



OPEN ACCESS

EDITED BY

Laura Cella,
National Research Council (CNR), Italy

REVIEWED BY

Roberto Pacelli,
University of Naples Federico II, Italy
Cynthia Terrones Campos,
University of Copenhagen, Denmark

*CORRESPONDENCE

Łukasz Kuncman
✉ lukasz.kuncman@umed.lodz.pl

†These authors have contributed equally to this work

RECEIVED 01 May 2024

ACCEPTED 16 July 2024

PUBLISHED 01 August 2024

CITATION

Kuncman Ł, Pajdziński M, Smółka K, Bilski M, Socha J, Stando R, Peszyńska-Piorun M, Korab K, Jereczek-Fossa BA and Fijuth J (2024) Early lymphocyte levels and low doses radiation exposure of lung predict lymphopenia in radiotherapy for lung cancer. *Front. Immunol.* 15:1426635. doi: 10.3389/fimmu.2024.1426635

COPYRIGHT

© 2024 Kuncman, Pajdziński, Smółka, Bilski, Socha, Stando, Peszyńska-Piorun, Korab, Jereczek-Fossa and Fijuth. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Early lymphocyte levels and low doses radiation exposure of lung predict lymphopenia in radiotherapy for lung cancer

Łukasz Kuncman^{1,2*}, Matusz Pajdziński^{1,2}, Krzysztof Smółka³, Mateusz Bilski^{4,5,6}, Joanna Socha⁷, Rafał Stando⁸, Magdalena Peszyńska-Piorun⁹, Katarzyna Korab⁶, Barbara Alicja Jereczek-Fossa^{10,11†} and Jacek Fijuth^{1,2†}

¹Department of Radiotherapy, Medical University of Lodz, Lodz, Poland, ²Department of External Beam Radiotherapy, Copernicus Memorial Hospital in Lodz Comprehensive Cancer Center and Traumatology, Lodz, Poland, ³Institute of Mechatronics and Information Systems, Lodz University of Technology, Lodz, Poland, ⁴Department of Radiotherapy, Medical University of Lublin, Lublin, Poland, ⁵Department of Brachytherapy, Lublin Cancer Center, Lublin, Poland, ⁶Department of Radiotherapy, Lublin Cancer Center, Lublin, Poland, ⁷Department of Radiotherapy, Regional Oncology Center, Czestochowa, Poland, ⁸Department of Radiation Oncology, Holycross Cancer Center, Kielce, Poland, ⁹Radiotherapy Planning Department, Copernicus Memorial Hospital in Lodz Comprehensive Cancer Center and Traumatology, Lodz, Poland, ¹⁰Department of Radiation Oncology, European Institute of Oncology IRCCS, Milan, Italy, ¹¹Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy

Introduction: Radiation induced lymphopenia (RIL) deteriorate survival and diminishes the benefit of immune checkpoint inhibitors in combined treatment of lung cancer. Given the inconsistent data across various studies on the predictors of RIL, we aim to methodically elucidate these predictors and formulate a practical guide for clinicians.

Methods: We conducted observational cohort study in four tertiary cancer centers. Patients with non-small cell lung cancer and small cell lung cancer, without lymphopenia grade >1, who underwent standalone radiotherapy (RT) in minimum 15 fractions were eligible. Dose-volume parameters of structures and clinical factors were comprehensively analyzed using various predictors selection methods and statistical models (Linear Regressors, Elastic Net, Bayesian Regressors, Huber Regression, regression based on k-nearest neighbors, Gaussian Process Regressor, Decision Tree Regressor, Random Forest Regressor, eXtreme Gradient Boosting, Automated Machine Learning) and were ranked to predict lymphocytes count nadir (alc_nadir).

Results: Two hundred thirty eight patients (stage I-3.4%, II-17.6%, III-75.2%, IV-3.8%) who underwent RT to median dose of 60 Gy were analyzed. Median alc_nadir was 0.68K/mm³. The 60 feature sets were evaluated in 600 models (RMSE 0.27-0.41K/mm³). The most important features were baseline lymphocyte count (alc_1), mean lung_dose, lung v05, lung v10, heart v05 and effective dose to immune cells (edic). In patients with alc_1 ≤ 2.005K/mm³, median alc_nadir

predictions were 0.54K/mm³ for lung_v05p > 51.8% and 0.76K/mm³ for lung_v05p ≤ 51.8%. Lymphopenia was rare in patients with alc_1 > 2.005K/mm³.

Discussion: RIL was most severe in patients with low early lymphocyte counts, primarily triggered by low RT doses in the heart and lungs.

KEYWORDS

lymphopenia, radiotherapy, immunotherapy, immune checkpoint inhibitors, radiation induced lymphopenia, effective dose to immune cells, lung cancer

1 Introduction

Landscape of treatment of lung cancer is evolving in recent years with increasing role of immunotherapy in treatment of all stages of non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) (1–3). Immune checkpoint inhibitors (programmed cell death protein 1 (PD-1)/programmed death-ligand 1 (PD-L1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4)) have become standard components in both the first line and subsequent lines of systemic therapy for metastatic NSCLC and SCLC (1, 3). Building on successes in metastatic cancer, those discoveries have been integrated into concurrent chemo-radiotherapy of advanced NSCLC, with results for SCLC anticipated in the coming years (4, 5). For patients ineligible for concurrent chemo-radiotherapy due to performance status, comorbidities, or loco-regional tumor burden, radiation therapy alone (RT) and sequential chemo-radiotherapy remain the alternatives, with ongoing trials also exploring the incorporation of immunotherapy in these scenarios (6, 7).

The incorporation of immune checkpoint inhibitors into the management of unresectable lung cancer has shifted the focus towards the immune system, especially highlighting the significance of lymphocytes. In the context of combined CRT and immune checkpoint inhibitors for treating unresectable, loco-regionally advanced lung cancer, the concern of lymphopenia becomes even more critical (8). RT is the primary culprit of lymphopenia due to the high radiosensitivity of lymphocytes, which are the most radiosensitive among all peripheral blood cells; this leads to radiation-induced lymphopenia (RIL) that impairs the effectiveness of consolidative immunotherapy (9–11). Additionally, treatment-related lymphopenia was found to reduce OS in lung cancer patients receiving immune checkpoint inhibitors (12).

Despite the compelling nature of the aforementioned data, unified guidelines for lymphocyte-sparing radiotherapy have yet to be established (13). Various studies present inconsistent data on how different dose-volume indicators influence treatment-related lymphopenia (11, 13). Two primary theories explain RIL: the first connects it to dose to immune cells in “blood rich” critical organs (heart, lungs, vessels) and the second associates it with bone marrow doses (14–19). Importantly, these theories, though separate, do not

contradict each other in elucidating RIL. In numerous scientific publications, a variety of statistical models have been employed, complicating the comparison and evaluation of methods, and hindering their clinical implementation (13).

Our objective is to methodically evaluate the aforementioned hypotheses using diverse statistical modeling techniques across four distinct patient cohorts who underwent RT or sequential chemo-radiotherapy in tertiary cancer centers to find features which predict and potentially prevent RIL most accurate and create physician friendly algorithm.

2 Materials and methods

2.1 Study design and participants

In this retrospective multicenter observational cohort study, we included adult patients diagnosed with histopathological confirmed SCLC or NSCLC. Patients were eligible if they were qualified for standalone radical RT or sequential chemo-radiotherapy between September 2019 and December 2022 by multidisciplinary board in four tertiary cancer centers in Poland. The RT was administered with curative intent with at least 15 fractions. Patients were excluded from the study if they presented with an absolute lymphocyte count (alc_1) of Common Terminology Criteria for Adverse Events (CTCAE) v5.0 grade >1 (less than 800 cells/cm³) at the initiation of RT or if they experienced a break in radiotherapy exceeding seven days for reasons not related to treatment complications. Data for this study were collected retrospectively from medical records and RT planning systems, while contours for the accessed target volume of total vertebral body bone marrow (vb) and the bone marrow volume (bm) were created prospectively for the purposes of the current study according to provided protocol. The vb area is defined as the volume of vertebrae body Th1–Th10. The bm volume encompasses all bones in the cranial-caudal dimension of Th1–Th10, including the vb area. All RT plans were recalculated after delineation of above structures.

