# Predicting prostate cancer radiotherapy complications: An integrated approach using radiomics, dosiomics, and machine learning

## E. Sadati[1], B. Hashemi[1*], S.R. Mahdavi[2], A. Nikoufar[3], H. Abdollahi[4]

*[1]Department of Medical Physics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran*
*[2]Department of Medical Physics, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran*
*[3]Department of Radiation Oncology, Faculty of Medicine, Iran University of Medical Sciences, Tehran, Iran*
*[4]Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC, Canada*

► *Original article*

## ABSTRACT

**Background:** We aimed to develop a robust prognostic model for assessing the risk of complications associated with radiotherapy in prostate cancer patients using radiomics and dosiomics feature and machine learning. **Materials and Methods:** A cohort of 60 patients undergoing pelvic radiation therapy was analyzed. The patients' radiomics and dosiomics features were extracted from segmented bladder and rectum regions in CT images, as well as 3D dose distribution data, respectively. Classifier algorithms, such as eXtreme Gradient Boosting (XGBoost), Decision Tree (DT), Support Vector Machines (SVM), K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), and Multilayer Perceptron (MLP) were used for prediction modeling. A 5-fold cross-validation method was used to evaluate the predictive classification of patients with and without proctitis and cystitis. The area under the receiver operating characteristic curve (AUC) was used for comparing models' performance, as well assessing their specificity and accuracy metrics. **Results:** Various combinations of feature selection and classifier algorithms evaluated on both training and test datasets revealed that for bladder toxicity, the Relief+KNN dosiomics model, Boruta+SVM radiomics model, and the combined radiomics and dosiomics model with ANOVA+XGBoost show the highest AUCs of 0.76, 0.68, and 0.67, respectively. Regarding the rectal toxicity, the best-performing models were Boruta+KNN for dosiomics (AUC 0.83), ANOVA+RF for radiomics (AUC 0.72), and ANOVA+XGBoost for the combined radiomics and dosiomics (AUC 0.71). **Conclusion:** Our study demonstrated the effectiveness of diverse algorithms leveraging quantitative features extracted from CT imaging and 3D dose distribution data in predicting post-radiotherapy complications in prostate cancer patients.

## INTRODUCTION

Prostate cancer is a global health concern affecting men, and radiotherapy is a pivotal component of its treatment [1, 2]. However, radiotherapy can lead to complications that significantly impact a patient's well-being and treatment outcomes. These complications encompass various aspects, including damage to the bladder, prostatic urethra, and rectum, necessitating precise data and management for effective treatment planning and enhanced patient care.

In recent years, the burgeoning fields of radiomics and dosiomics have garnered substantial attention due to their ability to extract quantitative features from medical imaging data and dose distribution information, respectively [1-4]. These features offer valuable insights into tumor characteristics and the radiation dose delivered to the tumor and surrounding tissues [5, 6]. The integration of advanced machine learning techniques, such as logistic regression and k-nearest neighbors, with radiomics and dosiomics, holds great promise in predicting treatment outcomes and identifying patients at higher risk of complications [7, 8].

Numerous studies have attempted to develop models for these complications based on dosimetric and clinical parameters [9-11], but these models often face limitations, including variations in patient radio sensitivity and uncertainties in dosimetric and planning parameters.

From both clinical and radiobiological perspectives, it is evident that patients' responses to radiotherapy are subject to individualization. Incorporating a patient's inherent radio sensitivity into the radiation treatment process, from patient selection to planning, may positively impact treatment outcomes [12]. Ongoing research in the field of radiogenomics aims to tailor treatments based on a patient's genomic characteristics [12, 13].

Through these approaches, our objective was to set up robust predictive models that can facilitate personalized treatment planning and relieve the risks related to complications. The findings of our study hold significant potential for enhancing treatment planning and personalized care for prostate cancer patients undergoing radiotherapy. By precisely recognizing patients at a higher risk of complications, clinicians can actualize custom-fitted methodologies to moderate these risks, driving to made strides in treatment outcomes and improving patient quality of life [14].

