed on the MRI-only test set, was conducted to support the good geometrical agreement between sCT and MRI: in general, sCT and MRI segmentations showed the best match, confirming the expected performance of the net in reproducing the MRI anatomy, while sCTs showed higher deviations with respect to the target CT scans. This evaluation was performed on segmentations of the kidneys, which are well contrasted organs, and can be considered representative of the geometrical accuracy in the reproduction of soft tissues.

Due to the lack of a proper ground truth on patients' data, we supported our results with a validation performed on a single CT-MRI volume pair obtained through a computational phantom (i.e., XCAT [37] for CT and correspondent ComBAT [38] for MRI). This approach, which guaranteed the perfect anatomical match between the two volumes, provided promising outcomes (i.e., MAE of 73.3 HU on the whole volume) aligned with the presented results and the literature. The MAE on the three channels described a discrete reproduction of the air fillings and soft tissues, with poorer performance on bones, as noticed also in the literature. NCC reached 0.89, describing an acceptable reproduction of the anatomical structures (see Supplementary Material S4).

For dose analysis, most patients (Figure 4a and Supplementary Materials S6–S8) presented DVH metrics within clinical tolerances (i.e., below 3% of the prescribed dose) and aligned with studies in the literature on protons (relative error < 3% on PTV [25] and relative error < 2% on ITV—the internal target volume [24]). Moreover, *p*-values gave statistical evidence of the equivalence between the DVH metrics from sCT and CT ($p < 0.05$, Table S7 in Supplementary Material S8); however, a larger test dataset would make these considerations more solid.

Concerning DVH metrics, Liu Y. et al. for protons [25] as well as Liu L. for photons [28] obtained promising results, with errors on the maximum dose on the PTV being in the range [−1,+1] Gy[RBE], also thanks to the use of lateral beams that avoided most of the air-filled organs. Florkow et al. [24] obtained acceptable errors on the OARs, with $D_{2\%}$ being in the range [−2.7,3.7]% of the prescribed dose. Notwithstanding the use of DIR, the work by Florkow et al. suffered from inter-scan variations (i.e., air fillings), that may have caused an overestimation of the actual differences between the planning CT and sCT [24]. In our work, the IQR on the $D_{2\%}$ relative error was [−0.24,0.22]%, but high errors were highlighted on the colon (37.03%) and the duodenum (16.25%, Figure 4c), which were highly affected by inter-acquisition motion of air fillings between sCT (representative of MRI anatomy) and planning CT. In this regard, the work by Knäusl et al. [32] showed that the constraints on OARs are very challenging for compliance, presenting errors of up to 28%, mostly due to the incorrect representation of bones or air cavities. Discrepancies in the dose distributions were also confirmed by the gamma analysis, which in our work showed a peak value of $94.88 \pm 4.9\%$ against the $99.37 \pm 0.99\%$ reported in the literature [25]. Table S8 in Supplementary Material S9 shows a comparison of gamma pass rates from relevant studies on proton and photon applications, which were, therefore, not fully comparable to our application and highlighted the need for more studies on CIRT. The range shift analysis presented median RS within the clinical threshold (i.e., 5 mm), but with critical results for patients P27 and P31 (median (IQR), 5.69 (6.97) mm and 3.09 (2.60) mm, respectively), because of the presence of unmatched air pockets (Figure 5b). In proton plans, a maximum RS value of 5.6 mm (5.68%) was reported [25], whereas in our case, the inconsistencies of air cavities led to RS values of up to 15.37 mm (i.e., RRS = 9.09%, patient P27). This aspect may be critical for CIRT application, and needs to be analyzed on a wider population.

Similarly, the highest dose discrepancies were mostly found in correspondence of the different disposition of air pockets between sCT and planning CT (Figure S5 in Supplementary Material S7). The maximum error for patient P27 was due to an incomplete reproduction of the kidney (Figure S6), which caused an overdose for duodenum (+5.7% D_{02}) with respect to the prescribed dose (Supplementary Material S6). The dose artefacts on recalculation for patient P21 reported regions of high dose differences (Figure S7), but this issue was not correlated with the quality of the sCT and could be easily overcome by acquiring wider volumes.