Potential clinical and dosimetry predictors of RIL were derived based on the review of literature from the PubMed, Scopus, Web of Science, databases performed by ŁK and supervised by JF. Those

factors were all in line with recently published LymphoTEC recommendations and metanalysis (8, 11). The data collected in accordance with the protocol included: Age in Years (age), Sex (female/male), Disease Stage (Stage; stages 1–4), Eastern Cooperative Oncology Group Performance Status (ECOG) (ecog_ps; 0–4), Prophylactic cranial irradiation before Lung RT (pci; No/Yes), absolute Lymphocyte Count from 1 Week of Radiotherapy (alc_1; in thousand/mm³), Nadir Absolute Lymphocyte Count during and up to Two Weeks Post-RT (alc_nadir; in thousand/mm³), Total Dose of Radiotherapy (rt_total_dose; in Gy), Fractional Dose of RT (rt_fraction_dose; in Gy), Duration of Radiotherapy (rt_duration; in days), Prior Chemotherapy (cht_before_rt; 0-No, 1-≤4 weeks, 2->4 weeks), Vertebral Body Volume (vb_v; in cm³), Mean Vertebral Body Dose Th1-Th10 (vb_md; in Gy), Percent Volume of Vertebral Body receiving 5, 10, 20, 30, and 40 Gy (vb_v05p, vb_v10p, vb_v20p, vb_v30p, vb_v40p), Volume of Bones (Bone Marrow) Th1-Th10 including Vertebral Body (bm_v; in cm³), Mean Dose in Bones Th1-Th10 including Vertebral Body (vb_md; in Gy), Percent Volume of Bone Marrow receiving 5, 10, 20, 30, and 40 Gy (bm_v05p, bm_v10p, bm_v20p, bm_v30p, bm_v40p), Volume of Planning Target Volume (ptv_v; in cm³), Heart Volume (heart_v; in cm³), Medium Heart Dose (heart_mhd; in Gy), Percent Volume of Heart receiving 5 and 10 Gy (heart_v05p, heart_v10p), Lung Volume (lung_v; in cm³), Medium Lung Dose (lung_mld; in Gy), Percent Volume of Lung receiving 5 and 10 Gy (lung_v05p, lung_v10p), Medium Body Dose Th1-Th10 (mbd; in Gy). The features calculated based on the collected data were: Absolute Lymphocyte Count Classification per CTCAE v5.0 (ctcae), Effective Dose to Immune Cells (edic) (15), RT-induced lymphopenia ≥G3, which was defined as Nadir Absolute Lymphocyte Count less than 0.5 (thousand/mm³). Full statistical description is provided in the [Supplementary Material](#). Lymphocytes are extremely radiosensitive and as proposed in previous papers linear model (not linear-quadratic) was used for assessment dose-volume effect for primary endpoint (lymphocytes counts) (20–22). Additionally, this approach facilitates comparison of different dose-levels used for modeling.

The study adhered to ethical standards and received approval from the relevant institutional review boards/director of participating hospitals. According to the local regulation, the consent of the Bioethics Committee was not required due to the retrospective nature and anonymization of the data.

2.2 Primary endpoint

The primary endpoint of the study was to methodically determine the dose-volume parameter(s) in radiotherapy treatment plans and clinical factors that most accurately predict alc_nadir, across four independent lung cancer patient's cohorts undergoing standalone RT or sequential chemo-radiotherapy.

2.3 Statistical analysis

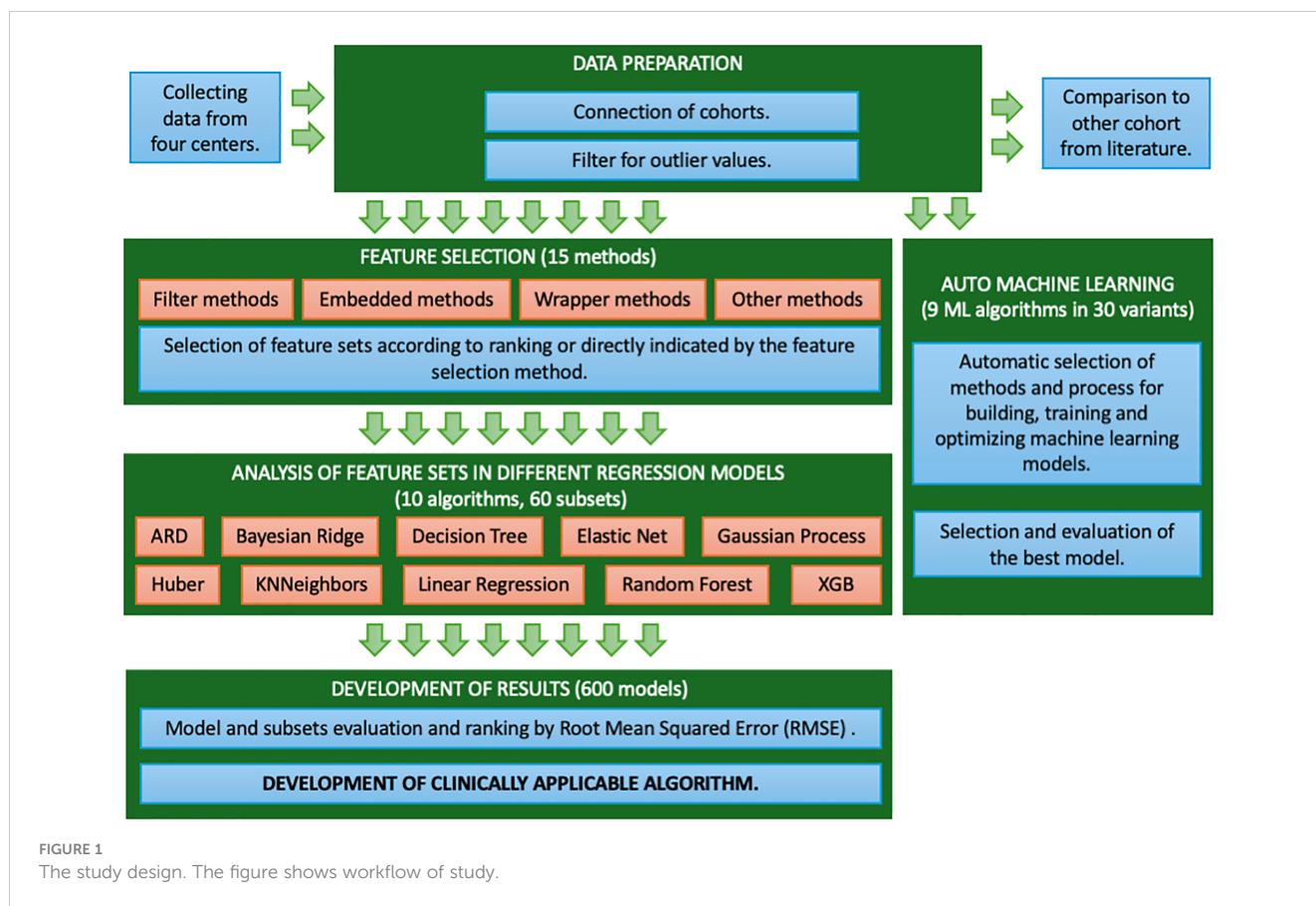
The study design is summarized in [Figure 1](#). Briefly data was compiled from four tertiary cancer centers, where rules based on boxplot procedures were used to remove outliers (the maximum

alc_nadir value was set at 1.65 thousand/mm³, the maximum alc_1 value was set at 3.57 thousand/mm³). Of 258 patients, 238 were included to further analysis.

Firstly, we generated models for prediction of ≥3 CTCAE using commonly used algorithms (Random Forest, ADABost, XGBoost and Logistic Regression) to compare the dataset to the data in the existing literature. The Receiver Operating Characteristic (ROC) was used for analysis. Second step aimed to find dose-volume metrics and clinical feature which predicts alc-nadir with best accuracy. The continuous variable (alc_nadir) was chosen for evaluation as a more sensitive measure compared to categorical metrics (ctcae).

Subsequently, feature selection methods were applied, recognizing the significant impact of variable selection on model efficiency (23). Given the lack of the universal feature selection method, multiple diverse methods were utilized to identify key features (24, 25). The filter-based methods (Variance Threshold, Chi squared, ANOVA, Information gain, Correlation Coefficient, Fisher score, Information Value (IV)), embedded methods (Lasso Regularization, Random Forest Importance), wrapper methods (Forward Feature Selection, Backward Feature Elimination, Exhaustive Feature Selection, Recursive Feature Elimination, Recursive Feature Elimination with Cross Validation) and other methods (Shapley Values, hybrid methods, arbitrary features selection) as shown in [Figure 1](#). Filter-based methods identify features by their statistical properties, such as Variance Threshold for low variance. Embedded methods like Lasso Regularization and Random Forest Importance select features during model training by shrinking coefficients or measuring feature impact on accuracy. Wrapper methods like Forward Feature Selection, Backward Feature Elimination, Exhaustive Feature Selection, Recursive Feature Elimination iteratively build and assess models to evaluate feature importance, using cross-validation for optimal feature selection. Other methods, such as Shapley Values, explain each feature's contribution to predictions, while hybrid and arbitrary selection methods leverage multiple techniques and domain knowledge.

