Atom Indonesia

Journal homepage: http://atomindonesia.brin.go.id

Dose Distribution Prediction by Machine Learning (Random Forest): Impact of Data Specificity in Cervical Cancer IMRT

Y. A. Lestari[1], A. Munandar[2], W. E. Wibowo[2], B. B. Patrianesha[1,3], P. Prajitno[1], D. S. K. Sihono[1*]

[1] *Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, West Java, 16424, Indonesia*
[2] *Department of Radiation Oncology, Faculty of Medicine, Universitas Indonesia, dr. Cipto Mangunkusumo General Hospital, Jakarta, 10430, Indonesia*
[3] *Research Center for Safety, Metrology, and Nuclear Quality Technology, National Research and Innovation Agency (BRIN), Tangerang Selatan, 15314, Indonesia*

ARTICLE INFO

Article history:
Received 8 February 2025
Received in revised form 30 April 2025
Accepted 7 May 2025

Keywords:

Treatment planning
Model performance
Random search
Dose distribution
Data quality

ABSTRACT

Creating optimal radiotherapy plans is time-consuming and relies heavily on expert judgment to balance target coverage and Organ at Risk (OAR) protection. This study aims to fill the gap in existing radiotherapy approaches by integrating the Random Forest (RF) to predict dose determination in cervical cancer using Intensity Modulated Radiotherapy (IMRT). A retrospective analysis was conducted using 173 randomly selected cases and 102 specific data (stage I to IIIC1r cervical cancer, without prior surgery, and Whole Pelvic Non-Extended Field). Predictions were based on geometric relationships between organs and absorbed doses, with the model trained using decision trees and hyperparameter tuning via Random Search (RS). Model performance was evaluated using Mean Squared Error (MSE), Nash Sutcliffe Efficiency (NSE), and P-values. The ML-RF model showed improved performance with specific data, achieving the largest MSE reduction in bladder Dmax (0.017 Gy to 0.007 Gy). All parameters showed P-values above 0.05, indicating no significant differences in mean values between the predicted and clinical data. However, NSE values varied across parameters, with good performance for Right Femoral Dmax (NSE = 0.530) and Left Femoral Dmax (NSE = 0.554), while lower agreement was observed for PTV CI (NSE = –0.167) and PTV HI (NSE = –0.137), suggesting challenges in capturing distribution patterns. These results demonstrate the model's ability to closely predict clinical dose distributions and underscore the value of incorporating specific clinical criteria. This approach may help reduce clinician workload and support planning standardization, while maintaining the importance of expert clinical judgment. Future studies should explore larger datasets, refined inclusion criteria, diverse treatment approaches, and beam weighting prediction to further improve model accuracy.

INTRODUCTION

Radiotherapy is one of the primary methods in the treatment of cervical cancer. This technique involves the use of a source of ionizing radiation, which is focused on a specific region of the body. In 1982, Brahme et al. introduced an innovative technology called "Intensity Modulated Radiation Therapy (IMRT)" that allows for the modulation of the beam's intensity, resulting in a more conformal dose distribution [1]. IMRT planning primarily relies on an optimization process, which involves a series of computations to determine the dose distribution that the targeted organ will absorb. The main objective of the current study is to deliver a high dose to the target volume while minimizing exposure to surrounding Organs at Risk (OAR) [2]. However, treatment plans are not always optimally designed and often require a trial-and-error approach

*Corresponding author.
E-mail address: dwi.seno@sci.ui.ac.id
DOI: https://doi.org/10.55981/aij.2026.1630

with dose constraints and trade-offs. OARs may receive unavoidable doses using a Treatment Planning System (TPS), which calculates the potential absorbed dose [3]. Adjustments may be influenced by the clinical experience of the medical physicist. Consequently, this process can be time-consuming, leading to variations in accuracy [2]. Hence, studies aimed at improving the efficiency of IMRT optimization procedures are highly desirable.

