Radiotherapy and Oncology 184 (2023) 109692

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Comprehensive dose evaluation of a Deep Learning based synthetic Computed Tomography algorithm for pelvic Magnetic Resonance-only radiotherapy

Jonathan J Wyatt [a,b,*], Sandeep Kaushik [c,d], Cristina Cozzini [c], Rachel A. Pearson [a,b], Steven Petit [e], Marta Capala [e], Juan A Hernandez-Tamames [f], Katalin Hideghéty [g], Ross J Maxwell [a], Florian Wiesinger [c], Hazel M. McCallum [a,b]

[a] Translational and Clinical Research Institute, Newcastle University; [b] Northern Centre for Cancer Care, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle, UK; [c] GE Healthcare, Munich, Germany; [d] Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland; [e] Department of Radiotherapy, Erasmus MC Cancer Institute; [f] Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, the Netherlands; [g] Department of Oncotherapy, University of Szeged, Szeged, Hungary

ARTICLE INFO

Article history:
Received 10 November 2022
Received in revised form 19 April 2023
Accepted 24 April 2023
Available online 6 May 2023

Keywords:
MR-only radiotherapy
Magnetic Resonance
Synthetic Computed Tomography
Deep Learning
Pelvic cancers

ABSTRACT

Background and Purpose: Magnetic Resonance (MR)-only radiotherapy enables the use of MR without the uncertainty of MR–Computed Tomography (CT) registration. This requires a synthetic CT (sCT) for dose calculations, which can be facilitated by a novel Zero Echo Time (ZTE) sequence where bones are visible and images are acquired in 65 seconds. This study evaluated the dose calculation accuracy for pelvic sites of a ZTE-based Deep Learning sCT algorithm developed by GE Healthcare.

Materials and Methods: ZTE and CT images were acquired in 56 pelvic radiotherapy patients in the radiotherapy position. A 2D U-net convolutional neural network was trained using pairs of deformably registered CT and ZTE images from 36 patients. In the remaining 20 patients the dosimetric accuracy of the sCT was assessed using cylindrical dummy Planning Target Volumes (PTVs) positioned at four different central axial locations, as well as the clinical treatment plans (for prostate (n = 10), rectum (n = 4) and anus (n = 6) cancers). The sCT was rigidly and deformably registered, the plan recalculated and the doses compared using mean differences and gamma analysis.

Results: Mean dose differences to the PTV D98% were $\leq$ 0.5% for all dummy PTVs and clinical plans (rigid registration). Mean gamma pass rates at 1%/1 mm were 98.0 ± 0.4% (rigid) and 100.0 ± 0.0% (deformable), 96.5 ± 0.8% and 99.8 ± 0.1%, and 95.4 ± 0.6% and 99.4 ± 0.4% for the clinical prostate, rectum and anus plans respectively.

Conclusions: A ZTE-based sCT algorithm with high dose accuracy throughout the pelvis has been developed. This suggests the algorithm is sufficiently accurate for MR-only radiotherapy for all pelvic sites.

Magnetic Resonance (MR)-only radiotherapy enables the superior soft-tissue contrast of MR to be used for delineation without the uncertainty of an MR-Computed Tomography (CT) registration, improving the geometric accuracy of treatments and potentially reducing patient side-effects [1]. MR cannot be used directly for radiotherapy dose calculations and so a method of generating a synthetic CT (sCT) from the MR needs to be developed [2]. Bulk-density and atlas-based methods have been developed [3] with several commercial solutions produced based on these approaches and evaluated on patients with prostate cancer with very small mean dose differences to CT ($\leq$0.5%) [4–6]. More recently a number of Deep Learning methods have been reported in the literature [7,8]. These potentially offer improved dose accuracy and faster image reconstruction times. However only a few studies have evaluated pelvic Deep Learning sCT models for dose calculation accuracy [9–11], which is the only clinically relevant parameter [9].

Most Deep Learning sCT algorithms have used conventional MR images (T1-weighted, T2weighted or Dixon sequences) as the input image [8]. Unlike conventional MR images, Zero Echo Time (ZTE) images capture signal from cortical bone, which potentially facilitates improved bone generation in the sCT [12–14]. Leynes et al. evaluated a Deep Learning sCT algorithm using a combination of ZTE and Dixon images for Positron Emission Tomography (PET)-MR attenuation correction [15]. Using a combination of images is time-consuming however, which reduces scanning efficiency and

* Corresponding author at: Northern Centre for Cancer Care, Newcastle upon Tyne Hospitals NHS Foundation Trust, Freeman Hospital, Freeman Road, Newcastle, NE7 7DN, UK.
E-mail address: jonathanwyatt@nhs.net (J.J Wyatt).

https://doi.org/10.1016/j.radonc.2023.109692
0167-8140/

can produces challenges in registering images due to changes in bladder filling during the imaging session. In contrast, a sCT algorithm using a single ZTE input would be highly efficient since the ZTE image is rapid to acquire. sCTs generated using a single ZTE image as input have not been dosimetrically evaluated for pelvic MR-only radiotherapy previously.

Most dose evaluations of sCT images in the pelvis have focused on using clinical prostate treatment plans [1,3]. This has limited relevance to other treatment sites in the pelvis where Planning Target Volumes (PTVs) may extend significantly superiorly and/or inferiorly of prostate PTVs. The dose calculation accuracy of the sCT is not robustly assessed in these areas because it is in the low dose region of the prostate treatment plan. A challenge for evaluating sCTs dosimetrically for pelvic radiotherapy sites apart from the prostate is that there can be significant variability between patients in the size and location of the PTV(s). This requires large patient evaluation cohorts to ensure the sCT is accurate for all potential treatment sites, which can be difficult, especially in rarer cancers. This study aimed to avoid this challenge by developing a novel dosimetric analysis which assessed sCTs throughout the pelvis, regardless of clinical indication. This enables a smaller number of patients to be used to evaluate sCTs for all pelvic radiotherapy sites.

