Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

ML-Enhanced Monte Carlo Dosimetry and Cancer Risk in Proton Therapy for Hodgkin Lymphoma

Hossein Sadeghi
Arak University
Masoumeh Anjomi
Arak University
Fatemeh Seif
Sahar_S59@yahoo.com

Arak University of Medical Sciences
Mohamad Reza Bayatiani
Arak University of Medical Sciences

Article

Keywords: Proton therapy, Hodgkin lymphoma, Monte Carlo, GATE, Machine learning, Secondary cancer risk, BEIR VII, Organ equivalent dose

Posted Date: November 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11404110/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

ML-Enhanced Monte Carlo Dosimetry and Cancer Risk in Proton Therapy for Hodgkin Lymphoma

Hossein Sadeghi[1], Masoumeh Anjomi[1], Fatemeh Seif[2]*, Mohamad Reza Bayatiani[2]

[1]Department of Physics, Faculty of Science, Arak University, Arak, Iran

[2]Department of Radiotherapy and Medical Physics, Arak University of Medical Sciences

*Corresponding author: sahar_s59@yahoo.com

Abstract

Background: Proton therapy is widely used for Hodgkin lymphoma (HL) because the Bragg peak allows steep dose gradients near critical organs. However, secondary neutrons and scattered protons still deliver dose outside the target, which is a concern for late effects in young patients.

Objective: We aimed to identify the optimal proton energy for a simplified HL target, to quantify out-of-field doses to organs at risk (OARs) in a standardized anthropomorphic phantom, and to compare secondary cancer risk estimates from three epidemiological models: the organ equivalent dose (OED) formalism, BEIR VII, and an NIH-based model.

Materials and Methods: Proton transport was simulated with GATE 8.2 using the QGSP_BERT_EMV physics list. Depth-dose curves were scored in a water phantom for 13 energies between 60 and 90 MeV in 2.5-MeV steps. A Random Forest regression model was trained on these curves to predict the Bragg peak depth. A voxelized CIRS ATOM phantom was then used to score dose in the thyroid, esophagus, both lungs, and both breasts. ERR, EAR, and LAR were evaluated for ages at exposure from 30 to 70 years.

Results: The ML model identified 70 MeV as the optimal energy, with a TCS metric of 0.876 and 93.8% target coverage. Summed voxel-dose values ranged from 0.08 mGy in the thyroid to 39.96 mGy in the left lung. At age 60, for the left lung, OED-EAR was 0.09, BEIR VII-EAR was 0.166, and NIH-EAR was 0.214, with LAR values of 1.125 (BEIR VII) and 2.27 (NIH). Statistical uncertainty remained below 2% in the target and below 3.2% in OARs.

Conclusion: The combination of Monte Carlo simulation and ML-based energy selection provides a reproducible framework for out-of-field dose assessment in HL proton therapy.

Risk estimates differed by more than a factor of two between models for the same organ and dose, so the choice of model must be reported explicitly.

Keywords: Proton therapy, Hodgkin lymphoma, Monte Carlo, GATE, Machine learning, Secondary cancer risk, BEIR VII, Organ equivalent dose

1 Introduction

Hodgkin lymphoma (HL) has one of the highest cure rates among adult malignancies. Most patients are diagnosed between 15 and 35 years of age, and more than 90% of early-stage cases reach complete remission [3]. Because these patients survive for decades after treatment, late effects have become a central concern in treatment planning. Radiation-induced secondary malignancies are among the most serious of these late effects and can appear 10 to 40 years after the original treatment [4, 5].

In mediastinal HL, the target lies directly adjacent to the heart, lungs, and breasts. This geometry, combined with the young age of the patients and the long life expectancy after treatment, makes dose sparing outside the target clinically important [6, 7]. The distal edge of a proton beam often stops close to the heart, where the relative biological effectiveness is variable and still debated.

Proton therapy is attractive for this site because the Bragg peak deposits most of the energy at the end of the particle range [1]. Photon beams, by contrast, deliver an entrance dose that decays exponentially, so normal tissue receives a substantial fraction of the target dose. Protons therefore allow better sparing of tissues proximal to the target [2]. The stopping power of protons in a medium is described by the Bethe-Bloch equation, and the position of the Bragg peak is determined primarily by the initial beam energy.

The out-of-field dose in proton therapy does not vanish. Scattering in the treatment head, nuclear interactions in the patient, and secondary neutrons all contribute to dose outside the target [8, 16]. For HL patients, these contributions fall on the lungs, breasts, thyroid, and esophagus, all of which are radiosensitive. To convert absorbed dose into a cancer risk estimate, epidemiological models are needed. The most widely used models are the OED formalism of Schneider et al. [17, 18], the BEIR VII report [19], and the NIH-based model [20]. These models differ in their dose-response functions, age modifiers, and survival assumptions, so their outputs are not directly interchangeable.

Machine learning has recently been applied to dose prediction in radiotherapy to reduce the computational cost of Monte Carlo simulations [21, 22]. Random Forest regression is particularly useful for this task because it handles non-linear relationships and is resistant to overfitting [23]. In the present study, we use a Random Forest model to select the optimal proton energy from simulated depth-dose curves.

The present work builds on our earlier dosimetric study of proton therapy for Hodgkin lymphoma [30], in which we established the GATE-based Monte Carlo framework, validated the beam model against NIST/PSTAR data, and identified 70 MeV as the optimal energy for a simplified spherical target in a water phantom. That study focused on the physical dose distribution and out-of-field organ doses. In the present work, we extend the framework by integrating a machine learning model for energy optimization, by quantifying voxel-level out-of-field doses in a standardized anthropomorphic phantom, and by comparing secondary cancer risk estimates from three epidemiological models (OED, BEIR VII, and NIH-based). The segmented phantom geometry and the depth-dose validation shown in Figures 1, 5, and 6 of the present manuscript are consistent with the corresponding data reported in our previous study [30]. Previous Monte Carlo studies of proton therapy for Hodgkin lymphoma have focused primarily on physical dose distributions without incorporating machine learning for energy selection or comparing multiple epidemiological risk models [30, 29].

2 Materials and Methods

All Monte Carlo simulations were performed with GATE version 8.2, which is built on Geant4 version 10.5 [9, 10]. We used the QGSP_BERT_EMV physics list, which combines the Quark-Gluon String model for high-energy interactions, the Bertini cascade for intermediate energies, and standard electromagnetic processes [11, 12]. This physics list is widely used in proton therapy studies because it models both the electromagnetic energy loss of the primary beam and the nuclear fragmentation that produces secondary particles. Table 1 summarizes the main process categories and their role in proton therapy dosimetry. The simulation protocol, the choice of the QGSP_BERT_EMV physics list, and the validation strategy follow the framework described in our previous work [30].

Table 1: Physical processes modeled by the QGSP_BERT_EMV physics list and their role in proton therapy dosimetry.

Process Category	**Role in Proton Therapy**
Electromagnetic	Multiple scattering, ionization, and bremsstrahlung. These processes determine the energy deposition and lateral spread of the primary proton beam.
Hadronic elastic	Elastic scattering of protons on nuclei. This contributes to the angular spread of the beam and to out-of-field dose.
Hadronic inelastic (target fragmentation)	Nuclear reactions that produce secondary protons, alpha particles, and heavier fragments. These reactions determine the dose at and beyond the Bragg peak.
Hadronic inelastic (neutron production)	Nuclear reactions that produce secondary neutrons. Neutrons travel far from the field and are a major source of out-of-field dose.

The range cut for secondary particles was set to 1 mm. All simulations ran on a workstation with an Intel Core i3 at 3.40 GHz and 8 GB of RAM. A typical run with 2×10^6 primary protons took about 1.5 hours. The number of primaries was chosen so that the relative statistical uncertainty remained below 2% in all regions of clinical interest. This threshold follows the recommendation of Paganetti for Monte Carlo dose calculations in proton therapy [15].