Machine Learning (ML) is a technique that enables automated predictions of events or values by analyzing patterns from previous instances. This process involves examining the relationships between predictive features in the data to make accurate forecasts about future occurrences. ML has been applied across various fields, including infrastructure, business, medicine, and more. In radiotherapy, studies on distribution prediction ML have been extensively undertaken [4]. Most of these studies have focused on cases such as prostate cancer [5-7], head and neck cancer [5,8] and cervical cancer [5,6]. ML predicts dose distribution by learning the relationship between patient geometry and the radiation dose absorbed by each organ [3]. It is important to note that neither of these cases are identical, as each patient has a unique organ anatomy. Therefore, even when ML might boost productivity, clinical professionals' treatment plan design still takes precedence [9]. There are various types of ML models, one of which is the Random Forest (RF) model, constructed using numerous decision trees, and its final output is a combination of all the decisions made by the trees. A study by Zhang et al. (2021) demonstrated the robust performance of the RF model in predicting the occurrence of radiation pneumonitis in cancer patients [10].

However, the studies conducted using ML with an RF model to estimate the dose distribution were limited in cervical cancer [11-13]. Cervical cancer was selected as the focus of this study due to its high prevalence and significant impact on public health worldwide [14]. As ML systems are exposed to more data, the accuracy may be increased [7]. In radiotherapy, dose prediction with ML models involves extracting quantitative features from medical imaging data and dose distributions. These features function as inputs that guide the model in predicting dose distribution patterns. Features such as Elongation, Flatness and Least Axis Length detailed geometric information about the patient, offering insights from different aspects of the patient's geometry. The inclusion of more features increases the potential for high accuracy and generalization, but it also raises the risk of overfitting and adds complexity to the model. In this study, the use of the RF model is expected to address overfitting and model complexity. The prediction process is distributed across multiple decision trees, which collectively generate the final outcome [15]. This approach utilizes the patient's geometric information derived from radiomics data, previously extracted based on studies in nuclear medicine by Liu et al. (2021), who also contributed to advancements in dose prediction methodologies [16]. In the field of radiotherapy, this approach has been applied to cases of lung cancer [11]. Therefore, it would be valuable to investigate whether studies on different types of cancer yielded accurate results. This study aimed to evaluate the effectiveness of the RF model for predicting radiotherapy dose distribution in cases of cervical cancer. We compared the dose distribution predictions between an ML with RF model and treatment planning conducted by medical physicists in the clinical setting.

METHODOLOGY

Workflow

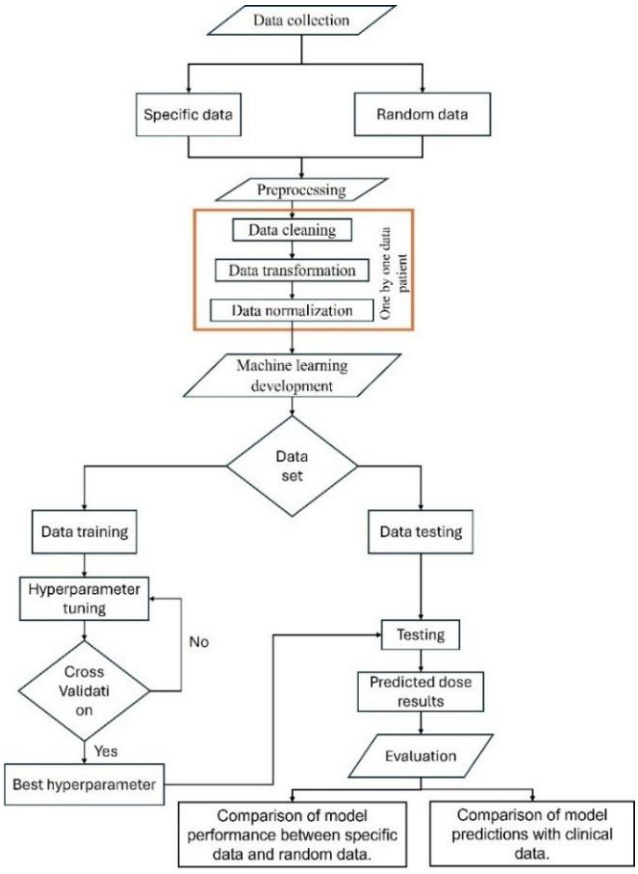

Fig. 1. The workflow of the study on predicting dose distribution in cervical cancer cases using ML with an RF model. There were 4 main procedure stages: data collection, preprocessing, ML development, and evaluation.

