

Deep learning-based optimization of field geometry for total marrow irradiation delivered with volumetric modulated arc therapy

Nicola Lambri^{1,2}  | Giorgio Longari³  | Daniele Loiacono³  |
 Ricardo Coimbra Brioso³  | Leonardo Crespi^{3,4}  | Carmela Galdieri²  |
 Francesca Lobefalo²  | Giacomo Reggiori²  | Roberto Rusconi^{1,2}  |
 Stefano Tomatis²  | Luisa Bellu^{1,2}  | Stefania Bramanti⁵  | Elena Clerici²  |
 Chiara De Philippis⁵  | Damiano Dei^{1,2}  | Pierina Navarra²  |
 Carmelo Carlo-Stella^{1,5}  | Ciro Franzese^{1,2}  | Marta Scorsetti^{1,2}  |
 Pietro Mancosu² 

¹Department of Biomedical Sciences, Humanitas University, Pieve Emanuele, Milan, Italy

²Radiotherapy and Radiosurgery Department, IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy

³Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milan, Italy

⁴Health Data Science Centre, Human Technopole, Milan, Italy

⁵Department of Oncology and Hematology, IRCCS Humanitas Research Hospital, Rozzano, Milan, Italy

Correspondence

Pietro Mancosu, Radiotherapy and Radiosurgery Department, IRCCS Humanitas Research Hospital, via Manzoni 56, 20089 Rozzano, Milan, Italy.
 Email: pietro.mancosu@humanitas.it

Funding information

Ministero della Salute, Grant/Award Number: GR-2019-12370739

Abstract

Background: Total marrow (lymphoid) irradiation (TMI/TMLI) is a radiotherapy treatment used to selectively target the bone marrow and lymph nodes in conditioning regimens for allogeneic hematopoietic stem cell transplantation. A complex field geometry is needed to cover the large planning target volume (PTV) of TMI/TMLI with volumetric modulated arc therapy (VMAT). Five isocenters and ten overlapping fields are needed for the upper body, while, for patients with large anatomical conformation, two specific isocenters are placed on the arms. The creation of a field geometry is clinically challenging and is performed by a medical physicist (MP) specialized in TMI/TMLI.

Purpose: To develop convolutional neural networks (CNNs) for automatically generating the field geometry of TMI/TMLI.

Methods: The dataset comprised 117 patients treated with TMI/TMLI between 2011 and 2023 at our Institute. The CNN input image consisted of three channels, obtained by projecting along the sagittal plane: (1) average CT pixel intensity within the PTV; (2) PTV mask; (3) brain, lungs, liver, bowel, and bladder masks. This “averaged” frontal view combined the information analyzed by the MP when setting the field geometry in the treatment planning system (TPS). Two CNNs were trained to predict the isocenters coordinates and jaws apertures for patients with (CNN-1) and without (CNN-2) isocenters on the arms. Local optimization methods were used to refine the models output based on the anatomy of the patient. Model evaluation was performed on a test set of 15 patients in two ways: (1) by computing the root mean squared error (RMSE) between the CNN output and ground truth; (2) with a qualitative assessment of manual and generated field geometries—scale: 1 = not adequate, 4 = adequate—carried out in blind mode by three MPs with different expertise in TMI/TMLI. The Wilcoxon

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Authors. *Medical Physics* published by Wiley Periodicals LLC on behalf of American Association of Physicists in Medicine.

signed-rank test was used to evaluate the independence of the given scores between manual and generated configurations ($p < 0.05$ significant).

Results: The average and standard deviation values of RMSE for CNN-1 and CNN-2 before/after local optimization were $15 \pm 2/13 \pm 3$ mm and $16 \pm 2/18 \pm 4$ mm, respectively. The CNNs were integrated into a planning automation software for TMI/TMLI such that the MPs could analyze in detail the proposed field geometries directly in the TPS. The selection of the CNN model to create the field geometry was based on the PTV width to approximate the decision process of an experienced MP and provide a single option of field configuration. We found no significant differences between the manual and generated field geometries for any MP, with median values of 4 versus 4 ($p = 0.92$), 3 versus 3 ($p = 0.78$), 4 versus 3 ($p = 0.48$), respectively. Starting from October 2023, the generated field geometry has been introduced in our clinical practice for prospective patients.

Conclusions: The generated field geometries were clinically acceptable and adequate, even for an MP with high level of expertise in TMI/TMLI. Incorporating the knowledge of the MPs into the development cycle was crucial for optimizing the models, especially in this scenario with limited data.

KEYWORDS

autoplanning, deep learning, total marrow irradiation

1 | INTRODUCTION

Total marrow irradiation (TMI) and total marrow lymphoid irradiation (TMLI) are recent radiotherapy (RT) treatments used as conditioning regimen for allogeneic hematopoietic stem cell transplantation in acute leukemia.¹ Over the past two decades, TMI and TMLI have been developed to selectively target the bone marrow and lymph nodes, while reducing the exposure of normal tissues and the late toxicities associated to traditional total body irradiation (TBI), where the whole body is irradiated.²

Several authors have conducted studies investigating the clinical feasibility of TMI/TMLI with intensity-modulated techniques, including intensity-modulated radiation therapy (IMRT), tomotherapy, and volumetric modulated arc therapy (VMAT). These studies have consistently shown that these techniques provide adequate target coverage and reduce doses to normal tissues, without compromising treatment delivery efficiency.^{3–6}

Phase I/II clinical trials have been or currently are under investigation to improve disease control with acceptable toxicity.¹ Although there is a growing clinical interest in TMI/TMLI worldwide,⁷ its widespread implementation is hindered by technological gaps during the RT process. Manual contouring of target and normal tissues can take up to 16 h,⁸ while plan optimization requires prolonged computation time and extensive work by an experienced medical physicist (MP).

Recent studies have shown the successful application of artificial intelligence (AI) algorithms for automating the treatment planning of TMI/TMLI. Watkins et al.

evaluated an AI segmentation software for total body contouring of 27 organs at risk (OARs) and four planning target volumes (PTVs), finding good spatial and dosimetric agreement with manual contours.⁹ Ahn et al. developed an AI-based model for plan optimization of VMAT-TMI, which improved dosimetric quality while reducing dependence on planner experience.¹⁰

However, other technological difficulties remain unanswered. To cover the large target volume of TMI/TMLI with VMAT, typically five isocenters and ten overlapping fields are needed for the upper body. For patients with large anatomical conformation, two specific isocenters are placed on the arms. Therefore, the positioning of isocenters and jaws apertures requires a complex field geometry designed by an MP with high level of expertise in TMI/TMLI.

A few authors have explored optimization methods for determining the isocenter location and jaws apertures in radiosurgeries, prostate, and rectal cases.^{11–15} In these studies, researchers employed brute-force methods, similarity scores based on the patient's anatomy, and AI algorithms, such as k-means clustering and deep learning (DL), to find the optimal solution.