After the above-described selection, obtained feature sets were evaluated using ten different regression models, including classical Linear Regressors, Linear Regressor with variable selection (Elastic Net), Bayesian Regressors (Automatic Relevance Determination – ARD and Bayesian Ridge), Outlier-robust Regressor (Huber Regression), regression based on k-nearest neighbors (KNN), Gaussian Process Regressor (GPR), Decision Tree Regressor, ensemble methods (Random Forest Regressor and eXtreme Gradient Boosting - Xgboost). Five-group cross-validation was used to prevent overfitting and provide a stable and more reliable assessment. Obtained models were ranked based on Root Mean Square Error (RMSE). Parallel to the feature selection process, we optimized machine learning models using automated machine learning (AutoML) tools. In the model verification process for machine learning, the dataset was divided into three groups: training and validation data (collectively 75%) using five-fold cross-validation, and test data (25%). The training data were utilized for constructing and optimizing the model, the validation data facilitated the tuning of hyperparameters and prevention of



overfitting through the selection of the optimal model configuration, while the test data enabled the assessment of the model’s final performance. After selection of features (max depth=2) decision tree was generated.

Calculations were performed using Python, scikit-learn (sklearn) machine learning (ML) library, dtreeviz library to visualized and interpreted decision trees, XGBoost Python package and statistical analysis package - statsmodels. An automatic machine learning library, AutoGluon, was used to optimize machine learning models. A Jupyter notebook was used as the integrated development environment (IDE).

3 Results

The 238 patients with NSCLC and SCLC (stage I-3.4%, II-17.6%, III-75.2%, IV-3.8%) in ECOG 0–3 performance status were enrolled. Patients were treated with hypo-fractionated RT in 62 cases and normo-fractionation in 176 cases to median RT dose of 60 Gy. Fraction doses ranged from 1.8 to 3.0 Gy. Total duration of treatment was 19–76 days (median 40). Lymphopenia G1, G2, G3 and G4 occurred in 31.5%, 35.3%, 26.9%, 1.3% of patients, respectively. The median alc_nadir was 0.68K/mm³. Clinical data are summarized in [Table 1](#). The description statistics of all collected dose-volume and clinical data are summarized in [Supplementary Table S1](#) in [Supplementary Material](#).

TABLE 1 Patient characteristics.

Feature	Total	CEN 1	CEN 2	CEN 3	CEN 4
Patients	238	141	25	34	38
Sex					
female	93	57	8	15	13
male	145	84	17	19	25
Age					
min	39	50	39	54	45
max	85	85	84	85	76
median	67	67	70	67	67
Disease stage (I-IV)					
I	8	4	1	3	0
II	42	19	5	6	12
III	179	115	18	20	26
IV	9	3	1	5	0
ALC from 1 week RT (thousand/mm³)					
min	0.81	0.81	0.94	1.10	0.81
max	3.43	3.43	3.00	3.38	3.24
median	1.76	1.63	1.97	1.97	1.72

(Continued)

TABLE 1 Continued

Feature	Total	CEN 1	CEN 2	CEN 3	CEN 4
ALC nadir (thousand/mm³)					
min	0.15	0.15	0.27	0.18	0.19
max	1.58	1.51	1.58	1.56	1.53
median	0.68	0.66	1.13	0.58	0.68
Total RT dose (Gy)					
min	40	40	42	53	56
max	68	66	66	68	66
median	60	60	60	60	66
Hypofractionation (YES: RT_fraction_dose>2, NO: RT_fraction_dose ≤ 2)					
YES	62	34	12	12	4
NO	176	107	13	22	34
V PTV - Volume planning target (cm³)					
min	55	55	70	66	105
max	1149	1058	860	1149	666
median	292	257	452	360	372
Lymphopenia CTCAE v5.0					
G0	12	4	4	2	2
G1	75	40	15	11	9
G2	84	58	2	8	16
G3	64	38	4	12	10
G4	3	1	0	1	1
G5	0	0	0	0	0
Chemotherapy before RT?					
No	73	46	6	8	13
CHT ≤ 4 weeks	36	0	11	13	12
CHT>4 weeks	129	95	8	13	13

As anticipated, we identified significant positive correlations among dose-volume parameters across distinct categories, specifically within regions including the heart (heart_mhd, heart_v05, heart_v10), lungs (lung_mld, lung_v05, lung_v10), bone marrow (bm_md, bm_v05, bm_v10, bm_v20, bm_v30, bm_v40), and vertebral bodies (vb_md, vb_v05, vb_v10, vb_v20, vb_v30, vb_v40), as delineated in the correlation matrix presented in Figure 2. The variables within these groups exhibited strong collinearity.

According to the calculations performed using four classifiers (AdaBoost, XGBoost, Random Forest, Logistic Regression), shown on Figure 3, models yield good performance in prediction of ctcae with area under curve (AUC) ranging from 0.71 to 0.76, with best performance for Random Forest and Logistic Regression. The base details of model were as follows: 5 features (alc_1, lung_mld, lung_v05p, heart_v05p, sex) and 5 fold cross validation.

The variance in the ROC metric incorporating individual results from cross-validation analyzes of best model was depicted in Supplementary Figure S1 in Supplementary Material. Our models had comparable performance in prediction of ctcae toxicity to those described in the literature thus we moved further to the next step described in Figure 1 (26–29).

Feature selection revealed 60 feature sets which were evaluated using ten different regression models and yield 600 models of alc_nadir prediction. Models consisting of a larger number of features did not demonstrate greater efficacy as shown in Figure 4. The detailed comparison of individual and grouped feature selection methods, as well as the performance of the corresponding models, is presented in Supplementary Figure S2. The Table 2 displays the efficacy of the five leading types of models (best one from each type) according to Root Mean Squared Error (RMSE), best 30 models are shown in Supplementary Table S2 in Supplementary Material. The RMSE and accuracy ranged 0.27–0.41 thousand/mm³ and 47–62% respectively. Leading models presented similar efficacy (best three were Random Forest Regressor as shown in Supplementary Table S2). The scatter plots and histograms of the models from Table 2 are shown in Figures 5A–J and represents a comparison between actual and predicted values. The clustering of points along the dashed line suggests a good model fit, especially for AutoML model. The most important features consistently observed across the models were: alc_1, lung_mld, lung_v05p, lung_v10p, heart_v05p and edic (calculated from lung_mld, heart_mld, mbd, number of fractions).

Final step aimed to generate clinically applicable algorithm to guide physicians. Decision tree from variables from best performing models included in Table 2 is shown on Figure 6. The alc_nadir was predicted based on alc_1, lung_v05p, and lung_mld. The decision tree model (RMSE = 0.29) identified alc_1 of 2.005K/mm³ as the threshold in the first step of selection. For patients with alc_1 values less than or equal to 2.005K/mm³ and lung_v05p ≤ 51.8%, a median alc_nadir of 0.76K/mm³ was estimated, while for those with lung_v05p > 51.8%, a median alc_nadir of 0.54K/mm³ was predicted. For patients with alc_1 greater than 2.005K/mm³, a lung_mld threshold of 10.67Gy segregated patients into groups with predicted median alc_nadir values of 1.11K/mm³ and 0.87K/mm³.