The workflow of this study can be seen in Fig. 1. There were 4 main procedure stages, which were data collection, preprocessing, ML development, and evaluation. The selected patient data were then preprocessed. In that stage, the data was processed first before building the ML model. There were two groups of datasets used: the dataset of all patients and the dataset of patients with specific criteria. In the first stage, the RF model was built, and the ML first learned the input dataset. During the development process, hyperparameter tuning was performed. The technique sought optimal hyperparameter values based on the validation scores obtained. Validation was performed using the Cross Validation (CV) technique. After the best hyperparameters were obtained, testing on previously unseen data could be performed. The final stage was the evaluation of the built ML model. We compared the performance of ML using the dataset of all patients and patients with specific criteria. Finally, we compared the dose distribution predictions made by the ML with clinical planning outcomes.

Ethics approval and consent to participate

This study was approved by the Medical Ethics Committee of Dr. Cipto Mangunkusumo National General Hospital, under protocol number 23-10-1659. The data used in this study is secondary data, specifically cervical cancer treatment planning data utilizing the IMRT technique.

Pre-processing

First, we checked the organs within the Digital Imaging and Communications in Medicine (DICOM) files by loading them into the 3D Slicer application. The application displayed a visualization and a list of organs from the DICOM files. In this study, the selected organs were the bladder, bowel, right femoral, left femoral, rectum, and Planning Target Volume (PTV). In this study, radiomics data was used as geometric information for the patients, while dosiomics data provided dose distribution information for each organ [13]. Both datasets were extracted and converted into numerical form, using the 3D Slicer application.

For dosiomics data extraction, we selected features such as total volume (VTotal), volume receiving 95% of the prescribed dose (V95%), volume receiving 50% of the prescribed dose (V50Gy), dose in 2% of the organ volume (D2%), dose in 50% of the organ volume (D50%), dose in 98% of the organ volume (D98%), and dose in 0.01 cc of the organ volume (D0.01cc). These features were selected based on the Conformity Index (CI), Homogeneity Index (HI), and OAR dose limits [13], which are commonly used to confirm that the radiotherapy treatment plan meets clinical standards.

Data normalization was applied to scale numerical features, ensuring consistent magnitudes and minimizing the impact of excessive variance on ML model performance. Normalization was applied only to the dosiomics data, while the radiomics data was not normalized. The dosiomics data, including the dose in a specific organ volume (Dx) and the organ volume receiving a certain dose (Vx), were normalized to the prescribed dose and the maximum volume of each organ for each patient [13].

Random forest model

After the data was pre-processed, the dataset was input into the Python application for model development. Radiomics data were used as the independent variable (X), while dosiomics data served as the dependent variable (Y) [11]. Seventy percent of the entire dataset was used for training, and the rest was used for testing.

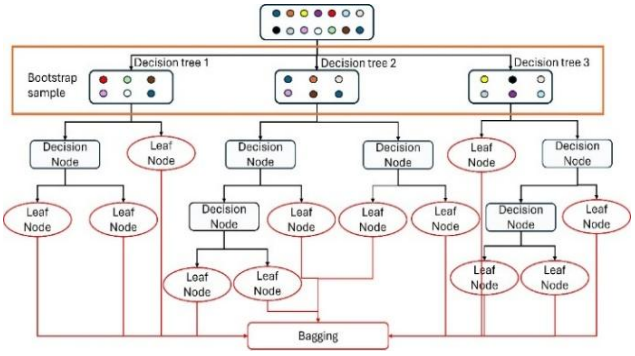

Fig. 2. The Random Forest model consists of three decision trees. The process begins with the bootstrap sampling stage (orange solid lines), where random subsets of the data are selected for each tree. Each tree constructs a structure of decision nodes (black solid lines), where splits are made based on randomly selected features. This process continues recursively until reaching the leaf nodes (red ovals), which represent the final outcomes of each tree. Finally, the outputs of all trees are combined using a bagging technique (red rectangular box), where the predictions are averaged to produce the final result.

Random Forest was a type of supervised ML, built using multiple decision trees [17]. Each tree was built with a randomly selected dataset and test features [18]. The analysis process within the RF model is illustrated in Fig. 2, which consists of four main stages. In the initial phase, bootstrap sampling is performed to generate multiple training subsets by randomly selecting samples with replacement from the original dataset. Each subset is subsequently used to train an individual decision tree. At each decision node within a tree, the algorithm evaluates candidate features and corresponding thresholds to

identify the split that produces the greatest reduction in prediction error, as measured by the Mean Square Error (MSE). This recursive splitting continues until stopping criteria are met, typically when a node reaches a predefined depth, contains too few samples, or achieves sufficient purity to form a leaf node. The final prediction is generated by aggregating the outputs of all trees through bagging, which improves generalization and mitigates overfitting.