Recently, two studies have reported the automation of field geometry for TBI delivered with VMAT using heuristic rules based on structures volumes and distances.^{16,17} Although this approach can find near-optimal solutions, it might not be fully adequate to generate complex field geometries required for TMI/TMLI.

To address this clinical challenge, we developed DL-based models to automatically optimize the field geometry—that is, isocenters positions and jaws

apertures—of TMI/TMLI. To the best of our knowledge, this is the first study presenting the application of DL methods for generating a field configuration in a multi-isocentric VMAT setting.

2 | MATERIALS AND METHODS

2.1 | Dataset

In this study, 117 patients treated with TMI/TMLI at our Institute from 2011 to 2023 were retrospectively selected. All patients signed an informed consent in accordance with the Declaration of Helsinki and approved by the Institutional Ethics Committee of IRCCS Humanitas Research Hospital (ID 2928, January 26, 2021 - ClinicalTrials.gov identifier: NCT04976205).

Computed tomography (CT) images used for simulation were acquired on a CT Big Bore (Philips Healthcare, Best, Netherlands) and on a SOMATOM go.Sim (Siemens Healthcare GmbH, Erlangen, Germany) for 107 and 10 patients, respectively. The resolution of the images was heterogeneous. Specifically, the pixel spacing was 1.17 mm for 98 cases, while it ranged between 1.31 and 1.37 mm for the remaining 17 cases. The slice thickness was 2.5, 3, 5, and 7.5 mm for, respectively, 2, 5, 109, and 1 CT series. All acquisitions had the same number of pixels, 512×512 , with different number of slices, from 184 to 534, depending on the patient's height and slice thickness.

Treatment planning was conducted using the Eclipse (Varian Medical Systems, Palo Alto, CA, USA) treatment planning system (TPS). All plans were delivered with VMAT on a TrueBeam (Varian Medical Systems, Palo Alto, CA, USA) equipped with a Millennium multileaf collimator. We refer to our previous studies for a detailed description of the protocol followed to delineate the PTV.^{18,19}

The field geometry was created by an experienced MP by positioning five isocenters—two full-arc fields per isocenter with collimator angle set at 90° —along the body of the patient. For each field, asymmetric jaw settings were used to guarantee a field overlap for at least 2 cm on each side, such that differences between the planned and delivered dose due to small patient misalignments were minimized.¹⁸ Isocenters locations and jaws apertures were decided by taking into consideration the specific anatomy of the patient.²⁰ In particular, the MP maximized the freedom of motion of the multileaf collimator for each field aperture. Notably, the most critical region was the abdomen, where fields could not include the ribs and iliac crests within the same beam's eye view (BEV) to avoid overexposing normal tissue.

For patients with large anatomical conformation, the maximum jaw aperture (40 cm) was not sufficient to provide an adequate target coverage. Using only isocenters along the body, the BEVs would contain the arms

mostly during a limited portion of the gantry rotation and from a lateral direction, with consequent unnecessary irradiation of medial healthy tissues. Therefore, in these cases, two specific isocenters—each with one field—were placed on the arms, while only four isocenters were used on the body to maintain a reasonable treatment time. This configuration allowed to extend the anterior/posterior BEVs on the arms, ensuring safe irradiation with no OARs overlap. Figure 1 shows two representative cases of field configuration with and without isocenters on the arms.

Since 2021, a different configuration of the collimator angle was introduced for the pelvic fields, where the collimator angle was changed from 90° to $5^\circ/355^\circ$ to align the fields along the femurs and irradiate one leg at a time. However, to homogenize the data and reduce the degrees of freedom, the collimator angle was set to 90° for all cases and the jaws apertures were adjusted accordingly (see Section S1 and Figure S1 of the Supporting Information).²¹

Table S1 provides a complete description of the dataset in terms of acquisition specifications and field geometry.

2.2 | Preprocessing

For each patient, the CT, RT structure set, and RT plans were exported in DICOM format. A three-channel image was constructed by projecting along the sagittal plane: (1) average CT pixel intensity within the PTV; (2) PTV mask; (3) brain, lungs, liver, bowel, and bladder masks. This “averaged” frontal view was built to aggregate the information analyzed by the MP when setting the field geometry in the TPS.

To facilitate the model training, a min-max normalization was applied to rescale the first channel of the images (CT pixel intensities) in the range $[0, 1]$:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

where the minimum and maximum were computed for each image.

The masks of the PTV and OARs were weighted by 0.3, and 0.5, respectively, except for the intestine, with a weight of 1. This decision was made to increase the pixel intensities between the most caudal ribs and the iliac crests, where the positioning of isocenters and jaws apertures was crucial to obtain an acceptable field geometry according to the anatomy of the patient (see Section 2.1).

The field geometries were extracted from the RT plans and matched to the pixel space of the images by transforming isocenters coordinates and jaws apertures from the DICOM patient coordinate system to the DICOM voxel coordinate system.

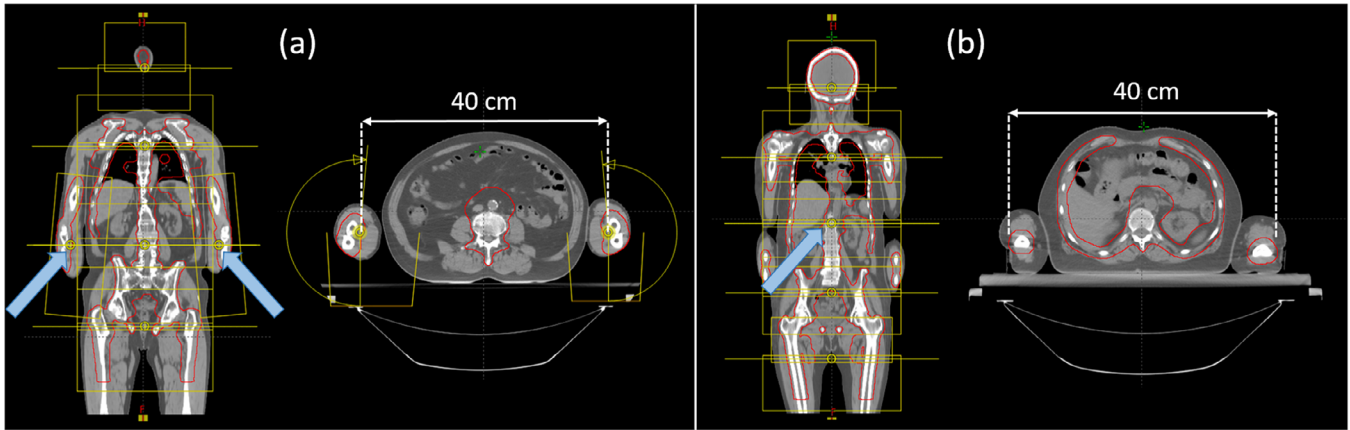


FIGURE 1 Field geometry designed by an expert MP for patients with (A) and without (B) isocenters on the arms. For case (A), the maximum jaw aperture (40 cm) cannot adequately cover the entire PTV. MP, medical physicist; PTV, planning target volume.