The overall aim of this prospective study was to comprehensively evaluate a Deep Learning sCT algorithm based on a rapid ZTE MR sequence including Deep Learning-based MR image reconstruction for all pelvic radiotherapy sites.

Materials and methods

Patients and treatment characteristics

A prospective study was performed to train and evaluate the Deep Learning sCT algorithm. Patients planned for radical/neoadjuvant (chemo)radiotherapy for cancers in the pelvis were included. Exclusion criteria included contraindicated for MR scanning, medical implants in the pelvic area (eg hip prostheses) and external contour greater than the MR scanner field of view. The study was approved by a research ethics committee (reference 20/LO/0583) and all patients gave informed consent.

Patient imaging

Patients received a planning CT scan (Sensation Open, Siemens, Erlangen, Germany) using a combined customisable foot and knee rest (Civco Medical Solutions, Coralville, Iowa, USA) with a voxel size of $1.1 \times 1.1 \times 3$ mm^3 and a tube voltage of 120 kVp. Patients being treated for anal and rectal cancers received a contrast-enhanced CT scan. Patients were imaged following routine bladder preparation consisting of an empty bladder 30 minutes prior to the scan, followed by drinking 400 ml of water, and bowel preparation consisting of the application of a micro-enema 60 minutes prior to the scan followed by bowel emptying.

All patients received an MR scan on a SIGNA PET/MR 3 T scanner (software version MP26 GE Healthcare, Waukesha, USA) after their radiotherapy planning CT scan and before their first treatment fraction. MR images were acquired within a median time of 6 days from the CT (range 1–15 days). Patients were scanned in the radiotherapy treatment position on a flat couch-top with a coil bridge for the anterior MR coil. Patients were positioned to match their radiotherapy planning CT scan using the same model of foot and knee rest and external lasers matched to patient tattoos. Patients followed the same bladder preparation protocol.

MR images were acquired using a rapid novel 3D radial ZTE sequence [16] with sequence parameters: flip angle 1°, receive bandwidth 694 Hz pixel^{-1}, nominal field of view $360 \times 360 \times 300$ mm^3, resolution $2.0 \times 2.0 \times 2.0$ mm^3, repetition time *TR* = 1.06 ms, nominal echo time *TE* = 0.016 ms, 59,392 3D center-out radial-spokes and acquisition time 65 s. Fat-water chemical shift effects were minimized using in-phase ZTE by adjusting the centre frequency between fat and water [13]. Image reconstruction was based on 3D gridding, including two-fold field of view extension to $720 \times 720 \times 600$ mm^3 (enabled by two-fold radial oversampling), DL-based de-noising and de-ringing [17] and 3D geometry correction. The scanner geometric accuracy was tested monthly during a radiotherapy quality assurance programme [18].

Synthetic CT generation

The sCT was generated from the ZTE image using a Deep Learning sCT algorithm by GE Healthcare. This used a 2D Convolutional Neural Network U-NET model with a bone focused loss function, as described in [19]. The model had three tasks: whole image regression, bone segmentation and image value regression within the bone region. The logic was to separate segmentation and regression tasks and optimise the model to simultaneously reduce errors in both, each task implicitly reinforcing the other [20]. Each task was driven by an individualised loss function and generated an associated output image. The three output images were then combined using the voxel intensity values from the whole image regression output, except for voxels within the bone segmentation output, which were assigned values from the bone regression output. The model was trained with patients from the training cohort, with ZTE and CT images registered together using an affine transformation. The ZTE images were bias corrected and normalised to ensure consistency across the dataset.

Hounsfield Unit analysis

Two registrations were calculated between the CT and sCT, a rigid and a deformable registration. Both were generated automatically in RayStation (v9, RaySearch Laboratories, Stockholm, Sweden) and involved registering and resampling the sCT to the CT matrix. The rigid registration ensured the sCT was evaluated as it would be used in a clinical MR-only pathway [10] and the deformable registration removed the confounding effect of small differences in external contour arising from the images being acquired in two imaging sessions [21].

The Mean Absolute Error (MAE) was calculated between the deformed and resampled sCT and CT as the absolute difference in Hounsfield Unit (HU) for each voxel. Any air pockets in the CT were over-ridden with water density. The MAE was calculated within the intersection of the external contours of the CT and the deformed sCT to account for discrepancies in the longitudinal extent of the images using a threshold of − 250 HU. Within this external contour the soft-tissue and bone regions were automatically contoured using thresholds of [−250 HU, 250 HU] and [250 HU, 4000 HU] respectively. The MAE was calculated within the external contour, soft-tissue and bone regions separately.

Comprehensive sCT dose evaluation with artificial plans

This aimed to evaluate the sCT at different points along the superior-inferior axis using all patients, providing dose accuracy measurements that would be relevant to all pelvic radiotherapy treatments. The sCT was calibrated using a HU-mass density curve measured on the CT scanner used for the training CT cases. The CT was calibrated using data measured on the clinical CT scanner. Any air pockets were over-ridden to water density. Patients who had received a contrast-enhanced CT had the contrast delineated and set to a mass density of 1.015 g cm^{-3}, the density of artery tissue.

The external contour was automatically determined for both sCT and CT using a threshold of − 250 HU.

Four separate isocentres were positioned on the CT, longitudinally separated by 5 cm. The first isocentre was positioned midpoint between the femoral heads and labelled FH. A further three points were determined with the same left–right and anterior-posterior position and 5 cm inferior (FH-5), 5 cm (FH + 5) and 10 cm (FH + 10) superior (see Fig. 1). One patient had the most inferior CT slice copied 10 times, lengthening the CT by 3 cm to ensure sufficient lateral scatter for accurate dose calculation on the CT for the FH-5 isocentre. The CT extent was not sufficient for the FH-5 isocentre in one patient and for the Fh + 10 isocentre in 8 patients. A $2 \times 2 \times 2$ mm^3 dose grid was used which covered the external contours of both CT and sCT.