This study innovates in predicting prostate cancer radiotherapy complications by integrating radiomics, dosiomics, and machine learning. Unlike traditional methods relying solely on dose-volume histograms, this approach leverages richer data and addresses limitations by capturing spatial information. It further emphasizes the strengths, limitations, and future directions for improved clinical utility in prostate cancer radiotherapy, ultimately aiming to optimize treatment and improve patient outcomes.

## MATERIALS AND METHODS

### Study design and patient cohort

This study was conducted using analysis of computed tomography (CT) images and treatment planning data from 60 patients who received external beam radiotherapy for prostate cancer. Patient data were anonymized, for the study, ensuring that all analyses were performed on de-identified datasets.

### Data acquisition

Planned CT images were obtained from a 16-slice SIEMENS scanner (Somatom Scope). Using special parameters, including a tube voltage of 120 kV, an exposure range of 225 mA, and a slice thickness of 3 mm, to optimize treatment planning. The MONACO (*version* 5.11) treatment planning system was used for all patients, ensuring the consistency of treatment plans. Intensity Modulated Radiation Therapy (IMRT) was chosen as the treatment modality, with standardized beam arrangements across all patients. This uniform approach allows for accurate comparisons and analysis of extracted radiomics and dosiomics features derived from the treatment plans.

### Toxicity assessment

Radiation-induced proctitis and cystitis were assessed based on patient records and documented per the Common Terminology Criteria for Adverse Events (CTCAE) version 5. Specifically, radiation-induced complications of grade ≥ 2 cystitis and proctitis were designated as primary toxicities for the bladder and rectum, respectively. Patients were classified as experiencing toxicity (Class 1) or not (Class 0).

### Segmentation

The Region of Interest (ROI) was meticulously segmented under the guidance of an experienced oncologist. This involved precise delineation of the anatomical area or tumor region of interest on the medical images used in our study. The oncologist, possessing specialized proficiency in prostate cancer imaging, drew upon their clinical expertise and skill to intricately outline the ROI, following established guidelines and protocols [15].

### Radiomics and dosiomics feature extracting

We employed the Pyradiomics library for feature extraction in this study, which is specifically designed to extract a wide range of features from segmented CT and 3D dose distribution (DD) data [16]. This comprehensive extraction encompassed various feature sets, including shapes, first-order, second-order, and higher-order features.

To ensure consistency in feature extraction, preprocessing steps were executed. Before the extraction, it was necessary to resample the CT images to standardize voxel sizes, allowing for meaningful comparisons. The 3D dose distribution data were characterized by an isotropic voxels of $1 \times 1 \times 1$ mm$^3$, which used to resample the CT images with a voxel size of $1 \times 1 \times 1$ mm$^3$ using B-spline interpolation. This standardization ensured that the extracted features were independent of the original voxel dimensions, empowering a more reliable analysis and comparison of the radiomics and dosiomics features over the patient cohort [17]. By minimizing the impact of image noise and variability, these discretization techniques aim to improve the accuracy and reliability of extracted features [18].

### Feature selection

Following feature extraction, three distinct datasets were generated: CT radiomics features, DD dosiomics features, and a combined dataset containing both CT and dose features. To facilitate model development and assessment, each dataset was partitioned into training (70%) and testing (30%) subsets. Feature values were standardized using z-scores on the training dataset and then applied to the test data.

To address the challenge of overfitting, it is essential to remove irrelevant and redundant features that do not contribute to the predictive accuracy of a model. This approach not only leads to faster and more cost-effective models but also enhances model performance. To achieve this, we employed various feature selection methods, including Recursive Feature Elimination (RFE), Analysis of Variance (ANOVA), Maximum Relevance Minimum Redundancy (MRMR), Boruta, and Relief. These feature selection techniques refined the feature sets, improving the models' performance by retaining relevant and non-redundant features [19-21].