Tuning hyperparameter

The RF algorithm has hyperparameters that control the growth of the trees within it. Therefore, hyperparameter tuning was performed to help determine the optimal combination of hyperparameters. There were many hyperparameter tuning methods, such as Random Search (RS) [19]. Hyperparameter tuning in the RS technique was performed by exploring parameters and selecting them randomly. Random Search would have the hyperparameter combination with the highest validation score, from the combinations that are randomly selected [20].

Table 1. The upper and lower bounds used for hyperparameter tuning between the RS technique were set to the values.

Hyperparameter	Description	Range
n_estimator	Number of decision tree	500 - 1000
max_features	Maximum feature of trees	5 - 40
min_sample_split	Minimum sample of split	2 - 10
min_sample_leaf	Minimum sample of leaf	1 - 20
Iteration: 100		

Hyperparameter optimization was performed using the RS technique within the ranges listed in Table 1. Each hyperparameter serves a specific function and can influence both model performance and computational efficiency. The n_estimators hyperparameter determines the number of decision trees in the RF; increasing the number of trees can enhance model convergence but also raises computational cost. The max_features hyperparameter determines the number of features considered at each split; using more features can strengthen individual trees. The min_samples_ leaf parameter defines the minimum number of samples required to form a leaf node, influencing tree complexity and generalization ability. The min_samples_split parameter specifies the minimum number of samples needed to split an internal node, affecting tree depth and model flexibility. The optimal hyperparameters were selected based on the configurations that achieved the highest validation scores [15].

Cross validation

Cross Validation (CV) was a technique to validate the performance of the models [21]. This technique was performed by dividing the dataset into several folds (k). Each fold formed smaller subsets, denoted as $D_1, D_2, \ldots, D_k$, with equal sizes. Additionally, the validation process was repeated based on the number of folds determined. This process was applied to the subsequent iterations, we chose a few 5 folds. CV provided a validation score after each training and testing. The MSE served as the metric for evaluating the extent of prediction errors in the model.

During training, validation was conducted to identify the optimal combination of hyperparameter values, with selection based on those yielding the lowest MSE. This study developed two models based on different datasets. Table 2 summarizes the optimal hyperparameter combinations identified for each model. Since different training datasets were used to build the two models, the selected hyperparameters also differed. Therefore, hyperparameter tuning is essential to ensure optimal model development. After identifying the optimal hyperparameters, the developed model was tested using data that had not been previously seen or used during training.

Table 2. The best hyperparameter was determined the RS technique in ML using random data and specific data.

Hyperparameter	Random data	Specific data
n_estimator	768	747
max_features	27	28
min_sample_split	4	2
min_sample_leaf	1	1
Iteration	100	100

Model evaluation

Our predictions were limited to specific dose levels, commonly used to assess the feasibility of the treatment plan generated. The evaluation of dose distribution for targets used CI and HI parameters, while for OAR, it relied on data from the Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC), are shown in Table 3.

Table 3. Dose limit parameters for OAR for cervical cancer based on QUANTEC data [22].

Organ	Parameter	Dose Limit / Volume
Bladder	Dmax	$< 65\ Gy$
	V50Gy	$< 50\%$
Bowel	V50Gy	$< 5\%$
	D0.01cc	$< 58\ Gy$
Femoral	Dmax	$< 5\%$
Rectum	V50Gy	$< 38\%$

The CI was a parameter used to assess how well the dose distribution matched the target volume. The CI value increased as it approached one, indicating that the dose distribution more effectively covered the target. The CI value was calculated using Eq. (1), with V95% representing the target volume receiving 95% of the prescribed dose and VPTV representing the total PTV volume [23].

$$CI = \frac{V95\%}{VPTV} \quad (1)$$

The HI was a parameter used to assess the level of dose uniformity within the target. The HI value improved as it approached zero, which was calculated using Eq. (2). Dmax represented the maximum dose received by the PTV, and Dp represented the prescribed dose [23].

$$HI = \frac{Dmax}{Dp} \quad (2)$$

Statistical analysis was conducted using a paired t-test, with the P-value calculated for each parameter. The hypothesis decision about whether the dose distribution generated by the ML model with RF aligned with the clinical planning results was based on the significance level. A P-value ≤ 0.05 indicated a significant difference between the dose distribution predictions from ML and the clinical planning results. A P-value > 0.05 indicated no significant difference [24]. The statistical tests were conducted using the GraphPad software.