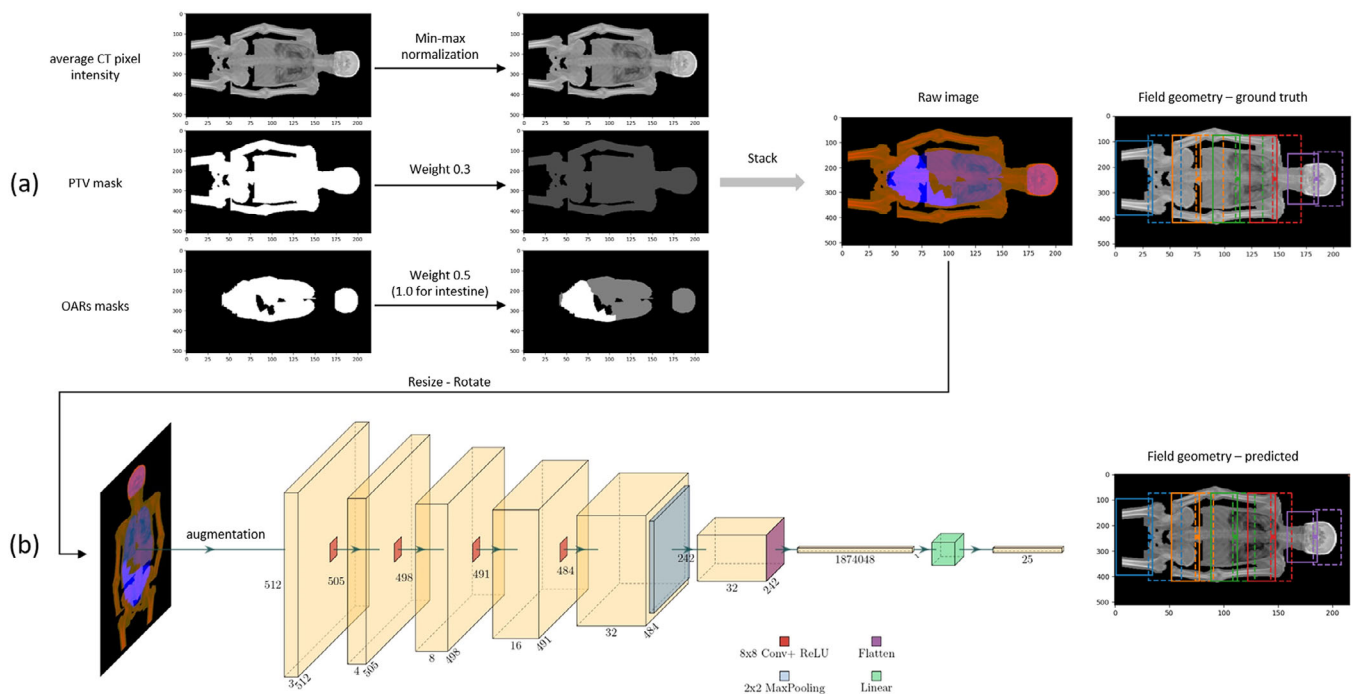


FIGURE 2 (A) Preprocessing pipeline implemented to create a raw image, with the associated ground truth field geometry. (B) Model diagram of the CNN used for predicting isocenters positions and jaws apertures for patients without isocenters on the arms. Dots and rectangles on the images denote isocenters positions and jaws apertures, respectively. Fields belonging to the same isocenters group are represented using the same color. CNN, convolutional neural network.

Figure 2A shows the pipeline implemented for preprocessing the dataset. By construction, the raw images were rectangular with resolution given by the pixel spacing and slice thickness of the acquisition, with horizontal patient orientation. To standardize their shape and achieve a vertical patient orientation as seen by the MP in the TPS, the images were resized to 512×512 and rotated 90° counterclockwise.

The same transformations were applied also to the coordinates of the isocenters—landmarks on the images—to preserve the correct field geometries. The jaws apertures, instead, were only resized, being the

rotation an isometry. Finally, the isocenters coordinates and jaws apertures were scaled by the image width to constrain the model output to values less than 1.

2.3 | Deep learning models

In this study, two convolutional neural networks (CNNs) were implemented to predict the field geometry for patients with (CNN-1) and without (CNN-2) isocenters on the arms. Both models were trained using 2D images

of shape $(3 \times) 512 \times 512$ representing the frontal view of the patients' anatomy. Among the 117 patients, 31 and 86 were planned with and without isocenters on the arms. These subgroups were split into training, validation, and test sets comprising 22/69, 3/8, and 6/9 patients, respectively.

To increase the number of examples in the training set and improve model regularization, the raw images were augmented without major alterations to the patients' anatomy, as it is tightly coupled to the corresponding field geometry. The following transformations, each one having a 70% chance of occurrence, were applied in sequence: horizontal flip, translation, elastic deformation, and cutout. We refer to Table S2 for a comprehensive description of the parameters used for data augmentation. The training sets for CNN-1 and CNN-2 were increased by 100% and 50% in size using augmentation. These ratios were selected after several experiments as the models performance reached a plateau beyond these values.

Figure 2B shows the architecture implemented for CNN-2: four convolutional layers to extract relevant patterns from the input, a max pooling layer to downsample the feature maps and summarize information, and a fully connected head to perform the regression of isocenters coordinates and jaws apertures. The dimensionality of the output was reduced from 60 to 25 by exploiting the symmetry of the field geometry. The same architecture was used also for CNN-1, with the only difference on the output shape, equal to 30. The detailed description is provided in Section S3 of the Supporting Information.

For training, a weighted mean squared error (MSE) was used as loss function:

$$loss_{CNN} = \frac{1}{N} \sum_i^N w_i (x_i - y_i)^2$$

where N is the dimensionality of the output, w_i is a weight factor to penalize the CNN prediction x_i , and y_i is the target value. The weights w_i were assigned a value of 3 for the pelvis and abdomen isocenters coordinates, critical for a correct field geometry on the ribs and iliac crests, and set to 1 for the remaining terms.

The Adam algorithm was used to optimize the CNN parameters, with a learning rate of 0.00001, and momentum parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. A scheduler was employed to reduce the learning rate by a factor of 10 when the training loss did not decrease once in 3 epochs. The training loop was run with a batch size of 10, using an early stopping criterion which monitored the MSE loss calculated on the validation set with a patience of 7 epochs.