Four cylindrical artificial PTVs 5 cm long with a diameter of 10 cm were drawn on the CT centred on each isocentre (Fig. 1) ensuring continuous coverage in the superior-inferior direction. The diameter of 10 cm was selected as approximately representative of pelvic PTVs. A 6 MV single 360° arc volumetric modulated arc therapy plan was optimised on the CT to deliver.

50 Gy in 25 fractions to the PTV centred on the FH isocentre. A general dose fall-off function was used to ensure conformality of the high dose region to the PTV. This plan was recalculated on the sCT. The difference in dose to the PTV median dose (D50%), near-minimum (D98%) and near-maximum (D2%) were calculated as a percentage of the prescription dose [22]. The treatment plan were then copied to the other three isocentres and recalculated on the CT and sCT. This analysis was repeated for both rigid and deformable registrations.

Site-specific sCT dose evaluation with clinical plans

An additional dose comparison was carried out using the clinical treatment plan to ensure clinical relevance using both rigid and deformable registrations. The clinical PTV(s) and relevant clinical OARs were copied from the CT to the sCT and the clinical treatment plan recalculated on the sCT. The difference in dose to the PTV D2%, D50% and D98% and relevant OAR constraints were calculated. A 3D global gamma analysis was performed using the Medical Interactive Creative Environment Toolkit (version 2021.1.2, Umeå University, Sweden) [23] within the following structures: the external contour, the union of all patient PTVs (primary, nodal and elective if present) and the volume enclosed by the 50% isodose line of the prescription dose. The union of patient PTVs ensured all the high dose regions were included within the PTV evaluation. The gamma analysis was performed twice with criteria 1%/1 mm and 2%/2 mm, with all points below 10% of the prescription dose excluded.

Results

56 patients were included within the study, with patients planned for anal (n = 12), rectal (n = 11) and prostate (n = 33) cancers. There were 13 female and 43 male patients. Patients were divided into training (n = 36) and evaluation (n = 20) cohorts, with female/male split of 7/29 and 6/14 for each cohort respectively. The treatment sites were similarly split between training and evaluation cohorts: 6/6 (anus), 7/4 (rectum) and 23/10 (prostate).

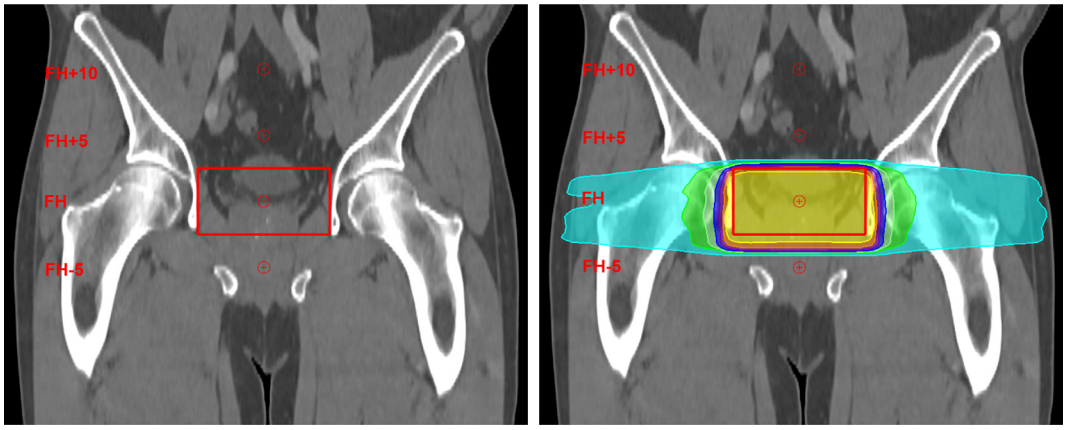
sCTs were successfully generated for all patients. Representative reconstructed ZTE and corresponding sCT and CT images are shown in Fig. 2. The mean MAE within the external contour was 36 ± 3 HU (± standard deviation). The MAE in the soft-tissue region was 7 ± 1 HU and in the bone region 98 ± 14 HU.

The comprehensive sCT dose evaluation with artificial plans had small mean dose differences at the different FH isocentre points. The dose differences were $\leq$ 0.7% using the rigid registration and the deformable registration reduced the mean dose differences to $\leq$ 0.6% and also reduced the range of dose differences (Table 1 and Fig. 3).

The mean dose differences for the site-specific sCT dose evaluation with clinical plans at clinically relevant DVH points for the PTV(s) and clinically relevant OARs are shown in Table 2. Gamma pass rates at 2%/2 mm within the external contour for the rigid registration were highest for the prostate patients, 99.0 ± 0.2% (97.4%, 99.9%), followed by rectum, 98.7 ± 0.5% (97.6%, 99.5%), and anus patients 98.3 ± 0.3% (97.2%, 99.4%), (Fig. 4). Gamma pass rates at 1%/1 mm were lower: 98.0 ± 0.4% (95.6%, 99.4%), 96.5 ± 0.8% (94.9%, 98.0%) and 95.4 ± 0.6% (93.8%,98.3%) for the prostate, rectum and anus patients respectively. The equivalent results for the deformable registration were much higher, 100.0 ± 0.0% (100.0%, 100.0%), 100.0 ± 0.0% (99.9%, 100.0%) and 99.9 ± 0.1% (99.4%, 100.0%) [2%/2 mm] and 100.0 ± 0.0% (99.9%, 100.0%), 99.8 ± 0.1% (99.6%, 99.9%) and 99.4 ± 0.4% (97.4%, 100.0%) [1%/1 mm] respectively for the prostate, rectum and anus patients.