RESULT AND DISCUSSION

Most IMRT planning procedures involve dose optimization. This process is influenced by clinical experience, which can result in variations in accuracy and may also be time-consuming [2]. Therefore, several studies have introduced automated approaches for dose optimization using ML [5,6]. To the best of our knowledge, no study has applied model to determine dose distribution in IMRT cases for cervical cancer, which could improve planning efficiency and validation of the results. This approach could serve as an alternative method for future treatment planning.

The quality of sample data affects the ability of ML to make predictions [25]. The error levels during training and testing were indicated by the MSE values, which were used to evaluate the model's performance. The MSE measurement the average squared difference between the predicted and actual dose values, where a lower MSE indicates higher prediction accuracy in dose distribution.

The study results, as shown in Table 4, indicate that most MSE values for each parameter are lower for the model trained with specific data. A lower MSE value reflects a lower prediction error by the model. Most parameters, such as bladder Dmax, bowel V50, femoral right Dmax, and femoral left Dmax, showed lower MSE values in the model built with specific data, suggesting improved model performance when using more specific datasets.

The quality of the sample data can significantly influence the model's predictive ability. The inclusion of specific cancer stage selection can enhance the consistency of the sample data. The geometric characteristics of organs and their radiation response are related to the cancer stage, leading to more consistent target area sizes across the samples. As a result, the model benefits from a narrower and more stable data range, which supports better prediction performance.

This issue was observed in the study by Ramlah et al. (2023), which predicted dose distribution using model on 60 lung cancer patient data without specifying the cancer type [11]. As a result, the prediction error for the OAR was higher compared to the PTV. For example, the MSE for PTV HI was only 0.02, while for right lung mean, heart mean, and spinal cord mean, the values were 0.104, 0.088, and 0.069, respectively. This occurred because the geometric information of the PTV also contained parts of the OAR, reducing the model's predictive capability. Therefore, adding cancer type specifications in patient data selection can improve model performance [11].

Table 4. Performance evaluation between ML-RS using random data and the specific data was based on MSE values.

Organ	Parameter	MSE [Dx (Gy) & Vx (%)]	
		Random data	Specific data
PTV	CI	0.000	0.000
	HI	0.005	0.005
Bladder	D_{max}	0.017	0.007
	V50Gy	0.001	0.000
Bowel	V50Gy	0.003	0.001
	D0.01cc	0.003	0.005
Right Femoral	Dmax	0.019	0.008
Left Femoral	Dmax	0.014	0.008

In this study, we investigated the feasibility of the ML-RF method for predicting dose distribution in cervical cancer cases by comparing it with clinically approved planning data. The mean ± standard deviation (mean ± SD) values, Nash Sutcliffe Efficiency (NES), and P-value for each parameter are shown in Table 5 to compare the prediction results of the ML-RF model on specific data with clinical planning data.

Table 5. The mean ± SD, NSE, and P-values for each parameter were compared between the model predictions using specific data and clinical planning outcomes.

Organ	Parameter	$Mean \pm SD$ [Dx (Gy) & Vx (%)]		NSE	$P-value$
		Clinic	Prediction		
PTV	CI	0.953 ± 0.004	0.953 ± 0.001	-0.167	0.727
	HI	0.144 ± 0.069	0.121 ± 0.009	-0.137	0.078
Bladder	D_{max}	49.5 ± 12.2	50.4 ± 12.5	0.362	0.121
	V50	17.73 ± 15.1	24.3 ± 14.6	0.157	0.738
Bowel	V50	5.30 ± 4.41	5.67 ± 2.26	0.052	0.962
	D0.01cc	51.11 ± 12.8	51.6 ± 12.6	0.233	0.346
Right Femoral	Dmax	43.7 ± 11.3	43.4 ± 11.1	0.530	0.997
Left Femoral	Dmax	43.53 ± 10.7	43.9 ± 11.1	0.554	0.786
Rectum	V50	12.3 ± 11.4	14.6 ± 7.71	0.034	0.644

Overall, all parameters exhibited P-values > 0.05, indicating no statistically significant differences between the predicted results and the clinical outcomes. These results indicate that, on average, the model achieved predictive performance comparable to the actual clinical outcomes.