2.4 | Local optimization

The models were not able to consistently create an optimal field geometry on the abdomen which did not include the most caudal ribs and iliac crests within the same BEV to avoid overexposing normal tissue. Therefore, a local optimization approach was applied to find the positions of the most caudal ribs and iliac crests along the cranial caudal (CC) axis and to adjust the field geometry predicted by the CNNs. Figure 3 shows a representative example where the field geometry was suboptimal in the abdominal region and the final field geometry obtained with local optimization.

The predicted location of the abdominal isocenter was used as reference to define two patches, P_1 and P_2 , over the PTV mask image (see Figure 3B). The search space consisted of the x-pixel indices spanning through these patches, as knowing the x-pixel location is sufficient to adjust the jaws apertures along the CC axis. For each patch, a parallel tempering algorithm was used to maximize the following loss function:

$$loss = \alpha \sum_{\substack{y \in P \\ x_{iliac} \leq x \leq x_{ribs}}} f(x, y) + \beta \sum_{y \in P} g(x, y) \quad x = \{x_{iliac}, x_{ribs}\}$$

where x_{iliac} and x_{ribs} are the candidate x-pixel indices for iliac crests and ribs at the current step of the optimization, P denotes either P_1 or P_2 , while α and β are coefficients which were set to 2 and 60, respectively. The functions $f(x, y)$ and $g(x, y)$ are defined as:

$$f(x, y) = \begin{cases} -1, & (x, y) \in PTV \\ 1, & otherwise \end{cases} \quad g(x, y) = \begin{cases} 0, & (x, y) \in PTV \\ 1, & otherwise \end{cases}$$

The first term of the loss computes the number of background pixels minus the number of pixels inside the PTV mask within the subset of P where $x_{iliac} \leq x \leq x_{ribs}$, that is, it maximizes the background while simultaneously minimizing the PTV included between two candidate ribs and iliac crests positions. The second term computes the number of background pixels along the y-axis at fixed x_{iliac} and x_{ribs} pixel locations, that is, it ensures that the ribs and iliac crests positions are found adjacent to the maximum extent of the PTV mask.

The optimization was performed for P_1 and P_2 separately to account for potential asymmetrical locations of ribs and iliac crests between the patches. The final positions x_{iliac}^* and x_{ribs}^* were selected by taking, respectively, the minimum and maximum value between the locations of iliac crests and ribs found with the local optimization on P_1 and P_2 .

After altering the isocenter position and jaws apertures on the abdomen according to the optimal positions

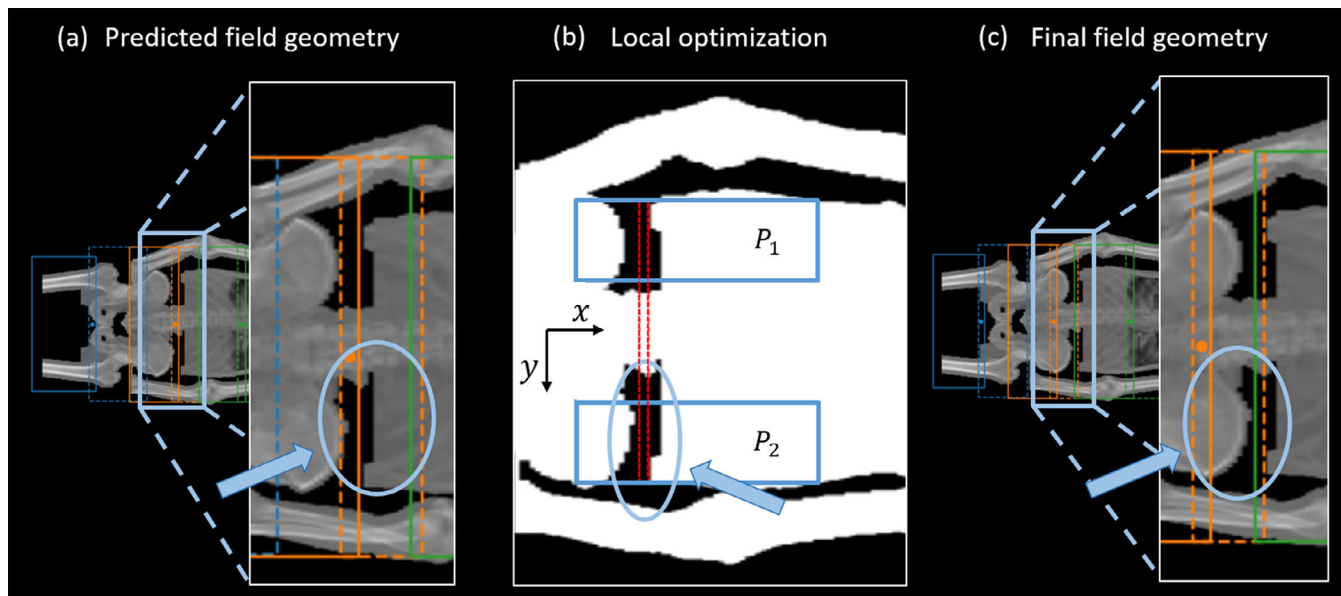


FIGURE 3 (A) Field geometry predicted by CNN-2 with suboptimal configuration on the abdomen: the ribs and iliac crests are under the same BEV of the jaw aperture denoted with the yellow dashed line. (B) Patches P_1 and P_2 defined over the PTV mask to search for the x-pixel indices of most caudal ribs and iliac crests. The red dashed lines denote the best locations found by the local optimization. (C) Final field geometry with adjusted isocenters and jaws apertures: ribs and iliac crests are under a single BEV. BEV, beam's eye view.

x_{ilic}^* and x_{ribs}^* , additional changes based on geometric rules were used to adjust the overall configuration.

2.5 | Field geometry evaluation

The quality of the generated field geometry was evaluated by computing the root mean squared error (RMSE) between the predicted and ground truth isocenters positions and jaws apertures of the test patients:

$$RMSE = \frac{1}{M} \sum_j \sqrt{\frac{1}{N} \sum_i (x_i^{(j)} - y_i^{(j)})^2}$$

where N is the dimensionality of the output and M the number of patients. A channel-wise ablation study was performed by switching off one channel at a time of the input image to investigate its effect on the CNNs error, and to interpret the importance of each channel for the CNNs predictions.

The quantitative assessment of the generated field geometry can be challenging to interpret as slight changes in configuration may still be acceptable. Therefore, a qualitative evaluation was performed by three MPs with different experience with TMI/TMLI. The MPs were asked to blindly inspect, for each test patient, two candidate field geometries, one which was created by an expert MP, and a second one generated by either CNN-1 or CNN-2, based on the width of the PTV. In case the PTV was larger than 47.5 cm, then CNN-1 was used.

This approach was employed to approximate the decision process of an experienced MP and simulate what would happen in clinical practice, where, for edge cases, the use of isocenters on the arms is subject to the MP's experience. Furthermore, this method provided a single option of field configuration which avoided ambiguity in case the MP was not an expert in TMI/TMLI.

