Clinical and Translational Radiation Oncology 58 (2026) 101122

Contents lists available at ScienceDirect

Clinical and Translational Radiation Oncology

journal homepage: www.sciencedirect.com/journal/clinical-and-translational-radiation-oncology

Original Research Article

Automated MR-only radiotherapy outperforms CT-based radiotherapy and decreases hands-on time for head-and-neck cancer treatment

Iris Lauwers [a,*], Marta Capala [a], Sandeep Kaushik [b,c], László Ruskó [d], Cristina Cozzini [b], Eszter Szabó [d], Ádám Kékesi [d], Borbála Deák-Karancsi [d], Jonathan Wyatt [e,f], Rachel Pearson [f], Gerda Verduijn [a], Florian Wiesinger [b], Juan Hernandez-Tamames [g,h], Steven Petit [a]

[a] Department of Radiotherapy, Erasmus MC Cancer Institute, University Medical Center Rotterdam, Rotterdam, Netherlands (the)
[b] GE HealthCare, Munich, Germany
[c] Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland
[d] GE Healthcare Magyarország Kft., Budapest, Hungary
[e] Translational and Clinical Research Institute, Newcastle University, Newcastle, UK
[f] Northern Centre for Cancer Care, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle, UK
[g] Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, Netherlands (the)
[h] Department of Imaging Physics, TU Delft, Delft, Netherlands (the)

ARTICLE INFO

Keywords:
MR-only
ZTE
Head and neck
Automatic OAR delineations
Synthetic CT
Dose comparison

ABSTRACT

Introduction: Cancer incidence is expected to increase in Europe by 18% in eighteen years. To account for the increasing patient numbers, the workload per patient needs to be reduced. One step towards future-proof radiotherapy is automated MR-only radiotherapy as it could eliminate the need for (i) a planning CT and (ii) for manual organ at risk (OAR) delineations. The aim of this study was to evaluate the feasibility of an automated MR-only workflow for head-and-neck radiotherapy.

Method: Automated MR-only radiotherapy consisted of a Zero-Echo-Time-based synthetic CT for dose calculations and automated T2w-based OAR delineations. Automated MR-Only RT was compared to the clinical workflow consisting of CT-based dose calculation and CT-based OAR delineations. Both approaches were benchmarked to a gold standard consisting of the planning CT for dose calculations and manual delineations on the T2w MR scan. Dice similarity coefficients (DSC), 95% Hausdorff distances and absolute DVH metrics were compared between the clinical and MR-only workflow using a linear mixed-effect model. A p-value < 0.05 was deemed significant.

Results: Seventeen head-and-neck cancer patients were included. The automated MR-only delineations were more accurate compared to the clinical CT delineations (DSC of 0.79 vs. 0.67; 95% Hausdorff distance 4.0 vs 5.8 mm (p-values < 0.001)). The average dose calculation errors of the automated MR only RT were smaller than the clinical workflow ($+0.34$ Gy vs. -1.39 (p-value < 0.01)).

Discussion: The automated MR-only head-and-neck radiotherapy workflow was more accurate than the standard CT based clinical workflow, demonstrating the feasibility of automated MR-only RT to decrease the workload for head-and-neck RT treatment preparation.

Introduction

The number of cancer patients is expected to increase with 18.4% worldwide between 2022 and 2040 [1]. Meanwhile, the medical workforce has to increase by 66% between 2013 and 2030 to keep up with the increasing health demand, according to the World Health Organization [2]. However, the workforce in healthcare is not expected to match such rapid growth [2]. This means there is an enormous need to reduce the hands-on time per patient to be able to keep on treating all patients and ensuring the wellbeing of both patients and medical personnel.

One strategy to reduce hands-on time per patient is MR-only

* Corresponding author at: Department of Radiotherapy, Erasmus MC Cancer Institute, P.O. Box 2040, 3000 CA Rotterdam, Netherlands (the).

https://doi.org/10.1016/j.ctro.2026.101122
Received 20 November 2025; Received in revised form 6 January 2026; Accepted 7 February 2026
Available online 13 February 2026

radiotherapy (RT) for treatment sites that require acquisition of both a computed tomography (CT) and magnetic resonance (MR) scan. With MR-only radiotherapy the need for a planning CT scan is eliminated by replacing it with an MR-derived synthetic CT (synCT) scan. Therefore, MR-only RT decreases the treatment preparation workload and cost. Other anticipated advantages of MR-only RT could include: improved accuracy by eliminating the registration uncertainty between the MR and CT images, as it may negatively impact target delineation accuracy; accelerating the start of treatment; eliminating the concomitant dose caused by the planning CT; decreasing the burden on the patient as the patients do not need a planning CT scan; and benefiting adaptive real-time MR-guided radiotherapy treatment, such as using in MRI-linac system (when using low-field-MR-based synCT [3,4]) [5–11].

To introduce MR-only RT into the clinic, both target and organs at risk (OARs) contours should be delineated on the MR images. This presents another opportunity to improve patient care and to reduce the workload. Due to the increased soft tissue contrast of the T2 compared to the CT, the visibility of most OARs is comparable or better on an MR image compared to CT [12]. Therefore, the MR-based delineation could potentially enable completely automated OAR delineations without requiring adaptations from the radiotherapy staff, thus decreasing the workload and potentially improving the OAR delineations. In this study, we refer to the combination of a synCT for dose planning and automated OAR delineation based on the MR as automated MR-only RT. A recent advancement facilitating the automated MR-only RT has been the development of MR contouring guidelines based on the DAHANCA and RTOG CT guidelines and adapted for MR in a consensus-based manner [12].

To the best of our knowledge, the feasibility of the fully automated MR-only radiotherapy workflow for the head-and-neck region has never been investigated. An automated MR-only workflow feasible for clinical implementation and workload reduction, should meet the following requirements: i) the scanning sequences used for the synCT generation should not vastly increase the clinical MR scanning time; ii) the automated MR-based OAR delineations (without manual adaption) should have at least a similar accuracy compared to the clinical OARs, which are in many centers delineated on the planning CT; iii) the dose accuracy of the complete automated MR-only workflow should be at least equal to the accuracy of the current clinical workflow.