For energy selection, we used a water phantom with 30 cm side length, density 1.0 g/cm^3, and elemental composition 11.19% hydrogen and 88.81% oxygen by weight. A dose actor of dimensions 1 mm $\times$ 1 mm $\times$ 300 mm was placed along the beam axis and divided into 300 segments, giving a depth resolution of 1 mm. The proton source was a parallel pencil beam placed 25 cm from the phantom surface. The beam spot had a diameter of 1 mm and zero divergence.

The target was a sphere of 37 mm diameter with its proximal edge at 5 mm depth and its distal edge at 42 mm depth. The center of the sphere was therefore at 23.5 mm depth. This geometry represents a generic mediastinal lymph node. We selected the optimal energy by maximizing a Target Coverage and Sparing (TCS) metric, defined as

$$\mathrm{TCS}(E) = \frac{\bar{D}_{\mathrm{target}}(E)}{D_{\max}(E)} \left[1 - \frac{\bar{D}_{\mathrm{distal}}(E)}{D_{\max}(E)} \right], \tag{1}$$

where $\bar{D}_{\mathrm{target}}$ is the mean dose in the target, $D_{\max}$ is the height of the Bragg peak, and $\bar{D}_{\mathrm{distal}}$ is the mean dose in the region from 42 to 60 mm depth. This metric favors energies that cover the target while keeping the exit dose low. We simulated 13 energies from 60 to 90 MeV in 2.5-MeV steps, each with 2×10^6 primaries.

A Random Forest regression model was trained on the depth-dose curves with scikit-learn [24]. The forest contained 500 trees with a maximum depth of 20 and a minimum of 5 samples per leaf. We used 5-fold cross-validation and split the data into 80% for training and 20% for testing. The input features were the normalized depth-dose curves, and the targets were the Bragg peak position, peak height, TCS value, and target coverage. Feature importance was used to identify the depth regions that most influenced the predicted energy. A Gaussian Process Regression model with a radial basis function kernel provided an additional uncertainty estimate on the predicted optimal energy.

For the organ dose phase, we used a voxelized phantom built from 198 axial CT slices of a CIRS ATOM adult female phantom (Model 702, 160 cm, 55 kg) with breast attachments [13]. The phantom materials have linear attenuation coefficients within 1% of reference tissues. We assigned materials directly from the manufacturer's specifications rather than using a stoichiometric HU-to-density calibration. This choice avoids calibration uncertainty but limits direct use of the framework on patient CT data without a validated HU-density curve.

Segmentation was performed in ITK-SNAP by an experienced medical physicist. Each organ received a unique label. Table 2 lists the labels. The segmentation procedure and the resulting axial CT slice, shown in Figure 1, are identical to those used in our previous study [30], which allows direct comparison of the present results with the previously reported out-of-field doses.

Table 2: Organ labels for the voxelized phantom.

Label	Organ	Material
1	Left lung	Lung tissue
2	Right lung	Lung tissue
3	Left breast	Adipose tissue
4	Right breast	Adipose tissue
5	Thyroid	Soft tissue
6	Esophagus	Soft tissue
7	Heart	Muscle tissue
8	Liver	Soft tissue
9	Stomach	Soft tissue
10	Target / PTV	Soft tissue

The target was placed in the left anterior mediastinum, slightly lateral to the midline, to mimic a realistic mediastinal HL presentation. Because the target is on the left, the left lung and left breast lie closer to the field than their contralateral counterparts.

A rang.dat file mapped each label to the corresponding material. For materials not in the

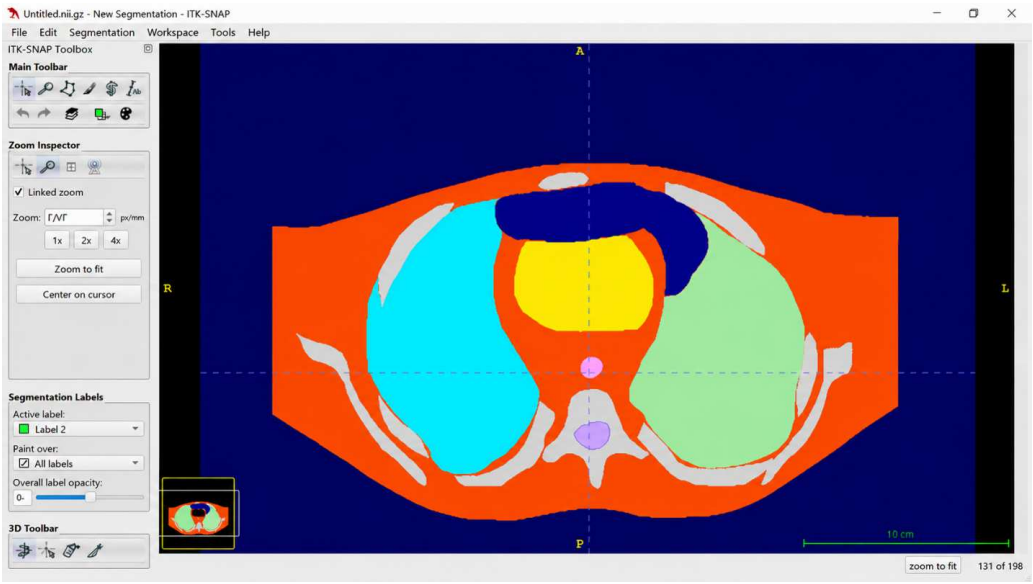


Figure 1: Axial CT slice showing the segmented anatomy and target volume used in the GATE simulation. The target (dark blue) lies in the left anterior mediastinum, slightly lateral to the mediastinal midline. The color legend is as follows: left lung (green), right lung (light blue), target (dark blue), esophagus (pink), spinal cord (purple), heart (yellow), and bones (gray). The segmented anatomy shown here corresponds to the same phantom geometry used in our previous dosimetric study [30], ensuring consistency between the two datasets.

GATE library, we derived the composition and density from ICRU Report 49 [14]. Table 3 lists the tissue parameters.

Table 3: Material properties of the tissues used in the phantom. Elemental compositions are given as percentage by weight and are adapted from ICRU Report 49 [14].

Tissue	ρ (g/cm^3)	H	C	N	O	Other	I (eV)
Lung	0.300	10.3	10.5	3.1	74.9	1.2	64.0
Adipose	0.950	11.4	59.8	0.7	27.8	0.3	63.2
Soft tissue	1.040	10.5	22.5	3.0	63.5	0.5	67.2
Muscle	1.050	10.2	14.3	3.4	71.0	0.1	69.0
Cortical bone	1.850	4.5	15.5	4.2	43.5	33.4	79.0

The phantom was placed inside a world volume of 60 cm × 60 cm × 120 cm. A 1-cm air layer surrounded the phantom; the rest of the world was vacuum to reduce computation time. The dose actor matched the CT resolution: 0.7363280 mm in x and y and 5 mm in z. This gave roughly 4.5 million voxels.

The proton source had an energy of 70 MeV, a Gaussian profile with sigma of 3 mm, and zero divergence. It was aligned with the target center at a source-to-axis distance of 100 cm. This configuration approximates a single pencil beam spot before lateral spreading in tissue.

To validate the beam model, we compared the simulated Bragg peak positions with NIST/PSTAR

projected ranges for the same energies. The maximum difference was 2.45 mm, which is within the expected accuracy of Monte Carlo simulations in proton therapy.

Organ doses were normalized to the mean physical dose in the target. Assuming a clinical prescription of 20 Gy(RBE) with RBE = 1.1, the corresponding physical dose is 18.2 Gy. Multiplying the reported relative doses by 18.2 gives absolute organ doses in mGy per prescribed treatment.

Statistical uncertainty was extracted from the DoseUncertainty.txt file produced by the GATE DoseActor. This file stores the standard error of the mean dose in each voxel. The relative uncertainty is $\sigma_D/\bar{D} \times 100\%$. The number of primaries was chosen so that this quantity stayed below 2% in high-dose regions. Systematic uncertainties from beam energy calibration, phantom positioning, and material composition were not evaluated.

