

Research on Deep Learning-Based Relative Biological Effectiveness Prediction Method

Yiying Wang

Shandong University

Zhao Peng

Xiangya Hospital of Central South University

Yunzhou Xia

Chinese Academy of Science Heavy Ion Medicine (CASHIM) Co. Ltd

Xianghong Jia

Shandong University

Quanqi Shi

Shandong University

Kun Wang

National Institute of Metrology

Chunyang Li

Beijing Institute of Radiation Medicine

Xiaomin Zhang

Beijing Institute of Radiation Medicine

Yaping Qi

qiyp@nim.ac.cn

National Institute of Metrology

Research Article

Keywords: Deep learning, Relative biological effectiveness, Regression prediction, Particle Radiotherapy

Posted Date: June 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10052695/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This study proposes a deep learning-based regression prediction method utilizing the PIDE3.4 radiation biological effects database for predicting Relative Biological Effectiveness (RBE). The method employs a matrix-based data processing pipeline, applying MinMax normalization to numerical features and One-Hot encoding to categorical features. The experimental dataset comprises 999 samples, with 800 allocated for training and 199 reserved as an independent test set. Results demonstrate that the method achieves high accuracy and stability in RBE prediction. During training, a 5-fold cross-validation approach yielded an average coefficient of determination (R^2) of 0.8959 and a Mean Absolute Error (MAE) of 0.2935. Predicted values closely align with ground truth values within the primary RBE range of 0–7, providing a reliable data-driven tool for predicting radiation biological effects and optimizing particle therapy treatment planning. Furthermore, applying the method to a clinical nasopharyngeal carcinoma case in carbon ion radiotherapy, unlike existing RBE prediction methods that focus solely on a single biological endpoint, this method overcomes this limitation. By integrating physical parameters—such as physical dose and linear energy transfer—along with biological parameters like region-specific cell types, the method achieves spatially heterogeneous prediction of RBE. Additionally, this study innovatively constructs a cross-cell RBE prediction framework based on ten-dimensional continuous cell-biological features, enabling RBE prediction for cell types absent from the training data, and achieved a generalization performance of $R^2 = 0.7600$ and $MAE = 0.4055$ on unseen cell types. This provides a novel approach for precisely assessing biological effects on tumors and organs at risk and for optimizing individualized treatments.

1. Introduction

The assessment of Relative Biological Effectiveness (RBE) is an essential issue in particle therapy and radiation protection. RBE is defined as the ratio of the dose required to produce the same biological effect as a reference radiation (typically X-rays or gamma rays) by a specific radiation (e.g., protons or carbon ions) [1]. RBE evaluation is a complex, multifactorial process that requires comprehensive consideration of both biological and physical factors, involving both quantitative parameters (e.g., physical dose and linear energy transfer (LET)) and categorical variables (e.g., radiation particles and cell types) [2].

In clinical proton radiotherapy, the RBE constant of 1.1 is traditionally used to calculate biological doses. However, recent white papers [31][32] and research from multiple centers in Europe have highlighted the limitations of fixed RBE values, leading to the gradual development of mechanistic models [26][27][28] [29]. For heavy ion radiotherapy, clinical practice adopts RBE models based on experimental data accumulation. Existing RBE models include Local Effect Model (LEM) [23][30][34] and Microdosimetric Kinetic Model (MKM) [24][35], both of which have been implemented in heavy ion therapy planning systems [3]. The Logistic Nanodosimetric Model (LNDM) [25], integrating nanoscale energy spectrum statistics with logistic regression to develop a more precise and personalized RBE expression was also proposed. These RBE models are typically validated using only 2–3 cell lines (e.g., HSG, V79). Although these models perform well on individual cell lines, their predictions vary substantially across experimental conditions and cell line types. Furthermore, the RBE varies with cell type, but its variation pattern remains unclear, making parametric modeling unfeasible.

The long-standing bottlenecks that have hindered breakthroughs in RBE modeling are precisely where the potential of deep learning technology lies. In recent years, deep learning has demonstrated tremendous potential in biomedical applications, particularly in medical imaging. Models like U-Net have achieved tumor segmentation with accuracy rivaling that of experts [4][5]. For tumor prognosis, models such as DeepSurv integrate radiomics and genomic data to provide personalized predictions of patient survival and post-radiotherapy toxicity assessment [6][9]. In radiotherapy planning optimization, deep learning enables automatic target delineation and dose distribution prediction [7] (e.g., DoseGAN generates proton therapy plans [8]). The end-to-end learning and automatic feature extraction capabilities of deep learning enable it to automatically learn large-scale experimental data across complex scenarios involving multiple cell types and radiation conditions, offering a novel prediction method to overcome the limitations of traditional RBE models in cell lines and data scale.

Therefore, this study pioneers the development of a region-specific heterogeneous RBE prediction method using deep learning. Building upon published radiation biology experimental data, we constructed a deep learning method demonstrating high prediction accuracy. This method can integrate multi-dimensional features including categorical variables, effectively identifying critical factors often overlooked in traditional RBE models—such as cell line type and radiation particle characteristics—that influence predicted RBE values. It is not intended to replace RBE prediction models such as LEM and MKM, but rather to provide a complementary data-driven approach. Furthermore, this method is applied to a nasopharyngeal carcinoma clinical case, predicting RBE distribution through region-specific heterogeneous cell type configurations, providing a new perspective to explain why biological effects predicted by a single cell type seriously deviate from actual clinical outcomes, thereby offering an interpretation of this clinical phenomenon that differs from traditional RBE models. The developed deep learning RBE method not only provides a novel pre-computational tool for radiation biology experiment design but also establishes a new pathway for personalized RBE assessment and treatment plan optimization in particle therapy.

2. Methods

2.1 Data sources

The source data for deep learning were derived from PIDE3.4 [10, 11] and the cell survival curve experimental data reported by Paganetti et al. in 2014 [12]. There is partial overlap between the PIDE3.4 database and the Paganetti 2014 dataset in terms of cell lines and radiation conditions; however, the two datasets are not entirely consistent in their feature dimensions. Specifically, PIDE3.4 contains complete linear-quadratic (LQ) model parameters (α and β values), whereas the Paganetti 2014 dataset directly provides RBE values. In this study, the two datasets were merged after deduplication based on cell line names and radiation conditions, resulting in a total of 999 samples. The merged dataset was then randomly divided into a training set (800 samples) and an independent test set (199 samples) at a ratio of 8:2. The training set was used for 5-fold cross-validation, while the test set was reserved solely for final performance evaluation and was not involved in method training.

PIDE3.4 is the latest version of the particle irradiation dataset (www.gsi.de/bio-pide) published by the German Heavy Ion Research Center (GSI). This study uses 885 sets of sample data, which include 119 types of cells (V79, T1, NB1RGB, HSG, CHO, etc.), 20 types of radiation particles (light ions such as protons, photons, and heavy ions such as carbon ions), and 2 irradiation methods (single-energy and Bragg broadened beam). PIDE includes in vitro cell survival experiment results collected through literature review, all of which were conducted in pairs after photon and certain ion types of irradiation. The dose response curves are represented by parameters of a linear quadratic (LQ) model, providing original experimental data including cell information, radiation information, dose, linear energy transfer (LET), and LQ curve parameters. Compared to single RBE biological experiment data, the large dataset in PIDE provides high-quality source data for deep learning to accurately predict RBE. Since some features in PIDE3.4 have similar meanings, this paper only selects part of the features from PIDE3.4 for deep learning input, as shown in Table 1.

Furthermore, this study utilized experimental data from the cell survival curve reported by Paganetti et al. in 2014 [12]. Given the limited sample size of cell types, only 114 experimental datasets were selected. To ensure consistency in experimental conditions, the analysis specifically examined key parameters in the Paganetti dataset, including cell lines, radiation source types, and dose rates, thereby guaranteeing the comparability and scientific validity of the findings.