Recent advances in synCT generation for the head-and-neck region show synCTs based on various MR sequences have clinically acceptable dose calculation accuracy when assessed on the CT-based clinical contours (*e.g.* [13–18]). When selecting an MR sequences for synCT generation, it is important to minimize the increase in scanning times, as scanning times are proportional to the number of patients that can be imaged during a shift, intra-scan motion, and the experienced discomfort of the head-and-neck immobilization masks [6]. An MRI sequence that fulfils requirement (i) is the Zero Echo Time (ZTE) scan proposed by Lauwers et al, of which the scanning time can be as low to 0:56 min for the head-and-neck region.

Currently, multiple MR-based automated OAR delineations models are described in the literature that compare similarity metrics between the automated and manual MR delineations [19–26]. Moreover, the study of Dinkla *et al* assessed automatic CT delineation on synCT images [27]. To the best of our knowledge, MRI-based automated OAR delineation for the head-and-neck region has not yet been compared to clinical OAR delineation, requirement (ii).

To this date, only few studies evaluated the dosimetric impact of automated delineation methods [23,28]; however, no comparison with the clinical workflow was made. Moreover, requirement (iii), the combination of automated MR-based OAR delineations with synCT yielding accurate dose calculations, has never been demonstrated for head-and-neck RT. Note that the performance of the combination of two methods, one for OAR delineation and the other for synCT generation, cannot be assessed thoroughly by their individual performance in isolation. For instance, similarity metrics alone are not directly translatable to the effect on treatment accuracy, as a small delineation error near a high dose gradient may have a large dosimetric impact. Therefore, the aim of this study was to determine the feasibility of automated MR-only RT for head-and-neck cancer based on the three requirements defined above, with the goal of reducing the radiotherapy treatment preparation workload.

Methods

Patients and treatment

This current research was part of the *Deep learning based MR Only Radiotherapy for head-and-neck cancer* clinical study approved by the Erasmus MC Institution Review Board [MEC-2019-0805] [29]. The inclusion criteria consisted of head-and-neck cancer patients who underwent a planning MRI for primary or post-operative RT and were at least eighteen years old. The exclusion criteria included any physical or mental health condition that interfered with the informed consent process or any contraindications for MRI, such as claustrophobia or metal implants. N = 19 head-and-neck patients treated were consecutively included between January 2022 and July 2024 to test the feasibility and accuracy of automated MR-only radiotherapy pipeline. Patients were treated using Volumetric-Modulated Arc Therapy (VMAT) according to standard clinical practice with the isocenter located in the center of the high dose PTV. Treatment planning was performed on the planning CT using Monaco 6.00.01 (Elekta AB, Stockholm, Sweden) using our automated treatment planning approach as described by Breedveld et al. [30]. The prescribed dose varied between 59 and 70 Gy and was delivered in 20–35 fractions.

CT and MRI scanning

Before the start of treatment, planning CT scans were acquired in treatment position with a radiotherapy immobilization mask according to the clinical practice on a Siemens SOMATOM Confidence (Siemens Healthineers, Erlangen, Germany). Next, patients were scanned in the treatment position with a radiotherapy immobilization mask on a 1.5 T GE MR450w using the GEM RT Head & Neck coil suite (GE HealthCare, Chicago, IL). A 2D T2 weighted (T2w) fast spin-echo (FSE) scan and a proton-density weighted (PDw) 3D Zero TE (ZTE) scan were acquired as described by Lauwers et al [14].

Automated MR-only workflow, clinical workflow, and the gold standard

The automated MR-only workflow consisted of a synCT for dose calculations and automated delineations based on the T2w FSE. The clinical workflow consisted of the CT for dose calculations and clinical OARs were generated by a manual correction of automated OAR delineations on the planning CT. The MR could be used to check the delineations if desired. Note that the clinical OAR delineations were influenced by uncertainty caused by the inferior soft tissue contrast compared to MRI. Therefore, the clinical workflow should not be considered as the most suitable benchmark to evaluate the performance of the automated MR-only workflow. Instead, we considered as gold standard: the planning CT in combination with manual delineations on the T2, as the planning CT represent the electron density better than the synCT, and completely manual T2 delineations are believed to be more accurate than both the automated T2 delineations and the clinical (mostly CT-based) delineations [12]. Both the clinical and MR-only workflow were compared to the gold standard.

Synthetic CT generation

The method to convert the ZTE MR scan to synCT scan was described by Kaushik et al and Lauwers et al [14,31]. The ZTE MR scan was selected as it can be acquired with a minimal addition in scanning time

and it provides measurable bone signal [31–34]. In short, a multi-task 2D U-net with weighted losses was used to convert the ZTE images to synCT images. The model was trained on 127 patients from different institutions and tested on the Erasmus MC patients used in this study. The mean absolute error and mean error on these synCTs were 94 ± 11 HU and +29 ± 11 HU respectively in the adapted total body contour on the test set. This is comparable to other state of the art synCT models. The gamma analysis of the clinical plans had an acceptance rate of 96.6 ± 1.7% and 99.6 ± 0.4% for the 1%1mm and 2%2mm respectively, thus outperforming most other synCT models [14].

Three sets of OAR delineations

For the automated MR-only workflow, T2-based automated OAR delineations were performed as described and trained by Czipczer et al. [35,36]. This method used a 2D U-net to localize structures and a 3D U-Net for the segmentation. A method based on the T2 scan was selected as it is part of the clinical workflow and the OAR are better visible on the T2 compared to the T1 [12].

Clinical delineations were made according to the standard clinical workflow in Erasmus MC. First, the automated delineation algorithm of MIM was used for auto contouring of OARs on the planning CT. This was followed by manual checking and editing by a radiotherapy technologist and an experienced radiation oncologist according to the guidelines [37].

For the gold standard, manual OAR delineations of the T2 scan were delineated by a clinician according to the previously established guidelines based on the RTOG and DAHANCA guidelines and adapted MR-guided contouring in a consensus-based manner and checked by an experienced radiation oncologist [12].