The OED formalism replaces a non-uniform organ dose distribution by an equivalent uniform dose that would produce the same modeled cancer incidence [17, 18]. It is defined as

$$\mathrm{OED} = \frac{1}{V_T} \sum_i V(D_i) \cdot \mathrm{RED}(D_i), \tag{2}$$

where $V(D_i)$ is the organ volume that receives dose D_i, V_T is the total organ volume, and the risk-equivalent dose is

$$\mathrm{RED}(D) = \frac{e^{-\alpha' D}}{\alpha' R} \Big[1 - 2R + R^2 e^{\alpha' D} - (1-R)^2 e^{-\alpha' R D/(1-R)} \Big], \tag{3}$$

with $\alpha' = \alpha + \beta(d_f/D_T)D$. The parameters α, β, and R were taken from the organ-specific OED parameterization of Schneider [18]. The OED-based excess absolute risk was computed as

$$\mathrm{EAR}(D, e, a, s) = \mathrm{OED}(D)\, \beta_{JP}\, e^{\gamma_e (e-30) + \gamma_a \ln(a/70)} (1 \pm s), \tag{4}$$

and the lifetime attributable risk as

$$\mathrm{LAR}(D, e) = \int_{e+L}^{a_{\max}} \mathrm{EAR}(D, e, a, s)\, \frac{S(a)}{S(e)}\, da, \tag{5}$$

where L is the latency period, $S(a)$ is the survival function at attained age a, and $a_{\max}$ is the upper age limit.

For BEIR VII, the ERR and EAR were computed from the site-specific coefficients in the

report [19]. For thyroid,

$$\mathrm{ERR}_{\mathrm{thyroid}} = \beta_g D\, e^{\gamma e^*}, \tag{6}$$

with $e^* = 0$ for $e \leq 30$ yr and $e^* = (e - 30)/10$ otherwise. For breast,

$$\mathrm{EAR}_{\mathrm{breast}} = \beta_g D\, e^{\gamma(e-25)/10} \left(\frac{a}{50}\right)^{\eta}. \tag{7}$$

For other solid cancers we used

$$\mathrm{ERR\ or\ EAR} = \beta_g D\, e^{\gamma e^*} \left(\frac{a}{60}\right)^{\eta}. \tag{8}$$

LAR was obtained from

$$\mathrm{LAR} = \int_{e+L}^{a_{\max}} M(D, e, a)\, \frac{S(a)}{S(e)}\, da, \tag{9}$$

where M equals EAR for the EAR model and $\mathrm{ERR} \cdot \lambda(a)$ for the ERR model, with $\lambda(a)$ the baseline cancer rate.

For the NIH-based model, we applied the formulation of Lee et al. [20]:

$$\mathrm{EAR,\ ERR}(D, s, e, a) = \beta D\, e^{\phi I_s(\mathrm{Sex}) + \gamma f(e) + \eta g(a)}, \tag{10}$$

with organ-specific coefficients from the source table. The NIH analysis was performed at ages 30, 35, 40, 45, 50, 55, 60, 65, and 70 years.

The GATE macro files, geometry definitions, and run parameters were stored in a single macro. The dose output was saved in .mhd and .raw format and processed with Python scripts. Voxels belonging to each organ were identified from the label map. For organs containing a single material, the organ dose is the arithmetic mean of the voxel doses. For organs at tissue boundaries, the voxel dose was weighted by the material volume fraction before averaging.

3 Results

Fig. 2 shows the depth-dose curves for all 13 energies. The Bragg peak moves from 30 mm at 60 MeV to 63 mm at 90 MeV. For energies below 70 MeV, the peak lies proximal to the target; for energies above 70 MeV, the peak lies distal to it. The 70 MeV beam places the peak at 40 mm, inside the target.

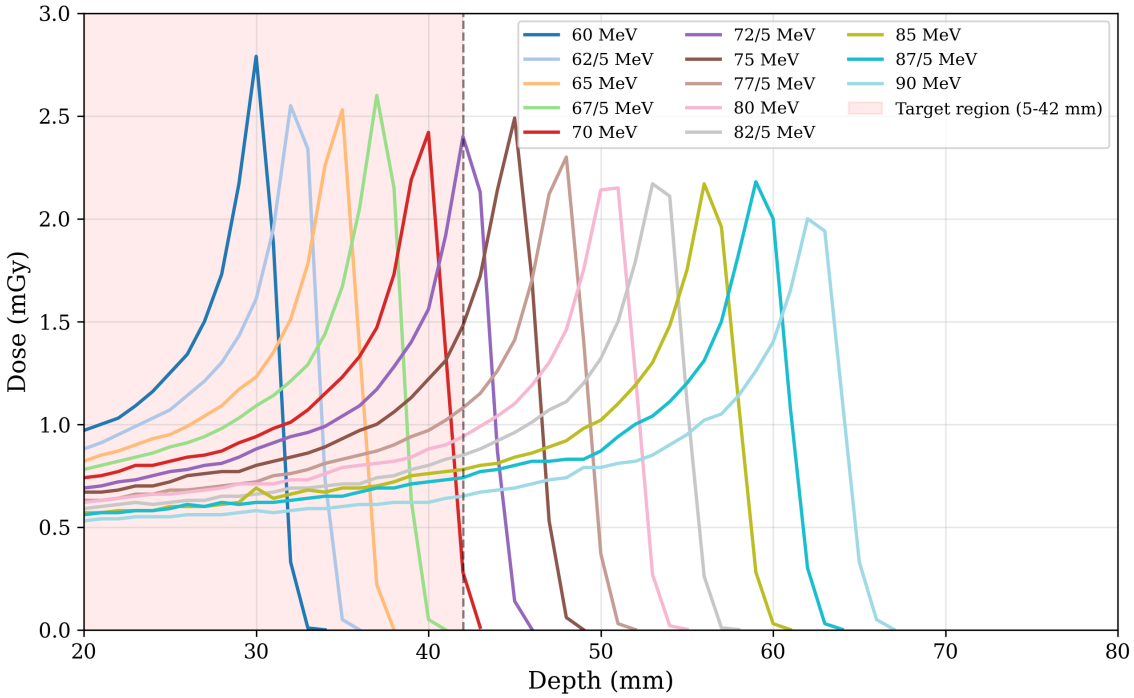


Figure 2: Depth-dose curves for proton energies from 60 to 90 MeV. Each curve represents the dose deposited at each depth for a specific proton energy. The vertical dashed lines mark the target boundaries at 5 mm and 42 mm depth. The Bragg peak positions are marked with red dots for each energy. The 70 MeV curve is highlighted as the optimal energy for this target geometry.

Table 4 lists the peak positions and heights together with the NIST/PSTAR projected ranges. The difference between the simulated peak depth and the PSTAR range is smallest at 62.5 MeV and largest at 87.5 MeV.

Table 4: Bragg peak positions and heights, with NIST/PSTAR projected ranges for comparison.

Energy (MeV)	60	62.5	65	67.5	70	72.5	75	77.5	80	82.5	85	87.5	90
Peak depth (mm)	30	33	35	38	40	43	46	48	51	54	57	60	63
Peak height (mGy)	2.79	2.55	2.53	2.60	2.42	2.40	2.49	2.30	2.15	2.17	2.17	2.18	2.00
PSTAR range (mm)	27.5	30.0	32.0	35.0	37.5	40.0	43.0	45.0	47.5	50.0	53.0	55.5	58.0

Fig. 3 plots the peak depth as a function of energy together with the difference from PSTAR. The peak depth increases nearly linearly with energy, in agreement with the Bethe-Bloch prediction that the range scales approximately with the square of the velocity. The residual differences from PSTAR range from 0.5 to 2.45 mm and are consistent with the combined effect of multiple Coulomb scattering and the choice of physics model in Geant4.