4 Discussion

Lymphocytes are the most radiosensitive blood cells, and although often underreported, lymphopenia is the most prevalent hematologic toxicity, with grade ≥ 2 and ≥ 3 occurring in our cohort in 63% and 28% of cases respectively (8). Lymphopenia is widely recognized as an unfavorable prognostic factor for both progression-free survival (PFS) and overall survival (OS) in patients with various cancers including NSCLC and SCLC (30, 31). Similarly, RIL has been shown to negatively impact OS and PFS in NSCLC and SCLC patients (8, 11, 13, 31). It's important to note that most studies linking RIL with lower survival were done before immunotherapy was used for lung cancer. The introduction immune checkpoint inhibitors might further amplify this negative impact, especially given that lymphopenia induced by treatment is a recognized as unfavorable prognostic factor for both OS and PFS during immune

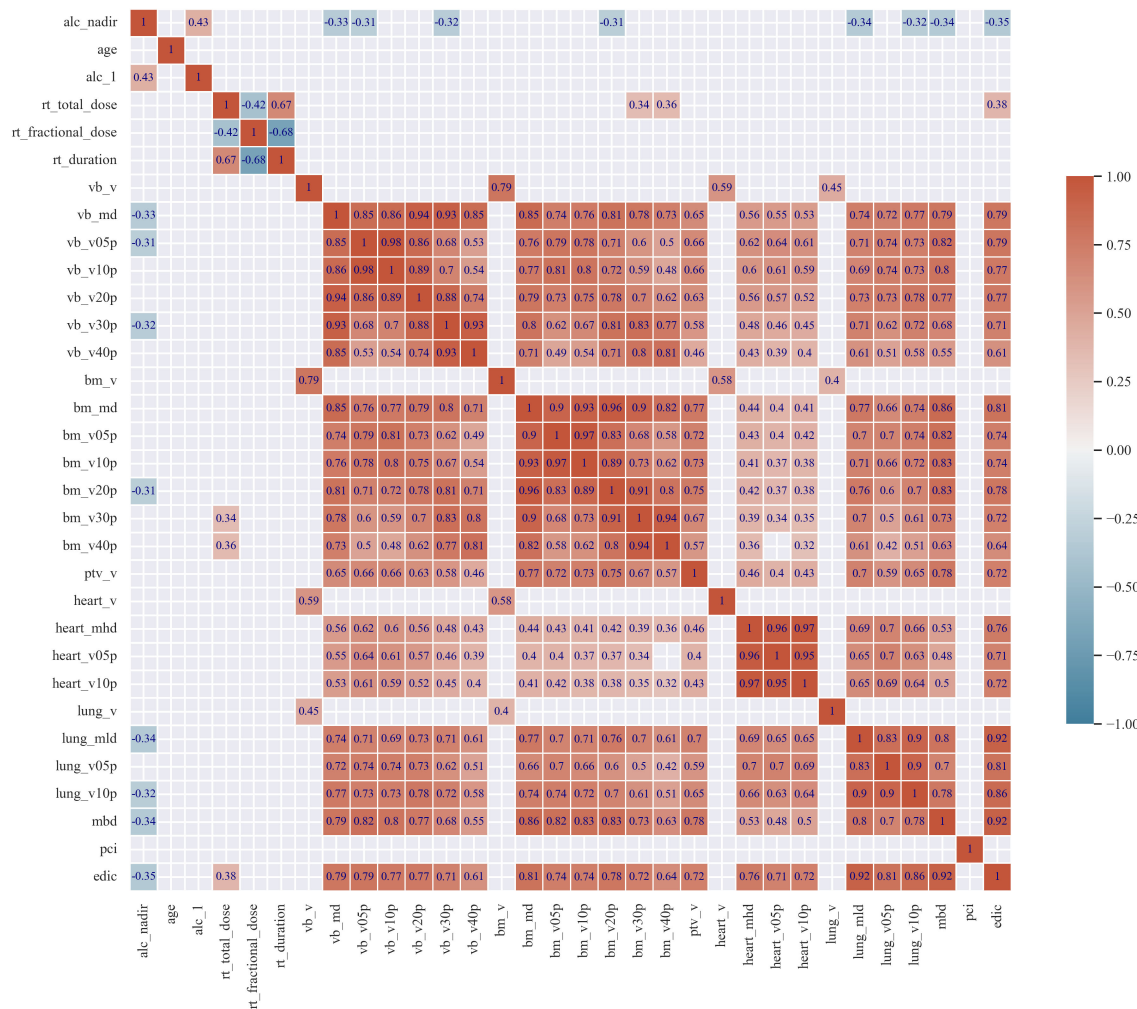


FIGURE 2
The correlation matrix. The correlation matrix heatmap for the values of the Spearman correlation coefficients (r) for analyzed features. For clarity of the correlogram, only coefficients (r) greater than 0.3 are shown.

checkpoint inhibitors therapy (12). Indeed, severe RIL diminishes the survival advantages of durvalumab following concurrent chemoradiotherapy in NSCLC as shown recently (9). Additionally, the Real World Evidence (RWE) studies confirm efficacy of consolidative immune checkpoint inhibitors after CRT in population with more elderly and poor performance status patients and may double overall survival in that population (32, 33). Therefore, developing strategies to minimize RIL could be crucial in improving the effectiveness of comprehensive lung cancer treatment approaches as never before.

Our analysis showed that from non-modifiable factors lymphocyte count at the beginning of treatment (alc_1) was predictive on occurrence of lymphocyte nadir (alc_nadir) further during treatment which stays in line with some of previously published papers (13, 34, 35). Moreover, in our analysis, alc_1 proved to be the main factor, appearing in all top 30 models. Based on our analysis in patients with an $alc_1 > 2.005K/mm^3$, clinically significant lymphopenia should not be anticipated. The value of our analysis is enhanced by excluding patients with lymphopenia (often caused by previous treatment).

In terms of clinical application, controlling the RT dose to critical structures like the lungs and a heart offer an opportunity to modify

risk factors and reduce risk of RIL what is possible with use of modern irradiation techniques(including protons) (8, 36). This adjustment could lead to a reduction in RIL, especially valuable for patients with non-modifiable risk factors for lymphopenia, such as advanced age, lower pre-RT lymphocyte count, and larger tumor size (12). Considering the immune system as an organ-at-risk (OAR) in RT planning is complex, as it is not confined to a specific anatomical area (37). Immune cells circulate throughout the body, often moving in and out of the RT field, which challenges traditional RT planning approaches. From clinical point of view, there are two prevailing theories regarding the negative impact of RT on the count of immune system cells. The first concept emphasizes the significant role of the dose in the circulating pool of immune cells with mostly used edic model in causing RIL, and it's directly related to the dose in critical organs such as the lungs, heart, and median body dose (14–17, 38). The second theory focuses on the impact of the dose in bone marrow on hematologic toxicity (18, 19). These two theories are not mutually exclusive but are rarely accessed together in studies; high doses in the lungs and heart strongly correlate with the dose in the chest bones, as evidenced by our correlation matrix.

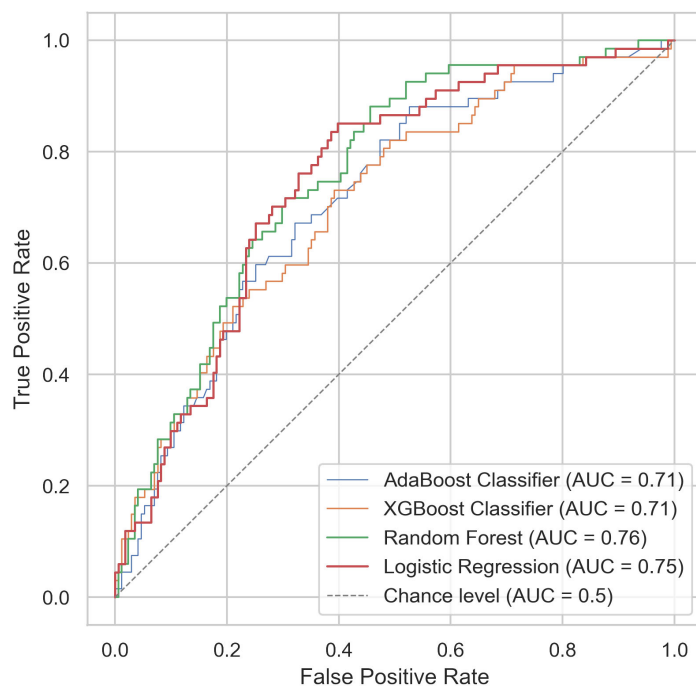


FIGURE 3

Prediction of Common Terminology Criteria for Adverse Events (CTCAE) \geq grade 3 lymphopenia. Receiver Operator Curves for prediction of Common Terminology Criteria for Adverse Events (CTCAE) \geq grade 3 lymphopenia.