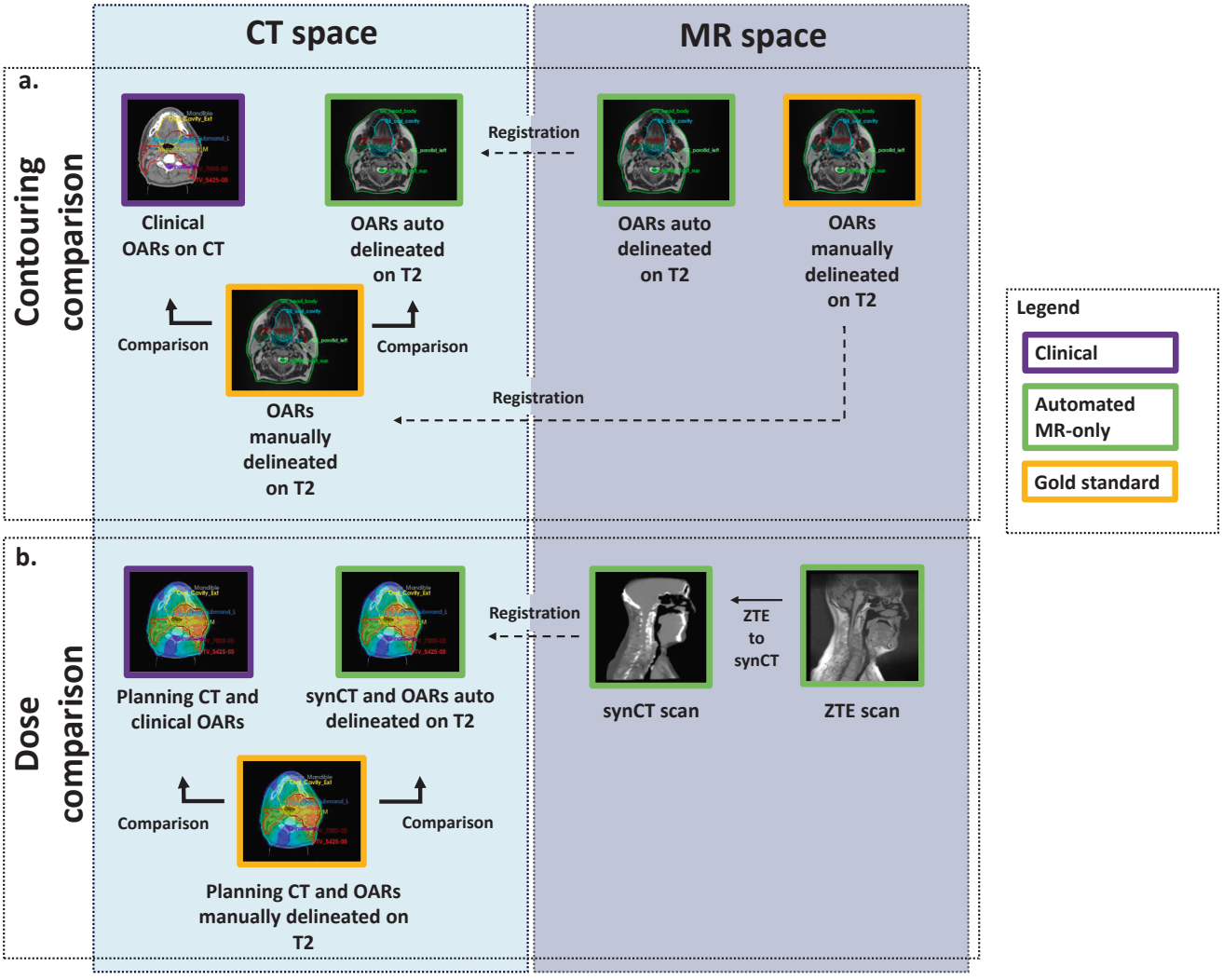

Fig. 1. A schematic overview of the comparison methods of the clinical (purple edge), automated MR-only (green edge) and gold standard (orange edge) workflows. The solid thin lines signify the deep learning model to convert the Zero Echo Time to a synthetic CT; the solid thick lines the comparisons that were made and the dotted arrows registration from the MR space (right) to the CT space (left). **a)** OAR contouring: the planning CT in the CT space was used for the clinical organ at risk delineations. The automated MR-only and gold standard organ at risk delineation were performed on the T2 in the MR space and registered to the CT space. Afterwards, the gold standard delineations were compared to the clinical and automated MR-only delineations. **b)** Dose comparison: the dose calculation for the clinical workflow and the gold standard were both based on the CT in the CT space. For the automated MR-only pipeline, the Zero Echo Time MRI was used to create a synthetic CT in the MR space. The synthetic CT was registered to the CT space and subsequently the dose was calculated on the synthetic CT. The calculated dose and the organ at risk delineations per workflows were combined in the CT space. Subsequently, the dose in the organs at risk of the gold standard was compared with the clinical and automated MR-only workflows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Comparing the automated MR-only and the clinical workflow with the gold standard

The comparisons are schematically described in Fig. 1. To be able to compare all delineations and doses in the same frame of reference, the manual T2, the automated OAR delineations, and the synCT were converted to the planning CT by non-rigid deformation of the T2w image to the planning CT. This means that there was a registration error between the clinical OAR and the gold standard. The influence of the registration error on the synthetic CT is deemed acceptable as the study of Lauwers et al. showed that the dose difference between in the OARs between the CT and the registered synthetic CT are small (on average 0.13 Gy) [14]. Moreover, the registered clinical OAR were checked by an experienced radiation oncologist and also deemed acceptable. The effect of the registration on the OAR was further investigated by comparing the error in centre of mass between the clinical and the manual T2 delineations (with registration) and between the automatic T2 and manual T2 delineations (no registration). If the registration error would be substantial, the error in centre of mass between the clinical and T2-based manual delineations would also be substantial.

OAR delineation comparison

Both the clinical and automated delineations were compared to the T2 manual gold standard delineation using the volumetric Dice similarity coefficient (DSC) (Equation (1)) and the 95% Hausdorff distance (requirement ii). To enable comparison, the T2w images were non-rigidly deformed to the planning CT to project the manual T2 OAR delineation on the frame of reference of the planning CT.

$$DSC = \frac{2|alternative\ OAR \cap gold\ standard\ OAR|}{|alternative\ OAR| + |gold\ standard\ OAR|} = \frac{2TP}{2TP + FN + FP} \quad (1)$$

where TP, FN and FP represent the true positive, false negative and false positive voxels respectively.