2.2 Data preprocessing

2.2.1 RBE calculation

The Paganetti 2014 dataset contains target RBE values, whereas the PIDE3.4 sample dataset lacks such values. Therefore, this study requires the RBE calculation model to derive deep learning target values from feature data. The LQ model, with its concise mathematical formulation, has been extensively validated in mammalian cell experiments and demonstrates high consistency with deep learning experimental data [13][14][21]. This paper employs the LQ model, the biological endpoint of cell survival rate, to compute RBE values for the aforementioned experimental data, as shown in Formula [33].

$$RBE = \frac{D_X}{D_p} = \frac{\alpha_X - \sqrt{\alpha_X^2 - 4\beta_X \ln(S)}}{-2\beta_X} \times \frac{-2\beta_p}{\alpha_p - \sqrt{\alpha_p^2 - 4\beta_p \ln(S)}}$$

In the formula, S denotes the cell survival rate, D represents the reference radiation dose (in Gy), while α and β indicate the initial linear slope and quadratic curvature of the survival curve, measured in Gy⁻¹ and Gy⁻² respectively, which are positively correlated with cellular radiation sensitivity. By fitting the α and β values from cell survival data, the irradiation dose can be derived. Combining this with the RBE definition, the RBE value for a specific particle radiation can be calculated. Here, X denotes reference radiation such as X-rays or photons, and p represents the particle radiation in question.

2.2.2 Systematization of input data

This deep learning method, implemented in Python, employs matrix-based data processing for workflow execution. The method begins by dividing the 999 samples from the sample database into a training set (800 samples) and an independent test set (199 samples). From the training set, continuous variables (DNA content, Charge, LET, Dose, α_X , β_X , α_p , β_p), categorical variables (Cells, CellClass, Ion, Irradiation Conditions), and the target variable (RBE) are extracted. Using ColumnTransformer, numerical features are normalized via MinMax scaling [16], while categorical features undergo One-Hot encoding [17]. Continuous and categorical variables are subsequently integrated at the input stage [18], aligning with the output target variable RBE in this study. The statistical analysis of the continuous and categorical variables is presented in Table 1. It should be noted that the PIDE3.4 database does not contain Dose values; therefore, this study assigned Dose values within a reasonable range and calculated the corresponding RBE, whereas the Paganetti 2014 dataset contains both Dose and RBE values, which accounts for the large discrepancy between the maximum and minimum Dose values in Table 1. It should be noted that the PIDE3.4 database does not contain Dose values. Therefore, this study assigned Dose values within a reasonable range and calculated the corresponding RBE. In contrast, the Paganetti 2014 dataset contains both Dose and RBE values, which accounts for the large discrepancy between the maximum and minimum Dose values in Table 1. Furthermore, under the LQ curve framework, as long as the overall exponent ($-\alpha D - \beta D^2$) remains negative, the biological meaning is not violated; hence, negative β values can occur.

Table 1
Statistical analysis of input variables for deep learning.

Type of variable	Variable name	Statistical analysis
continuous variable	DNAcontent	5.793 ± 0.195 ^[a] ; MIN: 5.6; MAX: 6.0
	Charge	8.564 ± 11.782 ^[a] ; MIN: 1.0; MAX: 92.0
	LET(keV/μm)	165.0 ± 1080.6 ^[a] ; MIN: 0.5; MAX: 15800
	Dose(Gy)	2.013 ± 0.299 ^[a] ; MIN: 2.0; MAX: 11.13
	alpha_x	0.242 ± 0.284 ^[a] ; MIN: 0; MAX: 2.728
	beta_x	0.046 ± 0.035 ^[a] ; MIN: -0.27; MAX: 0.347
	alpha_p	0.887 ± 0.704 ^[a] ; MIN: 0; MAX: 3.740
	beta_p	0.080 ± 0.150 ^[a] ; MIN: -0.87; MAX: 1.869
categorical variable Total data volume: 999	Cells	V79 (15.46%) ^[b] , T1 (4.66%), NB1RGB (4.38%), CHO (2.89%), HSG (2.70%), AG01522(2.24%), others (67.67%)
	CellClass	n (normal/non-tumor) ^[b] : 57.17%; t (tumor): 42.83%
	Ion	¹² C (35.94%) ^[b] , ¹ H (16.29%), ⁴ He (11.64%), ²⁰ Ne (7.91%), ⁵⁶ Fe (4.28%), ⁴⁰ Ar (4.19%), ²⁸ Si (2.89%), ⁶⁰ Co (2.79%), others (14.07%)
	IrradiationConditions	m (single energy) ^[b] : 74.77%; s (extended Bragg peak): 25.23%

*[a] represents the mean ± SD of the variable; [b] represents the percentage of the variable.

2.3 Method architecture and training

As described above, in the training of deep learning, this study first employed the train_test_split method [18] to partition the preprocessed feature data (X_processed) and target variables (y) into training (800 groups) and test sets (199 groups) at a specified ratio, with a fixed random_state to ensure experimental reproducibility.

This method employs Keras to construct a three-layer Sequential neural network [19]: the first hidden layer contains 64 neurons with ReLU activation [20]; the second hidden layer is reduced to 32 neurons, also employing ReLU activation; and the output layer consists of a single linear neuron for regression prediction. The choice of a three-layer neural network architecture was primarily based on comprehensive considerations of problem complexity and computational efficiency. For most regression tasks, the relationship between input features and target variables can be effectively captured by a shallow network, whereas excessively deep architectures not only increase the risk of overfitting but also substantially raise training costs. A three-layer structure is sufficient to model nonlinear relationships via the ReLU activation function, meeting the requirements of most regression tasks. Furthermore, the progressive reduction of the number of hidden layer neurons from 64 to 32 forms a "funnel-shaped" structure, which facilitates the gradual extraction of higher-order features while avoiding parameter redundancy, thereby ensuring the method's generalization capability.

Furthermore, the method was compiled using the Adam optimizer with a learning rate of 0.001, employing mean squared error (MSE) as the loss function while monitoring mean absolute error (MAE) as an additional metric [22]. MAE served solely as a monitoring indicator and was not involved in the loss function calculation. During training, mini-batch gradient descent was applied for 100 epochs with a batch size of 2, and 20% of the training data was reserved as a validation set. The batch size of 2 was selected based on the following consideration: with only 800 samples in the training set, a smaller batch size helps increase the frequency of parameter updates, thereby enhancing convergence stability under small-sample conditions. Empirical results confirmed that this batch setting enables stable convergence of the method. To prevent overfitting, early stopping was employed to monitor the validation loss, with a patience of 10—that is, training was terminated early if the validation loss did not decrease for 10 consecutive epochs.

2.4 Evaluation indicators

After establishing the AI deep learning method, this study conducted method validation by constructing an evaluation system comprising six metrics: MAE, MSE, Root Mean Squared Error (RMSE), R², Mean Absolute Percentage Error (MAPE), and Explained Variance. The research also performed visual analysis through true-predicted value scatter plots and residual plots for the test set.

The complete workflow of this deep learning method is illustrated in Fig. 1.

3.Results

3.1 Visualization of method prediction results

The prediction-real value scatter plot and residual plot of the test set are visualized in Fig. 2 (a) and Fig. 2 (b) respectively.

The scatter plot of predicted vs. actual values demonstrates a clear linear clustering pattern, with most data points densely clustered near the theoretical diagonal ($y = x$), indicating high consistency between the method's predictions and real-world values. In the overall sample set of the method, RBE values predominantly cluster between 0–7, with only 24 sample sets exceeding 7 (RBE values between 7–17). Minor deviations were observed in high-RBE regions, likely due to insufficient training data for extreme values leads to reduced generalization capabilities. However, the overall trend confirms the method's ability to accurately capture inherent data patterns, further corroborating the high R2 and low RMSE metrics further validate its strong generalization and predictive accuracy, making it reliable for practical applications. These results indicate robust predictive performance across the evaluated RBE range.