However, a detailed analysis of the NSE values highlighted notable inconsistencies in the model's predictive performance across different parameters. NSE assesses the model's ability to replicate the pattern of clinical data, where values close to 1 indicate strong agreement, and negative values indicate poor predictive ability. In our findings, while the majority of parameters achieved positive NSE values, a few negative NSEs were observed, highlighting challenges in capturing complex dose patterns. These results underscore the need for model refinement and potential expansion of the training dataset to improve generalization.

The model exhibited relatively good predictive performance for certain parameters, such as right femoral Dmax (NSE = 0.530) and Left Femoral Dmax (NSE = 0.554). However, for parameters like PTV CI (NSE = -0.167) and PTV HI (NSE = -0.137), negative NSE values were observed, suggesting limited accuracy in replicating the overall data distribution patterns, despite no statistically significant differences in mean values compared to clinical results. This indicates that P-value and NSE assess different aspects of model performance, where the P-value focuses on the similarity of mean values, while NSE evaluates the overall pattern agreement of the data.

When considering the mean ± SD values, the absolute differences between the clinical data and the predicted results across all parameters were generally small. For example, in the case of bowel V50Gy, the mean ± SD in the clinical data was 5.30% ± 4.41%, while in the predicted data it was 24.3% ± 14.6%. Similarly, for the right femoral Dmax, the mean ± SD in the clinical data was 43.7 Gy ± 11.3 Gy compared to 43.3 Gy ± 11.1 Gy in the predicted data. These findings show that both parameters exhibited minor differences between clinical and predicted mean values.

Overall, the findings indicate that the model has a satisfactory capability to predict dose distributions from patient treatment planning data. However, there remains room for improvement, particularly in modeling the variation patterns of dose distribution within target volumes, such as the PTV.

Reijtenbagh et al. (2022) developed a method to predict the $D2cm^3$ value to improve planning beyond the dosimetry constraints of multi-center analysis [26]. The prediction was based on the individual patient geometry using model. The results showed that the MSE values between clinical data and the model's predicted results for $D2cm^3$ for OARs (bladder, rectum, sigmoid, and small bowel) ranged from 0.13 Gy to 0.40 Gy. These values indicate that the model's prediction errors were very small [26].

Based on the results obtained and comparisons with previous studies, the simpler operation of the RS technique makes it more efficient for hyperparameter tuning. The predictive ability of ML can be enhanced by adding specificity to the training data while still considering the amount of data used. A decline in ML performance may be caused by overly varied radiomics data, which can lead ML to predict higher values. Further research could be conducted by taking the aforementioned factors into account.

In this study, the sample size and data consistency were key factors in developing a ML based dose distribution prediction method. We used 173 random data and 102 specific data, which may

not be considered large enough. However, the ML-RF model was still able to predict dose distributions with results that were comparable to clinical planning. The predictions were only made for specific dose parameters, such as Dmax, D0.01cc, D2%, D50%, D98%, VTotal, V95%, and V50, rather than for the entire dose distribution across the organ volume. These dose parameters were selected based on the requirements for CI, HI, and OAR dose limits, which are commonly used to assess whether the treatment plan meets clinical standards. Although the dose distribution was not predicted for the entire organ volume, the analyzed doses remain relevant as a reference in the dose distribution optimization process.

CONCLUSION

The model successfully predicted dose distributions in cervical cancer cases, with results closely matching clinical plans. In addition, selecting more specific patient data may enhance the model's predictive performance. However, the size of the dataset remains a critical factor in ensuring the model's prediction accuracy.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to Department of Radiation Oncology, Faculty of Medicine, Universitas Indonesia, dr. Cipto Mangunkusumo General Hospital for providing technical support, including access to TPS facilities and data resources. We also extend our heartfelt thanks to all the staff and employees for their assistance and for sharing their knowledge throughout the course of this research.