The MPs assigned to each field geometry a score from 1, that is, not adequate for planning and to be recreated from scratch, to 4, that is, adequate for planning without further modifications. The evaluations were performed directly in the Eclipse TPS, such that the MPs could analyze in detail the proposed field geometries. For this purpose, the CNNs were integrated into a planning automation software for TMI/TMLI presented in our previous study.²² The Wilcoxon signed-rank test was used to evaluate for each MP the independence of the given scores between manual and generated configurations, with significance level set at $p < 0.05$.

2.6 | Software

The code used in this study was written in Python 3.11.1 and is publicly available at <https://github.com/nlambriilCH/tmi-isocenter>. The main libraries used for preprocessing, model training and optimization were imgaug-0.4.0, pytorch-2.0.1, lightning-2.0.1, and gradient-free-optimizers-1.3.0, while the statistical analysis was performed with scipy-1.8.1.

TABLE 1 Quantitative evaluation for CNN-1 and CNN-2 on the test set, before and after the local optimization, and for the channel-wise ablation study (without local optimization).

Model	Target	RMSE [mm]				
		Before local optimization	After local optimization	Ablated channel 1	Ablated channel 2	Ablated channel 3
CNN-1	Isocenters coordinates	18 ± 5	17 ± 5	18 ± 5	46 ± 10	16 ± 6
	Jaws apertures	13 ± 2	11 ± 3	12 ± 2	27 ± 5	13 ± 3
	Overall field geometry	15 ± 2	13 ± 3	15 ± 3	35 ± 6	15 ± 3
CNN-2	Isocenters coordinates	19 ± 5	22 ± 7	22 ± 8	21 ± 6	29 ± 13
	Jaws apertures	13 ± 2	15 ± 3	15 ± 3	14 ± 3	21 ± 6
	Overall field geometry	16 ± 2	18 ± 4	18 ± 4	17 ± 3	25 ± 5

Note: Results are reported for isocenters coordinates and jaws apertures separately, and for the overall field geometry. Channels were obtained by projecting along the sagittal plane: (1) average CT pixel intensity within the PTV; (2) PTV mask; (3) OARs masks.

Abbreviations: CNN, convolutional neural network; OAR, organ at risk; PTV, planning target volume; RMSE, root mean squared error.

3 | RESULTS

The training and validation $loss_{CNN}$ for CNN-1 and CNN-2 are provided in Figure S2. Both training curves reached a value of $\approx 2 \times 10^{-6}$ before early stopping was triggered. Due to the smaller dataset, the validation loss of CNN-1 reached soon a plateau at $\approx 9 \times 10^{-4}$, compared with a value of $\approx 3 \times 10^{-4}$ for CNN-2. Accordingly, the training took 35 and 64 min for CNN-1 and CNN-2, respectively.

Table 1 reports the models RMSE obtained on the test set before and after local optimization, and RMSE from the channel-wise ablation study, without local optimization. The errors were computed separately for isocenters coordinates, jaws apertures, and overall field geometry.

For CNN-1, the local optimization slightly decreased the average error (15 ± 2 mm vs. 13 ± 3 mm), while an opposite trend was observed for CNN-2 (16 ± 2 mm vs. 18 ± 4 mm). The removal of channel 1 had the lowest impact on the models error, where, for the overall field geometry, we did not observe any relevant changes (CNN-1: 15 ± 2 mm vs. 15 ± 3 mm; CNN-2: 16 ± 2 mm vs. 18 ± 4). The ablation of channel 2 and channel 3 greatly affected the precision of CNN-1 and CNN-2, respectively, with a percentage increase in RMSE for the overall field geometry of 133% (CNN-1, ablated channel 2) and 56% (CNN-2, ablated channel 3). In particular, we found the greatest loss in performance for CNN-1 on isocenters positioning, for which the RMSE increased from 18 ± 5 mm to 46 ± 10 mm. By inspecting the predicted field geometry, we observed that the largest contribution to the model error was due to incorrectly predicted positions of the isocenters on the arms. On the other hand, the decrease in performance of CNN-2 with the removal of channel 3 resulted from a decay in the overall quality of the generated field geometry. Figure S3 and Figure S4 provide two representative cases of field geometry predicted by CNN-1 and CNN-2 after channel-wise ablation.

Figure 4 shows two representative cases of local optimization of unacceptable (a) and near optimal (b) field geometry predicted by the CNNs. In the former case, the lower (red) jaw aperture of the abdominal isocenter covered both ribs and iliac crests, while in the latter case the upper (blue) jaw aperture of the abdominal isocenter exceeded on the ribs for a few millimeters.

Figure 5 shows the scores given by the three MPs for the fifteen test patients. MP 1 was an expert in TMI/TMLI, MP 2 had intermediate experience, and MP 3 had little to no experience with TMI/TMLI planning. The evaluators were free to change between the proposed field geometries and employ all available views in the TPS (frontal, transversal, sagittal, and 3D) with no restrictions. MP 3, being the least experienced with TMI/TMLI, was briefly instructed before the start of the assessment regarding the overall field configuration they would have encountered.

We found no significant differences between manual and generated field geometries for any MP, with median scores of 4 versus 4 (MP 1, $p = 0.92$), 3 vs. 3 (MP 2, $p = 0.78$), and 4 vs. 3 (MP 3, $p = 0.48$), respectively. Out of 15 patients, the three MPs evaluated the generated field geometry to be at least as adequate as the manual counterpart for 12 (MP 1), 9 (MP 2), and 8 (MP 3) instances. None of the automatic configurations was rated “not adequate” (i.e., score = 1). Table S3 reports the complete evaluations for each patient, while Figure S5 provides two representative cases of generated field geometries as shown within the Eclipse TPS. The pipeline for field geometry generation took a total of 56.2 s, allocated as follows: contours extraction (27.4 s), preprocessing (27.6 s), model inference (0.1 s), and local optimization (1.1 s).