The main limitation of the study was the lack of a proper ground truth to validate the proposed approach, which could not be fully compensated due to the poor performance of multi-modal DIR in the abdominal site. As such, the use of computational phantoms [37,38] to ensure the correspondence between CT and MRI scans will be considered in the future as an effective approach to validate the proposed network, as anticipated in this work (Supplementary Material S4), and enlarge the training dataset. In addition, although the three channel implementation allowed good performance of the net in such a complex anatomical site with limited data, manual segmentation can be demanding, especially for not well-contrasted structures in VIBE acquisitions, such as bone, and is definitely not suited for clinical application. Further steps could be, therefore, to include the acquisition of specific MRI sequences (e.g., ultrashort echo time, UTE) to facilitate bone segmentation or avoid channel separation in an improved version of the net [26]. We also expect that our results could be improved by increasing the dataset; in future analyses, we intend to include different respiratory phases to (i) achieve higher accuracy and limit errors in the reproduction of tissues, (ii) eliminate separation in the three channels, and (iii) derive

synthetic respiratory-correlated 4DCT [10,41]. Finally, although our results were mainly affected by air-filling effects, the absence of the thermoplastic mask on the sCT could also have an impact [32]; as such, a uniform and pre-defined outline to the sCT could be applied [32], although it would not be an optimal countermeasure.

Despite the above-mentioned limitations, this work showed that the three-channel cGAN can generate accurate sCTs of the abdominal site that can support treatment planning, evaluation, and adaptation in CIRT. To the authors' best knowledge, this work is the first analysis applied to the abdomen for CIRT, and thus represents a starting point for future in-depth analyses of the feasibility of MRI-only workflows in CIRT.