AUTHOR CONTRIBUTION

Yolandita Ayu Lestari: investigation, data analysis, writing – original draft, and writing – review and editing the manuscript. Dwi Seno Kuncoro Sihono: Conceptualization, Methodology, Resources, Writing – Review & Editing. Prawito Prajitno: Conceptualization, Methodology, Resources, Writing – Review & Editing. Wahyu Edy Wibowo: Conceptualization, Resources, Validation, Writing – Review & Editing. Arie Munandar: Conceptualization, Resources, Validation, Writing – Review & Editing. Bisma B. Patrianesha: Data curation, Validation, Writing – Review & Editing.

REFERENCES

1. A. Brahme, J. -E. Roos, and I. Lax, *Solution of an integral equation encountered in rotation therapy*, in: Physics in Medicine & Biology (1982) 27.
2. C. G. Orton, T. R. Bortfeld, A. Niemierko *et al*., Med. Phys. **35** (2008) 4911.
3. J. J. Boutilier, T. Lee, T. Craig *et al*., Med. Phys. **42** (2015) 1586.
4. R. Lu, R. J. Radke, L. Hong *et al*., IEEE Trans. Biomed. Eng. **53** (2006) 908.
5. X. Kui, F. Liu, M. Yang *et al*., Meta Radiol. **2** (2024) 100057.
6. X. Mao, J. Pineau, R. Keyes *et al*., Int. J. Radiat. Oncol. Biol. Phys. **108** (2020) 802.
7. C. McIntosh, L. Conroy, M. C. Tjong *et al*., Nat. Med. **27** (2021) 999.
8. T. Giaddui, H. Geng, Q. Chen *et al*., Adv. Radiat. Oncol. **5** (2020) 1342.
9. H. Wang, P. Dong, H. Liu *et al*., Med. Phys. **44** (2017) 389.
10. X. Zhang, D. Yang, Y. Jiang *et al*., J. Inflamm. Res. **14** (2021) 4205.
11. R. Ramlah, M. Fadli, J. Valerian *et al*., Iran. J. Med. Phys. **20** (2023) 298.
12. G. Zhang, Z. Jiang, J. Zhu *et al*., Radiat. Oncol. **17** (2022) 179.
13. C. Fiandra, F. Cattani, M.C. Leonardi *et al*., Adv. Radiat. Oncol. **8** (2023) 101228.
14. C. A. Burmeister, S. F. Khan, G. Schäfer *et al*., Tumour. Virus. Res. **13** (2022) 200238.
15. Y. Li, C. Zou, M. Berecibar *et al*., Appl. Energy **232** (2018) 197.
16. D. Liu, X. Zhang, T. Zheng *et al*., Arch. Gynecol. Obstet. **303** (2021) 811.
17. J. D. Ospina, J. Zhu, C. Chira *et al*., Int. J. Radiat. Oncol. Biol. Phys. **89** (2014) 1024.
18. C. C. M. Chen, H. Schwender, J. Keith *et al*., IEEE/ACM Trans. Comput. Biol. Bioinf. **8** (2011) 1580.
19. S. Albahli, Multimedia Tools Appl. **83** (2024) 52711.
20. M. H. Rizky, M. R. Faisal, I. Budiman *et al*., Indones. J. Comput. Cybern. Syst. **18** (2024) 95.
21. F. Pedregosa, G. Varoquaux, A. Gramfort *et al*., J. Mach. Learn. Res. **12** (2011) 2825.

22. L. B. Marks, E. D. Yorke, A. Jackson *et al*., Int. J. Radiat. Oncol. Biol. Phys. **76** (2010) S10.

23. K. Kartutik, W. E. Wibowo, S. A. Pawiro, *Comparison of radiotherapy dosimetry for 3D-CRT, IMRT, and SBRT based on electron density calibration*, in: Journal of Physics: Conference Series (2016) 012017.

24. I. Miyaji, H. Fukui, Eur. J. Educ. Res. **9** (2020) 1663.

25. L. Judijanto, D. M. Priyangan, H. N. Muthmainah *et al*., Eastasouth J. Inf. Syst. Comput. Sci. **1** (2023) 75.

26. D. Reijtenbagh, J. Godart, A. de Leeuw *et al.*, Radiother. Oncol. **170** (2022) 169.