Starting from October 2023, the CNN models have been introduced in our clinical practice for prospective patients. After 6 months of clinical implementation, CNN-1 was found less robust and requiring more corrections than CNN-2. Based on our internal preliminary

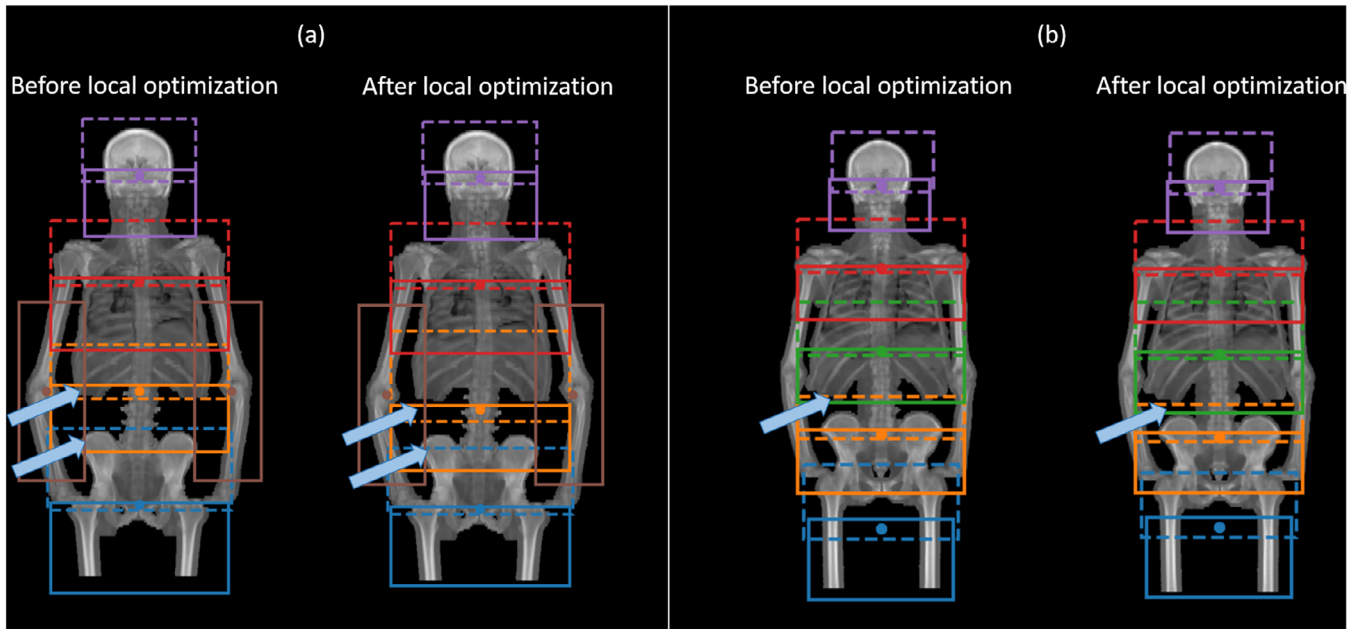


FIGURE 4 Representative cases of local optimization of unacceptable (A) and near optimal (B) field geometry predicted by the CNNs. Dots and rectangles denote isocenters positions and jaws apertures, respectively. Fields belonging to the same isocenters group are shown with the same color, using dashed and solid lines for the upper and lower jaws apertures, respectively. CNN, convolutional neural network.

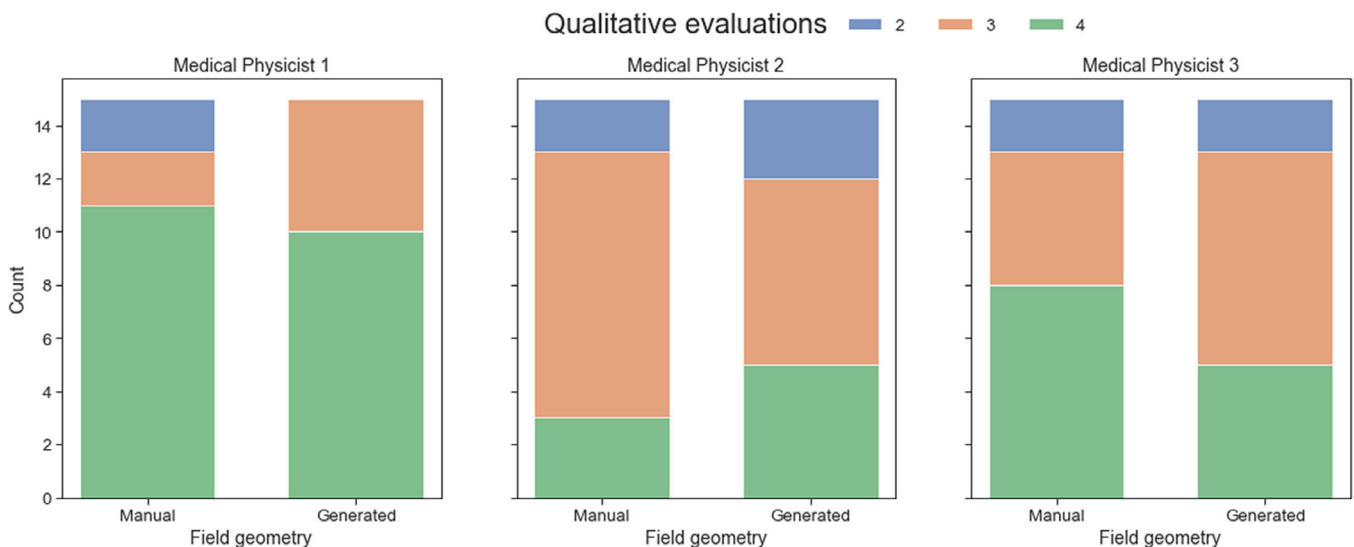


FIGURE 5 Results of the qualitative evaluations performed by the three MPs for the 15 test patients. Scale: 1 = not adequate, 2 = major changes, 3 = minor changes, 4 = adequate. MP, medical physicist.

assessments, the efficiency gains for an MP expert in TMI/TMLI planning amount to a few minutes, increasing to tens of minutes for less experienced MPs.

4 | DISCUSSION

In this study, we developed CNN models able to generate a clinically acceptable field geometry for TMI/TMLI delivered with VMAT. Incorporating the knowledge of the

MPs was fundamental to achieve results which were comparable in quality to those crafted manually by an MP expert in TMI/TMLI planning. Specifically, the information analyzed by the MP during the creation of the field geometry was captured in the input image using a frontal view projection of the patient's anatomy, while a local optimization refined the predictions near critical regions—ribs and iliac crests—as performed by an expert MP. The local optimization did not introduce any relevant changes in the RMSE of the models,

suggesting the solutions generated by the CNNs required little tweaking to reach an optimal field geometry.

The performance of CNN-1 and CNN-2 on the test set was similar, with RMSE of 15 ± 2 mm and 16 ± 2 mm, respectively. The local optimization produced opposite effects on the error of the models, reducing the error of CNN-1 (13 ± 3 mm) and increasing that of CNN-2 (18 ± 4 mm). We hypothesize such difference was due to the limited degrees of freedom available for positioning the isocenters along the body once two isocenters were used specifically for the arms. As a result, the field geometries created by the MP were more uniform among cases, and the local optimization was likely to produce minor changes in the field geometry generated by CNN-1. Nonetheless, we highlight that the objective of the local optimization was not to minimize the RMSE. This metric was evaluated to ensure that the procedure did not introduce relevant changes. Thus, despite the potential increase in RMSE, local optimization remains crucial if the generated configurations are utilized directly for dose optimization without prior scrutiny by the planner.