The residual plot demonstrates that the prediction errors (residuals) are uniformly distributed within the target variable range, showing no significant systematic bias or heteroscedasticity. This indicates the method maintains consistent data fitting. Although some residual areas exhibit a slight upward trend, the overall error fluctuates near the zero line without apparent nonlinear patterns, further validating the strong explanatory power reflected in the method's evaluation metrics R² and variance explained. Additionally, outliers are concentrated in specific regions (e.g., increased residuals in high-value areas), suggesting the method's insufficient feature learning in these zones. The upward trend in residuals at high-value areas may reflect limited sample sizes in sparse data regions, leading to inadequate fitting and naturally increased prediction uncertainty. The random distribution characteristics of residuals demonstrate that the method effectively captures primary data patterns without overlooking significant features, demonstrating good generalization and robustness, making it suitable for practical prediction tasks.

3.2 Results from 5-Fold Cross-Validation

Due to the relatively small sample size used in deep learning, 5-fold cross-validation enables more effective utilization of limited data, thereby improving the reliability of method evaluation. The 5-fold cross-validation was performed using stratified sampling based on cell type, ensuring that the distribution of cell types in each fold remained consistent with that of the original training set, thus avoiding bias in validation results caused by uneven cell type distribution. Specifically, the 800 samples in the training set were first grouped by cell type. Within each cell type group, samples were randomly divided into five subsets, and the corresponding subsets from each cell type were then combined to form five mutually exclusive folds, ensuring that the cell type composition in each fold was consistent with that of the overall training set. This approach not only reduces evaluation bias introduced by random data partitioning but also provides a more comprehensive assessment of the method's generalization performance through five independent training-validation cycles. After cross-validation, the method was retrained on the full dataset (all 5 folds combined) using the same hyperparameter settings for actual prediction. That is, the 5-fold cross-validation does not retain the intermediate five methods but only records their performance metrics for algorithm evaluation and hyperparameter tuning. The final deployed method is trained on the full dataset. Particularly in cases with limited sample sizes, this data-reuse strategy more accurately identifies whether the method suffers from overfitting or underfitting compared to single-split validation methods. Moreover, the performance variation across different folds helps analyze the uniformity of data distribution. Therefore, for small-sample deep learning tasks, 5-fold cross-validation serves as an effective strategy to mitigate data scarcity.

The 5-fold cross-validation results on the training set are summarized in Table 2. It should be emphasized that the evaluation metrics in this table (such as MAE, R², etc.) represent the average or range performance across the validation sets during 5-fold cross-validation, used to assess the method's generalization capability rather than the performance of the final deployed method. The method exhibited consistent predictive performance across all data subsets. The best-performing fold achieved a MAE of 0.2308, indicating high prediction accuracy under specific data distributions. MAE values across all folds ranged from 0.2308 to 0.3692, showing limited fluctuation and reflecting the method's fundamental stability and adaptability to data variability. Both MSE and RMSE remained at low levels, suggesting that overall prediction bias was well controlled. The highest R² reached 0.9312, while the explained variance peaked at 0.9320, indicating that the method captures approximately 93% of the variance in the data. Consistently low MAPE values further confirm that prediction errors remain within acceptable margins relative to the actual values. In summary, the comprehensive evaluation metrics demonstrate that the method possesses high predictive accuracy and robustness, effectively capturing complex relationships in the target variable.

Table 2
Prediction results from 5-Fold Cross-Validation.

Metric	MAE	MSE	RMSE	R ²	MAPE	Explained Variance
Fold 1	0.2616	0.3559	0.5966	0.8984	0.1114	0.9010
Fold 2	0.2308	0.1834	0.4282	0.9312	0.0922	0.9320
Fold 3	0.2881	0.3529	0.5941	0.8950	0.1109	0.8952
Fold 4	0.3692	0.6961	0.8344	0.8646	0.1296	0.8674
Fold 5	0.3176	0.4543	0.6740	0.8905	0.1206	0.8906
average value	0.2935	0.4085	0.6255	0.8959	0.1129	0.8972
SD(%)	4.8300	17.5500	14.3400	2.1500	1.3700	2.1900

3.3 Simulation prediction results

After the evaluation, to better understand the deep learning results, the study implemented input windows for continuous and categorical variables in Python based on the original code. The predicted results of the deep learning objective value (RBE) are shown in Table 3, with all samples drawn from the database used in this study.

Under identical12C ion incidence conditions, the predicted RBE values for Sample 1 (LET = 12.16 keV/μm), Sample 2 (LET = 482.7 keV/μm), and Sample 3 (LET = 1245 keV/μm) exhibited dose-dependent growth (from 2.19 to 3.35), consistent with the measured RBE variation trend (from 2.24 to 3.41), validating

the correlation between LET and RBE. When using 20Ne heavy ions, Sample 4 showed complete agreement between predicted RBE (1.85) and measured RBE (1.85). Under low LET conditions, Sample 5's predicted RBE for 60Co photon radiation (1.14) closely matched the measured value (1.26), with a relative standard deviation of 9.52%. Analysis of Sample 6 (1H, LET = 27.6 keV/μm), Sample 7 (1H, LET = 37.8 keV/μm), and Sample 8 (12C, LET = 60 keV/μm) demonstrated that the method's relative standard deviation for proton and other light ions' RBE predictions ranged between 1.76% and 2.19%, indicating good accuracy.

The results of sample 9 (T1 cells) and sample 10 (HSG cells) demonstrated higher relative standard deviations (RSD) under limited sample sizes compared to V79 cells (T1: 5.36%; HSG: 4.07%), reflecting the method's performance limitations with insufficient data. Additionally, all samples showed a decreasing trend in RSD as LET increased. When analyzing 10 sample groups, all RSD values remained below 13.33% (median: 2.19%), indicating that this deep learning method demonstrates reliability and scientific validity in RBE prediction, providing an effective methodological reference for radiobiological research.

Furthermore, to validate the reliability of deep learning, this study expanded beyond the original database by incorporating in vitro experimental data from literature [15]. The predicted results in Table 4 demonstrate radiation response parameters of three distinct cell samples (A-172, M/10, and U-251MG) under carbon ion irradiation across six validation sets. Data revealed that all predicted RBE values exceeded experimental truth values, indicating potential systematic overestimation of RBE by the current method, with an average relative standard deviation of 23.89%. This discrepancy may stem from insufficient representation of these cell types in the deep learning dataset or incomplete consideration of cellular repair mechanisms and microenvironmental factors. The current method's prediction bias suggests the need to introduce additional biological variables or expand sample sizes to enhance prediction accuracy, thereby more comprehensively revealing cell-specific radiation response mechanisms.

Table 3
RBE Prediction Results based on Deep Learning

Feature	DNA Content	Charge	LET	alpha_x	beta_x	alpha_p	beta_p	Dose	Cells	Cell Class	Ion	Irradiation Conditions	RBE predicted value	RBE ground truth
Sample 1	5.6	6	12.16	0.067	0.037	0.284	0.035	2.0	V79	n	¹² C	s	2.19	2.24
Sample 2	5.6	6	60	0.191	0.020	0.657	0.066	2.0	V79	n	¹² C	m	3.08	3.03
Sample 3	5.6	6	482.7	0.104	0.025	0.516	0.015	2.0	V79	n	¹² C	m	3.35	3.41
Sample 4	5.6	10	1245	0.183	0.017	0.432	-0.028	2.0	CHO	n	²⁰ Ne	m	1.85	1.85
Sample 5	5.6	27	1.16	0.051	0.025	0.129	0.030	8.6	V79	n	⁶⁰ Co	m	1.14	1.26
Sample 6	5.6	1	2.33	0.132	0.021	0.064	0.042	2.0	V79	n	¹ H	s	1.02	0.90
Sample 7	5.6	1	27.6	0.133	0.050	0.775	0.008	2.0	V79	n	¹ H	m	3.41	3.35
Sample 8	5.6	1	37.8	0.106	0.052	0.680	-0.021	2.0	V79	n	¹ H	m	3.26	3.19
Sample 9	6	2	1.8	0.170	0.026	0.202	0.026	2.0	T1	n	⁴ He	m	1.18	1.12
Sample 10	6	6	30	0.348	0.057	0.434	0.082	2.0	HSG	t	¹² C	m	1.28	1.23

Table 4
Comparison of deep learning-predicted RBE values with experimental RBE values. The input data from columns 2 to 13 are originated from literature[15], the columns 14 are deep learning-predicted RBE values.