Dosimetric comparison

Next, the combined effect of synCT and automated OAR contouring was assessed (requirement (iii)). For that purpose, the performance of a combination of two methods, one for OAR delineation and the other for synCT generation, cannot be assessed thoroughly by their individual performance in isolation. The clinical plan was recalculated on the planning CT and on the synCT using the Monte Carlo dose algorithm of Scimoca 1.5.0.2821 (ScientificRT GmbH, Munich Germany) with a 2x2x2 mm grid spacing and a computational uncertainty of 0.5% ("extra fine"). Scimoca was used as it allows recalculation of the clinical plans with high accuracy in batch processing mode on both the planning CT and synCT. The dose distributions of the automated MR-only workflow and clinical workflow were compared to the gold standard using the following DVH dose metrics: Dmean for the parallel OARs and near maximum dose (D2%) for the serial OARs.

Statistics

The following statistical analyses were performed to assess the performance of the clinical workflow and the automated MR-only workflow compared to the gold standard. For the OAR delineations comparison, the volumetric DSC and the 95% Hausdorff distance of the clinical and automated MR-only workflow workflows were compared. For the dose comparison, the absolute dose errors of the two workflows were compared. These comparisons were performed using the linear mixed-effect model as this model can handle testing with repeating measurements (multiple patients with multiple OARs). Analysis was performed in R (version 4.5.0, https://www.r-project.org/) using the lme4 (version 1.1.37) and lmerTest (version 3.1.3) packages. Supplementary information A provides more information about the application of the linear mixed-effect model. A p-value < 0.05 was deemed significant.

The correlation was assessed between the geometric metrics and the dose differences compared to the gold standard for both the clinical and automatic MR-only, using a Spearman Rank Correlation.

Results

Nineteen patients were enrolled in this study. Two patients had to be excluded due to a coil defect for one patient and raw imaging data that could not be retrieved for the other. The remaining seventeen patients were used to compare the automated MR-only workflow with the clinical workflow. Table 1 shows the results of the OAR delineation comparison (requirement ii). The automated MR-only workflow performed better than the clinical workflows as demonstrated by a higher average DSC (0.791 ± 0.126 versus 0.673 ± 0.181, $p < 0.001$) and a lower Hausfdorff distance (4.0 ± 1.8 mm versus 5.8 ± 1.5 mm, $p < 0.001$). Several OARs are less well visible on the CT. Especially, the constrictor muscles, brain stem and spinal cord are poorly visible CT, while their visibility on the MR is excellent [12]. In CT-based clinical contours, these OARs had an average DSC of 0.380, 0.775 and 0.687 respectively. In the automatic MR-based contours, these OARs had an average DSC of 0.581, 0.905 and 0.856 respectively.

The registration error was assessed by assessing the error in centre of mass between the clinical (registration) and T2-based manual delineations (no registration) compared to the gold standard. The average absolute centre of mass error was merely 1 mm higher for the clinical OAR (registration) compared to the MR-only OAR (no registration) (see Table 2).

The average of the mean DVH error per organ at risk of the clinical and MR-only workflow were -1.31 ± 2.48 Gy (average $\pm$ standard deviation) and $+0.41 \pm 1.96$ Gy respectively (Table 3 and Fig. 2). The absolute DVH error (mean dose and near maximum dose) was significantly smaller for the MR-only workflow compared to the clinical workflow ($p < 0.01$).

Fig. 2 presents boxplots of the DVH metrics for the clinical and automated MR-only workflow (requirement iii). For the clinical workflow he average deviation of the mean dose from the ground truth was -1.47 ± 2.49 Gy for parallel organs. For the organs with near maximum dose constraints, the D2% deviated by -0.82 ± 2.44 Gy from the ground truth. For the MR-only workflow, the average deviation of the mean dose was 0.43 ± 2.18 Gy and 0.35 ± 1.30 Gy for the D2%.

In Fig. 3, shows the relation between the OAR similarity metrics and the absolute dose differences. A clear inverse correlation was seen for

Table 1
Comparison between the clinical contours and automated T2 contours with the manual gold standard contours. The Dice coefficient and 95% Hausdorff Distance are given for the difference organs at risk and the average of all organs of risks.

	Dice Coefficient [-]		95% Hausdorff Distance [mm]	
	Clinical RT	Automated MR-only RT	Clinical RT	Automated MR-only RT
Brainstem	0.775	0.905	6.1	2.6
Glnd_submand_L	0.768	0.842	5.0	3.9
Glnd_submand_R	0.789	0.840	3.8	3.7
Larynx	0.690	0.855	5.3	1.0
Mandible	0.781	0.890	3.7	1.9
Musc_constrict_I	0.259	0.532	8.3	6.5
Musc_constrict_M	0.481	0.630	6.1	5.0
Musc_constrict_S	0.399	0.582	8.4	5.7
Oral_cavity	0.864	0.922	6.7	4.0
Parotid_L	0.791	0.820	6.4	5.8
Parotid_R	0.793	0.816	5.4	6.3
Spinalcord	0.687	0.856	4.1	1.8
Average ± standard deviation	**0.673 ± 0.181**	**0.791 ± 0.126**	**5.8 ± 1.5**	**4.0 ± 1.8**

Table 2
The absolute difference in centre of for the clinical organ at risk (OAR) delineations and MR-only OAR delineations compared to the manual T2 OAR delineations.

	Center of Mass Difference [mm]	
	Clinical	MR-only
Brainstem	2.911	1.485
Glnd_submand_L	2.754	1.687
Glnd_submand_R	2.130	1.704
Mandible	2.323	1.574
Musc_constrict_I	4.548	4.072
Musc_constrict_M	4.260	3.873
Musc_constrict_S	5.489	3.884
Oral_cavity	3.266	1.531
Parotid_L	3.006	2.327
Parotid_R	2.711	3.015
Spinal_cord	4.735	2.746
Larynx	4.537	2.167
Average	**3.556**	**2.506**

the DSC ($R = 0.721$; $p < 0.0001$)and a clear positive correlation for the 95% Hausdorff distance ($R = 0.802$; $p < 0.0001$) For most OARs, both the OAR similarity metrics and dose differences are better with the automated MR-only compared to the clinical ones. This is especially noticeable for the constrictor muscles contours (*i.e.* the three OARs on the most left-hand-side in Fig. 3a and the two OAR on the most right-hand-side in Fig. 3b).