### Classification

A diverse range of machine learning and classifiers (eXtreme Gradient Boosting (XGBoost), Decision Tree (DT), Support Vector Machines (SVM), K-Nearest Neighbor (KNN), Logistic Regression (LR), Random Forest (RF), and Multilayer Perceptron (MLP)) were employed to explore different approaches and identify the most effective ones for our specific problem. This allows us to :evaluate multiple classifiers, we can assess their strengths and weaknesses, and choose the one best suited to our dataset prediction task and exploring diverse algorithms helps us identify broader trends and reduces the risk of relying solely on an algorithm potentially specific to the given data.

Each classifier underwent hyperparameter tuning using the training dataset. This process optimizes the model's internal configuration to achieve the best possible performance [22].

### Model evaluation

We employed bootstrapping and tested models across 1000 bootstraps. This technique involves drawing multiple random samples with replacements from the original data and generating various training and testing sets. By evaluating performance across these resampled datasets, we could account for variability and gain a more robust understanding of the model's generalizability.

We used several metrics to assess model performance, including:

**Area under the curve (AUC)**: The Area Under the Curve of a Receiver Operating Characteristic (ROC) curve is essential for assessing binary classifier model effectiveness. The ROC curve visually demonstrates the model's ability to differentiate between two classes at varying thresholds .The ROC curve was constructed by plotting the true positive rate (TPR) against the false positive rate (FPR) at each threshold, providing insight into the model's performance across decision boundaries. TPR and FPR calculations across all thresholds facilitate a thorough evaluation of the model .The AUC measures the model's overall discriminative capacity. An AUC of 0.5 indicates no discrimination, while an AUC of 1.0 represents perfect classification. Consequently, a higher AUC signifies better classifier performance, making it an essential metric for model evaluation and comparison.

**Specificity (SPE)**: Specificity, often abbreviated as SPE, is a crucial metric for evaluating the performance of a binary classifier. It measures the ability of the model to correctly identify actual negative cases. In other words, specificity indicates the proportion of patients who do not experience an event (e.g., radiation-induced complications) and are accurately classified as not having the event .A high specificity means that the model minimizes false positives, ensuring that patients without complications are correctly identified. This is important in clinical settings, as high specificity helps prevent unnecessary treatments or interventions for patients who are not at risk.

**Accuracy (ACC):** Accuracy is a key metric for assessing binary classifier performance. It quantifies the ratio of correctly identified cases to the total cases. Essentially, accuracy reflects the model's prediction reliability for both the target event and its absence. Elevated accuracy indicates model dependability, critical in clinical contexts where precise patient classification influences management decisions.

**Sensitivity (SEN):** Sensitivity, or the true positive rate, quantifies a binary classifier's accuracy in identifying positive instances. It measures the ratio of correctly classified positive cases to the total actual positive cases .A high sensitivity is vital in clinical settings to ensure accurate detection of patients requiring intervention, thus minimizing false negatives. In conclusion, sensitivity serves as a critical metric for assessing the diagnostic performance of models in identifying at-risk individuals.

### Statistical validation

We employed k-fold cross-validation (k=5) to further validate the performance of our models. This technique divides the data into k folds, trains the model on k-1 folds, and tests it on the remaining fold. This process is repeated k times, and the average performance across all folds provides a more robust estimate of model generalizability compared to a single train-test split. An array of performance metrics and evaluation measures were calculated, and tailored to the specific problem and the type of model being employed. These measures included common results such as the AUC, SPE, ACC, and SEN. These metrics provided a reliable estimate of the model's ability to generalize to unseen data and significantly contributed to enhancing the validity of the research [23, 24].

## RESULTS

Among the 60 patients with prostate cancer, 15 (approximately 25%) developed proctitis of grade 2 or higher, and 21 patients (around 35%) experienced cystitis of grade 2 or higher. Figures 1 and 2 present the performance evaluation of various models that built on test data set in predicting rectal and bladder toxicities. We evaluated three different models, each using unique feature selection and classification methods, for predicting rectal toxicity in prostate cancer radiotherapy.