Feature	DNA Content	Charge	LET	alpha_x	beta_x	alpha_p	beta_p	Dose	Cells	Cell Class	Ion	Irradiation Conditions	RBE predicted value	RBE experime value [15]
Verify 1	6	6.0	100	0.405	0.059	1.855	0.019	1.40	A-172	t	¹² C	m	4.60	3.95
Verify 2	6	6.0	100	0.300	0.068	2.310	0	1.75	M/10	n	¹² C	m	6.39	5.25
Verify 3	6	6.0	100	0.031	0.055	1.098	0.027	3.15	U- 251MG	t	¹² C	m	6.21	5.02
Verify 4	6	6.0	100	0.405	0.059	1.855	0.019	3.75	A-172	t	¹² C	m	4.01	3.20
Verify 5	6	6.0	100	0.300	0.068	2.310	0	4.00	M/10	n	¹² C	m	4.94	3.98
Verify 6	6	6.0	100	0.031	0.055	1.098	0.027	6.10	U- 251MG	t	¹² C	m	3.96	3.00
average	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Table 5 compares the RBE fitting values reported by the PIDE team with the corresponding experimental RBE values from the literature. Literature analysis reveals that the RBE fitting values from the PIDE team are generally slightly higher than the corresponding experimental measurements. This systematic bias may be attributed to heterogeneity in the literature data, differences in the fitting methods used for the LQ model, or incomplete documentation of experimental conditions (such as cell culture environments and irradiation settings). Since the deep learning method in this study was trained directly using the PIDE team's fitted values as the target, the method inherently learned this systematic bias and incorporated it into deep learning predictions, which to some extent explains why the deep learning predictions in Table 4 are generally higher than the validation experimental values. As shown in Table 5, the relative standard deviations between the fitted and experimental values are largely consistent with those presented in Table 4. This finding suggests that the quality of training data directly affects the reliability of method outputs. In future research, refitting using raw experimental data or integrating data from multiple independent sources may help mitigate the impact of such biases on method generalization performance.

Table 5
Comparison of RBE fitting values from the PIDE team with experimental data.

Verify 1	PIDE team's RBE fitting value [15]	RBE experimental value	Relative Error(%)
	4.50	3.95	13.92
Verify 2	8.02	5.25	52.76
Verify 3	5.99	5.02	19.32
Verify 4	4.00	3.20	25.00
Verify 5	4.98	3.98	25.13
Verify 6	3.18	3.00	6.00
average value	--	--	23.69

3.4 Comparison of cell sample selection

Due to the limited sample size for certain cell types in the database used in this study, the performance of deep learning was inevitably compromised. To address this, five distinct datasets were selected from the original data, each containing different cell types with relatively large sample sizes, and deep learning was performed using the same architecture. The test sets were distinguished by cell type and sample size. For example, Test #1 comprised 11 cell types, each with no fewer than 10 samples (denoted as "N10 + T11"); Test #3 comprised 9 cell types, each with at least 12 samples (denoted as "N12 + T9"). The training results are shown in Table 6. As the sample size per cell type increased (from N10 to N12), method performance (R^2) improved from 0.9236 to 0.9737. However, as the number of cell types decreased (from T11 to T9), the total test sample size declined, and method performance did not continue to improve. This suggests that under data-sparse conditions, simply increasing the sample size per cell type while sacrificing cell type diversity does not necessarily lead to improved generalization capability. This trade-off indicates that constructing a training set requires balancing both the coverage of cell types and the sample size per cell type.

Table 6
Comparison of cell type tests.

name of index	Meaning representation	MAE	MSE	RMSE	R ²	MAPE	Explained Variance
Test #1	N10 + T11	0.2112	0.1370	0.3701	0.9236	12.58%	0.9237
Test #2	N11 + T10	0.2071	0.1888	0.4346	0.9236	9.33%	0.9285
Test #3	N12 + T9	0.1455	0.0480	0.2191	0.9737	10.73%	0.9737
Test #4	N14 + T8	0.2152	0.1150	0.3392	0.9367	12.95%	0.9487
Test #5	N16 + T7	0.2773	0.1896	0.4354	0.9275	20.65%	0.9314

3.5 Results of the ablation experiment

As shown in Fig. 3, the ablation study comparison clearly illustrates the performance differences of various method configurations during the training process, while Table 7 summarizes the final performance comparison of these configurations on the independent test set. The statistical analysis based on 10 repeated training runs was designed to systematically evaluate the influence of loss functions, optimizers, and network architectures on method performance.

Figure 3(a)(b) compares the training dynamics of different method configurations. The training set MAE curve (Fig. 3a) reflects the learning dynamics of the method, whereas the validation set MAE curve (Fig. 3b) monitors overfitting and convergence. Final performance is evaluated based on the independent test set (Table 7). Statistical results indicate that the method using MAE as the loss function achieved the lowest mean absolute error on the test set ($MAE = 0.37 \pm 0.03$); however, the difference compared to the original method ($MAE = 0.39 \pm 0.03$) did not reach statistical significance ($p = 0.0547$). The configuration without the MAE metric ($MAE = 0.40 \pm 0.04$) also showed no significant difference from the original method ($p = 0.5936$), indicating that under the current experimental conditions, the choice of loss function has limited impact on final prediction accuracy. Observations from the training process reveal that the configuration without MAE monitoring exhibited higher initial training loss but converged rapidly, primarily due to the quadratic penalty of the MSE loss function on large errors. However, this rapid decline in training loss did not translate into a corresponding improvement in test set performance, highlighting the disconnect between training loss optimization and final method performance. Notably, the configuration using the SGD optimizer (test set $MAE = 0.47 \pm 0.03$) not only converged significantly slower but also performed markedly worse than the original method with the Adam optimizer ($p < 0.001$), statistically validating the advantage of the adaptive optimizer Adam.

Figure 3(c)(d) compare the training processes of different neural network architectures. To evaluate the impact of network structure on method performance, the following architectural variants were defined: the original network (64 – 32 neurons, three hidden layers), a deeper network (64 – 32 neurons, four hidden layers), a shallower network (64 – 32 neurons, one hidden layer), a wider network (128 – 64 neurons, three hidden layers), and a narrower network (32 – 16 neurons, three hidden layers). All variants were trained under identical configurations to assess the effects of network depth and width on performance. Although the wider network achieved the lowest mean absolute error (MAE) on the test set (0.39 ± 0.02), the difference compared to the original method (0.41 ± 0.03) was not statistically significant ($*p = 0.1245$). Similarly, when compared with other architectures—including the deeper network (0.43 ± 0.02 , $*p = 0.1799$), the shallower network (0.43 ± 0.03 , $*p = 0.1039$), and the narrower network (0.40 ± 0.05 , $*p = 0.5926$)—the original method showed no significant differences in performance ($*p \geq 0.05$). These results indicate that, under the current dataset and experimental settings, variations in network architecture have a limited impact on the final generalization performance of the method. The original method, while maintaining a relatively simple structure, achieved predictive accuracy comparable to that of more complex architectures.

Table 7
Statistical performance comparison of method configurations and network structure ablation experiments.

network topology	Test set MAE (Mean $\pm$ Standard Deviation)	Difference from the original method MAE	t-statistic	p-value	statistical significance
MAE loss function	0.37 ± 0.03	-0.02	2.0545	0.0547	non-significant
original method	0.39 ± 0.03	\	\	\	benchmark method
No MAE indicator	0.40 ± 0.04	+ 0.01	-0.5432	0.5936	non-significant
SGD optimizer	0.47 ± 0.03	+ 0.08	-5.9245	< 0.001	significant
Wider(128 – 64)	0.39 ± 0.02	-0.02	1.6114	0.1245	non-significant
Narrower(32 – 16)	0.40 ± 0.05	-0.01	0.5448	0.5926	non-significant
Original(64 – 32,3layers)	0.41 ± 0.03	\	\	\	benchmark method
Deeper(4 layers)	0.43 ± 0.02	+ 0.02	-1.3953	0.1799	non-significant
Shallower(1 layer)	0.43 ± 0.03	+ 0.02	-1.7129	0.1039	non-significant

*Based on an independent samples t-test with 10 repeated training runs, significance level $\alpha = 0.05$. In MAE differences, '+' indicates worse than the original method, and '-' indicates better than the original method.