Discussion

A prospective study was performed to train and evaluate the dose calculation accuracy of a Deep Learning sCT algorithm for pelvic MR-only radiotherapy, both comprehensively for all pelvic sites



(a) Isocentre Points (b) Dummy Plan

Fig. 1. Example CT showing the four isocentre points marking the FH-5, FH, FH + 5 and FH + 10 levels (a) and the cylindrical PTV (red line) and dummy plan dose at the FH level (b).

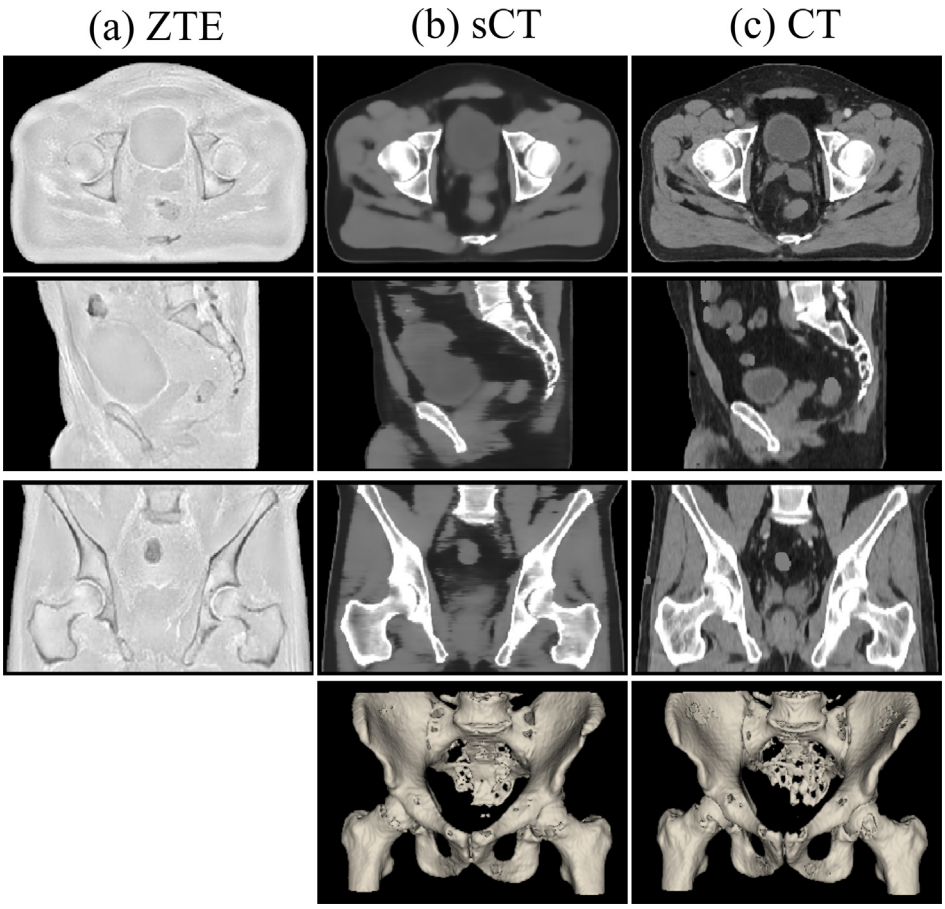


Fig. 2. Example ZTE images (a) and corresponding sCT images (b) for a patient in the evaluation cohort. CT images for the same patient are shown in (c).

Table 1
Dose differences for artificial plans. Cylindrical PTV dose differences at the DVH constraints for the FH isocentre plans. The number of patients included for the analysis at each FH isocentre is indicated. Dose differences using the rigid and deformable registrations reported as mean ± standard error (minimum, maximum) as percentage of prescribed dose (50 Gy).

Isocentre	Patients	Constraint	Mean Dose Difference / % *Rigid Registration*	*Deformable Registration*
FH + 10	12	D2%	−0.7 ± 0.1 (−1.7,0.1)	−0.6 ± 0.1 (−1.4,0.1)
		D50%	-0.6 ± 0.1 (−1.5,0.2)	−0.5 ± 0.1 (−1.2,0.1)
		D98%	−0.5 ± 0.1 (−1.2,0.1)	−0.5 ± 0.1 (−1.4,0.0)
FH + 5	20	D2%	−0.3 ± 0.2 (−1.8,1.8	−0.5 ± 0.1 (−1.5,0.1)
		D50%	−0.3 ± 0.2 (−1.6,1.7)	−0.5 ± 0.1 (−1.3,0.0)
		D98%	−0.3 ± 0.2 (−2.0,1.9)	−0.4 ± 0.1 (−1.3,0.1)
FH	20	D2%	0.1 ± 0.2 (−2.0,2.6)	−0.3 ± 0.1 (−1.1,0.5)
		D50%	0.0 ± 0.2 (−2.0,1.5)	−0.3 ± 0.1 (−1.2,0.4)
		D98%	−0.2 ± 0.1 (−1.8,1.0)	−0.3 ± 0.1 (−1.3,0.3)
FH-5	19	D2%	0.3 ± 0.2 (−1.7,2.3)	−0.1 ± 0.1 (−0.8,0.4)
		D50%	0.2 ± 0.2 (−1.7,2.4)	−0.2 ± 0.1 (−0.8,0.4)
		D98%	0.1 ± 0.2 (−1.8,1.6)	−0.2 ± 0.1 (−0.7,0.3)

and for the clinical prostate and ano-rectal patient plans included in this study. The mean PTV D98% dose differences were ≤ 0.5% for all four isocentre plans (comprehensive evaluation) as well as for the clinical plans when rigid registration was used for evaluation. The deformable registration reduced the range of D98% dose differences to within [−1.4%, 0.3%] and [−0.9%, 0.4%] for the comprehensive and clinical evaluations respectively.