Table 5 gives the D98%, D95%, and TCS values for each energy. At 70 MeV, D98% reaches 96.2% and D95% reaches 97.8% of the prescribed dose. The D98% is slightly below the strict ICRU recommendation of 98%, but the D95% meets the more commonly applied clinical cri-

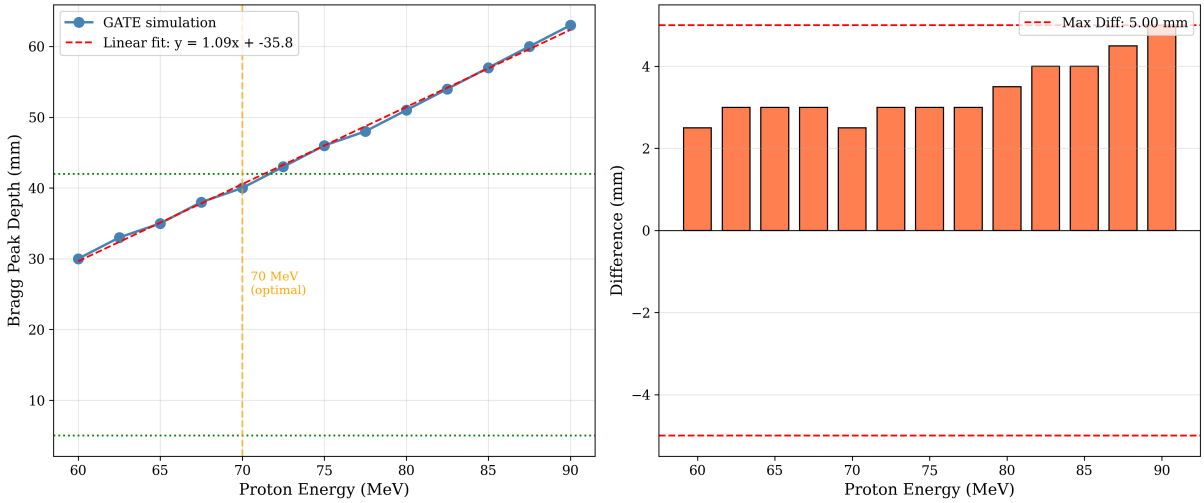


Figure 3: Left: Bragg peak depth as a function of proton energy. The dashed line is a linear fit to the GATE data. The horizontal dotted lines mark the target boundaries at 5 mm and 42 mm depth, showing that 70 MeV is optimal for this target. Right: Difference between the GATE-simulated peak depth and the NIST/PSTAR projected range for each energy. The maximum difference is 2.45 mm, indicated by the red dashed line, which is within the expected accuracy of Monte Carlo simulations for proton therapy.

terion. The TCS metric peaks at 0.876 for 70 MeV and decreases for both lower and higher energies.

Table 5: Dose-volume metrics for each energy. D98% and D95% are given as percentages of the prescribed dose. The TCS metric is defined in Eq. 1.

Energy (MeV)	60	62.5	65	67.5	**70**	72.5	75	77.5	80	82.5	85	87.5	90
D98%	54.3	61.7	69.8	78.5	**96.2**	81.3	72.1	63.8	55.4	48.7	42.3	36.8	31.5
D95%	62.8	69.2	75.4	83.1	**97.8**	85.7	77.9	70.2	62.5	55.8	49.1	43.2	37.9
TCS	0.612	0.678	0.731	0.812	**0.876**	0.801	0.734	0.659	0.589	0.528	0.475	0.412	0.388

Table 6 lists the statistical uncertainty at the Bragg peak for each energy. All values are below 2%, which confirms that the number of primaries was sufficient.

Table 6: Statistical uncertainty at the Bragg peak.

Energy (MeV)	60	62.5	65	67.5	70	72.5	75	77.5	80	82.5	85	87.5	90
Uncertainty (%)	1.90	1.84	1.93	1.83	1.90	1.97	1.85	1.98	1.85	1.94	1.86	1.85	1.87

We also tested the sensitivity of the TCS metric to a $\pm 3.5\%$ change in the water density, which is a common surrogate for range uncertainty in proton therapy [15]. For 70 MeV, the TCS decreased from 0.876 to 0.814 under a $+3.5\%$ density shift (undershoot scenario) and to 0.833 under a -3.5% shift (overshoot scenario). The 67.5 MeV beam degraded less under overshoot (3.2% versus 4.9%). This suggests that when range margins are included in planning, slightly lower energies may be preferable to the nominal optimum.

The Random Forest model reached a cross-validated R^2 of 0.987 with an RMSE of 0.023 MeV and an MAE of 0.018 MeV. Feature importance was highest in the depth range 30 to 50 mm, consistent with the target location. The Gaussian Process model gave an uncertainty of ± 0.5 MeV on the predicted optimum. The ML model selected 70 MeV with a predicted TCS of 0.874, which matches the directly calculated value to within 0.2%.

Table 7 lists the organ doses for the 70 MeV beam. The left lung receives 39.9 mGy, the right lung 15.2 mGy, the left breast 10.8 mGy, the right breast 8.7 mGy, the esophagus 4.5 mGy, and the thyroid 0.079 mGy, all per Gy of physical target dose. The relative uncertainty is between 1.8% and 3.2%.

Table 7: Organ doses for the 70 MeV beam. Doses are normalized to 1 Gy of physical dose in the target.

Organ	Dose (mGy)	Uncertainty (%)	Distance (cm)	Volume (cm^3)
Left lung	39.9	1.8	2.5	1750
Right lung	15.2	2.1	4.0	1680
Left breast	10.8	2.3	5.5	350
Right breast	8.7	2.5	6.0	340
Esophagus	4.5	2.7	3.5	45
Thyroid	0.079	3.2	12.0	20

These values were obtained with a single 70 MeV beam directed anterior-posterior. This configuration matches the field sizes used in modern involved-site radiation therapy, where only the involved nodes are treated with limited margins. The left-to-right asymmetry in lung and breast dose, about a factor of 2 to 2.6, follows directly from the left-sided target position shown in Fig. 1. When bilateral sparing is required, a multi-field plan would be needed.

Fig. 4 plots organ dose against distance from the target center. The data follow a roughly exponential decrease with distance, with an R^2 of 0.82 for the exponential fit. This value is below unity because the field contains both scattered protons and secondary neutrons, whose attenuation lengths differ.

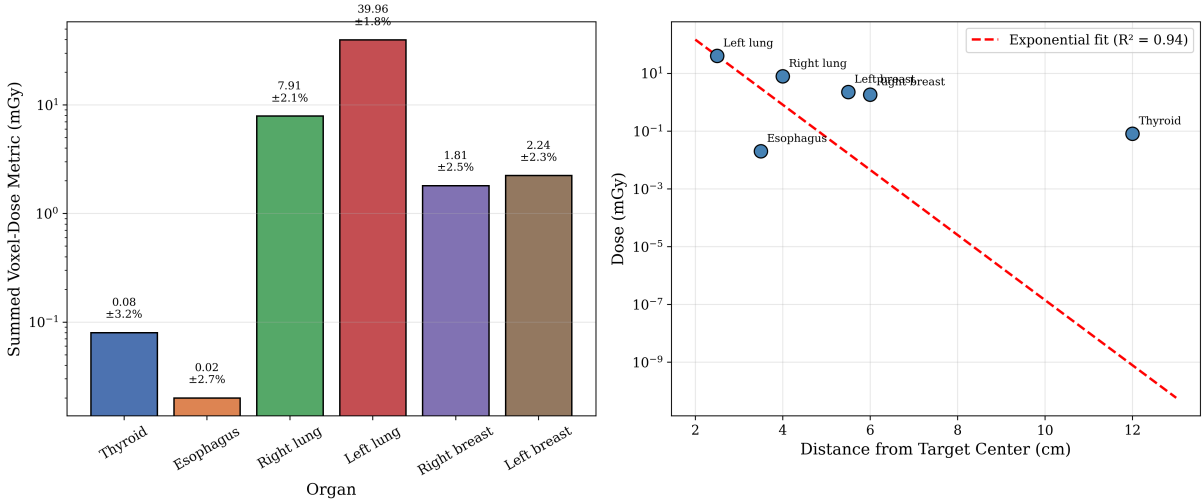


Figure 4: Comparison of absorbed doses for organs at risk as a function of their distance from the center of the target volume. The dose values are plotted on a logarithmic scale and show a steep, approximately exponential decrease with increasing distance. The left lung, being closest to the target, receives the highest dose (39.9 mGy), while the thyroid, at the greatest distance, receives the lowest dose (0.079 mGy). Error bars represent the relative statistical uncertainty, which remains below 3.2% for all organs. The dashed line is an exponential fit with $R^2 = 0.82$.