The ablation study revealed the model predictions were mostly driven by the channels containing the masks of the PTV and relevant OARs. Although channel 1 (i.e., CT pixel intensity) contained the richest information regarding the anatomy of the patient, its removal had almost no impact on the models error. On the other hand, the mask of the PTV was crucial for CNN-1 to accurately position the isocenters on the arms (see Figure S3). This dependence supports our approach, employed during the qualitative evaluation, to approximate the decision process of an expert MP for generating a field geometry with isocenters on the arms based on the patient's PTV width.

The aim of this study was to automatically optimize the field geometry of TMI/TMLI based our decade experience. The models were trained to generate configurations consistent with our internal approach for TMI/TMLI planning, which, although common, might differ from other Institutions' practice. Nonetheless, three MPs validated all the generated field geometries by considering them acceptable for planning, indicating that the models could be beneficial in case of clinical use by planners with varying levels of experience.

We did not compare the dose distributions between manual and generated field geometries because of two limitations: TMI/TMLI optimization is a time consuming process which takes 4–8 h using all computational resources of a TPS, and it requires high expertise to perform. As the manual and generated field geometries were similar in quality, the resulting dose distributions would have been planner dependent more than field configuration dependent.

A few studies have investigated the optimization of field geometry for site specific treatments. Salter et al.

quantified the impact of isocenter location on treatment plan quality for intensity-modulated stereotactic radiosurgery (SRS).¹¹ Authors conducted an exhaustive search of the optimal isocenter location by considering more than 3400 configurations of 18 clinical plans, using three different TPSs and collimators. Huang et al. applied a similar strategy to find the optimal jaw aperture for rectal cases treated with VMAT using one full arc.¹² Almost 600 configurations were automatically created starting from 10 clinical plans. Schreibmann and Fox implemented a feature selection search engine to generate a near-optimal plan configuration for prostate cancer.¹⁵ A similarity score based on geometrical measures was used to search the best matching case among 83 VMAT plans. The retrieved plan was then used to suggest isocenter, arc directions, MLC patterns, and optimization constraints. Yock and Kim applied k-means clustering on 30 patients to determine the number and position of isocenters in SRS for multiple metastases.¹³ More recently, Berdyshev et al. developed a DL model for isocenter selection in SRS delivered with Gamma Knife (Elekta, Stockholm, Sweden).¹⁴ Authors trained a ResNet on tumor shape descriptors of 533 plans to predict isocenter locations as a 3D heatmap.

Approaching the field geometry optimization of TMI/TMLI with an exhaustive search would have been unfeasible. AI algorithms allowed to optimize the field geometry based on past experience without the need of exploring all possible configurations. Despite the typical requirement for large datasets in training DL models, both Berdyshev et al. and the present study showed that satisfactory performance could be achieved even with datasets considerably smaller than the norm for DL models. In our case, given the limited data availability, local optimization methods were essential for refining the models output.

Although we did not perform a systematic evaluation of the models generalizability, we were able to test them on two external cases from City of Hope. The resulting field geometries, shown in Figure S6, were considered adequate according to our procedure, suggesting the broader applicability of the models to other patient cohorts. Furthermore, we highlight that our methodology can be reproduced locally with site-specific data, potentially resulting in the development of more specialized models.

The combination of DL and local optimization techniques with their respective strengths was crucial to reach our goal. Although there exist some general rules for the design of the field geometry of TMI/TMLI according to our procedure, their implementation would have been challenging to achieve. Even for an MP specialized in TMI/TMLI, it proved difficult to precisely explain the requirements and the logical steps involved in designing a field geometry from scratch. Thus, producing a configuration of comparable quality to that of the DL models was non-trivial, especially when placing the isocenters,

as their positions are interrelated. After the DL models provided a good starting point for the field geometry, we could establish well-defined rules for refining the field geometry through local optimization methods.

In this study, the CNNs were trained using a projected frontal view of the patients, aiming to closely mirror the planner's strategy in optimizing geometric parameters. The isocenters coordinate along the anterior/posterior direction was set to the anterior/posterior barycenter of the PTV for all cases, +4 cm toward the anterior for the isocenters on the arms. This decision allowed to reduce the complexity of the models by decreasing the number of degrees of freedom and focus on the patient view which was more relevant for the overall geometry. Importantly, we found this approach to be effective in all evaluated cases. Nonetheless, we acknowledge the possibility of introducing other patient views, especially in case of different approaches to the field geometry employing oblique beam angles.

This study was part of the AuToMI project, whose aim is to increase the accessibility of TMI/TMLI in clinical practice through process automation and implementation of AI algorithms.^{19,21–23} To this aim, the planning automation software we use in our clinic (a binary plug-in script based on the Eclipse Scripting API—ESAPI) is publicly available at <https://github.com/nlambri/CH/TMIAutomation>. We underline that the application of these models resembles that of DL segmentation, in which the output is validated by a human operator, and modified if necessary.

5 | CONCLUSIONS

We developed a DL-based method for the optimization of multiple isocenters positions and jaws apertures for TMI/TMLI. The CNNs produced field geometries of comparable quality to those crafted by an MP specialized in TMI/TMLI. Revisions are still necessary to adjust inconsistencies and refine the models output. However, these adjustments are subject to inter-observer variability and depend on the MP decision process and experience.

ACKNOWLEDGMENTS






This study was supported by the AuToMI project which is funded by Ministero della Salute (Rome, IT). GRANT_NUMBER: GR-2019-12370739. The authors also thank professor Jeffrey Y.C. Wong, professor Chunhui Han, and Dr Tyler Watkins from City of Hope (California, USA) for providing the clinical data.

Open access funding provided by BIBLIOSAN.

CONFLICT OF INTEREST STATEMENT

The authors do not have relevant conflicts of interest to disclose.

ORCID

Nicola Lambri  <https://orcid.org/0000-0001-8706-6480>
 Giorgio Longari  <https://orcid.org/0000-0002-2086-9091>
 Daniele Loiacono  <https://orcid.org/0000-0002-5355-0634>
 Ricardo Coimbra Brioso  <https://orcid.org/0000-0003-1635-0489>
 Leonardo Crespi  <https://orcid.org/0000-0001-6219-8444>
 Carmela Galdieri  <https://orcid.org/0000-0003-3293-2737>
 Francesca Lobefalo  <https://orcid.org/0000-0003-3634-2582>
 Giacomo Reggiori  <https://orcid.org/0000-0003-3336-5690>
 Roberto Rusconi  <https://orcid.org/0000-0002-8385-9296>
 Stefano Tomatis  <https://orcid.org/0000-0002-4791-0373>
 Luisa Bellu  <https://orcid.org/0000-0002-7286-4014>
 Stefania Bramanti  <https://orcid.org/0000-0002-4117-7991>
 Elena Clerici  <https://orcid.org/0000-0002-1765-863X>
 Chiara De Philippis  <https://orcid.org/0000-0002-7738-997X>
 Damiano Dei  <https://orcid.org/0000-0001-9873-7868>
 Pierina Navarra  <https://orcid.org/0000-0002-6913-1442>
 Carmelo Carlo-Stella  <https://orcid.org/0000-0003-3144-0124>
 Ciro Franzese  <https://orcid.org/0000-0001-6893-6284>
 Marta Scorsetti  <https://orcid.org/0000-0003-0320-559X>
 Pietro Mancosu  <https://orcid.org/0000-0002-0165-7931>