References

1. Durante, M.; Orecchia, R.; Loeffler, J.S. Charged-Particle Therapy in Cancer: Clinical Uses and Future Perspectives. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 483–495. [CrossRef]
2. Liermann, J.; Shinoto, M.; Syed, M.; Debus, J.; Herfarth, K.; Naumann, P. Carbon Ion Radiotherapy in Pancreatic Cancer: A Review of Clinical Data. *Radiother. Oncol.* **2020**, *147*, 145–150. [CrossRef] [PubMed]
3. Li, H.; Dong, L.; Bert, C.; Chang, J.; Flampouri, S.; Jee, K.; Lin, L.; Moyers, M.; Mori, S.; Rottmann, J.; et al. AAPM Task Group Report 290: Respiratory Motion Management for Particle Therapy. *Med. Phys.* **2022**, *49*, e50–e81. [CrossRef]
4. Jaffray, D.A. Image-Guided Radiotherapy: From Current Concept to Future Perspectives. *Nat. Rev. Clin. Oncol.* **2012**, *9*, 688–699. [CrossRef] [PubMed]
5. Otazo, R.; Lambin, P.; Pignol, J.P.; Ladd, M.E.; Schlemmer, H.P.; Baumann, M.; Hricak, H. MRI-Guided Radiation Therapy: An Emerging Paradigm in Adaptive Radiation Oncology. *Radiology* **2021**, *298*, 248–260. [CrossRef] [PubMed]
6. Keall, P.J.; Brighi, C.; Glide-Hurst, C.; Liney, G.; Liu, P.Z.Y.; Lydiard, S.; Paganelli, C.; Pham, T.; Shan, S.; Tree, A.C.; et al. Integrated MRI-Guided Radiotherapy—Opportunities and Challenges. *Nat. Rev. Clin. Oncol.* **2022**, *19*, 458–470. [CrossRef]
7. Kurz, C.; Buizza, G.; Landry, G.; Kamp, F.; Rabe, M.; Paganelli, C.; Baroni, G.; Reiner, M.; Keall, P.J.; van den Berg, C.A.T.; et al. Medical Physics Challenges in Clinical MR-Guided Radiotherapy. *Radiat. Oncol.* **2020**, *15*, 93. [CrossRef] [PubMed]
8. Hoffmann, A.; Oborn, B.; Moteabbed, M.; Yan, S.; Bortfeld, T.; Knopf, A.; Fuchs, H.; Georg, D.; Seco, J.; Spadea, M.F.; et al. MR-Guided Proton Therapy: A Review and a Preview. *Radiat. Oncol.* **2020**, *15*, 129. [CrossRef]
9. Paganelli, C.; Meschini, G.; Molinelli, S.; Riboldi, M.; Baroni, G. Patient-Specific Validation of Deformable Image Registration in Radiation Therapy: Overview and Caveats. *Med. Phys.* **2018**, *45*, e908–e922. [CrossRef] [PubMed]
10. Meschini, G.; Vai, A.; Paganelli, C.; Molinelli, S.; Maestri, D.; Fontana, G.; Pella, A.; Vitolo, V.; Valvo, F.; Ciocca, M.; et al. Investigating the Use of Virtual 4DCT from 4DMRI in Gated Carbon Ion Radiation Therapy of Abdominal Tumors. *Z. Med. Phys.* **2022**, *32*, 98–108. [CrossRef]
11. Johnstone, E.; Wyatt, J.J.; Henry, A.M.; Short, S.C.; Sebag-Montefiore, D.; Murray, L.; Kelly, C.G.; McCallum, H.M.; Speight, R. 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. [CrossRef]
12. Han, X. MR-Based Synthetic CT Generation Using a Deep Convolutional Neural Network Method. *Med. Phys.* **2017**, *44*, 1408–1419. [CrossRef]
13. Dinkla, A.M.; Florkow, M.C.; Maspero, M.; Savenije, M.H.F.; Zijlstra, F.; Doornaert, P.A.H.; van Stralen, M.; Philippens, M.E.P.; van den Berg, C.A.T.; Seevinck, P.R. Dosimetric Evaluation of Synthetic CT for Head and Neck Radiotherapy Generated by a Patch-Based Three-Dimensional Convolutional Neural Network. *Med. Phys.* **2019**, *46*, 4095–4104. [CrossRef]
14. Maspero, M.; Savenije, M.H.F.; Dinkla, A.M.; Seevinck, P.R.; Intven, M.P.W.; Jurgenliemk-Schulz, I.M.; Kerkmeijer, L.G.W.; van den Berg, C.A.T. Dose Evaluation of Fast Synthetic-CT Generation Using a Generative Adversarial Network for General Pelvis MR-Only Radiotherapy. *Phys. Med. Biol.* **2018**, *63*, 185001. [CrossRef]
15. Spadea, M.F.; Pileggi, G.; Zaffino, P.; Salome, P.; Catana, C.; Izquierdo-Garcia, D.; Amato, F.; Seco, J. Deep Convolution Neural Network (DCNN) Multiplane Approach to Synthetic CT Generation From MR Images—Application in Brain Proton Therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2019**, *105*, 495–503. [CrossRef]
16. Kazemifar, S.; Montero, A.M.B.; Souris, K.; Rivas, S.T.; Timmerman, R.; Park, Y.K.; Jiang, S.; Geets, X.; Sterpin, E.; Owrangi, A. Dosimetric Evaluation of Synthetic CT Generated with GANs for MRI-Only Proton Therapy Treatment Planning of Brain Tumors. *J. Appl. Clin. Med. Phys.* **2020**, *21*, 76–86. [CrossRef]
17. Koerkamp, M.L.G.; de Hond, Y.J.M.; Maspero, M.; Kontaxis, C.; Mandija, S.; Vasmel, J.E.; Charaghvandi, R.K.; Philippens, M.E.P.; van Asselen, B.; van den Bongard, H.J.G.D.; et al. Synthetic CT for Single-Fraction Neoadjuvant Partial Breast Irradiation on an MRI-Linac. *Phys. Med. Biol.* **2021**, *66*, 085010. [CrossRef]
18. Spadea, M.F.; Maspero, M.; Zaffino, P.; Seco, J. Deep Learning Based Synthetic-CT Generation in Radiotherapy and PET: A Review. *Med. Phys.* **2021**, *48*, 6537–6566. [CrossRef]
19. Isola, P.; Zhu, J.; Efros, A.A.; Ai, B.; Berkeley, U.C. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.
20. Cusumano, D.; Lenkowicz, J.; Votta, C.; Boldrini, L.; Placidi, L.; Catucci, F.; Dinapoli, N.; Antonelli, M.V.; Romano, A.; de Luca, V.; et al. A Deep Learning Approach to Generate Synthetic CT in Low Field MR-Guided Adaptive Radiotherapy for Abdominal and Pelvic Cases. *Radiother. Oncol.* **2020**, *153*, 205–212. [CrossRef]
21. Fu, J.; Singhrao, K.; Cao, M.; Yu, V.; Anand, P.S.; Yang, Y.; Guo, M.; Raldow, A.C.; Ruan, D.; Lewis, J.H. Generation of Abdominal Synthetic CTs from 0.35T MR Images Using Generative Adversarial Networks for MR-Only Liver Radiotherapy. *Biomed. Phys. Eng. Express* **2020**, *6*, 015033. [CrossRef]
22. Xu, L.; Zeng, X.; Zhang, H.; Li, W.; Lei, J.; Huang, Z. BPGAN: Bidirectional CT-to-MRI Prediction Using Multi-Generative Multi-Adversarial Nets with Spectral Normalization and Localization. *Neural Netw.* **2020**, *128*, 82–96. [CrossRef]
23. Xu, K.; Cao, J.; Xia, K.; Yang, H.; Zhu, J.; Wu, C.; Jiang, Y.; Qian, P. Multichannel Residual Conditional GAN-Leveraged Abdominal Pseudo-CT Generation via Dixon MR Images. *IEEE Access* **2019**, *7*, 163823–163830. [CrossRef]
24. Florkow, M.C.; Guerreiro, F.; Zijlstra, F.; Seravalli, E.; Janssens, G.O.; Maduro, J.H.; Knopf, A.C.; Castelein, R.M.; van Stralen, M.; Raaymakers, B.W.; et al. Deep Learning-Enabled MRI-Only Photon and Proton Therapy Treatment Planning for Paediatric Abdominal Tumours. *Radiother. Oncol.* **2020**, *153*, 220–227. [CrossRef] [PubMed]