Despite testing numerous sets of clinical and diametric features in our cohort of patients, most significant models demonstrated similar performance in prediction of $\text{alc}_{\text{nadir}}$. Reassuringly, in the era of artificial intelligence and advanced statistical methods, the subjective selection of features made by experts (heuristic methods) based on their experience and knowledge exhibited the best fit, as shown in [Supplementary Figure S2](#). However, it also showed high variability. This variability may arise because experts can make mistakes or simply not choose optimal solutions. Although expert knowledge is often undervalued in modern science, our study indicates that combining statistical methods with expert knowledge can yield the best results. Surprisingly, three-five variables models presented best performance in which lung_{mld} , $\text{lung}_{\text{v05p}}$, $\text{lung}_{\text{v10p}}$, $\text{heart}_{\text{v05p}}$ or edic played crucial role in prediction RIL. It's worth to emphasize that low doses variables ($\text{lung}_{\text{v05p}}$, $\text{lung}_{\text{v10p}}$, $\text{heart}_{\text{v05p}}$) shows a strong correlation with lung_{mld} and $\text{heart}_{\text{mld}}$ and edic . Similar results concerning heart and lung dose metrics were observed in study by Abravan et al., where data mining techniques were employed to discern regions where a significant correlation exists between RT dose and $\geq\text{G3}$ lymphopenia (31). Heart, lungs, and thoracic vertebrae were identified as regions linked to RIL, with key diametric parameters being mean doses to the lungs and heart and thoracic vertebrae V20 (31). In other studies, Tang et al. (lung_{v05} , lung_{v10}) and Xie et al. (lung_{v05}) and Kim et al. (lung_{v05}) and Kong et al. (lung_{v05}) observed that low doses in the lung were associated with RIL (39–42). Additionally, studies have demonstrated that low doses to the heart ($\text{heart}_{\text{v05}}$) are crucial in the induction of RIL during stereotactic body radiation therapy (SBRT) and CRT (43, 44). The metrics edic and EDRIC (Estimated Dose of Radiation to the

Immune System) were also predictive of RIL (14, 16). Similarly, Kim et al. noted that the dose to circulating blood cells (as a function of lung and heart dose) was associated with severe RIL during photon and proton radiotherapy (36). Other studies addressing various malignancies in thoracic region have identified relationship between mean and low doses in heart and lung (and low doses) and RIL (11, 26). Our analysis showed that the threshold values of $\text{lung}_{\text{v05p}} \leq 51.8\%$ and lung_{mld} of 10.67Gy can be significant in prediction of $\text{alc}_{\text{nadir}}$ for patients with $\text{alc}_1 \leq 2.005\text{K}/\text{mm}^3$ and $\text{alc}_1 > 2.005\text{K}/\text{mm}^3$ respectively.

In our most effective models, the features previously outlined in literature such as age and tumor volume, treatment time, hypofractionation and pci usage were absent; however, they were incorporated in subsequent (less efficient) models (39, 40, 44). Surprisingly, previously described correlation of low-medium dose metrics of vertebrae (v05-v20) and other of bones with hematologic toxicity have been rarely observed in our best models (18, 19, 31, 45). It is noteworthy that dose metrics for vertebrae and other bones can serve as proxies for lung and heart dose metrics, as demonstrated in our correlation matrix. Furthermore, in some of the aforementioned studies where bones doses were associated with RIL, heart and lung doses were not assessed which complicate comparison (18, 19, 45). Additionally, it may be assumed that since approximately 50% of active bone marrow is in the pelvis, the dose metrics of bone marrow in this region may be critical in the induction of RIL. This correlation has been observed in numerous studies focusing on gynecological, genitourinary, and lower gastrointestinal malignancies (11, 35). Another explanation for the lack of impact of the dose-volume parameters of vertebrae and bones on lymphopenia could be the

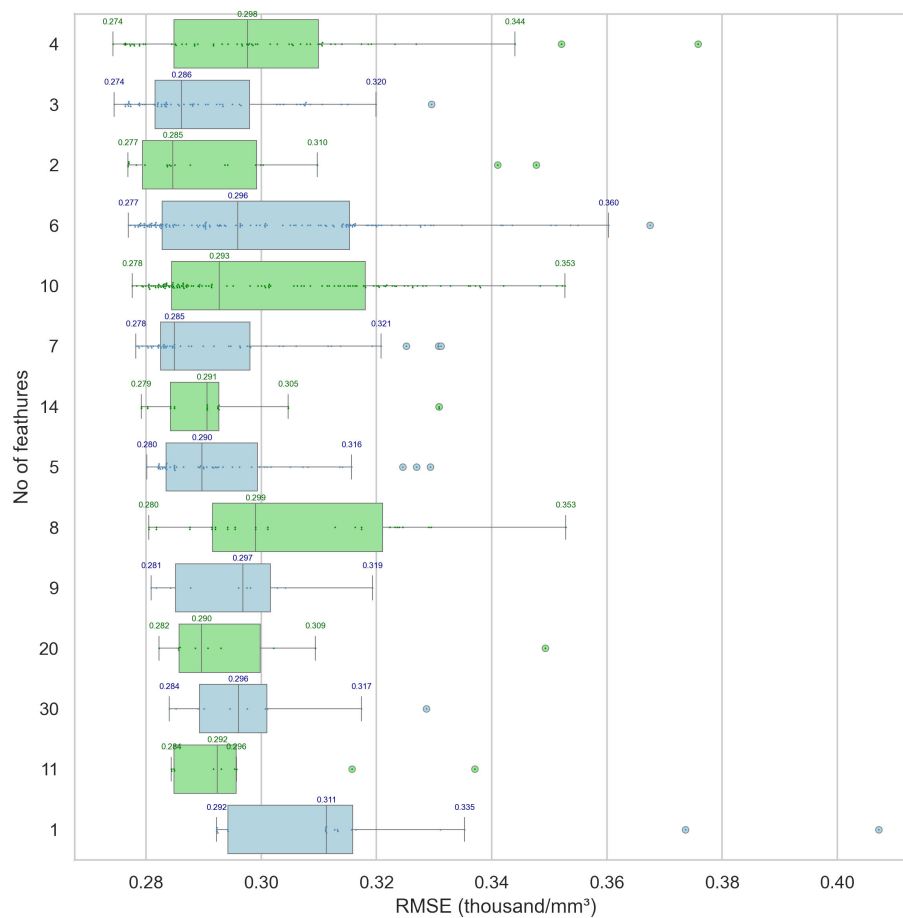


FIGURE 4

The optimal outcome, as indicated by the lowest Root Mean Square Error (RMSE) of individual models depending on the number of selected features. The models with specific number of features are grouped according to increasing RMSE. The dots represent the actual RMSE values of the models. The boxplots of the root mean square error (RMSE) values for individual models are displayed along the X-axis, as a function of the number of selected features, which are indicated on the Y-axis. The plots are organized in ascending order of the minimum RMSE value. Models incorporating fewer features proved to be the most effective; notably, the set containing three variables was an optimal selection across all computational models assessed.

TABLE 2 Best performing models for prediction of absolute lymphocyte count nadir.

Id	Method	Features	No of features	MAE (thousand/mm ³)	RMSE (thousand/mm ³)	Accuracy (%)
1	Random Forest cross validation, 5 fold	alc_1, heart_v05p, lung_mld, lung_v05p	4	0.213	0.274	60.58
2	ARD Regression cross validation, 5 fold	alc_1, lung_mld, lung_v05p	3	0.219	0.276	60.01
3	Bayesian Ridge cross validation, 5 fold	alc_1, lung_mld, lung_v05p, lung_v10p	4	0.219	0.277	60.13
4	Linear Regression cross validation, 5 fold	alc_1, edic	2 (5)	0.218	0.277	60.33
5	AutoML train (75%): cross validation, 5 fold, test (25%)	All	30	0.195 (validation) 0.201 (test score)	0.252 (validation) 0.272 (test score)	70.15 (test score)

Models selected according to Root Mean Squared Error (RMSE) and automatic machine learning method prediction is presented. RMSE, Mean Absolute Error (MAE), Accuracy = 100% - MAPE, where MAPE - Mean Absolute Percentage Error, also known as mean absolute percentage deviation (MAPD) is shown.