REFERENCES

1. Wong JYC, Filippi AR, Scorsetti M, Hui S, Muren LP, Mancosu P. Total marrow and total lymphoid irradiation in bone marrow transplantation for acute leukaemia. *Lancet Oncol*. 2020;21(10):e477–e487. doi:10.1016/S1470-2045(20)30342-9
2. Wong JYC, Filippi AR, Dabaja BS, Yahalom J, Specht L. Total body irradiation: guidelines from the International Lymphoma Radiation Oncology Group (ILROG). *Int J Radiat Oncol Biol Phys*. 2018;101(3):521–529. doi:10.1016/j.ijrobp.2018.04.071
3. Wilkie JR, Tiryaki H, Smith BD, Roeske JC, Radosevich JA, Aydogan B. Feasibility study for linac-based intensity modulated total marrow irradiation. *Med Phys*. 2008;35(12):5609–5618. doi:10.1118/1.2990779
4. Hui SK, Kapatoes J, Fowler J, et al. Feasibility study of helical tomotherapy for total body or total marrow irradiation. *Med Phys*. 2005;32(10):3214–3224. doi:10.1118/1.2044428
5. Fogliata A, Cozzi L, Clivio A, et al. Preclinical assessment of volumetric modulated arc therapy for total marrow irradiation. *Int J Radiat Oncol Biol Phys*. 2011;80(2):628–636. doi:10.1016/j.ijrobp.2010.11.028

6. Surucu M, Yeginer M, Kavak GO, Fan J, Radosevich JA, Aydogan B. Verification of dose distribution for volumetric modulated arc therapy total marrow irradiation in a humanlike phantom: total marrow irradiation with RapidArc. *Med Phys*. 2011;39(1):281-288. doi:10.1118/1.3668055
7. Hui SK, Storme G, Wong J, Aristei C, Al Malki MM, Aydogan B. Editorial: total marrow irradiation. *Front Oncol*. 2023;13:1240530. doi:10.3389/fonc.2023.1240530
8. Schultheiss TE, Wong J, Liu A, Olivera G, Somlo G. Image-guided total marrow and total lymphatic irradiation using helical tomotherapy. *Int J Radiat Oncol Biol Phys*. 2007;67(4):1259-1267. doi:10.1016/j.ijrobp.2006.10.047
9. Watkins WT, Qing K, Han C, Hui S, Liu A. Auto-segmentation for total marrow irradiation. *Front Oncol*. 2022;12:970425. doi:10.3389/fonc.2022.970425
10. Ahn KH, Rondelli D, Koshy M, et al. Knowledge-based planning for multi-isocenter VMAT total marrow irradiation. *Front Oncol*. 2022;12:942685. doi:10.3389/fonc.2022.942685
11. Salter BJ, Fuss M, Sarkar V, et al. Optimization of isocenter location for intensity modulated stereotactic treatment of small intracranial targets. *Int J Radiat Oncol Biol Phys*. 2009;73(2):546-555. doi:10.1016/j.ijrobp.2008.09.011
12. Huang Y, Yue H, Wang M, et al. Fully automated searching for the optimal VMAT jaw settings based on Eclipse Scripting Application Programming Interface (ESAPI) and RapidPlan knowledge-based planning. *J Applied Clin Med Phys*. 2018;19(3):177-182. doi:10.1002/acm2.12313
13. Yock AD, Kim G. Technical note: using k-means clustering to determine the number and position of isocenters in MLC -based multiple target intracranial radiosurgery. *J Applied Clin Med Phys*. 2017;18(5):351-357. doi:10.1002/acm2.12139
14. Berdyshev A, Cevik M, Aleman D, et al. Knowledge-based isocenter selection in radiosurgery planning. *Medical Physics*. 2020;47(9):3913-3927. doi:10.1002/mp.14305
15. Schreiber E, Fox T. Prior-knowledge treatment planning for volumetric arc therapy using feature-based database mining. *J Applied Clin Med Phys*. 2014;15(2):19-27. doi:10.1120/jacmp.v15i2.4596
16. Simiele E, Skinner L, Yang Y, et al. A step toward making VMAT TBI more prevalent: automating the treatment planning process. *Pract Radiat Oncol*. 2021;11(5):415-423. doi:10.1016/j.prro.2021.02.010
17. Teruel JR, Taneja S, Galavis PE, et al. Automatic treatment planning for VMAT-based total body irradiation using Eclipse scripting. *J Appl Clin Med Phys*. 2021;22(3):119-130. doi:10.1002/acm2.13189
18. Mancosu P, Navarra P, Castagna L, et al. Interplay effects between dose distribution quality and positioning accuracy in total marrow irradiation with volumetric modulated arc therapy. *Med Phys*. 2013;40(11):111713. doi:10.1118/1.4823767
19. Dei D, Lambri N, Stefanini S, et al. Internal guidelines for reducing lymph node contour variability in total marrow and lymph node irradiation. *Cancers*. 2023;15(5):1536. doi:10.3390/cancers15051536
20. Mancosu P, Navarra P, Castagna L, et al. Anatomy driven optimization strategy for total marrow irradiation with a volumetric modulated arc therapy technique. *J Appl Clin Med Phys*. 2012;13(1):138-147. doi:10.1120/jacmp.v13i1.3653
21. Lambri N, Dei D, Hernandez V, et al. Evaluation of plan complexity and dosimetric plan quality of total marrow and lymphoid irradiation using volumetric modulated arc therapy. *J Appl Clin Med Phys*. 2023. doi: 10.1002/acm2.13931. Published online.
22. Lambri N, Dei D, Hernandez V, et al. Automatic planning of the lower extremities for total marrow irradiation using volumetric modulated arc therapy. *Strahlenther Onkol*. 2022. doi:10.1007/s00066-022-02014-0. Published online November 3.
23. Lambri N, Antonetti SL, Dei D, et al. Impact of the extremities positioning on the set-up reproducibility for the total marrow irradiation treatment. *Curr Oncol*. 2023;30(4):4067-4077. doi:10.3390/curroncol30040309

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lambri N, Longari G, Loiacono D, et al. Deep learning-based optimization of field geometry for total marrow irradiation delivered with volumetric modulated arc therapy. *Med Phys*. 2024;51:4402–4412. <https://doi.org/10.1002/mp.17089>