The comprehensive dose evaluation demonstrated small dose differences (≤0.7%) using the rigid registration for each isocentre level. There were six D50% dose differences greater than ± 1.5%, with the FH-5 cm point having the largest differences (Fig. 3). However using the deformable registration reduced all these differences below 1.5%. This suggests that the primary reason for the larger dose differences for some patients was not incorrect HU assignment but differences in patient positioning between sCT and CT, which maybe particularly problematic at the FH-5 point due to differences in femur angle. The deformable registration results did suggest a small systematic bias to negative dose differences for all four isocentre plans, with means of − 0.5% and an asymmetric range around zero [−1.4%, 0.3%]. However this difference is small and all differences were within the ± 2.0% often suggested as the threshold above which dose differences are considered clinically significant [24].

The clinical plan dose differences were even smaller, with all mean dose differences ≤ 0.4% for the rigid registration (Table 2). Gamma pass rates with criteria 2%/2 mm were high, with all mean

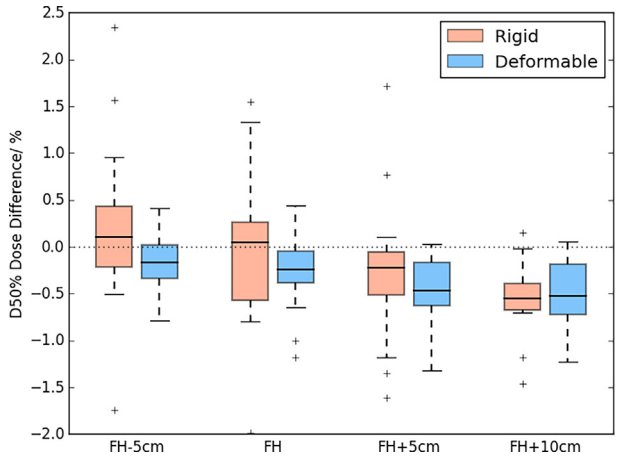


Fig. 3. Boxplots of D50% dose differences within the true external contour (orange) and intersection external contour (blue) for each FH point.The dotted line indicates zero dose difference. The rectangles indicate the interquartile range (IQR), with the horizontal black line the median value, the black whiskers the maximum (minimum) data point within $Q3 + 1.5IQR$ ($Q1 - 1.5IQR$) and the black crosses outlier data points.

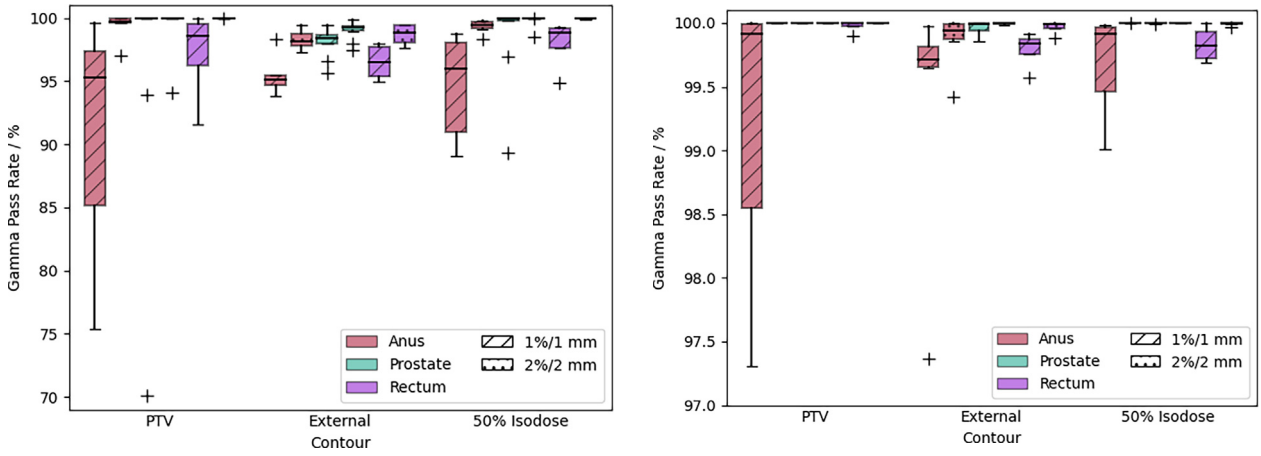
pass rates $\geq$ 98.3% within all contours and for all treatment sites using the rigid registration. At 1%/1 mm, pass rates were lower, especially within the PTV(s). The anus patients had the lowest pass rate, probably due to having the most inferior PTVs where there were the largest discrepancies between CT and sCT external contours. The deformable registration resulted in much higher gamma pass rates, especially at 1%/1 mm, with mean pass rates $\geq$ 99.4%. This again suggests most of the differences in dose distribution were due to differences in patient position rather than HU assignment.

There was good agreement between clinical PTV dose differences and the closest FH isocentre point to that clinical PTV. The prostate PTVs were located between FH-5 and FH isocentres, and the mean PTV D50% dose difference was within 0.3% of the mean dose difference for both those isocentres. The dose differences of the FH isocentres near the rectum primary (FH) and elective (FH + 5) PTVs, and the anus primary (FH-5), elective (FH + 5) and nodal (FH + 10) PTVs were all within 0.2%. This implies that the comprehensive FH plan evaluation was clinically relevant to the anus, prostate and rectum patients studied. This can be extended to suggest that this method has demonstrated that the sCT evaluated in this study would be appropriate for all pelvic radiotherapy sites, including patients not included in this study such cervix cancer and nodal prostate cancer. Accordingly, this method is a very efficient way of evaluating sCT dose accuracy for a number of clinical sites with a relatively small number of patients. This would have particularly benefit for rare treatment sites or smaller centres looking to validate sCT algorithms, which may struggle to produce large enough patient cohorts.