**Dosiomics model:** This model combined Boruta feature selection with the KNN classifier, ACC of 0.76, SPE of 0.77, and AUC of 0.83. It demonstrated balanced performance in both accuracy and

specificity, making it potentially valuable for identifying rectal toxicity cases.

**Radiomics model:** This model, utilizing ANOVA feature selection and RF classification, achieved an exceptional ACC of 0.85 and a high SPE of 0.92, though its AUC was 0.72. This highlights its robustness in predicting rectal toxicity.

**Combined radiomic and dosiomic features model:** This model, incorporating ANOVA and XGBoost, yielded promising results with an ACC of 0.77, SPE of 0.93, and AUC of 0.71. While the AUC may suggest room for improvement, the model exhibited high specificity alongside substantial predictive power.
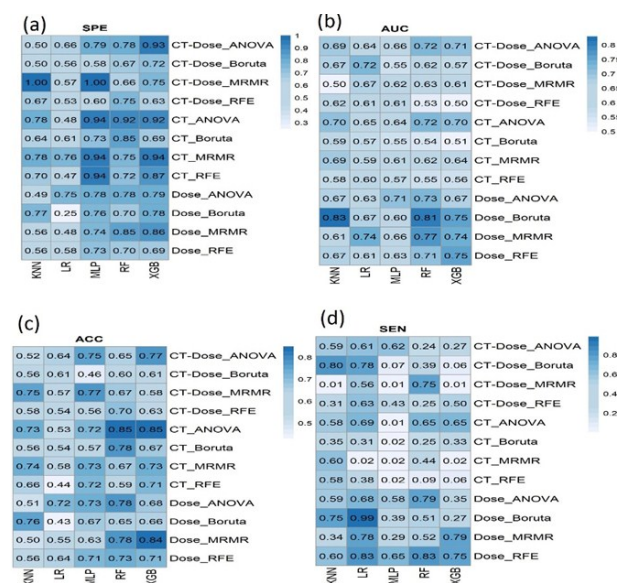


**Figure 1.** Performance comparison of various feature selection methods (ANOVA, Boruta, MRMR, RFE) and machine learning algorithms (KNN, LR, MLP, RF, XGB) on predicting rectal toxicity using CT, Dose, and CT-Dose datasets. The performance metrics - **(a)** SPE, **(b)** AUC, **(c)** ACC, and **(d)**- for each combination of feature selection and classification algorithm on the test dataset.

Three models were evaluated for their ability to predict bladder toxicity:

**Dosiomics model:** This model, utilizing Relief feature selection and KNN classification, achieved an ACC of 0.66, SPE of 0.64, and AUC of 0.76. While demonstrating reasonable predictive power (AUC), its overall performance suggests room for improvement.

**Radiomics model:** This model, incorporating Boruta feature selection and SVM classification, achieved an ACC of 0.69, SPE of 0.67, and AUC of 0.68. It displayed potential, particularly in accuracy and specificity, indicating its capability in bladder toxicity prediction, but further refinement might be needed.

**Combined radiomic and dosiomic features model:** This model, utilizing Relief feature selection and RF classification, achieved an ACC of 0.61, a high SPE of 0.75, and an AUC of 0.68. This model excelled in specificity, highlighting its ability to accurately

identify bladder toxicity cases, but its overall accuracy may require further optimization.