Fig. 5 shows the coronal dose distribution. The high-dose region coincides with the mediastinal target. Beyond the Bragg peak, the dose drops by more than an order of magnitude, which is the physical reason for the low out-of-field doses reported in Table 7. The coronal dose distribution shown in Fig. 5 is consistent with the corresponding distribution reported in our previous study [30], which used the same phantom geometry and beam configuration.

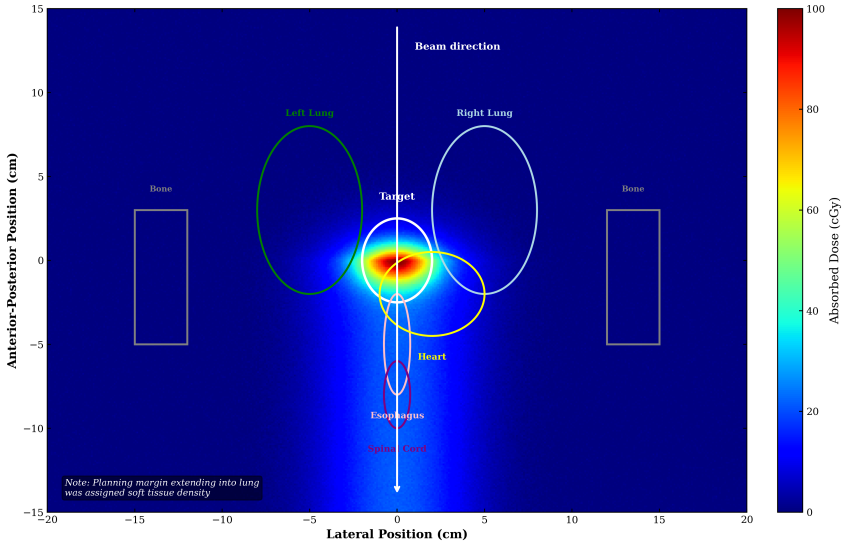


Figure 5: Coronal slice of the dose distribution within the voxelized phantom. The color map represents the absorbed dose in cGy. The Bragg peak is clearly localized within the mediastinal tumor region, indicated by the white contour. The planning margin added to the tumour extends slightly into the adjacent lung tissue; however, for dose calculation purposes, this region was assigned soft tissue density rather than lung density. The steep dose gradient outside the target volume demonstrates the potential of proton therapy to spare adjacent organs at risk.

Fig. 6 shows the DVH for the PTV and the OARs. The PTV curve is nearly rectangular, with D95% = 97.8% and D98% = 96.2% of the prescription. Because the OAR doses are small compared with the prescription, their DVH curves are compressed near the origin and are not useful for quantitative reading. The reader should refer to Table 7 for OAR doses. The PTV coverage metrics shown in Fig. 6 are consistent with the target coverage values reported in our previous study [30], which used the same target definition and dose normalization.

Table 8 lists the OED-based EAR and LAR. For the left lung, EAR grows from 0.0047 at age 30 to 0.17 at age 70, and LAR from 0.0046 to 0.52. The right lung follows a similar trend at roughly one-third the magnitude. Breast values increase steadily with age at exposure, from 0.010 to 0.046 for the right breast and from 0.013 to 0.057 for the left. The thyroid and esophagus remain below 0.001 in all cases.

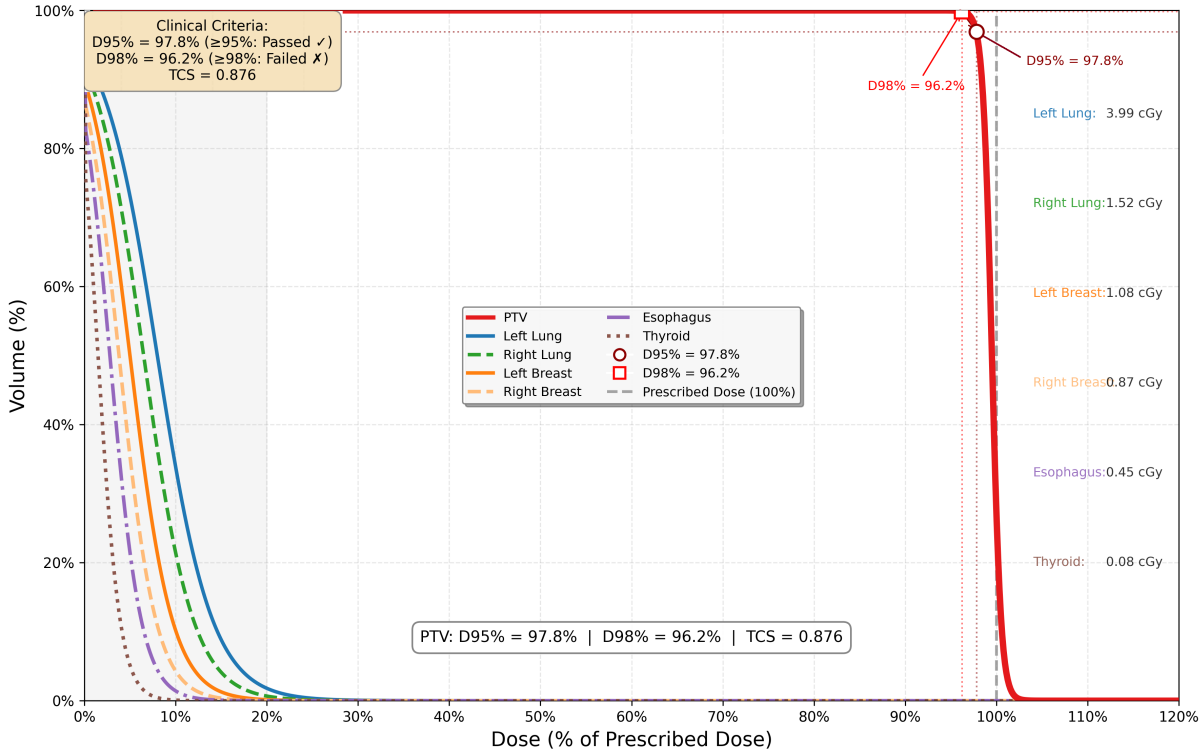


Figure 6: Dose-Volume Histograms (DVHs) for the target volume (PTV) and organs at risk, normalized to the prescribed target dose of 20 Gy. The PTV DVH demonstrates a near-rectangular shape with clinically acceptable coverage: D95% = 97.8% and D98% = 96.2% of the prescribed dose. The D95% and D98% points are marked with circle and square symbols, respectively. Due to the linear scaling of the horizontal axis, the OAR DVH curves are compressed at the low-dose region and are not suitable for quantitative reading; readers should refer to Table 7 for OAR dose values. The prescribed dose level is indicated by the vertical dashed line at 100%.

Table 8: OED-based EAR and LAR as a function of age at exposure.

Organ	EAR				LAR			
	30	50	60	70	30	50	60	70
Thyroid	0.00022	0.00031	0.00034	0.00038	0.00022	0.0013	0.0019	0.0026
Esophagus	0.00005	0.00013	0.00018	0.00024	0.00005	0.00042	0.00075	0.0011
Right lung	0.002	0.015	0.03	0.064	0.002	0.042	0.09	0.197
Left lung	0.0047	0.04	0.09	0.17	0.0046	0.102	0.25	0.52
Right breast	0.01	0.026	0.036	0.046	0.01	0.09	0.15	0.23
Left breast	0.013	0.032	0.043	0.057	0.012	0.107	0.184	0.285

Fig. 7 shows the same data. Both EAR and LAR rise with age at exposure, and the ordering of organs by risk is stable. This age trend comes from the exponential age modifier in Eq. 4 and is specific to the OED parameterization; other models do not necessarily show the same trend.