3.6 Medical application of the RBE prediction method

In this study, the aforementioned relative biological effectiveness prediction method was applied to a clinical carbon-ion radiotherapy case to achieve regionally heterogeneous RBE prediction. It should be noted that this section aims to demonstrate the technical feasibility of the method in a clinical setting and serves as a methodological validation, rather than representing real clinical practice. In actual clinical applications, the assignment of cell types should be individualized based on patient biopsy results.

A treatment plan of nasopharyngeal carcinoma, with prescribed doses of 54 GyRBE in 18 fractions for the clinical target volume (CTV) and planning target volume (PTV), was designed by a physicist in hospital. The dose constraints for organs at risk were brainstem Dmax < 54 GyRBE and left temporal lobe V40GyRBE < 7 cc. Specifically, RT-structure, RT-dose, and analytically calculated LETd were used as input parameters for the deep learning RBE method to predict RBE values on a voxel-by-voxel basis within the radiation field. In this process, each pixel represents an independent voxel unit, and its corresponding physical and biological features were extracted and fed into the method. The input data were categorized into three types: (1) spatially varying dose and LETd images, where each pixel possesses unique physical properties; (2) globally uniform parameters (Charge, α_x , β_x , α_p , β_p , Ion, IrradiationConditions), which remain constant across the entire prediction region; and (3) regionally assigned biological features (DNA content, Cells, CellClass), delineated by RT structure. Specifically, the CTV region (delineated in blue) was assigned U87 glioma cells, the left temporal lobe region (delineated in green) was assigned AG01522 cells, the brainstem region (delineated in red) was assigned NB1RGB cells, and all other regions were uniformly assigned MKT-Br cells, with corresponding cellular parameters. The selection of cell types here was based on cell lines available in the study database and was intended solely to demonstrate the method's regionally predictive capability, rather than representing a clinically recommended approach for nasopharyngeal carcinoma. These three categories of input were matched to the trained deep learning method, which then output the predicted RBE value for each location and subsequently generated the RBE distribution, achieving an end-to-end mapping from medical images to relative biological effects. The results are shown in Fig. 4.

As illustrated in Figs. 4(c)–4(d), the CTV region exhibited consistent RBE predictions, while the other three regions showed distinct results due to differences in cell types. In the regionally heterogeneous prediction, the RBE values for the brainstem and left temporal lobe were lower than the overall prediction, yet higher than those in the non-heterogeneous prediction. Similarly, other regions exhibited higher RBE values in the regionally heterogeneous prediction compared to the non-heterogeneous prediction, which is attributable to the use of U87 tumor cells throughout the target region in the non-heterogeneous setting. The overall trend in RBE prediction also correlated well with the PIDE database. This application case demonstrates that regional heterogeneous prediction of RBE is highly necessary, rather than merely proving the clinical effect of this case. The LEM model currently used in clinical practice lacks the ability to differentiate the cell types represented by the CTV and organs-at-risk regions when calculating RBE values. However, Figs. 4(c) and 4(d) demonstrate the necessity of differentiating cell types, providing a new perspective for further explaining clinical phenomena. Through this application case, it is demonstrated that regionally heterogeneous prediction of RBE is highly necessary, rather than merely proving the clinical effect of this case. The LEM model currently used in clinical practice lacks the ability to differentiate the cell types represented by the CTV and organs-at-risk regions when calculating RBE values, yet Figs. 4(c) and 4(d) demonstrate the necessity of differentiating cell types, providing a new perspective for further explaining clinical phenomena. Under this framework, the RBE prediction for each voxel is determined collectively by its local physical properties, global radiation conditions, and the biological characteristics of the region to which it belongs. During method operation, the system automatically identifies the region of the voxel, retrieves the corresponding biological parameters, and performs integrated calculations incorporating dose and LETd values at that location. This approach ensures that the resulting RBE distribution reflects not only the spatial effects of physical properties on RBE but also the biological differences in radiation response across distinct cellular regions. Notably, at the boundaries between regionally defined cell types, the system clearly demonstrates the gradient in RBE values arising from transitions in cell type. This regionally predictive capability provides a valuable tool for accurately assessing biological dose distributions within complex tumors, enabling clinicians to visually observe differences in radiation response across cellular regions. In summary, this section demonstrates the technical feasibility of the method for regionally heterogeneous RBE prediction using a nasopharyngeal carcinoma case, thereby establishing a methodological foundation for future clinical exploration; however, real-world applications would require calibration based on individualized patient biological information. Thus, the potential clinical value of this targeted, regionally heterogeneous RBE prediction warrants further investigation to elucidate its guiding significance for particle radiotherapy practice.

3.7 Validation of Cross-Cell RBE Prediction Capability Based on Continuous Cell Features

To overcome the dependence of traditional RBE prediction models on specific cell lines, this study proposes a novel approach that converts cell types into continuous feature vectors based on biological knowledge, constructing a deep learning method capable of predicting RBE for unseen cell types. To this end, a ten-dimensional feature space integrating multi-source biological information was first constructed. Four fundamental biological parameters—cell cycle time, α/β ratio, plating efficiency, and radiation sensitivity index—were obtained from classic radiobiology literature (e.g., Barendsen 1982, Paganetti 2014) and the ATCC cell bank. A systematic inference rule was then established: for common cell lines (e.g., V79, HSG, T1), literature-reported values were directly adopted; for unreported cells, reasonable assignments were made based on species and tissue origin information inferred from standardized nomenclature (e.g., "U87" indicates glioma, "CHO" indicates Chinese hamster ovary cells), referencing typical parameters of cells from the same category. Subsequently, six binary classification features were derived, including human-derived, hamster-derived, mouse-derived, brain tissue-derived, lung tissue-derived, and tumor cell markers. These ten features collectively constitute the input vector, with both continuous and categorical features standardized to ensure training stability.

To rigorously evaluate the method's cross-cell generalization capability, a cell type-based held-out validation strategy was adopted. All 123 cell types were stratified into a training cell type set and an unseen cell type test set at an 80%/20% ratio, ensuring no overlap in cell types between training and test sets, thereby providing a genuine assessment of the method's predictive capability for "unseen" cell types. Ultimately, the training set comprised 98 cell types (800 samples), and the unseen cell test set comprised 25 cell types (199 samples). The method achieved excellent performance on training cells ($R^2 = 0.9652$,

MAE = 0.1571) and maintained strong predictive accuracy on unseen cells ($R^2 = 0.7600$, MAE = 0.4055), demonstrating robust cross-cell generalization capability.

The predictive performance for different cell types in the test set is shown in Table 8. For most unseen cell types, the method's prediction errors were within a clinically acceptable range (MAE = 0.15–0.30). Cell types with larger sample sizes, such as T1 and NB1RGB, exhibited good predictive stability. Cells such as ONS76 and CGL1 showed excellent performance with prediction errors below 0.10. Notably, predictions for cells with extremely high RBE values, such as HE20 (MAE = 3.722), were suboptimal, suggesting the need for specialized prediction modules for unique radiation response mechanisms in future work.

To further validate the method's practical application value, predictions were made for eight new cell types entirely absent from the database, with results shown in Table 9. All predicted RBE values fell within a reasonable range of 1.18–4.27, with an average predicted value of 2.49 ± 1.56 , closely aligning with the clinically observed RBE range for common tumors (1.5–3.5), confirming the method's predictive capability for new cell types. This research can provide a feasible tool for individualized RBE assessment in clinically rare and emerging tumor types, further enhancing the method's clinical application value and advancing particle radiotherapy toward truly personalized and precision medicine.