This study contributes to understanding radiomics and dosiomics for predicting toxicities in prostate cancer radiotherapy, paving the way for personalized treatment strategies. Further validation and refinement are crucial for clinical implementation, potentially revolutionizing management of these toxicities and improving patient care.
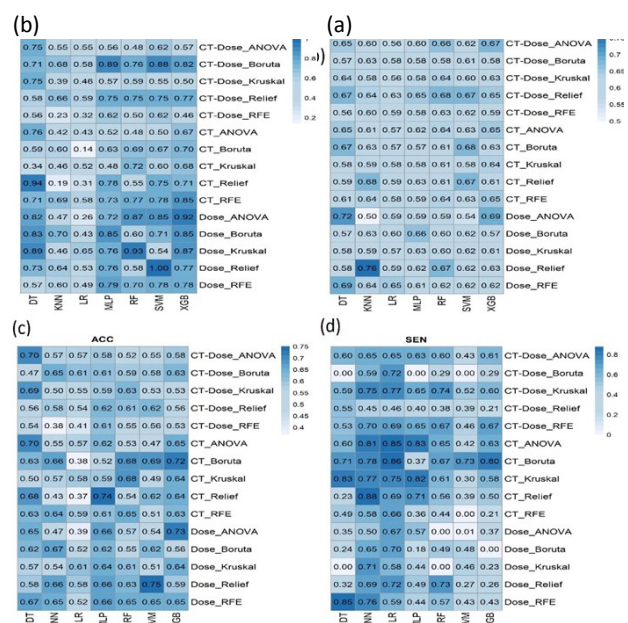


**Figure 2.** Performance comparison of various feature selection methods (ANOVA, Boruta, MRMR, RFE) and machine learning algorithms (KNN, LR, MLP, RF, XGB) on predicting bladder toxicity using CT, Dose, and CT-Dose datasets. The performance metrics - **(a)** SPE, **(b)** AUC, **(c)** ACC, and SEN **(d)** - for each combination of feature selection and classification algorithm on the test dataset.

## DISCUSSION

Predicting radiation-induced complications in prostate cancer radiotherapy remains challenging due to limitations in current dosimetric parameters, such as inadequate consideration of individual anatomical variations and tissue sensitivities. In our study, we explored the potential of CT radiomics and dosiomics to enhance prediction accuracy for acute bladder and rectal toxicities associated with radiotherapy.

Our findings highlight that specific radiomic features extracted from pre-treatment CT images, combined with 3D dose distribution data, yielded promising results in predicting these toxicities. This suggests that integrating radiomics and dosiomics alongside conventional dosimetry could improve the precision of complication prediction. Notably, we achieved an AUC of 0.75 and 0.83 for urinary and gastrointestinal injuries, respectively, underscoring

the predictive potential of CT and dose-derived features.

Several studies in radiomics and dosiomics have also demonstrated promising results in predicting complications induced by radiotherapy. For instance, Van Dijk *et al.* [25] developed predictive models for late xerostomia and sticky saliva post-radiotherapy using CT image features, showing superior performance over clinical models. Similarly, Kraus *et al.* [26] utilized dosiomics and radiomics features to predict pneumonitis following thoracic SBRT, further illustrating the utility of these approaches in different clinical contexts. Also, Qingying Yang *et al.* [27] created a strong radiomics model utilizing non-contrast CT scans for predicting pulmonary hypertension (PH). Their study compared this model's performance to prediction models based on clinical and radiological factors using ten different machine learning algorithms. The findings revealed that the SVM model had the highest prediction accuracy, with an AUC of 0.87 and an accuracy of 0.83. Moreover, the combined predictive model, which included radiomics features alongside clinical and radiological parameters, showed the best performance in forecasting pH.

In this study, SVM was utilized to develop models, as it is considered appropriate for small sample sizes [28]. Validation of the models was done through nested cross-validation. A recent study by Bourbonne *et al.* [29] indicated that relying on a single random split of data for training and testing with small sample sizes could yield unreliable results. They suggested using nested cross-validation in the absence of external validation. The use of nested cross-validation for small sample sizes is recommended [30] and has been implemented in various studies [31].

In recent years, advancements in radiotherapy outcome modeling have aimed to develop models that precisely predict radiotherapy endpoints while avoiding issues of overfitting and under fitting. This has been accomplished by assessing a range of contributing parameters. Pre-treatment factors, including radiomic features derived from patient imaging, play a crucial role in helping clinicians identify patients who may benefit from dose escalation or reduction [32, 33]. Our study adds to this field by utilizing outcome modeling to predict toxicity using high-quality CT images, offering a simple, non-invasive, and cost-effective approach.