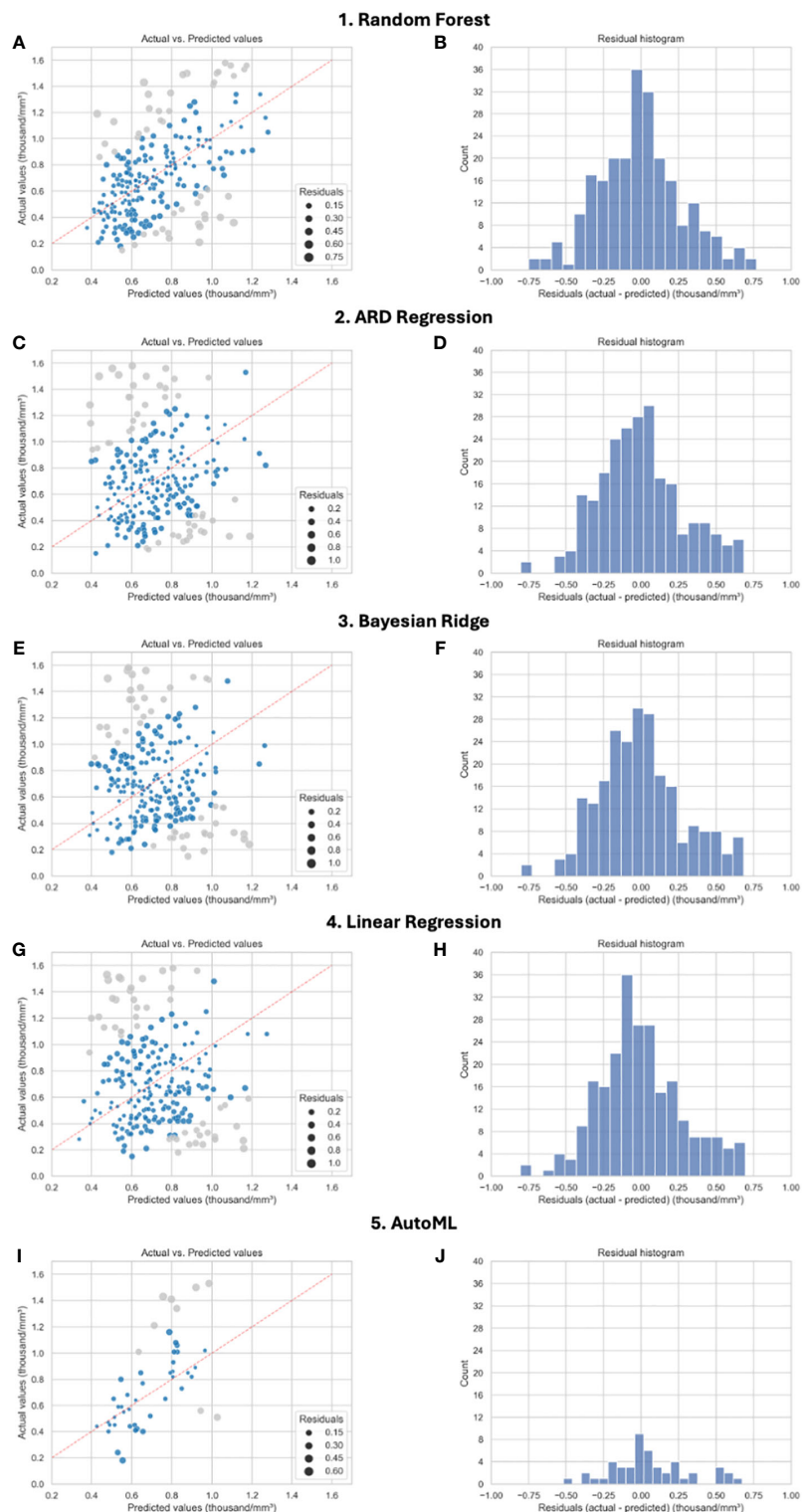
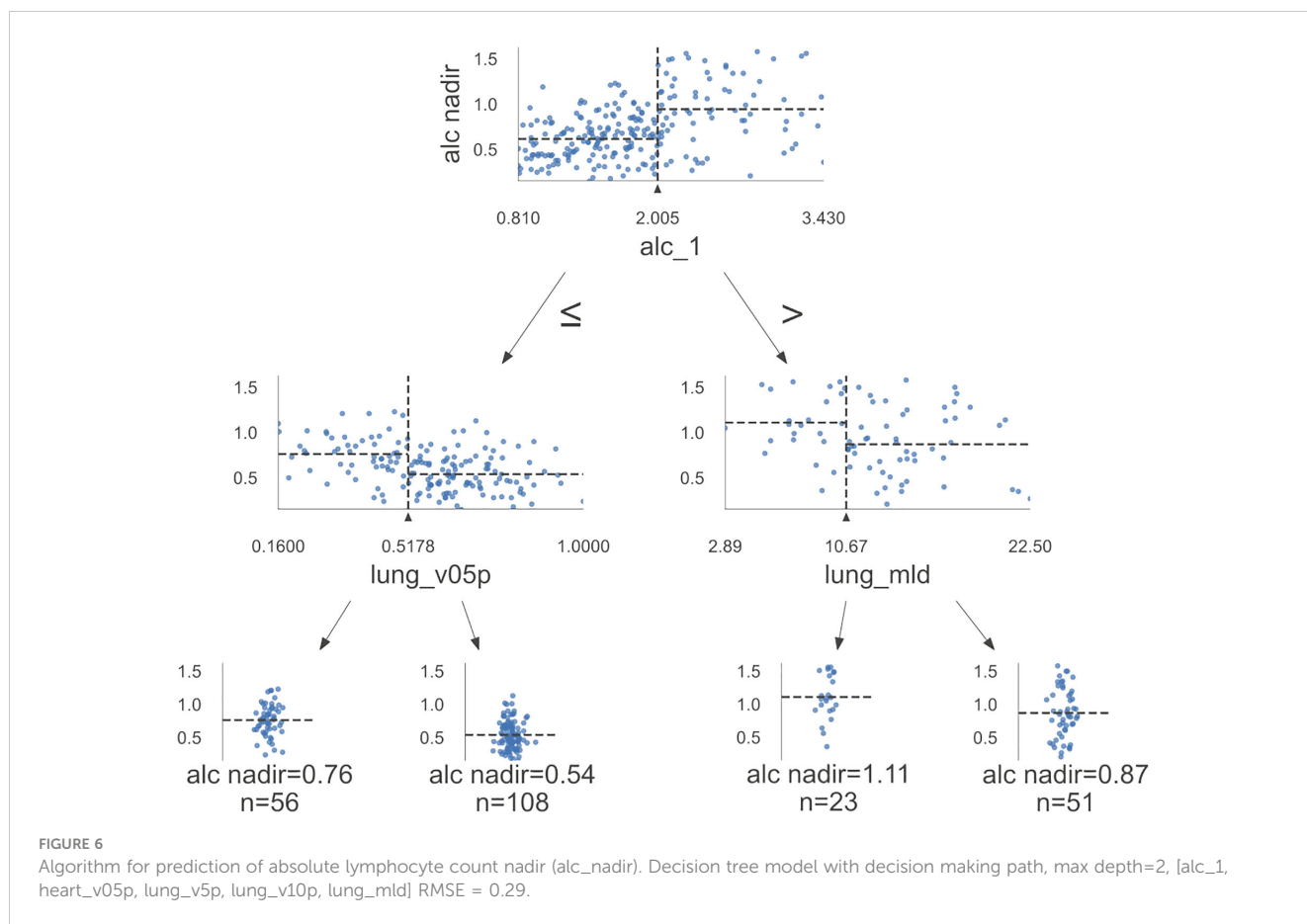


FIGURE 5 (A–J) Model performance in prediction absolute lymphocyte count (alc_nadir). Comparison of most efficient models for prediction alc_nadir. The plots on the left show the actual values vs predicted values given by the models. The darker color applies to the top 80% of results. A perfect regression model would display data points on the diagonal defined by predicted equal to actual values. The size of the points is related to the residuals. The plots on the right show residual histograms. For automatic machine learning (AutoML) the values come only from the test set (25%). The split is as follows: 75% of the set was used for training data, on which 5-fold cross-validation was again applied. This set therefore contains both training and validation data. The results testing was based on the remaining 25% of the data set. For all other models, cross-validation with 5-fold splits was used (all data were used for both training and testing). The final result is based on the average value obtained from the individual sets.



bone marrow regeneration observed during irradiation, which was assessed in patients undergoing chemoradiotherapy of the pelvic area (46).