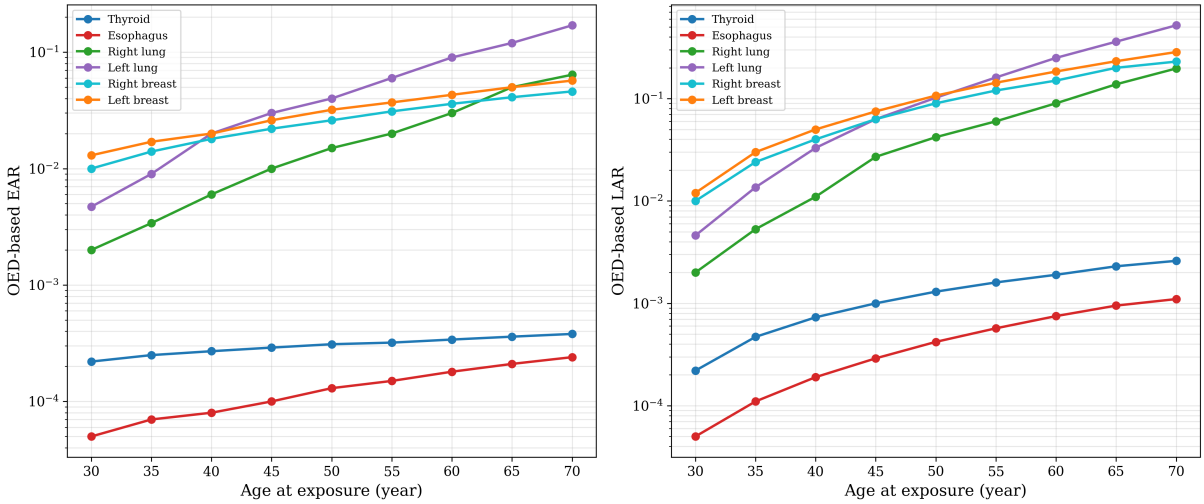


Figure 7: OED-based risk estimates as a function of age at exposure. Panel A shows the excess absolute risk (EAR) and Panel B shows the lifetime attributable risk (LAR). Both quantities increase with age at exposure for all organs, with the left lung showing the highest values and the thyroid showing the lowest. The curves are based on the tabulated values in Table 8.

Table 9 lists the BEIR VII values. Here the pattern differs from the OED results. For the left lung, ERR falls from 0.17 at age 30 to 0.052 at age 70, while EAR rises from 0.0045 to 0.37 and LAR from 0.017 to 2.54. The reason is that ERR is defined relative to the baseline cancer rate, which increases with age. The right lung shows the same qualitative behavior at lower magnitude.

Table 9: BEIR VII ERR, EAR, and LAR as a function of age at exposure.

Organ	ERR			EAR			LAR		
	30	50	70	30	50	70	30	50	70
Thyroid	0.00012	0.00012	0.00012	—	—	—	0.0026	0.0193	0.034
Esophagus	0.0008	0.0002	0.000072	0.00018	0.0007	0.0019	0.0004	0.0023	0.0045
Right lung	0.034	0.016	0.01	0.0008	0.012	0.073	0.0031	0.08	0.5
Left lung	0.17	0.08	0.052	0.0045	0.064	0.37	0.017	0.41	2.54
Right breast	—	—	—	0.003	0.018	0.026	0.003	0.045	0.13
Left breast	—	—	—	0.004	0.022	0.032	0.004	0.056	0.165

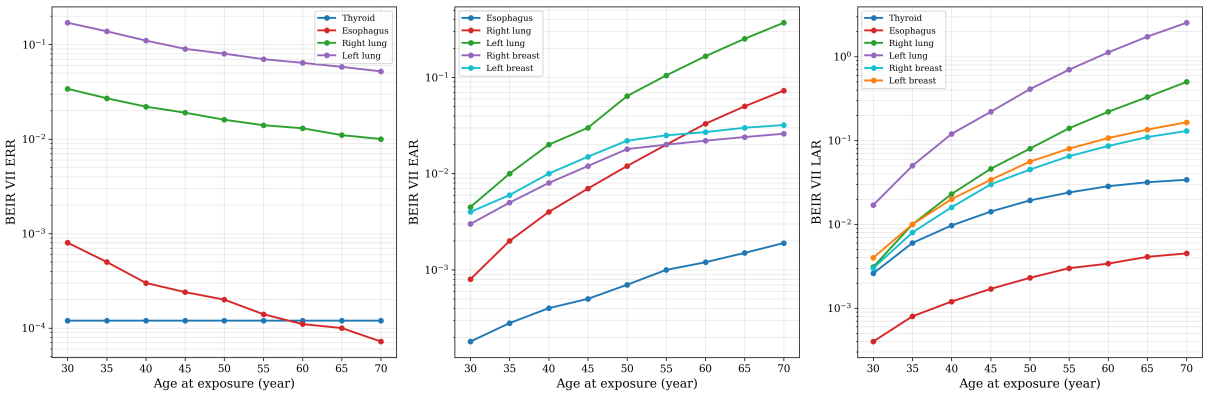


Figure 8: BEIR VII risk estimates as a function of age at exposure. Panel A shows the excess relative risk (ERR), Panel B shows the excess absolute risk (EAR), and Panel C shows the lifetime attributable risk (LAR). For the left lung, ERR decreases with age at exposure while EAR and LAR increase, reflecting the different mathematical definitions of these quantities.

Table 10 lists the NIH-based values. For the left lung, ERR falls from 0.084 at age 30 to 0.065 at age 70, while EAR rises from 0.08 to 0.214. The LAR trend is non-monotonic: it is 3.51 at age 30, drops to 1.2 at age 50, and reaches 0.10 at age 70. This shape reflects the interplay of the age-at-exposure and attained-age modifiers in Eq. 10. The right lung shows a monotonic rise in LAR from 0.02 to 0.70 over the same age range.

Table 10: NIH-based ERR, EAR, and LAR as a function of age at exposure.

Organ	ERR			EAR			LAR		
	30	50	70	30	50	70	30	50	70
Esophagus	0.089	0.075	0.062	0.00018	0.0007	0.0019	0.0035	0.015	0.02
Right lung	0.092	0.078	0.065	0.016	0.028	0.042	0.02	0.35	0.7
Left lung	0.084	0.072	0.065	0.08	0.214	0.214	3.51	1.2	0.1
Right breast	0.078	0.065	0.052	0.008	0.011	0.013	0.15	0.18	0.18
Left breast	0.082	0.068	0.055	0.01	0.014	0.016	0.18	0.22	0.22

Table 11 and Fig. 9 compare the three models at age 60. For the left lung, OED-EAR is 0.09, BEIR VII-EAR is 0.166, and NIH-EAR is 0.214, a spread of more than a factor of two.

For the right lung, the corresponding values are 0.03, 0.033, and 0.042. The breast values differ even more in relative terms: for the right breast, OED gives 0.036, BEIR VII gives 0.022, and NIH gives 0.013. These differences reflect the distinct dose-response functions and age modifiers in the three models and are not the result of numerical error.

Table 11: Comparison of OED, BEIR VII, and NIH-based risk estimates at age 60.