25. Liu, Y.; Lei, Y.; Wang, Y.; Wang, T.; Ren, L.; Lin, L.; McDonald, M.; Curran, W.J.; Liu, T.; Zhou, J.; et al. MRI-Based Treatment Planning for Proton Radiotherapy: Dosimetric Validation of a Deep Learning-Based Liver Synthetic CT Generation Method. *Phys. Med. Biol.* **2019**, *64*, 145015. [CrossRef] [PubMed]

26. Qian, P.; Xu, K.; Wang, T.; Zheng, Q.; Yang, H.; Baydoun, A.; Zhu, J.; Traughber, B.; Muzic, R.F. Estimating CT from MR Abdominal Images Using Novel Generative Adversarial Networks. *J. Grid Comput.* **2020**, *18*, 211–226. [CrossRef]

27. Olberg, S.; Chun, J.; Choi, B.S.; Park, I.; Kim, H.; Kim, T.; Kim, J.S.; Green, O.; Park, J.C. Abdominal Synthetic CT Reconstruction with Intensity Projection Prior for MRI-Only Adaptive Radiotherapy. *Phys. Med. Biol.* **2021**, *66*, 204001. [CrossRef]

28. Liu, L.; Johansson, A.; Cao, Y.; Dow, J.; Lawrence, T.; Balter, J. Abdominal Synthetic CT Generation from MR Dixon Images Using a U-Net Trained with 'Semi-Synthetic' CT Data. *Phys. Med. Biol.* **2020**, *65*, 125001. [CrossRef] [PubMed]