Table 8
Prediction Performance of Representative Unseen Cell Types

Cell Type	Sample Count	True RBE Mean	Predicted RBE Mean	MAE	Performance Grade
ONS76	1	1.00	1.07	0.075	Excellent
CGL1	2	1.08	1.16	0.078	Excellent
HFIB2	1	1.20	1.28	0.080	Excellent
T1	63	3.07	3.02	0.207	Good
NB1RGB	38	2.90	2.86	0.258	Good
Neo	6	2.18	2.36	0.256	Good
HE20	6	4.63	8.35	3.722	Average
TG98G	7	2.40	3.28	0.879	Average
U-251MG	2	4.76	3.53	1.223	Average
*Note: Excellent prediction (MAE < 0.15); Good prediction ($0.15 \leq \text{MAE} < 0.30$); Average prediction (MAE ≥ 0.30)					

Table 9
RBE Prediction Results for Non-Database Cell Types

Cell Type	Tissue Origin	α/β Value (Gy)	Radiosensitivity	Predicted RBE Range	Mean $\pm$ SD
PC-3	Prostate Cancer	3.6	High	1.20–4.22	2.49 ± 1.56
SKOV-3	Ovarian Cancer	3.3	Medium	1.19–4.22	2.49 ± 1.56
BT-474	Breast Cancer	3.4	High	1.19–4.22	2.49 ± 1.56
PANC-1	Pancreatic Cancer	3.2	Medium	1.19–4.22	2.49 ± 1.56
HT-29	Colon Cancer	3.1	Medium	1.18–4.22	2.49 ± 1.57
SH-SY5Y	Neuroblastoma	3.5	High	1.19–4.22	2.49 ± 1.56
MDA-MB-231	Breast Cancer	3.7	High	1.20–4.22	2.49 ± 1.56
C6	Rat Glioma	3.0	Medium	1.19–4.27	2.51 ± 1.59

4.Discussion

This study developed a Python-based deep learning method to predict the RBE in radiation biology. Although mechanistic models such as LEM and MKM are well-suited for specific experimental conditions, deep learning methods can automatically learn complex relationships among multiple factors from large-scale experimental data, enabling rapid RBE fitting across complex scenarios involving diverse cell types and radiation conditions. The two approaches have distinct strengths and can complement each other: mechanistic models provide theoretical references for deep learning, while deep learning helps extend the applicability of mechanistic models to different conditions. The deep learning method developed in this paper not only provides a new pre-computational tool for radiobiological experimental design, but also offers a novel solution pathway for individualized RBE assessment and treatment planning optimization in particle therapy.

The data used in this study were sourced from the PIDE3.4 database of the German Heavy Ion Research Center and Paganetti's cell experiments, with RBE values calculated using the LQ model serving as the target variable. We acknowledge that the current model is developed on a limited dataset (999 samples), which reflects the inherent scarcity of high-quality radiobiological data due to experimental constraints. Importantly, unlike traditional parametric models, our deep learning framework is inherently extensible: as new experimental data become publicly available, the model can be readily retrained or fine-tuned to improve its accuracy and generalizability, thereby overcoming the data limitation bottleneck progressively. For method construction, a neural network with

two hidden layers was built using the Keras framework, employing data preprocessing techniques such as normalization and One-Hot encoding. Experimental results demonstrated excellent predictive performance, with strong agreement between predicted and actual values. Additionally, the study utilized 5-fold cross-validation, ablation experiments, and sample prediction tests to validate the RBE prediction efficacy of the deep learning method, all of which yielded favorable outcomes. Furthermore, this research validated the method's capability for precise three-dimensional biological effect field reconstruction in complex human environments through clinical nasopharyngeal carcinoma cases. By integrating physical parameters with region-specific biological parameters, the method achieved mapping from medical images to relative biological effects, generating RBE distribution maps that intuitively reflect the radiation sensitivity of different tissues. This approach effectively captures inter-regional biological effect differences often overlooked by traditional models, which is particularly effective in characterizing the radiation response properties of organs at risk and tumors, thereby providing a basis for developing more effective individualized treatment plans in clinical practice. Additionally, this study proposed a cross-cell RBE prediction framework based on continuous cell features. By constructing ten-dimensional cell biology feature vectors, the method achieved RBE prediction for cell types not present in the training data, demonstrating strong generalization capability and clinical translation potential.

5.Conclusions

Overall, the method exhibited strong accuracy and reliability in RBE prediction, providing a practical data-driven tool for radiation biology research. Prediction precision could be further enhanced in the future by expanding sample sizes and optimizing features. Future research may explore the following directions: at the data level, integrating real biological and clinical data from more clinical centers, combined with findings from emerging disciplines such as radiogenomics, to construct more representative training datasets; at the algorithmic level, introducing advanced architectures such as attention mechanisms and graph neural networks to better capture complex feature relationships, while employing interpretability tools like SHAP and LIME to further improve the clinical applicability of the method.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the National Natural Science Foundation of China (No. 12305404) and the National Key R&D Program of China (2023YFF0613504).

Author Contribution

Y.W. and Y.Q. conceived this study. Y.W. developed the deep learning method and performed the data analysis. Y.Q. provided the biological training data, and Z.P. provided guidance on deep learning. Y.X. provided clinical case data, while X.J., Q.S., K.W., C.L., and X.Z. provided guidance on the research direction and revision of the paper. All authors read and approved the final manuscript.

Acknowledgement

We sincerely thank all team members for their contributions to data collection, method development, results analysis, and manuscript preparation and revision. Special thanks are given to the corresponding author for guidance on research direction and academic integrity, as well as to collaborating institutions for their support in data resources and technical platforms. Finally, we express heartfelt appreciation to the peer reviewers and all colleagues who provided valuable feedback. This research was supported by the National Natural Science Foundation of China (No.12305404) and the National Key R&D Program of China (2023YFF0613504).