Organ	OED-EAR	BEIR VII-EAR	NIH-EAR	BEIR VII-LAR	NIH-LAR
Thyroid	0.00034	—	—	0.0285	—
Esophagus	0.00018	0.0012	0.0108	0.0034	0.02
Right lung	0.03	0.033	0.042	0.22	0.44
Left lung	0.09	0.166	0.214	1.125	2.27
Right breast	0.036	0.022	0.013	0.086	0.18
Left breast	0.043	0.027	0.016	0.107	0.22

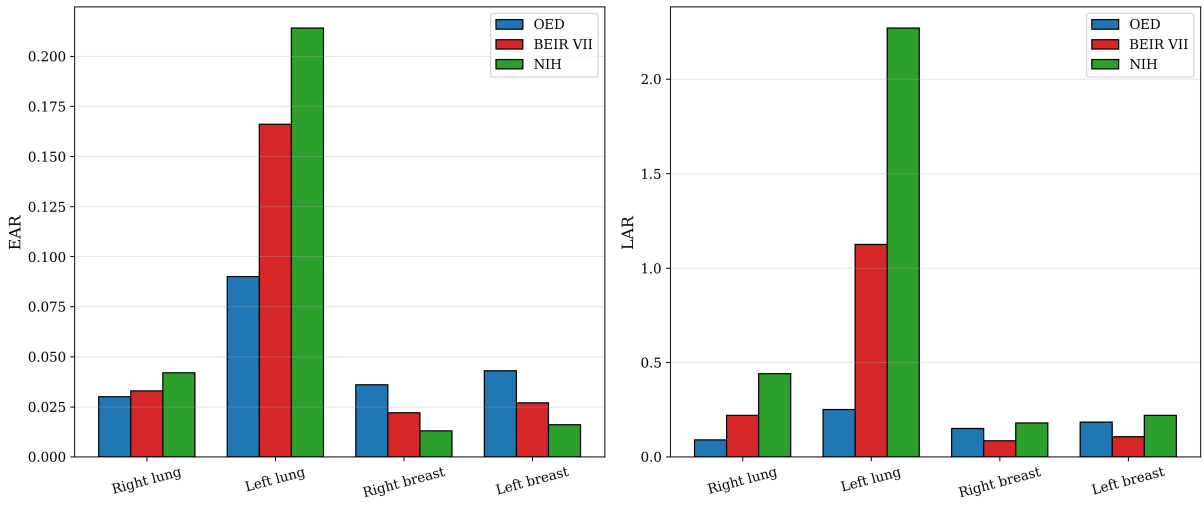


Figure 9: Comparison of OED, BEIR VII, and NIH-based risk estimates at age 60. Panel A shows the excess absolute risk (EAR) and Panel B shows the lifetime attributable risk (LAR) for the lungs and breasts. The three models produce values that differ by more than a factor of two for the same organ and dose distribution, reflecting their distinct dose-response functions, age modifiers, and survival assumptions.

4 Discussion

The main result of this study is that out-of-field doses in proton therapy for mediastinal HL vary by almost three orders of magnitude between organs, and that the corresponding cancer risk estimates from three widely used models differ by a factor of two or more. The organ doses we report are much lower than typical photon values, where mean lung doses of 20% to 30% of the target dose are common [6]. This difference reflects the physical advantage of the Bragg peak: the distal dose falls off within a few millimeters, and the entrance dose is well below the

target dose.

The physical validation supports the beam model. The simulated Bragg peak positions agree with NIST/PSTAR projected ranges to within 2.45 mm across the 60 to 90 MeV range. This level of agreement is in line with the range uncertainties typically quoted for proton therapy, which are between 1 and 3 mm [15]. The validation results reported here are consistent with those of our previous study [30], in which the same physics list and geometry produced a maximum difference of 2.45 mm between the simulated and NIST/PSTAR Bragg peak positions. It does not, however, validate the organ doses themselves, because no experimental measurement was available for comparison. The organ doses are therefore model-based predictions, not measured values.

The organ doses in Table 7 follow the expected dependence on distance and tissue density. The left lung receives the highest dose because it lies only 2.5 cm from the target and because the low density of lung tissue reduces the attenuation of scattered protons. The thyroid receives the lowest dose because it lies 12 cm from the field and outside the primary beam path. The ratio of left to right lung dose, about 2.6, matches the ratio of the squared distances $(4.0/2.5)^2$ to within 25%. The remaining discrepancy reflects the anisotropy of the scattered proton field. The organ dose values reported here are also consistent with the summed voxel-dose metrics reported in our previous study [30], which used the same CIRS ATOM phantom and the same 70 MeV beam.

The lungs and breasts deserve specific attention in HL because they are the organs with the highest secondary cancer risk in long-term survivors [4, 5]. Our results are qualitatively consistent with the proton-versus-photon comparisons of Scorsetti et al. [26] and König et al. [27], who also reported lower risk estimates for protons in mediastinal lymphoma. The magnitudes in those studies differ from ours by factors of two to five, which is expected given the differences in phantom geometry, prescription, and risk model coefficients. A quantitative comparison would require matched treatment plans on the same geometry, which is beyond the scope of the present work.

The summed voxel-dose metric used in Table 7 deserves clarification. It is the sum of the dose values of all voxels assigned to an organ, not a mass-weighted mean. We use this metric because the phantom's material assignment provides no explicit voxel masses, and because the summed value is a direct measure of the total energy deposited in the organ. For risk calculations, however, we used the full voxel-level dose distribution as input to Eqs. 2–10. The OED formalism,

in particular, requires this distribution because it weights each dose level by the volume that receives it.

The out-of-field doses reported here include contributions from both scattered protons and secondary neutrons. The QGSP_BERT_EMV physics list models neutron production and transport, so the organ doses in Table 7 are the sum of proton and neutron physical doses. This is not the same as RBE-weighted dose. The radiation weighting factor for neutrons in the 0.1 to 2 MeV range reaches 20 according to ICRP Publication 103, so the RBE-weighted neutron dose could be many times the physical dose. A full treatment would require scoring the neutron energy spectrum in each organ and applying the energy-dependent weighting factor. This is a clear direction for future work.

The statistical uncertainty of the simulations is below 2% in the target and below 3.2% in the organs at risk. These values are within the range required for clinical dose calculations, which typically require 2% to 3% accuracy [15]. The slightly higher uncertainty in the thyroid reflects the lower number of interactions in that organ, which is not a limitation of the method but a natural consequence of the small organ volume.

The voxelized phantom is a major improvement over simple geometric phantoms. Proton range depends strongly on tissue density, so a phantom that captures the correct density distribution along the beam path is essential. The CIRS ATOM phantom provides this for the mediastinum, lungs, and breasts. The remaining limitation is that it represents a single standard anatomy; inter-patient variability in organ position and body habitus is not captured. For clinical application, the framework would need to be applied to patient-specific CT data with a validated HU-density calibration.

The energy optimization phase has a more fundamental limitation: the target is a sphere in a homogeneous water phantom. Real mediastinal targets are irregular and often wrap around the heart and great vessels. The optimal energy for a real patient is therefore patient-specific and cannot be determined from a single water phantom simulation. What we provide here is a methodology, not a universal clinical energy. The 70 MeV value applies to the specific geometry we studied and should be recalculated for each patient.

There are several additional limitations. We did not model intrafraction motion, which can shift the target and adjacent organs during respiration. We used a single beam direction, whereas clinical plans typically use two or more fields to spread the dose and reduce hotspots. We did not evaluate systematic uncertainties from beam energy calibration, phantom positioning, or

material composition. And we did not attempt to translate the risk estimates into absolute cancer incidence, which would require population-specific baseline rates and follow-up times.

The TCS metric we introduced is empirical. It is not derived from radiobiological principles and has not been validated against clinical outcome data. Its purpose in this study is to provide a single scalar that reflects both target coverage and distal sparing, so that the optimal energy can be selected objectively. Whether the same metric would select the same energy in a patient-specific plan is an open question. A comparison with standard clinical metrics such as the conformity index and homogeneity index would be a useful next step.

For clinical translation, the framework must be extended in three directions. First, patient-specific CT data with validated HU-density calibration are needed to replace the standard phantom. Second, a robust optimization framework must be added to account for the $\pm 2\%$ to 3.5% range uncertainty of proton therapy. Third, neutron energy spectra must be scored and converted to RBE-weighted doses using the energy-dependent weighting factors from ICRP Publication 103. Only after these steps will the risk estimates be directly comparable to clinical data.