Discussion

Due to the increasing cancer incidence in combination with increasing workforce shortages there is an enormous need to reduce the hands-on time per patient to be able to keep treating patients where radiotherapy is indicated. For sites that require both an MR and CT scan for treatment preparation, automated MR-only RT could be an effective strategy to reduce hands-on time per patient. This strategy omits the need for a separate planning CT and exploits the superior soft tissue contrast of MRI, compared to CT, allowing accurate auto-delineations without manual adaptations. Here, we present an automated MR-only

Table 3
The average and standard deviation of the errors in dose in organ at risk DVH metrics (Dmean for the parallel organs at risk; D2% for the serial organs at risk) are given for the clinical and MR-only workflows compared to the gold standard.

		Clinical pipeline		Automated MR-only pipeline	
Organ at risks	**Metric**	**Average [Gy]**	**Standard deviation [Gy]**	**Average [Gy]**	**Standard deviation [Gy]**
Brainstem	D(2%)	−2.18	4.33	0.35	1.65
Glnd_Submand_L	Dmean	−0.08	1.82	0.55	0.71
Glnd_Submand_R	Dmean	−0.26	1.67	0.11	1.30
Larynx	Dmean	−0.59	3.29	−0.64	1.78
Mandible	D(2%)	0.32	1.31	0.54	1.42
Musc_Constrict_I	Dmean	−3.42	4.34	−1.25	2.60
Musc_Constrict_M	Dmean	−0.84	3.16	1.03	3.36
Musc_Constrict_S	Dmean	−3.57	2.99	0.69	4.01
Oral_Cavity_Ext	Dmean	−1.38	1.29	0.49	1.07
Parotid_L	Dmean	−2.43	1.98	0.95	2.12
Parotid_R	Dmean	−0.65	1.90	1.96	2.67
SpinalCord	D(2%)	−0.59	1.67	0.16	0.84
Average		**−1.31**	**2.48**	**+0.41**	**1.96**

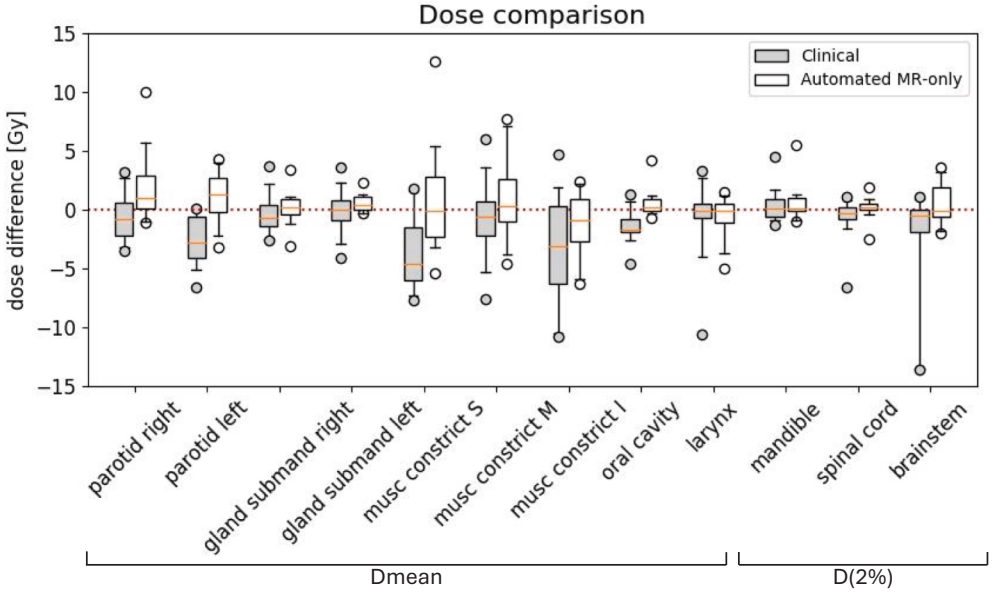

Fig. 2. Boxplot of the dose volume histogram (DVH) metrics (Dmean for the parallel organs at risk; D2% for the serial organs at risk). In grey, the dose errors between the DVH metrics of the clinical workflow and the gold standard. In white, the dose errors between the DVH metrics of the automated MR-only workflow and the gold standard. The red lines represent the medians, the boxes the 25% to 75% range, the whiskers the 5% to 95% range, and the circles datapoint outside of the range of the whiskers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

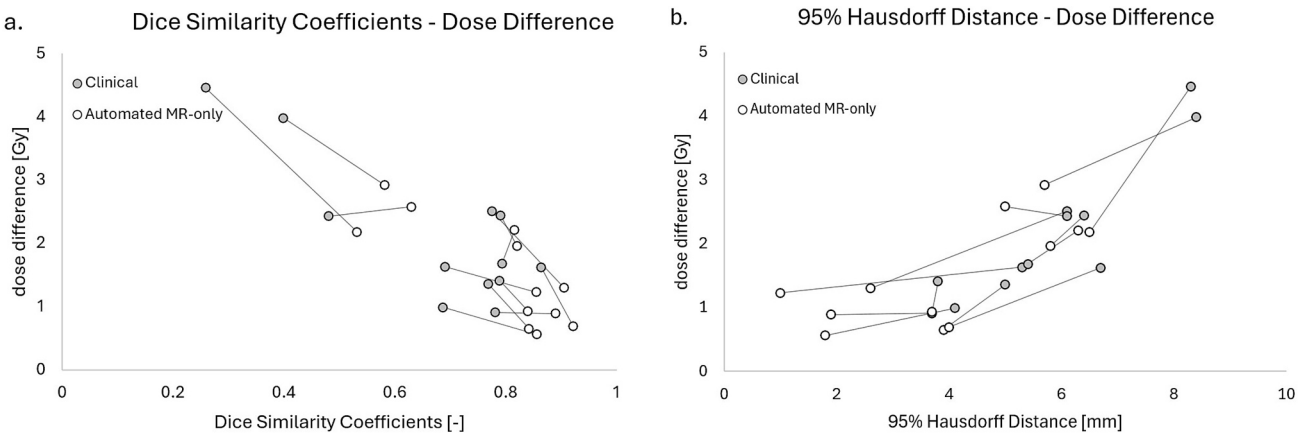

Fig. 3. The absolute dose differences as function of the OAR similarity metric for the clinical workflow in white and the automated MR-only workflow in grey compared to the ground truth. The lines between two points indicate that these points belong to the same organ at risk. **a)** the absolute dose difference as function of the Dice similarity coefficient. **b)** the absolute dose difference as function of the 95% Hausdorff distance.