A key question for MR-only radiotherapy is what magnitude of dose differences would be considered clinically significant. Over the whole treatment pathway the dosimetric uncertainty should be within 3% and the geometric uncertainty 2–4 mm [25]. This requires the uncertainty of any individual component to be < 1.0% and < 1.0 mm to not significantly increase the overall uncertainty [25]. The MR-CT registration uncertainty removed by the MR-only process is estimated at 2 mm for the pelvis [26] and the mean dose differences for both FH isocentre and clinical plans on both rigid and deformable registrations were $\leq$ 0.7%. This suggests that MR-only radiotherapy using this sCT algorithm would

Table 2
Dose differences for clinical plans. Mean dose differences at DVH constraints for different regions of interest. The number of patients is shown for each structure in brackers. It varies for the different PTVs for the ano-rectal patients as not all patients had elective or nodal volumes. All results given as mean ± standard error (minimum, maximum) as percentage of prescription dose.

Site	Structure	Constraint	Mean Dose Difference / *% Rigid RegistrationDeformable Registration*	
Anus	PTV Primary (6)	D2%	0.3 ± 0.4 (−1.1,2.0)	−0.3 ± 0.1 (−0.7,0.2)
		D50%	0.1 ± 0.4 (−1.5,1.1)	−0.3 ± 0.1 (−0.7,0.2)
		D98%	−0.1 ± 0.3 (−1.6,0.5)	−0.3 ± 0.1 (−0.8,0.2)
	PTV Nodal (4)	D2%	−0.3 ± 0.4 (−1.2,0.3)	−0.3 ± 0.2 (−0.8,0.1)
		D50%	−0.4 ± 0.3 (−1.1,0.3)	−0.3 ± 0.2 (−0.9,0.0)
		D98%	−0.4 ± 0.3 (−1.0,0.2)	−0.3 ± 0.2 (−0.9,0.0)
		D2%	−0.1 ± 0.2 (−1.2,0.4)	−0.3 ± 0.1 (−0.8,0.1)
		D50%	−0.1 ± 0.2 (−0.8,0.3)	−0.3 ± 0.1 (−0.5,−0.0)
		D98%	−0.2 ± 0.2 (−0.8,0.3)	−0.2 ± 0.1 (−0.4,0.0)
	Bladder (6)	D50%	0.0 ± 0.3 (−1.2,1.0)	−0.2 ± 0.1 (−0.6,−0.1)
	Small Bowel (6)	D30%	−0.4 ± 0.2 (−1.0,0.2)	−0.3 ± 0.1 (−0.6,−0.0)
Prostate	PTV (10)	D2%	0.3 ± 0.2 (−0.5,1.7)	−0.1 ± 0.1 (−0.4,0.7)
		D50%	0.3 ± 0.2 (−0.5,1.3)	−0.1 ± 0.1 (−0.4,0.6)
		D98%	0.3 ± 0.1 (−0.3,0.9)	−0.1 ± 0.1 (−0.3,0.4)
	Bladder (10)	D50%	0.0 ± 0.0 (−0.3,0.1)	0.0 ± 0.0 (−0.2,0.2)
	Rectum (10)	D30%	0.0 ± 0.1 (−0.3,0.8)	0.0 ± 0.0 (−0.1,0.1)
Rectum	PTV Primary (4)	D2%	0.0 ± 0.2 (−0.5,0.3)	−0.0 ± 0.1 (−0.2,0.1)
		D50%	−0.2 ± 0.2 (−0.6,0.2)	−0.1 ± 0.1 (−0.3,0.1)
		D98%	−0.2 ± 0.2 (−0.6, 0.1)	−0.2 ± 0.1 (−0.3, 0.0)
	PTV Elective (2)	D2%	−0.1 ± 0.4 (−0.6,0.3)	0.1 ± 0.1 (0.0,0.1)
		D50%	−0.2 ± 0.4 (−0.6,0.2)	0.0 ± 0.1 (−0.0,0.1)
		D98%	−0.4 ± 0.3 (−0.6,−0.1)	0.0 ± 0.0 (−0.1,0.0)
	Bladder (4)	D50%	−0.1 ± 0.1 (−0.4,0.2)	−0.1 ± 0.1 (−0.2,0.0)
	Bowel Bag (4)	D30%	−0.4 ± 0.1 (−0.7,−0.2)	−0.2 ± 0.1 (−0.3,−0.1)



(a) Rigid Registration

(b) Deformable Registration

Fig. 4. Boxplots of gamma pass rates for clinical plans within the PTV, external contour and 50% isodose contour for the anus (red bars, n = 6), prostate (green bars, n = 10) and rectum (purple bars, n = 4). Gamma pass rates with criteria 1%/1 mm are shown with right diagonal hatching and with 2%/2 mm with dotted hatching. One outlier result for the prostate pass rate with 1%/1 mm criteria is not shown (53.8%). The rectangles indicate the interquartile range (IQR), with the horizontal black line the median value, the black whiskers the maximum (minimum) data point within $Q3 + 1.5IQR$ ($Q1 - 1.5IQR$) and the black crosses outlier data points.

significantly reduce geometric uncertainty without a significant change in dosimetric uncertainty. This suggests that the sCT is dosimetrically sufficiently accurate for clinical use for most patients, although the larger dose difference for one patient implies some form of patient-specific QA of sCT dose calculation accuracy would be warranted [27].