What is important is that different validation techniques are used in the analyzed statistics methods. Only automated machine learning involves a split into training and test sets. The small number of patients is a limitation, particularly for the interpretability of machine learning models. Training on small datasets can lead to greater variability, making predictions less stable and more sensitive to input data changes. In our study, we used cross-validation techniques to reduce overfitting and enhance result reliability. We used various statistical methods to optimally select features affecting lymphopenia, which had not been done before. The limitation of our study arises from its retrospective nature and the relatively small, yet considerable, number of patients, given the stringent inclusion criteria (no initial lymphopenia, standalone RT treatment, and RT delivered in at least 15 fractions). These criteria resulted in a cohort comprised primarily of fragile patients, which does not fully represent the broader population eligible for concurrent chemo-radiotherapy with consolidative immune checkpoint inhibitors. Our analysis did not summarize late lymphopenia due to the difficulty in its assessment in a retrospective study (irregularity of blood parameter measurements post-chemoradiotherapy), which may constitute a limitation of the study.

The strength of our multicenter study is that we focused specifically on the patients undergoing RT alone, without concurrent systemic treatment and without lymphopenia at the

beginning of treatment. This approach aimed to minimize the influence of chemotherapy on study endpoint as patient who suffered from lymphopenia before RT were excluded. We selected the more sensitive numerical variable, alc_nadir, over ctcae toxicity, and subsequently conducted a comprehensive testing of various statistical models and variable selection strategies, generating 600 predictive models—a scope of analysis that, to the best of our knowledge, has never been done so comprehensively before.

5 Conclusions

The RIL was predominantly triggered by low and median RT doses in the heart and lungs (e.g., lung_mld, lung_v05p, lung_v10p, heart_v05p or edic), with lesser impact from bone marrow-dependent doses. RIL was most severe in patients who had lower initial lymphocyte counts. For patients with alc_1 values less or equal than 2.005/mm³ and lung_v05p ≤ 51.8%, a median alc_nadir of 0.76K/mm³ was estimated, whereas for those with lung_v05p > 51.8%, a median alc_nadir of 0.54K/mm³ was estimated.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession

number(s) can be found below: <https://1drv.ms/f/s!AqWHyMZlPESTg51KmmQc6Vo0OppBmQ?e=dBmAc5>.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

LK: Conceptualization, Formal Analysis, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. MP: Investigation, Resources, Writing – review & editing. KS: Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – review & editing. MB: Investigation, Resources, Writing – review & editing. JS: Investigation, Resources, Writing – review & editing. RS: Investigation, Resources, Writing – review & editing. MP-P: Investigation, Resources, Writing – review & editing. KK: Investigation, Resources, Writing – review & editing. BJ-F: Supervision, Writing – review & editing. JF: Supervision, Writing – review & editing.

References

- Russano M, La Cava G, Cortellini A, Citarella F, Galletti A, Di Fazio GR, et al. Immunotherapy for metastatic non-small cell lung cancer: Therapeutic Advances and Biomarkers. *Curr Oncol.* (2023) 30:2366–87. doi: 10.3390/curroncol30020181
- Leiter A, Veluswamy RR, Wisnivesky JP. The global burden of lung cancer: current status and future trends. *Nat Rev Clin Oncol.* (2023) 20:624–39. doi: 10.1038/s41571-023-00798-3
- Rossi S, Pagliaro A, Michelini A, Navarra P, Clerici E, Franceschini D, et al. The era of immunotherapy in small-cell lung cancer: More shadows than light? *Cancers.* (2023) 15:5761. doi: 10.3390/cancers15245761
- Spigel DR, Faivre-Finn C, Gray JE, Vicente D, Planchard D, Paz-Ares L, et al. Five-year survival outcomes from the PACIFIC trial: Durvalumab after chemoradiotherapy in stage III non-small-cell lung cancer. *J Clin Oncol Off J Am Soc Clin Oncol.* (2022) 40:1301–11. doi: 10.1200/JCO.21.01308
- Senan S, Okamoto I, Lee GW, Chen Y, Niho S, Mak G, et al. Design and rationale for a phase III, randomized, placebo-controlled trial of durvalumab with or without tremelimumab after concurrent chemoradiotherapy for patients with limited-stage small-cell lung cancer: The ADRIATIC study. *Clin Lung Cancer.* (2020) 21:e84–8. doi: 10.1016/j.clcc.2019.12.006
- Riely GJ, Wood DE, Ettinger DS, Aisner DL, Akerley W, Bauman JR, et al. Non-Small Cell Lung Cancer, Version 4.2024, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw.* (2024) 22:249–74. doi: 10.6004/jnccn.2204.0023
- Girard N, Bar J, Garrido P, Garassino MC, McDonald F, Mornex F, et al. Treatment characteristics and real-world progression-free survival in patients with unresectable stage III NSCLC who received durvalumab after chemoradiotherapy: findings from the PACIFIC-R study. *J Thorac Oncol.* (2023) 18:181–93. doi: 10.1016/j.jtho.2022.10.003
- Upadhyay R, Venkatesulu BP, Giridhar P, Kim BK, Sharma A, Elghazawy H, et al. Risk and impact of radiation related lymphopenia in lung cancer: A systematic review and meta-analysis. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2021) 157:225–33. doi: 10.1016/j.radonc.2021.01.034
- Jing W, Xu T, Wu L, Lopez PB, Grassberger C, Ellsworth SG, et al. Severe radiation-induced lymphopenia attenuates the benefit of durvalumab after concurrent chemoradiotherapy for NSCLC. *JTO Clin Res Rep.* (2022) 3:100391. doi: 10.1016/j.jtocrr.2022.100391

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1426635/full#supplementary-material>

- Friedes C, Chakrabarti T, Olson S, Prichett L, Brahmer JR, Forde PM, et al. Association of severe lymphopenia and disease progression in unresectable locally advanced non-small cell lung cancer treated with definitive chemoradiation and immunotherapy. *Lung Cancer Amst Neth.* (2021) 154:36–43. doi: 10.1016/j.lungcan.2021.01.022
- Venkatesulu B, Giridhar P, Pujari L, Chou B, Lee JH, Block AM, et al. Lymphocyte sparing normal tissue effects in the clinic (LymphoTEC): A systematic review of dose constraint considerations to mitigate radiation-related lymphopenia in the era of immunotherapy. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2022) 177:81–94. doi: 10.1016/j.radonc.2022.10.019
- Zhang Y, Huang C, Li S. Influence of treatment-related lymphopenia on the efficacy of immune checkpoint inhibitors in lung cancer: a meta-analysis. *Front Oncol.* (2023) 13:1287555. doi: 10.3389/fonc.2023.1287555
- Cella L, Monti S, Pacelli R, Palma G. Modeling frameworks for radiation induced lymphopenia: A critical review. *Radiother Oncol.* (2024) 190:110041. doi: 10.1016/j.radonc.2023.110041
- Ladbury CJ, Rusthoven CG, Camidge DR, Kavanagh BD, Nath SK. Impact of radiation dose to the host immune system on tumor control and survival for stage III non-small cell lung cancer treated with definitive radiation therapy. *Int J Radiat Oncol Biol Phys.* (2019) 105:346–55. doi: 10.1016/j.ijrobp.2019.05.064
- Qiu J, Lin H, Ke D, Yu Y, Xu J, Qiu H, et al. Higher radiation dose on immune cells is associated with radiation-induced lymphopenia and worse prognosis in patients with locally advanced esophageal squamous cell carcinoma. *Front Immunol.* (2023) 14:1066255. doi: 10.3389/fimmu.2023.1066255
- Jin JY, Hu C, Xiao Y, Zhang H, Paulus R, Ellsworth SG, et al. Higher radiation dose to the immune cells correlates with worse tumor control and overall survival in patients with stage III NSCLC: A secondary analysis of RTOG0617. *Cancers.* (2021) 13:6193. doi: 10.3390/cancers13246193
- Xu C, Jin JY, Zhang M, Liu A, Wang J, Mohan R, et al. The impact of the effective dose to immune cells on lymphopenia and survival of esophageal cancer after chemoradiotherapy. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2020) 146:180–6. doi: 10.1016/j.radonc.2020.02.015
- Fabian D, Ayan A, DiCostanzo D, Barney CL, Aljabban J, Diaz DA, et al. Increasing radiation dose to the thoracic marrow is associated with acute hematologic