29. Kumagai, M.; Hara, R.; Mori, S.; Yanagi, T.; Asakura, H.; Kishimoto, R.; Kato, H.; Yamada, S.; Kandatsu, S.; Kamada, T. Impact of Intrafractional Bowel Gas Movement on Carbon Ion Beam Dose Distribution in Pancreatic Radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* **2009**, *73*, 1276–1281. [CrossRef] [PubMed]

30. Rietzel, E.; Schardt, D.; Haberer, T. Range Accuracy in Carbon Ion Treatment Planning Based on CT-Calibration with Real Tissue Samples. *Radiat. Oncol.* **2007**, *2*, 14. [CrossRef] [PubMed]

31. Witt, M.; Weber, U.; Kellner, D.; Engenhart-Cabillic, R.; Zink, K. Optimization of the Stopping-Power-Ratio to Hounsfield-Value Calibration Curve in Proton and Heavy Ion Therapy. *Z. Med. Phys.* **2015**, *25*, 251–263. [CrossRef]

32. Knäusl, B.; Kuess, P.; Stock, M.; Georg, D.; Fossati, P.; Georg, P.; Zimmermann, L. Possibilities and Challenges When Using Synthetic Computed Tomography in an Adaptive Carbon-Ion Treatment Workflow. *Z. Med. Phys.* **2022**. [CrossRef]

33. Tustison, N.J.; Avants, B.B.; Cook, P.A.; Zheng, Y.; Egan, A.; Yushkevich, P.A.; Gee, J.C. N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* **2010**, *29*, 1310–1320. [CrossRef]

34. Udupa, J.K. On Standardizing the MR Image Intensity Scale. *Magn. Reson. Med. Off. J. Int. Soc. Magn. Reson. Med.* **1999**, *42*, 1072–1081.

35. Gupta, D.; Kim, M.; Vineberg, K.A.; Balter, J.M. Generation of Synthetic CT Images from MRI for Treatment Planning and Patient Positioning Using a 3-Channel U-Net Trained on Sagittal Images. *Front. Oncol.* **2019**, *9*, 964. [CrossRef]

36. Ulyanov, D.; Vedaldi, A.; Lempitsky, V. Instance Normalization: The Missing Ingredient for Fast Stylization. *arXiv* **2016**, arXiv:1607.08022.

37. Segars, W.P.; Sturgeon, G.; Mendonca, S.; Grimes, J.; Tsui, B.M.W. 4D XCAT Phantom for Multimodality Imaging Research. *Med. Phys.* **2010**, *37*, 4902–4915. [CrossRef]

38. Paganelli, C.; Summers, P.; Gianoli, C.; Bellomi, M.; Baroni, G.; Riboldi, M. A Tool for Validating MRI-Guided Strategies: A Digital Breathing CT/MRI Phantom of the Abdominal Site. *Med. Biol. Eng. Comput.* **2017**, *55*, 2001–2014. [CrossRef]

39. Pileggi, G.; Speier, C.; Sharp, G.C.; Izquierdo Garcia, D.; Catana, C.; Pursley, J.; Amato, F.; Seco, J.; Spadea, M.F. Proton Range Shift Analysis on Brain Pseudo-CT Generated from T1 and T2 MR. *Acta Oncol.* **2018**, *57*, 1521–1531. [CrossRef]

40. Vitolo, V.; Cobianchi, L.; Brugnatelli, S.; Barcellini, A.; Peloso, A.; Facoetti, A.; Vanoli, A.; Delfanti, S.; Preda, L.; Molinelli, S.; et al. Preoperative Chemotherapy and Carbon Ions Therapy for Treatment of Resectable and Borderline Resectable Pancreatic Adenocarcinoma: A Prospective, Phase II, Multicentre, Single-Arm Study. *BMC Cancer* **2019**, *19*, 922. [CrossRef]

41. Meschini, G.; Vai, A.; Paganelli, C.; Molinelli, S.; Fontana, G.; Pella, A.; Preda, L.; Vitolo, V.; Valvo, F.; Ciocca, M.; et al. Virtual 4DCT from 4DMRI for the Management of Respiratory Motion in Carbon Ion Therapy of Abdominal Tumors. *Med. Phys.* **2020**, *47*, 909–916. [CrossRef]