The results from this study compare well with other Deep Learning sCT approaches in the pelvis. Maspero et al. reported dose differences of 0.4% and 0.4% to the PTV D50% and gamma pass rates at 2%/2 mm of 95 ± 2% and 92 ± 3% for prostate and rectum patients respectively [9]. These were larger differences than found in this study. Similarly, Yoo et al. found PTV D50% dose differences for prostate patients of 0.4 ± 0.3% and gamma pass rate (2%/2 mm) of 93 ± 4% [11]. A study looking at ano-rectal patients reported mean PTV dose differences of − 0.1%, with a 1%/1 mm gamma pass rate of 96% [10]. These agreed well with the dose differences and gamma pass rates for the rectum and anus patients reported here. Fetty et al. evaluated five different Deep Learning algorithms on 0.35 T, 1.5 T and 3 T MR images [28]. The median PTV doses differences for all algorithms on all images were $\leq$ 0.5%, in agreement with the results found here. In a 2021 review of Deep Learning sCT algorithms for pelvic MR-only radiotherapy, only 4/8 studies reported doses differences within the $\leq$ 0.5% found in this study [8]. This suggests that the sCT algorithm performs at least as well as those presented in the literature as sufficiently accurate for clinical use.

The use of a ZTE image for the generation of the sCT requires using an additional dedicated MR sequence to be acquired which does not have any other use in the radiotherapy planning process. In contrast other solutions use T2-weighted images which are also considered clinically optimal for target and/or OAR delineation [21]. However, using a dedicated fast ZTE sequence allows the other clinical scans to be optimised for delineation tasks without additional constraints such as large volume coverage required for dose calculation. This can for example be used to improve the image resolution without prohibitively long scan times. And since the ZTE sequence is very fast, 65 s, effectively this might result in overall similar acquisition time but improved imaging results.

A limitation of this study is that only data from a single centre has been used. This is particularly an issue for Deep Learning algorithms where over-fitting is a recognised problem, leading to poor generalisation of the algorithm [29]. Future work could extend this evaluation to patient images acquired at different centres on different scanners to validate the generalisability of the results reported here. In addition, the model training data could be extended to include images from other centres to improve the robustness of the algorithm. An important aspect of MR-only radiotherapy that this paper has not considered is the choice of reference image for image matching with Cone Beam (CB)CT on the treatment machine. Image matching using sCT [30] or MR [31] as the reference image could be investigated. A potential advantage of ZTE-based sCT images is the improved accuracy of bone generation which may facilitate improved sCT matching to CBCT.

In conclusion, a ZTE-based Deep Learning algorithm successfully generated sCTs for all patients. Mean dose differences to the PTV D98% in the comprehensive dose analysis were $\leq$ 0.5% for all FH isocentre plans using the rigid registration. For the clinical plans PTV D98% dose differences were similarly small, $\leq$±0.4%, with mean gamma pass rates at the stringent criteria of 1%/1 mm being 98.0 ± 0.4%, 96.5 ± 0.8% and 95.4 ± 0.6% for the prostate, rectum and anus patients respectively. These mean dose differences were < 1%, ensuring there is not a significant increase in overall dosimetric uncertainty in the radiotherapy pathway. In addition, the comprehensive dose analysis method demonstrated the sCT accuracy throughout the pelvis, not just for the clinical sites included in this study. This suggests the sCT is sufficiently accurate for clinical use in MR-only radiotherapy for all pelvic sites.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is part of the Deep MR-only Radiation Therapy activity (project numbers: 19037, 20648, 210995) that has received funding from EIT Health. EIT Health is supported by the European Institute of Innovation and Technology (EIT), a body of the European Union and receives support from the European Union's Horizon 2020 Research and innovation program.

Conflicts of Interest

Sandeep Kaushik, Cristina Cozzini and Florian Wiesinger are employees of GE Healthcare.

References

[1] Bird D, Henry AM, Sebag-Montefiore D, Buckley DL, Al-Qaisieh B, Speight R. A systematic review of the clinical implementation of pelvic magnetic resonance imaging-only planning for external beam radiation therapy. Int J Radiat Oncol Biol Phys 2019;105:479–92. https://doi.org/10.1016/j.ijrobp.2019.06.2530.

[2] Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. Radiat Oncol 2017;12:28.

[3] Johnstone E, Wyatt JJ, Henry AM, Short SC, Sebag-Montefiore D, Murray L, et al. A 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.

[4] Persson E, Gustafsson C, Nordstr¨om F, Sohlin M, Gunnlaugsson A, Petruson K, et al. MROPERA–A Multi-center/multi-vendor validation of MRI-only prostate treatment planning using synthetic CT images. Int J Radiat Oncol Biol Phys 2017;99:692–700

[5] Tyagi N, Fontenla S, Zhang J, Cloutier M, Kadbi M, Mechalakos J, et al. Dosimetric and workflow evaluation of first commercial synthetic CT software for clinical use in pelvis. Phys Med Biol 2017;62:2961.

[6] Gonzalez-Moya A, Dufreneix S, Ouyessad N, Guillerminet C, Autret D. Evaluation of a commercial synthetic computed tomography generation solution for magnetic resonance imaging-only radiotherapy. J Appl Clin Med Phys 2021;22:191–7. https://doi.org/10.1002/acm2.13236.

[7] Spadea MF, Maspero M, Zaffino P, Seco J. Deep learning based synthetic-CT generation in radiotherapy and PET: a review. Med Phys 2021;48:6537–66. https://doi.org/10.1002/mp.15150.

[8] Boulanger M, Nunes JC, Chourak H, Largent A, Tahri S, Acosta O, et al. Deep learning methods to generate synthetic CT from MRI in radiotherapy: a literature review. Phys Med: Eur J Med Phys 2021;89:265–81. https://doi.org/10.1016/j.ejmp.2021.07.027.

[9] Maspero M, Savenije MHF, Dinkla AM, Seevinck PR, Intven MPW, Jurgenliemk-Schulz IM, 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:. https://doi.org/10.1088/1361-6560/aada6d185001.