workflow that uses a fast ZTE scan (scanning time can be as low as 56 s per scan) for the synCT generation (requirement i); automated T2-based OAR delineations (without manual correction) that were more accurate than the clinical delineations (requirement ii); and for the first time we demonstrated that the automated MR-only workflow (synCT and T2-based automated delineations) led to increased dosimetric accuracy compared to the current clinical CT-based workflows (requirement iii). Therefore, this automated MR-only workflow can be introduced in the clinic and is not only expected to reduce the hands-on-time, but it could also improve the treatment accuracy.

By omitting the need for a separate planning CT scan and manual modification of OAR contours, it was estimated that per patient automated MR-only radiotherapy could save approximately 90 min (two radiation technicians (RTTs) x 30 min per CT scan + 30 min manual OAR modification). Assuming the following additional tasks for RTTs per patient: 4 h per patient for treatment planning, 35 treatment fractions of 15 min with two RTT, the time saving of automated MR-only was approximately 7%. By itself 7% is not sufficient to solve the estimated workforce shortages, though it does contribute. Other measures to reduce RT workload include for instance hypofractionation for palliative and curative treatment where possible [38–41].

In most previous studies the automated MR-only OAR delineation is merely checked by comparing OAR similarity metrics; however the correlation between the OAR similarity metrics and the dose differences show that merely a part of the dose difference can be explained by the delineations. Other factors that influence the dose DVHs are the dose gradients and the difference in HU between the synthetic and planning CT. McDonald *et al* and Koteva et al performed a dosimetric comparison between the MR-only pipeline and the gold standard (manual T2 based OAR contours), for five and thirteen head-and-neck patients using T2 and a combination of T1 and T2 images respectively. Both studies gave more or less comparable results to ours [23,28]. However, no comparison was made with the current clinical workflow or in combination with a synCT. To the best of our knowledge the current study was the first to demonstrate that automated MR-only RT can even outperform the standard clinical workflow.

A previous study by Lauwers et al. showed that the dose errors caused by the ZTE-based synCT were merely 0.13 Gy on average, over all the OAR in the same patient cohort as utilized in this study [14]. Therefore, the dose errors found in the current study were largely caused by the differences in OAR delineations. This observation is backed by the strong correlations between DSC, Hausdorff Distance and dose errors. The larger dose errors for the clinical delineations can likely be explained, primarily by the worse visibility of the OARs on the CT scan. If the available MR scan was used and laid over the CT scan while delineating the OAR, the information from both scans could be used and the accuracy is expected to be similar to the gold standard.

Most OAR delineations were significantly improved by the MR-only workflow (average DSC 0.778) compared to the clinical workflow (average DSC 0.670). All average DSC scores of the automated OAR delineations were above 0.8, except for the pharyngeal constrictor muscles. Even though forthe constrictor muscles, the improvement was most apparent between the clinical and T2 automated delineations respectively, the DSC remained low. This might be explained by the small size of the structures and the high noise-to-signal ratio in those areas of the scan on the T2. All in all, the automated MR-only OAR delineations were more accurate than the clinical workflow; therefore, no structural manual OAR modification would be required with MR-only to reach at least similar performance compared to the clinical workflow. This improvement of automated MR compared to the clinical OAR contours could be explained by the improved soft tissue contrast on the MR. Note, that the current clinical OAR contours could be improved if the MR would be used in the contouring process.

The automated T2-based delineation algorithm used in the current study was previously developed, trained and tested by Czipczer et al [35]. This method was selected as it was extensively compared to the existing models. The geometrical accuracy was found to be similar or better than those found in literature [19,21,22,24–26,42,43]. Our geometrical accuracy was similar to that described by Czipczer et al when compared to manual T2 delineations, showing that the model could be used properly on our data.

This study has some limitations. First, the MRI was registered to the CT. This can cause a registration error between the gold standard and the clinical and MR-only workflow. In Supplementary Information A, this was analysed and deemed acceptable.

A second limitation is the gold standard delineation of the mandible. The gold standard delineations were made on the T2 as almost all OAR were clearly visible on this scan. However, the mandible is more clearly visible on the CT. Therefore, caution should be taking when interpreting the mandible results. However, a DSC of ~ 0.8 on the clinical mandible delineation shows a relatively high agreement between the delineation on the CT and T2. Moreover, comparing CT-based delineations with T2-based delineations might seem unjust as both the automated and gold standard delineations were made on the T2; however, this represents the comparison with the clinical workflow and is, therefore, very relevant.

It should also be noted that the automatic MR-only RT pipeline can currently not be used yet in a clinical setting. Although, at the time of writing, the automatic T2-based OAR delineation model has been approved by the American Food and Drug Administration (FDA) and can be used clinically in parts of the world, the synthetic CT algorithm has

not gone through the approval process yet.

Another notion to take into account is that the MR-based OAR delineation are in general different from CT-based OAR delineation [44]. As most OARs are better visual on the MR [12], it is stated here that the OAR delineations are more accurate on the T2 compared to the CT. By using MRI as the reference, the comparison is biased in favor of the MR-based workflow, and that the observed accuracy differences are partly due to modality differences rather than purely the effect of automation. It should be noted that most currently available toxicity models are based on CT delineations. A consequence of systematically using MRI based delineations may imply that such models may need to be revisited to account for more accurate delineations on the T2 MRI.