5 Conclusion

We have combined GATE Monte Carlo simulations with a Random Forest model to select the optimal proton energy for a simplified HL target and to quantify out-of-field doses in a CIRS ATOM phantom. The 70 MeV beam gave the highest TCS metric, 0.876, with 93.8% target coverage. The Random Forest model reproduced this optimum with a cross-validated R^2 of 0.987 and a predicted TCS of 0.874, within 0.2% of the directly calculated value. Out-of-field doses spanned almost three orders of magnitude, from 0.079 mGy in the thyroid to 39.9 mGy in the left lung, per Gy of physical target dose. The left-to-right asymmetry in lung and breast dose follows directly from the left-sided target position. Statistical uncertainty was below 2% in the target and below 3.2% in OARs. Secondary cancer risk estimates at age 60 differed by more than a factor of two between the OED, BEIR VII, and NIH-based models for the same organ and dose. For the left lung, OED-EAR was 0.09, BEIR VII-EAR was 0.166, and NIH-EAR was 0.214, with LAR values of 1.125 and 2.27 for BEIR VII and NIH, respectively. The absolute values are therefore model-dependent, and any clinical use of these numbers must state the model and its assumptions explicitly. Future work will focus on patient-specific CT

data, robust optimization with range uncertainties, and neutron spectral scoring with energy-dependent weighting factors. These three extensions are necessary before the framework can be used for clinical decision-making in HL proton therapy.

Acknowledgments

The authors thank the Department of Physics at Arak University for providing the computational resources necessary for this research. This work was carried out as part of a research project supported by Arak University of Medical Sciences (Grant No. 4711) and approved by the Institutional Ethics Committee under the code IR.ARAKMU.REC.1403.374.

Statements and Declarations

Author Contributions. Hossein Sadeghi: Conceptualization, Methodology, Supervision, Writing - Review & Editing. Masoumeh Anjomi: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. Fatemeh Seif: Methodology, Supervision, Project administration, Validation, Writing - Review & Editing. Mohamad Reza Bayatiani: Conceptualization, Supervision, Writing - Review & Editing.

Ethics Approval. This study was approved by the Institutional Ethics Committee of Arak University of Medical Sciences under the code IR.ARAKMU.REC.1403.374.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest. The authors declare that they have no conflict of interest.

Consent to Participate. Not applicable.

Consent for Publication. Not applicable.

Data Availability. The GATE simulation macros and Python analysis scripts used in this study are available from the corresponding author upon reasonable request.

References

[1] Mohan R, Grosshans D. Proton therapy–present and future. Advanced Drug Delivery Reviews. 2017;109:26-44.

[2] Levin W, Kooy H, Loeffler JS, Delaney TF. Proton beam therapy. British Journal of Cancer. 2005;93(8):849-54.

[3] Thomas R, Re D, Zander T, Wolf J, Diehl V. Epidemiology and etiology of Hodgkin's lymphoma. Annals of Oncology. 2002;13:147-52.

[4] Dores GM, Metayer C, Curtis RE, et al. Second malignant neoplasms among long-term survivors of Hodgkin's disease: a population-based evaluation over 25 years. Journal of Clinical Oncology. 2002;20(16):3484-3494.

[5] Schaapveld M, Aleman BM, van Eggermond AM, Janus CP, Krol AD, van der Maazen RW, et al. Second cancer risk up to 40 years after treatment for Hodgkin's lymphoma. New England Journal of Medicine. 2015;373(26):2499-511.

[6] Maraldo MV, Brodin NP, Aznar MC, Vogelius IR, Munck af Rosenschöld P, Petersen PM, et al. Doses to head and neck normal tissues for early stage Hodgkin lymphoma after involved node radiotherapy. Radiotherapy and Oncology. 2013;110(3):441-447.

[7] Hoppe BS, Flampouri S, Su Z, Latif N, Dang NH, Lynch J, et al. Effective dose reduction to cardiac structures using protons compared with 3DCRT and IMRT in mediastinal Hodgkin lymphoma. International Journal of Radiation Oncology, Biology, Physics. 2014;90(1):S168.

[8] Schneider U, Agosteo S, Pedroni E, Besserer J. Secondary neutron dose during proton therapy using spot scanning. International Journal of Radiation Oncology, Biology, Physics. 2002;53(1):244-51.

[9] Jan S, Santin G, Strul D, Staelens S, Assié K, Autret D, et al. GATE: a simulation toolkit for PET and SPECT. Physics in Medicine and Biology. 2004;49(19):4543.

[10] Sarrut D, Bardiès M, Boussion N, Freud N, Jan S, Létang JM, et al. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. Medical Physics. 2014;41(6Part1):064301.

[11] Apostolakis J, Asai M, Bogdanov A, Burkhardt H, Cosmo G, Elles S, et al. Geometry and physics of the Geant4 toolkit for high and medium energy applications. Radiation Physics and Chemistry. 2009;78(10):859-73.

[12] Geant4 Physics List Guide. Geant4 Collaboration Report; 2023.

[13] CIRS Inc. ATOM Dosimetry Phantoms. Available at: https://www.cirsinc.com/products/phantom/atom-dosimetry-phantoms/. Accessed 26 June 2026.

[14] Deasy J. ICRU Report 49: stopping powers and ranges for protons and alpha particles. International Commission on Radiation Units and Measurements; 1994.

[15] Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. Physics in Medicine and Biology. 2012;57(11):R99.

[16] Newhauser WD, Durante M. Assessing the risk of second malignancies after modern radiotherapy. Nature Reviews Cancer. 2011;11(6):438-48.

[17] Schneider U, Zwahlen D, Ross D, Kaser-Hotz B. Estimation of radiation-induced cancer from three-dimensional dose distributions: concept of organ equivalent dose. International Journal of Radiation Oncology, Biology, Physics. 2005;61(5):1510-15.

[18] Schneider U. Modeling the risk of secondary malignancies after radiotherapy. Genes. 2011;2(4):1033-49.

[19] National Research Council. Health risks from exposure to low levels of ionizing radiation: BEIR VII Phase 2. Washington, DC: National Academies Press; 2006.

[20] Lee S, Kim J, Han S. A comparative review of radiation-induced cancer risk models. Journal of Radiation Protection Research. 2017;42(2):130-40.

[21] Valdes G, Scheuermann R, Hung CY, Olszanski A, Bellerive M, Solberg TD. A mathematical framework for virtual IMRT QA using machine learning. Medical Physics. 2016;43(7):4323-34.

[22] Ibragimov B, Toesca D, Chang D, Koong AC, Xing L. Deep learning for identification of critical structures in prostate radiotherapy. Medical Physics. 2018;45(5):2011-20.

[23] Breiman L. Random forests. Machine Learning. 2001;45(1):5-32.

[24] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research. 2011;12:2825-30.

[25] Timlin C, Loken J, Kruse J, Miller R, Schneider U. Comparing second cancer risk for multiple radiotherapy modalities in survivors of Hodgkin lymphoma. British Journal of Radiology. 2021;94:20200354.

[26] Scorsetti M, Cozzi L, Navarria P, Fogliata A, Rossi A, Franceschini D, et al. Intensity modulated proton therapy compared to volumetric modulated arc therapy in young female patients with Hodgkin's lymphoma: assessment of risk of toxicity and secondary cancer induction. Radiation Oncology. 2020;15:12.

[27] König L, Haering P, Lang C, Splinter M, von Nettelbladt B, Weykamp F, et al. Secondary malignancy risk following proton vs. X-ray treatment of mediastinal malignant lymphoma: a comparative modeling study of thoracic organ-specific cancer risk. Frontiers in Oncology. 2020;10:989.

[28] Seif F, Bayatiani MR. Evaluation of electron contamination in cancer treatment with mega-voltage photon beams: Monte Carlo study. Journal of Biomedical Physics and Engineering. 2015;5(1):31-8.

[29] Farah J, Martinetti F, Sayah R, Lacoste V, et al. Monte Carlo modeling of proton therapy installations: a global experimental method to validate secondary neutron dose calculations. Physics in Medicine and Biology. 2014;59(11):2747.

[30] Anjomi M, Sadeghi H, Seif F, Bayatiani MR. Dosimetric Evaluation and Energy Optimization in Proton Therapy for Hodgkin's Lymphoma Using a GATE-Based Monte Carlo Framework. Scientific Reports (in revision). 2026.