Data Availability

The access URL for the PIDE 3.4 data is: www.gsi.de/bio-pide

References

1. Paganetti H, Blakely E, Carabe-Fernandez A, et al. Report of the AAPM TG-256 on the relative biological effectiveness of proton beams in radiation therapy. *Med Phys*. 2019;46(3):e53–78.
2. McNamara AL, Schuemann J, Paganetti H. A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published in vitro cell survival data. *Phys Med Biol*. 2015;60(21):8399–16.
3. Duan XZ. Research on RBE uncertainty and biological dose optimization in carbon ion radiotherapy [dissertation]. University of Chinese Academy of Sciences; 2023.
4. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *MICCAI 2015. LNCS. Volume 9351*. Springer; 2015.
5. Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods*. 2021;18:203–11.
6. Katzman JL, Shaham U, Cloninger A, et al. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol*. 2018;18(1):24.
7. Kandalan RN, Nguyen D, Rezaeian NH, et al. Dose prediction with deep learning for prostate cancer radiation therapy. *Radiother Oncol*. 2020;153:228–35.
8. Kearney V, Chan JW, Wang T, et al. DoseGAN: a generative adversarial network for synthetic dose prediction. *Sci Rep*. 2020;10(1):11073.
9. Dean J, Wong K, Gay H, et al. Incorporating spatial dose metrics in machine learning-based NTCP models. *Clin Transl Radiat Oncol*. 2018;8:27–39.
10. Friedrich T, Pfuhl T, Scholz M. Update of the particle irradiation data ensemble (PIDE) for cell survival. *J Radiat Res*. 2021;62:645–55.
11. Friedrich T, Scholz U, Elsasser T, et al. Systematic analysis of RBE and related quantities using a database of cell survival experiments with ion beam irradiation. *J Radiat Res*. 2013;54:494–514.
12. Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. *Phys Med Biol*. 2014;59(22):R419–72.
13. Van Leeuwen CM, Oei AL, Crezee J, et al. The alfa and beta of tumours: a review of parameters of the linear-quadratic model. *Radiat Oncol*. 2018;13(1):96.
14. Barendsen GW. Dose fractionation, dose rate and iso-effect relationships for normal tissue responses. *Int J Radiat Oncol Biol Phys*. 1982;8(11):1981–7.
15. Parisi A, Beltran CJ, Furutani KM. Clonogenic survival RBE calculations in carbon ion therapy. *Quantum Beam Sci*. 2023;7:3.
16. Han J, Kamber M, Pei J. Data mining: concepts and techniques. 3rd ed. Morgan Kaufmann; 2012.
17. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–30.
18. Zhang C, Yang Z, He X, Deng L. Multimodal intelligence: representation learning, information fusion, and applications. *IEEE J Sel Top Signal Process*. 2020;14(3):478–93.
19. Chollet F. Keras: the Python deep learning library. 2015. <https://keras.io>
20. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. In: *ICML'10*. Omnipress; 2010. pp. 807–14.
21. Wedenberg M, Lind BK, Hårdemark B. A model for the relative biological effectiveness of protons. *Acta Oncol*. 2013;52(3):580–8.
22. Heaton J. Deep learning (Goodfellow, Bengio, Courville). *Genet Program Evolvable Mach*. 2018;19:305–7.
23. Beuve M. Formalization and theoretical analysis of the local effect model. *Radiat Res*. 2009;172(3):394–402.
24. Hawkins RB. A microdosimetric-kinetic model for the effect of non-Poisson distribution of lethal lesions on the variation of RBE with LET. *Radiat Res*. 2003;160(1):61–9.
25. Yang JF, Liu XG, Zhang H, et al. Modified logistic nanodosimetry model for calculating relative biological effectiveness. *Nucl Sci Tech*. 2025;36:61.
26. Qi Y, Warmenhoven JW, Henthorn NT, et al. Mechanistic modelling of slow & fast NHEJ DNA repair pathways following radiation for G0/G1 normal tissue cells. *Cancers*. 2021;13(9):2202.
27. McMahon SJ, Schuemann J, Paganetti H, et al. Mechanistic modelling of DNA repair and cellular survival following radiation-induced DNA damage. *Sci Rep*. 2016;6(1):1–14.
28. McMahon SJ, McNamara AL, Schuemann J, et al. A general mechanistic model enables predictions of the biological effectiveness of different qualities of radiation. *Sci Rep*. 2017;7(1):1–14.
29. Hu A, Zhou W, Luo X et al. Correlation between DNA double-strand break distribution in 3D genome and radiation-induced cell death. 2024. 10.48550/arXiv.2501.07579
30. Wang C, Zhu YN, Lin Y, et al. Multi-modality biological dose optimization for combined carbon-ion and proton radiation therapy. *Med Phys*. 2025;52(11):e70077.
31. Paganetti H, Simone CB 2nd, Bosch WR, et al. NRG Oncology white paper on the relative biological effectiveness in proton therapy. *Int J Radiat Oncol Biol Phys*. 2025;121(1):202–17.
32. Hahn C, Ödén J, Dasu A, et al. Towards harmonizing clinical LET reporting in proton radiotherapy: a European multi-centric study. *Acta Oncol*. 2022;61(2):206–14.
33. Chaudhary P, Marshall TI, Perozziello FM, et al. Relative biological effectiveness variation along monoenergetic and modulated Bragg peaks of a 62-MeV therapeutic proton beam. *Int J Radiat Oncol Biol Phys*. 2014;90(1):27–35.
34. Scholz M, Kellerer AM, Kraft-Weyrather W, et al. Computation of cell survival in heavy ion beams for therapy. *Radiat Environ Biophys*. 1997;36(1):59–66.
35. Hawkins RB. A microdosimetric-kinetic model for the effect of non-Poisson distribution of lethal lesions on the variation of RBE with LET. *Radiat Res*. 2003;160(1):61–9.

Figures

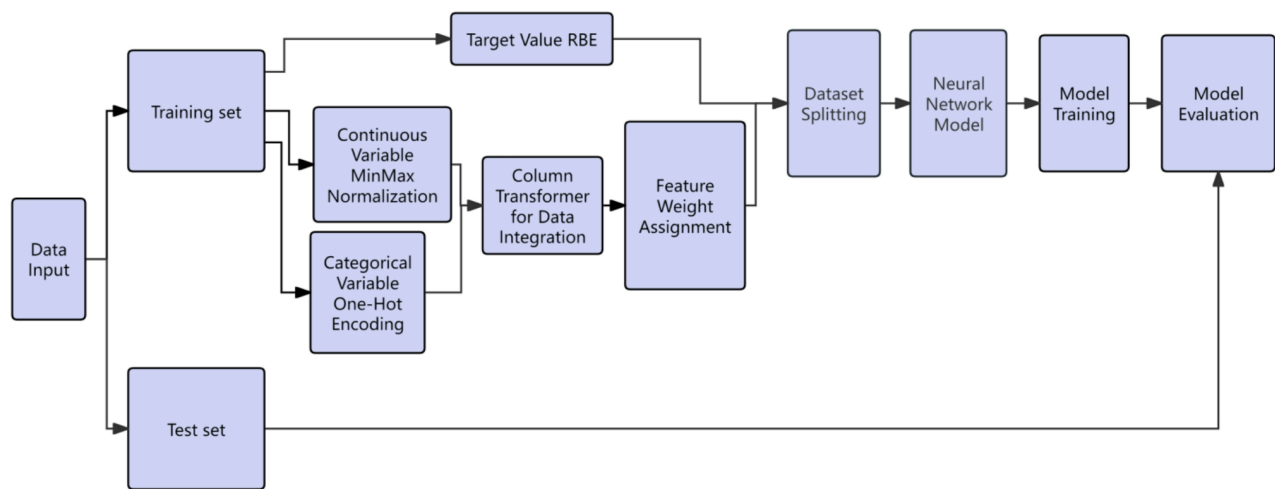


Figure 1
Flowchart of the Deep Learning RBE Prediction Method

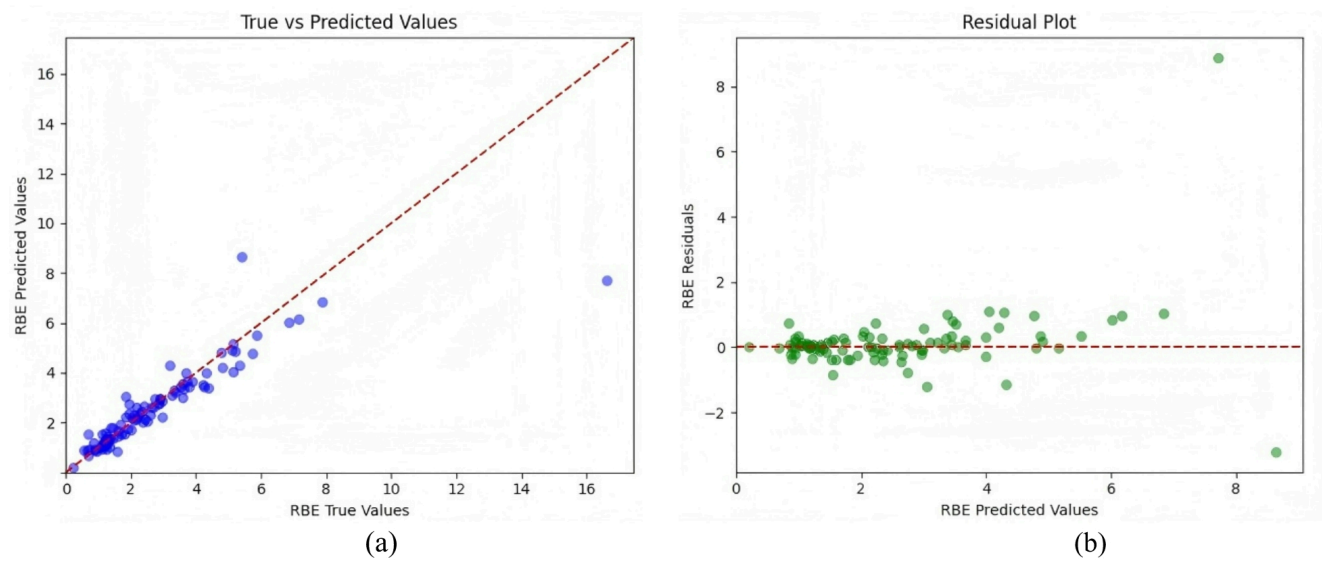


Figure 2
(a) shows the scatter plot of predicted versus actual values, while (b) shows the residual plot.