Finally, the workflow is called an automated MR-only workflow; however, it is not completely automated as the tumor could not be automatically delineated. This is a research topic with a lot of interest, in which the increased soft tissue contrast of the T2 could prove to be useful.

Conclusion

Due to the increasing cancer incidence in combination with increasing workforce shortages there is an enormous need to reduce the hands-on time per patient to be able to keep on treating all RT patients. This automated MR-only workflow utilizes a quick ZTE sequences for synCT generation for the head-and-neck region, thus barely prolonging the MR scanning time. Moreover, it can both improve the OAR delineation and dose accuracy and reduce the hands-on-time up to ~ 7% for the RTTs in the treatment preparation process. Therefore, it could contribute considerably to hands-on-time reduction for head-and-neck RT.

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.

Erasmus MC Cancer Institute has research collaborations with Accuray Inc., Sunnyvale, USA, Elekta AB, Stockholm, Sweden and with Varian, a Siemens Healthineers Company (Palo Alto, CA, USA). Author IL acknowledges support from The Dutch Cancer Society (project number 12141).

CRediT authorship contribution statement

Iris Lauwers: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft. **Marta Capala:** Conceptualization, Data curation, Methodology, Project administration, Resources, Validation, Supervision, Writing – review & editing. **Sandeep Kaushik:** Resources, Software, Methodology, Formal analysis, Investigation. **László Ruskó:** Conceptualization, Resources, Software, Methodology, Formal analysis, Supervision, Writing – review & editing, Investigation. **Cristina Cozzini:** Resources, Software, Methodology, Formal analysis, Investigation. **Eszter Szabó:** Investigation, Writing – review & editing. **Ádám Kékesi:** Investigation. **Borbála Deák-Karancsi:** Investigation, Methodology, Writing – review & editing. **Jonathan Wyatt:** Writing – review & editing, Methodology. **Rachel Pearson:** Methodology. **Gerda Verduijn:** Conceptualization, Methodology, Validation, Supervision, Writing – review & editing. **Florian Wiesinger:** Conceptualization, Methodology, Software, Resources, Project administration, Funding acquisition, Writing – review & editing. **Juan Hernandez-Tamames:** Software, Resources, Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Methodology. **Steven Petit:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Prof. Juan A. Hernandez-Tamames has received research funding from GE Healthcare. Dr. Marta Capala has received consulting fees for consultancy for GE on behalf of the Department of Radiotherapy, Erasmus Medical Center Rotterdam, though not related to the research described in the manuscript. Sandeep Kaushik, László Rusko, Cristina Cozzini and Florian Wiesinger are employees at GE Healthcare. Erasmus MC Cancer Institute has research collaborations with Accuray Inc., Sunnyvale, USA, Elekta AB, Stockholm, Sweden and with Varian, a Siemens Healthineers Company (Palo Alto, CA, USA).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ctro.2026.101122.

References

[1] Commission E. European Cancer Information System: Long term estimates of cancer incidence and mortality, for all cancers. European Union.
[2] Secretariat WHO. Global strategy on human resources for health: Workforce 2030. 2016.
[3] Cusumano D, Lenkowicz J, Votta C, Boldrini L, Placidi L, Catucci F, 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;153:205–12.
[4] Lenkowicz J, Votta C, Nardini M, Quaranta F, Catucci F, Boldrini L, et al. A deep learning approach to generate synthetic CT in low field MR-guided radiotherapy for lung cases. Radiother Oncol 2022;176:31–8.
[5] 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 2021;89:265–81.
[6] Johnstone E, Wyatt JJ, Henry AM, Short SC, Sebag-Montefiore D, Murray L, 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.
[7] Karlsson M, Karlsson MG, Nyholm T, Amies C, Zackrisson B. Dedicated magnetic resonance imaging in the radiotherapy clinic. Int J Radiat Oncol Biol Phys 2009;74:644–51.
[8] Lagendijk JJ, Raaymakers BW, van Vulpen M. The magnetic resonance imaging-linac system. Semin Radiat Oncol 2014;24:207–9.
[9] McKenzie EM, Santhanam A, Ruan D, O'Connor D, Cao M, Sheng K. Multimodality image registration in the head-and-neck using a deep learning-derived synthetic CT as a bridge. Med Phys 2020;47:1094–104.
[10] Owrangi AM, Greer PB, Gilde-Hurst CK. MRI-only treatment planning: benefits and challenges. Phys Med Biol 2018;26.
[11] 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.
[12] Paczona VR, Capala ME, Deak-Karancsi B, Borzasi E, Együd Z, Vegvary Z, et al. Magnetic resonance imaging–based delineation of organs at risk in the head and neck region. Adv Radiat Oncol 2022;8.
[13] Klages P, Benslimane I, Riyahi S, Jiang J, Hunt M, Deasy JO, et al. Patch-based generative adversarial neural network models for head and neck MR-only planning. Med Phys 2020;47:626–42.
[14] Lauwers I, Capala M, Kaushik S, Rusko L, Cozzini C, Kleijnen JP, et al. Synthetic CT generation using Zero TE MR for head-and-neck radiotherapy. Radiother Oncol 2025;205:110762.
[15] Olin AB, Hansen AE, Rasmussen JH, Ladefoged CN, Berthelsen AK, Hakansson K, et al. Feasibility of multiparametric positron emission tomography/magnetic resonance imaging as a one-stop shop for radiation therapy planning for patients with head and neck cancer. Int J Radiat Oncol Biol Phys 2020;108:1329–38.
[16] Olin AB, Thomas C, Hansen AE, Rasmussen JH, Krokos G, Urbano TG, et al. Robustness and generalizability of deep learning synthetic computed tomography for positron emission tomography/magnetic resonance imaging-based radiation therapy planning of patients with head and neck cancer. Adv Radiat Oncol 2021;6:100762.
[17] Palmer E, Karlsson A, Nordstrom F, Petruson K, Siversson C, Ljungberg M, et al. Synthetic computed tomography data allows for accurate absorbed dose calculations in a magnetic resonance imaging only workflow for head and neck radiotherapy. Phys Imaging Radiat Oncol 2021;17:36–42.
[18] Qi M, Li Y, Wu A, Lu X, Zhou L, Song T. Multisequence MR-generated sCT is promising for HNC MR-only RT: a comprehensive evaluation of previously developed sCT generation networks. Med Phys 2022;49:2150–8.
[19] Dai X, Lei Y, Wang T, Zhou J, Rudra S, McDonald M, et al. Multi-organ auto-delineation in head-and-neck MRI for radiation therapy using regional convolutional neural network. Phys Med Biol 2022;67.