[10] Bird D, Nix MG, McCallum H, Teo M, Gilbert A, Casanova N, et al. Multicentre, deep learning, synthetic-CT generation for ano-rectal MR-only radiotherapy treatment planning. Radiother Oncol 2021;156:23–8. https://doi.org/10.1016/j.radonc.2020.11.027.

[11] Yoo GS, Luu HM, Kim H, Park W, Pyo H, Han Y, et al. Feasibility of synthetic Computed Tomography images generated from Magnetic Resonance imaging scans using various Deep Learning methods in the planning of radiation therapy for prostate cancer. Cancers 2022;14:40. https://doi.org/10.3390/cancers14010040.

[12] Wiesinger F, Sacolick LI, Menini A, Kaushik SS, Ahn S, Veit-Haibach P, et al. Zero TE MR bone imaging in the head. Magn Reson Med 2016;75:107–14. https://doi.org/10.1002/mrm.25545.

[13] Engstr¨om M, McKinnon G, Cozzini C, Wiesinger F. In-phase zero TE musculoskeletal imaging. Magnetic Resonance in Medicine 2020;83:195–202. https://doi.org/10.1002/mrm.27928

[14] Ljungberg E, Damestani NL, Wood TC, Lythgoe DJ, Zelaya F, Williams SCR, et al. Silent zero TE MR neuroimaging: current state-of-the-art and future directions. Prog Nucl Magn Reson Spectrosc 2021;123:73–93. https://doi.org/10.1016/j.pnmrs.2021.03.002.

[15] Leynes AP, Yang J, Wiesinger F, Kaushik SS, Shanbhag DD, Seo Y, et al. ZeroEcho-time and dixon deep pseudo-CT (ZeDD CT): direct generation of PseudoCT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. J Nucl Med 2018;59:852–8. https://doi.org/10.2967/jnumed.117.198051.

[16] Wiesinger F, Bylund M, Yang J, Kaushik S, Shanbhag D, Ahn S, et al. Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning. Magn Reson Med 2018;80:1440–51. https://doi.org/10.1002/mrm.27134.

[17] Lebel RM. Performance characterization of a novel deep learning-based MR image reconstruction pipeline. arXiv:200806559 [cs, eess] 2020;.

[18] Wyatt JJ, McCallum HM, Maxwell RJ. Developing quality assurance tests for simultaneous Positron Emission Tomography – Magnetic Resonance imaging for radiotherapy planning. Phys Imaging Radiat Oncol 2022;22:28–35. https://doi.org/10.1016/j.phro.2022.03.003.

[19] Kaushik S, Bylund M, Cozzini C, Shanbhag D, Petit SF, Wyatt JJ, et al. Region of Interest focused MRI to Synthetic CT Translation using Regression and Classification Multi-task Network. arXiv:220316288 [physics] 2022;.

[20] Caruana R. Multitask Learning. Mach Learn 1997;28:41–75. https://doi.org/10.1023/A:1007379606734.

[21] Siversson C, Nordstro¨m F, Nilsson T, Nyholm T, Jonsson J, Gunnlaugsson A, et al. Technical Note: MRI only prostate radiotherapy planning using the statistical decomposition algorithm. Med Phys 2015;42:6090–6097

[22] Gŕegoire V, Mackie T, De Neve W, Gospodarowicz M, Purdy J, van Herk M, et al. State of the art on dose prescription, reporting and recording in Intensity-Modulated Radiation Therapy (ICRU report No. 83). Cancer Radiother 2011;15:555–559

[23] Nyholm T, Berglund M, Brynolfsson P, Jonsson J. EP-1533: ICE-Studio - an Interactive visual research tool for image analysis. Radiother Oncol 2015;115:S837. https://doi.org/10.1016/S0167-8140(15)41525-7.

[24] Korsholm ME, Waring LW, Edmund JM. A criterion for the reliable use of MRI-only radiotherapy. Radiat Oncol 2014;9:16.

[25] Thwaites D. Accuracy required and achievable in radiotherapy dosimetry: Have modern technology and techniques changed our views? J Phys Conf Ser 2013;444:012006.

[26] Nyholm T, Nyberg M, Karlsson MG, Karlsson M. Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments. Radiat Oncol 2009;4.

[27] Wyatt JJ, Pearson RA, Walker CP, Brooks RL, Pilling K, McCallum HM. Cone beam computed tomography for dose calculation quality assurance for magnetic resonance-only radiotherapy. Phys Imaging Radiat Oncol 2021;17:71–6. https://doi.org/10.1016/j.phro.2021.01.005.

[28] Fetty L, L¨ofstedt T, Heilemann G, Furtado H, Nesvacil N, Nyholm T, et al. Investigating conditional GAN performance with different generator architectures, an ensemble model, and different MR scanners for MR-sCT conversion. Phys Med Biol 2020;65:105004. https://doi.org/10.1088/1361-6560/ab857b

[29] Shen C, Nguyen D, Zhou Z, Jiang SB, Dong B, Jia X. An introduction to deep learning in medical physics: advantages, potential, and challenges. Phys Med Biol 2020;65:05TR01. https://doi.org/10.1088/1361-6560/ab6f51.

[30] Bird D, Beasley M, Nix MG, Tyyger M, McCallum H, Teo M, et al. Patient position verification in magnetic-resonance imaging only radiotherapy of anal and rectal cancers. Phys Imaging Radiat Oncol 2021;19:72–7. https://doi.org/10.1016/j.phro.2021.07.005.

[31] Wyatt JJ, Brooks RL, Ainslie D, Wilkins E, Raven E, Pilling K, et al. The accuracy of Magnetic Resonance – Cone Beam Computed Tomography soft-tissue matching for prostate radiotherapy. Phys Imaging Radiat Oncol 2019;12:49–55. https://doi.org/10.1016/j.phro.2019.11.005.