Moving forward, our ongoing research focuses on developing robust models to predict toxicity using high-quality CT images and 3D dose distribution through outcome modeling [32]. Future efforts should include validating our findings with larger and more diverse patient populations to strengthen the generalizability and clinical utility of our models [25, 34].

## CONCLUSION

This study demonstrates the promising advances in using radiomics and dosiomics for toxicity prediction in prostate cancer radiotherapy, further investigation with comprehensive validation studies and broader demographic inclusion is essential. These efforts will facilitate informed treatment decisions and ultimately improve patient outcomes.

## REFERENCES

1. Liang B, Yan H, Tian Y*, et al.* (2019) Dosiomics: extracting 3D spatial features from dose distribution to predict incidence of radiation pneumonitis. *Frontiers in Oncology, 9: 269.*
2. Murakami Y, Soyano T, Kozuka T*, et al.* (2022) Dose-based radiomic analysis (dosiomics) for intensity modulated radiation therapy in patients with prostate cancer: correlation between planned dose distribution and biochemical failure. *Int J Radiat Oncol Biol Phys, 112(1): 247-59.*
3. Rossi L, Bijman R, Schillemans W*, et al.* (2018) Texture analysis of 3D dose distributions for predictive modelling of toxicity rates in radiotherapy. *Radiotherapy and Oncology, 129(3): 548-53.*
4. Wu A, Li Y, Qi M*, et al.* (2020) Dosiomics improves prediction of locoregional recurrence for intensity modulated radiotherapy treated head and neck cancer cases. *Oral Oncology, 104: 104625.*
5. Huang Y, Feng A, Lin Y*, et al.* (2022) Radiation pneumonitis prediction after stereotactic body radiation therapy based on 3D dose distribution: dosiomics and/or deep learning-based radiomics features. *Radiation Oncology, 17(1): 188.*
6. Shevach J, Weiner A, Morgans AK (2019) Quality of life–focused decision-making for prostate cancer. *Current Urology Reports, 20: 1-7.*
7. Leger S, Zwanenburg A, Pilz K*, et al.* (2017) A comparative study of machine learning methods for time-to-event survival data for radiomics risk modelling. *Scientific Reports, 7(1): 13206.*
8. Chan MF, Witztum A, Valdes G (2020) Integration of AI and machine learning in radiotherapy QA. *Frontiers in Artificial Intelligence, 3: 577620.*
9. Thor M, Olsson C, Oh JH*, et al.* (2016) Urinary bladder dose–response relationships for patient-reported genitourinary morbidity domains following prostate cancer radiotherapy. *Radiotherapy and Oncology, 119(1): 117-22.*
10. Viswanathan AN, Yorke ED, Marks LB*, et al.* (2010) Radiation dose–volume effects of the urinary bladder. *Int J Radiat Oncol Biol Phys, 76(3): S116-S22.*
11. Falahatpour Z, Geramifar P, Mahdavi SR*, et al.* (2022) Potential advantages of FDG-PET radiomic feature map for target volume

delineation in lung cancer radiotherapy. *Journal of Applied Clinical Medical Physics, 23(9): e13696*.

12. Fahrig A, Koch T, Lenhart M*, et al.* (2018) Lethal outcome after pelvic salvage radiotherapy in a patient with prostate cancer due to increased radiosensitivity. *Strahlentherapie und Onkologie, 194 (1): 60-6*.

13. El Naqa I, Kerns SL, Coates J*, et al.* (2017) Radiogenomics and radiotherapy response modeling. *Physics in Medicine & Biology, 62 (16): R179*.

14. Parker C, Castro E, Fizazi K*, et al.* (2020) Prostate cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology, 31(9): 1119-34*.

15. Gibson E, Bauman GS, Romagnoli C*, et al.* (2016) Toward prostate cancer contouring guidelines on magnetic resonance imaging: dominant lesion gross and clinical target volume coverage via accurate histology fusion. *Int J Radiat Oncol Biol Phys, 96(1): 188-96*.