[20] Hague C, McPartlin A, Lee LW, Hughes C, Mullan D, Beasley W, et al. An evaluation of MR based deep learning auto-contouring for planning head and neck radiotherapy. Radiother Oncol 2021;158:112–7.
[21] Kawahara D, Tsuneda M, Ozawa S, Okamoto H, Nakamura M, Nishio T, et al. Deep learning-based auto segmentation using generative adversarial network on magnetic resonance images obtained for head and neck cancer patients. J Appl Clin Med Phys 2022;23:e13579.
[22] Korte JC, Hardcastle N, Ng SP, Clark B, Kron T, Jackson P. Cascaded deep learning-based auto-segmentation for head and neck cancer patients: Organs at risk on T2-weighted magnetic resonance imaging. Med Phys 2021;48:7757–72.
[23] Koteva V, Eiben B, Dunlop A, Gupta A, Gangil T, Wong KH, et al. Clinical acceptance and dosimetric impact of automatically delineated elective target and organs at risk for head and neck MR-Linac patients. Front Oncol 2024;14:1358350.
[24] Lei Y, Zhou J, Dong X, Wang T, Mao H, McDonald M, et al. Multi-organ segmentation in head and neck MRI using U-Faster-RCNN. In proc of SPIE. 2020; 11313.
[25] Podobnik G, Ibragimov B, Tappeiner E, Lee C, Kim JS, Mesbah Z, et al. HaN-Seg: the head and neck organ-at-risk CT and MR segmentation challenge. Radiother Oncol 2024;198:110410.
[26] Zhong Z, He L, Chen C, Yang X, Lin L, Yan Z, et al. Full-scale attention network for automated organ segmentation onhead and neck CT and MR images. IET Image Proc 2023;17:660–73.
[27] Dinkla AM, Florkow MC, Maspero M, Savenije MHF, Zijlstra F, Doornaert PAH, et al. Dosimetric evaluation of synthetic CT for head and neck radiotherapy generated by a patch-based three-dimensional convolutional neural network. Med Phys 2019;46:4095–104.
[28] McDonald BA, Cardenas CE, O'Connell N, Ahmed S, Naser MA, Wahid KA, et al. Investigation of autosegmentation techniques on T2-weighted MRI for off-line dose reconstruction in MR-linac workflow for head and neck cancers. Med Phys 2024; 51:278–91.
[29] CCMO. Deep learning based MR Only Radiotherapy for head-and-neck cancer. CCMO; 2019.
[30] Breedveld S, Storchi PR, Voet PW, Heijmen BJ. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. Med Phys 2012;39:951–63.
[31] Kaushik SS, Bylund M, Cozzini C, Shanbhag D, Petit SF, Wyatt JJ, et al. Region of interest focused MRI to synthetic CT translation using regression and segmentation multi-task network. Phys Med Biol 2023;68.
[32] 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.
[33] Wiesinger F, Ho ML. Zero-TE MRI: principles and applications in the head and neck. Br J Radiol 2022;95:20220059.
[34] 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.
[35] Czipczer V, Kolozsvári B, Deák-Karancsi B, Capala ME, Pearson RA, Borzási E, et al. Comprehensive deep learning-based framework for automatic organs-at-risk segmentation in head-and-neck andpelvis forMR-guidedradiation therapy planning. Front Phys 2023;11–2023.
[36] Ruskó L, Capala M, Czipczer V, Kolozsvári B, Deak-Karancsi B, Czabány R, et al. Deep-learning-based segmentation of organs-at-risk in the head for MR-assisted Radiation Therapy Planning BIOIMAGING. 2021;2:31-4.
[37] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. Radiother Oncol 2015;117:83–90.
[38] de Vasconcellos Ferreira PM, Gomes M, Almeida A, Cornelio JS, Arruda TJ, Mafra A, et al. Evaluation of oral mucositis, candidiasis, and quality of life in patients with head and neck cancer treated with a hypofractionated or conventional radiotherapy protocol: a longitudinal, prospective, observational study. Head Face Med 2023;19:7.
[39] Elbers JBW, Gunsch PA, Debets R, Keereweer S, van Meerten E, Zindler J, et al. HYpofractionated, dose-redistributed RAdiotherapy with protons and photons to combat radiation-induced immunosuppression in head and neck squamous cell carcinoma: study protocol of the phase I HYDRA trial. BMC Cancer 2023;23:541.
[40] Piras A, Boldrini L, Menna S, Venuti V, Pernice G, Franzese C, et al. Hypofractionated radiotherapy in head and neck cancer elderly patients: a feasibility and safety systematic review for the clinician. Front Oncol 2021;11: 761393.
[41] Vreugdenhil M, Fong C, Sanghera P, Hartley A, Dunn J, Mehanna H. Hypofractionated chemoradiation for head and cancer: data from the PET NECK trial. Oral Oncol 2021;113:105112.
[42] Chen H, Lu W, Chen M, Zhou L, Timmerman R, Tu D, et al. A recursive ensemble organ segmentation (REOS) framework: application in brain radiotherapy. Phys Med Biol 2019;64:025015.
[43] Mlynarski P, Delingette H, Alghamdi H, Bondiau PY, Ayache N. Anatomically consistent segmentation of organs at risk in MRI with convolutional neural networks. arXiv preprint arXiv. 2019;1907.02003.
[44] Lekshmi R, Gupta M, Gupta S, Joseph D, Krishnan AS, Sharma P, et al. Comparison of magnetic resonance imaging and CT scan-based delineation of target volumes and organs at risk in the radiation treatment planning of head and neck malignancies. J Med Imaging Radiat Sci 2023;54:503–10.