- toxicities in patients receiving chemoradiation for esophageal cancer. *Front Oncol.* (2019) 9:147. doi: 10.3389/fonc.2019.00147
19. Barney CL, Scoville N, Allan E, Ayan A, DiCostanzo D, Haglund KE, et al. Radiation dose to the thoracic vertebral bodies is associated with acute hematologic toxicities in patients receiving concurrent chemoradiation for lung cancer: results of a single-center retrospective analysis. *Int J Radiat Oncol Biol Phys.* (2018) 100:748–55. doi: 10.1016/j.ijrobp.2017.11.025
20. Sung W, Cho B. Modeling of radiation effects to immune system: a review. *J Korean Phys Soc.* (2022) 81:1013–9. doi: 10.1007/s40042-022-00574-z
21. Sung W, Grassberger C, McNamara AL, Basler L, Ehrbar S, Tanadini-Lang S, et al. A tumor-immune interaction model for hepatocellular carcinoma based on measured lymphocyte counts in patients undergoing radiotherapy. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2020) 151:73–81. doi: 10.1016/j.radonc.2020.07.025
22. Jin JY, Mereniuk T, Yalamanchali A, Wang W, Machtay M, Spring Kong FM, et al. A framework for modeling radiation induced lymphopenia in radiotherapy. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2020) 144:105–13. doi: 10.1016/j.radonc.2019.11.014
23. Seth S, Principe JC. A conditional independence perspective of variable selection. In: *2010 IEEE International Workshop on Machine Learning for Signal Processing*. IEEE, Kittila, Finland (2010). p. 456–61. Available at: <http://ieeexplore.ieee.org/document/5588682/>.
24. Zhao J, Wang GY, Wu ZF, Tang H, Li H. (2002). The study on technologies for feature selection, in: *Proceedings International Conference on Machine Learning and Cybernetics*. Beijing, China: IEEE. pp. 689–93. Available at: <https://ieeexplore.ieee.org/document/1174435>.
25. Chen CW, Tsai YH, Chang FR, Lin WC. Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results. *Expert Syst.* (2020) 37:e12553. doi: 10.1111/exsy.12553
26. Terrones-Campos C, Ledergerber B, Forbes N, Smith AG, Petersen J, Helleberg M, et al. Prediction of radiation-induced lymphopenia following exposure of the thoracic region and associated risk of infections and mortality. *Clin Oncol.* (2023) 35:e434–44. doi: 10.1016/j.clon.2023.04.003
27. Zhu C, Lin SH, Jiang X, Xiang Y, Belal Z, Jun G, et al. A novel deep learning model using dosimetric and clinical information for grade 4 radiotherapy-induced lymphopenia prediction. *Phys Med Biol.* (2020) 65:035014. doi: 10.1088/1361-6560/ab63b6
28. Kim Y, Chamseddine I, Cho Y, Kim JS, Mohan R, Shusharina N, et al. Neural network based ensemble model to predict radiation induced lymphopenia after concurrent chemo-radiotherapy for non-small cell lung cancer from two institutions. *Neoplasia N Y N.* (2023) 39:100889. doi: 10.1016/j.neo.2023.100889
29. Ebrahimi S, Lim G, Hobbs BP, Lin SH, Mohan R, Cao W. A hybrid deep learning model for forecasting lymphocyte depletion during radiation therapy. *Med Phys.* (2022) 49:3507–22. doi: 10.1002/mp.15584
30. Venkatesulu BP, Mallick S, Lin SH, Krishnan S. A systematic review of the influence of radiation-induced lymphopenia on survival outcomes in solid tumors. *Crit Rev Oncol Hematol.* (2018) 123:42–51. doi: 10.1016/j.critrevonc.2018.01.003
31. Abravan A, Faivre-Finn C, Kennedy J, McWilliam A, van Herk M. Radiotherapy-related lymphopenia affects overall survival in patients with lung cancer. *J Thorac Oncol.* (2020) 15:1624–35. doi: 10.1016/j.jtho.2020.06.008
32. Wang Y, Zhang T, Huang Y, Li W, Zhao J, Yang Y, et al. Real-world safety and efficacy of consolidation durvalumab after chemoradiation therapy for stage III non-small cell lung cancer: A systematic review and meta-analysis. *Int J Radiat Oncol Biol Phys.* (2022) 112:1154–64. doi: 10.1016/j.ijrobp.2021.12.150
33. Sankar K, Bryant AK, Green M, Ramnath N. Doubling of median overall survival in a nationwide cohort of veterans with stage III non-small cell lung cancer in the durvalumab era. *J Clin Oncol.* (2021) 39:8546–6. doi: 10.1200/JCO.2021.39.15_suppl.8546
34. Xu Z, Yang L, Yu H, Guo L. A machine learning model for grade 4 lymphopenia prediction during pelvic radiotherapy in patients with cervical cancer. *Front Oncol.* (2022) 12:905222. doi: 10.3389/fonc.2022.905222
35. Kuncman Ł, Stawiski K, Masłowski M, Kucharz J, Fijuth J. Dose–volume parameters of MRI-based active bone marrow predict hematologic toxicity of chemoradiotherapy for rectal cancer. *Strahlenther Onkol.* (2020) 196:998–1005. doi: 10.1007/s00066-020-01659-z
36. Kim N, Shin J, Ahn SH, Pyo H, Noh JM, Yang K, et al. Reduced radiation exposure to circulating blood cells in proton therapy compared with X-ray therapy in locally advanced lung cancer: Computational simulation based on circulating blood cells. *Front Oncol.* (2023) 13:1119173. doi: 10.3389/fonc.2023.1119173
37. Qian JM, Akama-Garren E, Shin J, Gunasti L, Bang A, Pike LRG, et al. Dosimetric modeling of lymphopenia in patients with metastatic cancer receiving palliative radiation and PD-1 immune checkpoint inhibitors. *Adv Radiat Oncol.* (2021) 7:100880. doi: 10.1016/j.adro.2021.100880
38. Chen F, Jin JY, Hui TSK, Jing H, Zhang H, Nong Y, et al. Radiation induced lymphopenia is associated with the effective dose to the circulating immune cells in breast cancer. *Front Oncol.* (2022) 12:768956. doi: 10.3389/fonc.2022.768956
39. Tang C, Liao Z, Gomez D, Levy L, Zhuang Y, Gebremichael RA, et al. Lymphopenia association with gross tumor volume and lung V5 and its effects on non-small cell lung cancer patient outcomes. *Int J Radiat Oncol • Biol • Phys.* (2014) 89:1084–91. doi: 10.1016/j.ijrobp.2014.04.025
40. Xie X, Lin SH, Welsh JW, Wei X, Jin H, Mohan R, et al. Radiation-induced lymphopenia during chemoradiation therapy for non-small cell lung cancer is linked with age, lung V5, and XRCC1 rs25487 genotypes in lymphocytes. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2021) 154:187–93. doi: 10.1016/j.radonc.2020.09.002
41. Kim N, Myoung Noh J, Lee W, Park B, Park H, Young Park J, et al. Proton beam therapy reduces the risk of severe radiation-induced lymphopenia during chemoradiotherapy for locally advanced non-small cell lung cancer: A comparative analysis of proton versus photon therapy. *Radiother Oncol J Eur Soc Ther Radiol Oncol.* (2021) 156:166–73. doi: 10.1016/j.radonc.2020.12.019
42. Kang BH, Li X, Son J, Song C, Kang HC, Kim HJ, et al. Prediction and clinical impact of delayed lymphopenia after chemoradiotherapy in locally advanced non-small cell lung cancer. *Front Oncol.* (2022) 12:891221. doi: 10.3389/fonc.2022.891221
43. Zhao Q, Li T, Chen G, Zeng Z, He J. Prognosis and risk factors of radiation-induced lymphopenia in early-stage lung cancer treated with stereotactic body radiation therapy. *Front Oncol.* (2020) 9:1488. doi: 10.3389/fonc.2019.01488
44. Zhao Q, Chen G, Ye L, Shi S, Du S, Zeng Z, et al. Treatment-duration is related to changes in peripheral lymphocyte counts during definitive radiotherapy for unresectable stage III NSCLC. *Radiat Oncol.* (2019) 14:86. doi: 10.1186/s13014-019-1287-z
45. Deek MP, Benenati B, Kim S, Chen T, Ahmed I, Zou W, et al. Thoracic vertebral body irradiation contributes to acute hematologic toxicity during chemoradiation therapy for non-small cell lung cancer. *Int J Radiat Oncol Biol Phys.* (2016) 94:147–54. doi: 10.1016/j.ijrobp.2015.09.022
46. Kuncman Ł, Orzechowska M, Stawiski K, Masłowski M, Ciężńska M, Gottwald L, et al. The kinetics of FMS-related tyrosine kinase 3 ligand (Flt-3L) during chemoradiotherapy suggests a potential gain from the earlier initiation of immunotherapy. *Cancers.* (2022) 14:3844. doi: 10.3390/cancers14163844