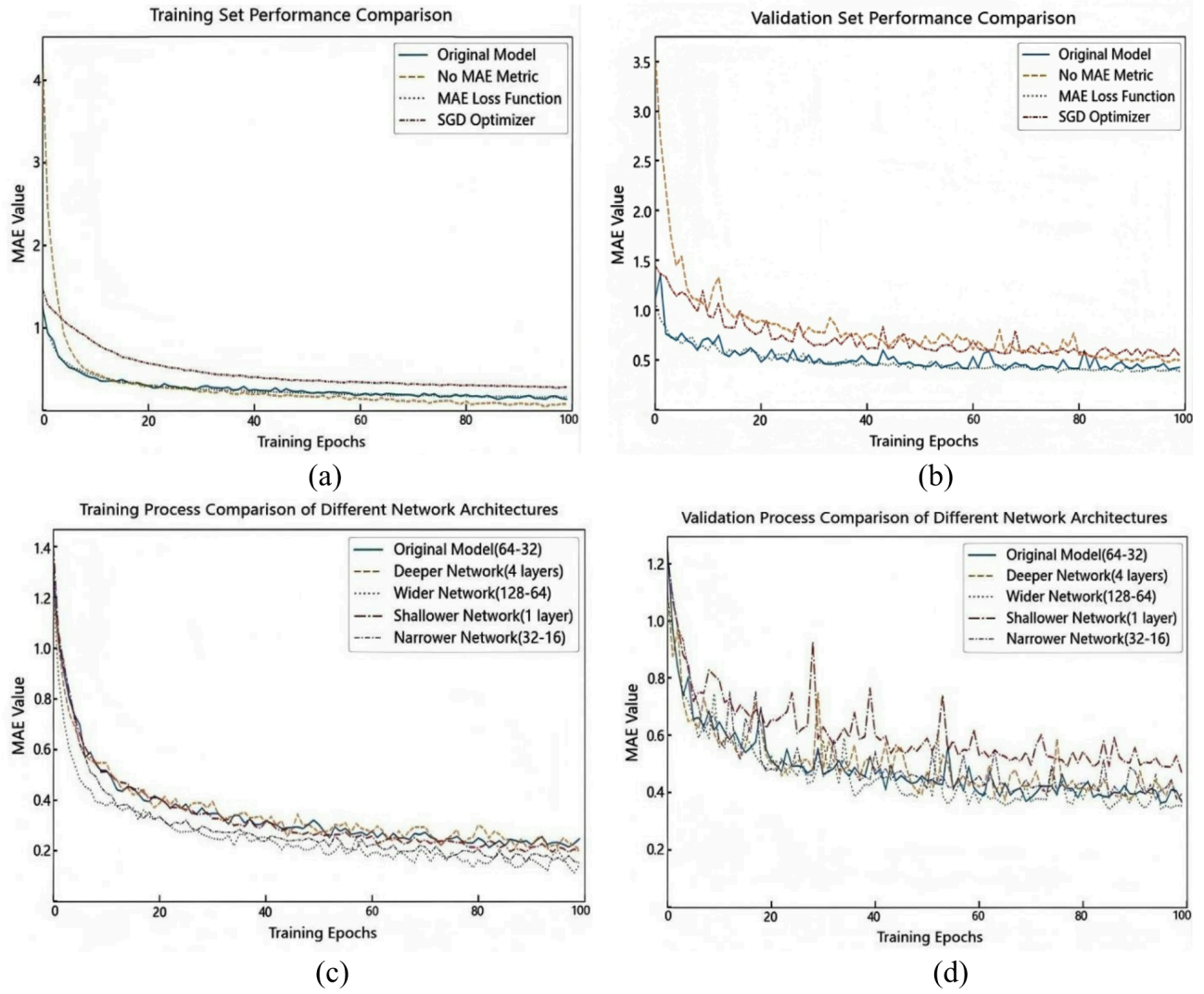


Figure 3

(a)-(b) show the performance comparison between training and validation sets for different method configurations in the ablation experiment, while (c)-(d) present the training and validation process comparisons for various network architectures.

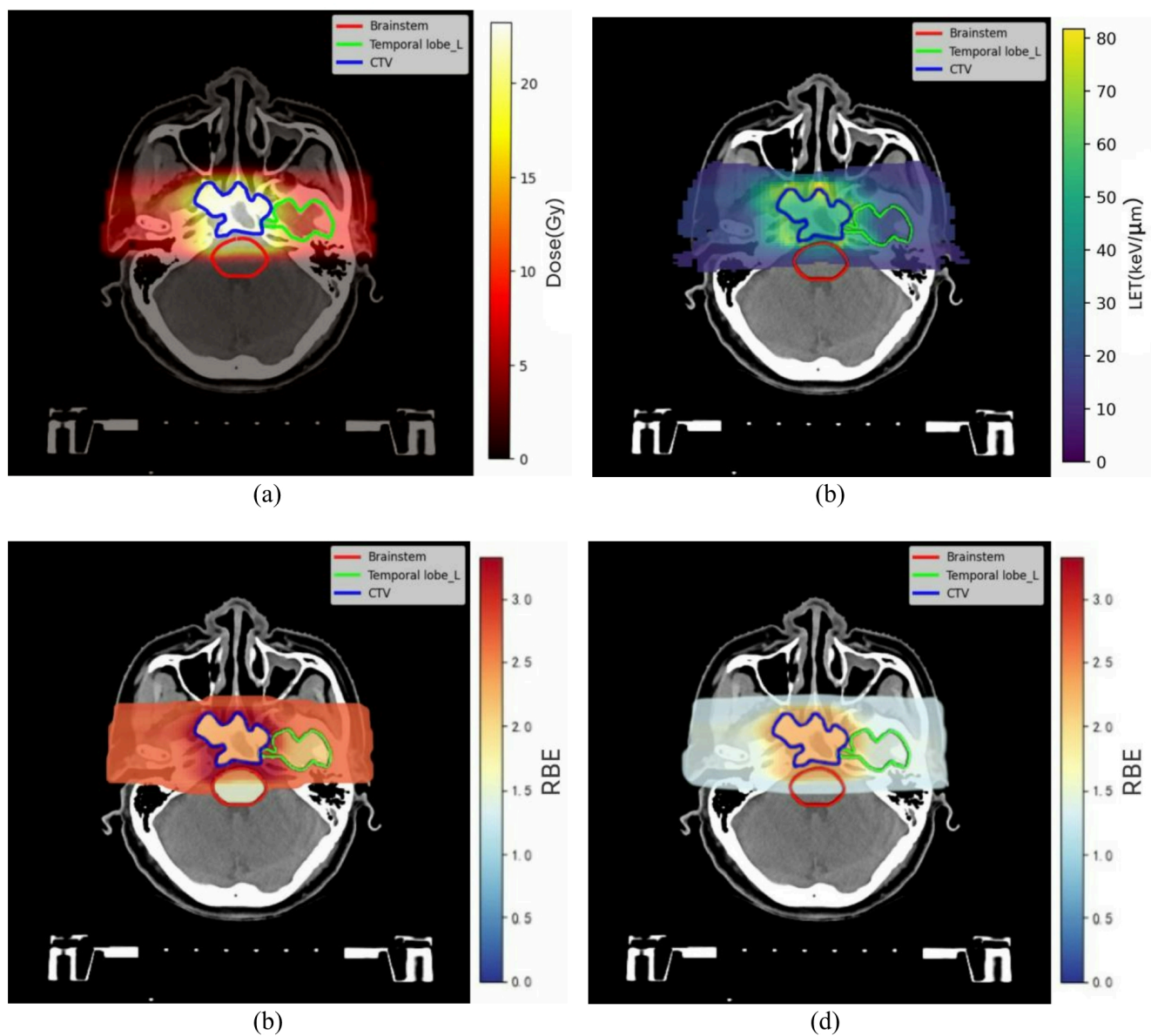


Figure 4

Application results of the deep learning method in nasopharyngeal carcinoma cases (CT slice #45). **(a)** Physical dose distribution **(b)** LETd distribution **(c)** Predicted distribution of RBE regional heterogeneity **(d)** Predicted distribution of RBE non-heterogeneity