16. Van Griethuysen JJ, Fedorov A, Parmar C*, et al.* (2017) Computational radiomics system to decode the radiographic phenotype. *Cancer Research, 77(21): e104-e7*.

17. Zhang X, Zhang Y, Zhang G*, et al.* (2022) Deep learning with radiomics for disease diagnosis and treatment: challenges and potential. *Frontiers in Oncology, 12: 773840*.

18. Limkin EJ, Sun R, Dercle L*, et al.* (2017) Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Annals of Oncology, 28(6): 1191-206*.

19. Bugata P and Drotar P (2020) On some aspects of minimum redundancy maximum relevance feature selection. *Science China Information Sciences, 63(1): 112103*.

20. Radovic M, Ghalwash M, Filipovic N, Obradovic Z (2017) Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. *BMC Bioinformatics, 18: 1-14*.

21. Abdulsalam SO, Mohammed AA, Ajao JF*, et al.* (2020) Performance evaluation of ANOVA and RFE algorithms for classifying microarray dataset using SVM. In: *Information Systems: 17th European, Mediterranean, and Middle Eastern Conference, EMCIS 2020, Dubai, United Arab Emirates*, p. 480-92. Springer.

22. Way TW, Sahiner B, Hadjiiski LM, Chan HP (2010) Effect of finite sample size on feature selection and classification: a simulation study. *Medical Physics, 37(2): 907-20*.

23. Hsu WC, Liu CC, Chang F, Chen SS (2013) Selecting genes for cancer classification using SVM: an adaptive multiple features scheme. *International Journal of Intelligent Systems, 28(12): 1196-213*.

24. Wainer J and Cawley G (2021) Nested cross-validation when selecting classifiers is overzealous for most practical applications. *Expert Systems with Applications, 182: 115222*.

25. Van Dijk LV, Brouwer CL, Van der Laan HP*, et al.* (2017) Geometric image biomarker changes of the parotid gland are associated with late xerostomia. *Int J Radiat Oncol Biol Phys, 99(5): 1101-10*.

26. Kraus KM, Oreshko M, Bernhardt D*, et al.* (2023) Dosiomics and radiomics to predict pneumonitis after thoracic stereotactic body radiotherapy and immune checkpoint inhibition. *Frontiers in Oncology, 13: 1124592*.

27. Yang Q, Sun J, Guo Y*, et al.* (2022) Radiomics Features on Computed Tomography Combined With Clinical-Radiological Factors Predicting Progressive Hemorrhage of Cerebral Contusion. *Frontiers in Neurology, 13: 839784*.

28. Song J, Yin Y, Wang H*, et al.* (2020) A review of original articles published in the emerging field of radiomics. *European Journal of Radiology, 127: 108991*.

29. Bourbonne V, Da-Ano R, Jaouen V*, et al.* (2021) Radiomics analysis of 3D dose distributions to predict toxicity of radiotherapy for lung cancer. *Radiotherapy and Oncology, 155: 144-50*.

30. Bradshaw TJ, Boellaard R, Dutta J*, et al.* (2022) Nuclear medicine and artificial intelligence: best practices for algorithm development. *Journal of Nuclear Medicine, 63(4): 500-10*.

31. Fan Z, Sun Z, Fang S*, et al.* (2021) Preoperative Radiomics Analysis of 1p/19q Status in WHO Grade II Gliomas. *Frontiers in Oncology, 11: 616740*.

32. Coates J and El Naqa I (2016) Outcome modeling techniques for prostate cancer radiotherapy: Data, models, and validation. *Physica Medica, 32(3): 512-20*.

33. Mostafaei S, Abdollahi H, Kazempour Dehkordi S*, et al.* (2020) CT imaging markers to improve radiation toxicity prediction in prostate cancer radiotherapy by stacking regression algorithm. *La Radiologia Medica, 125: 87-97*.

34. Saeedi E, Dezhkam A, Beigi J*, et al.* (2019) Radiomic feature robustness and reproducibility in quantitative bone radiography: a study on radiologic parameter changes. *Journal of Clinical Densitometry, 22(2): 203-13*.