Development and validation of a random forest model for predicting radiation pneumonitis in lung cancer patients receiving moderately hypofractionated radiotherapy: a retrospective cohort study

Yang Zhang[1#^], Zongjuan Li[2#], Han Xiao[1], Zhenjiang Li[3], Jian He[1], Shisuo Du[1], Zhaochong Zeng[1]

[1]Department of Radiation Oncology, Zhongshan Hospital, Fudan University, Shanghai, China; [2]Department of Medical Oncology, Shanghai Pulmonary Hospital and Thoracic Cancer Institute, Tongji University School of Medicine, Shanghai, China; [3]Department of Radiation Physics, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China

Contributions: (I) Conception and design: Y Zhang, Zongjuan Li, Z Zeng, S Du; (II) Administrative support: S Du; (III) Provision of study materials or patients: Zongjuan Li, H Xiao; (IV) Collection and assembly of data: Y Zhang, Zhenjiang Li; (V) Data analysis and interpretation: Y Zhang, Zongjuan Li, J He; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.
#These authors contributed equally to this work.

Correspondence to: Shisuo Du, PhD; Zhaochong Zeng, PhD. Department of Radiation Oncology, Zhongshan Hospital, Fudan University, 180 Fenglin Road, Shanghai 200030, China. Email: du.shisuo@zs-hospital.sh.cn; zeng.zhaochong@zs-hospital.sh.cn.

Background: Radiation pneumonitis (RP) is a type of toxicity commonly associated with thoracic radiation therapy. We sought to establish a random forest (RF) model and evaluate its ability to predict RP in patients with non-small cell lung cancer (NSCLC) receiving moderately hypofractionated radiotherapy (hypo-RT).

Methods: A total of 106 patients with stage II–IVa NSCLC who received moderately hypofractionated helical tomotherapy (2.3–3.0 Gy/fraction) at Zhongshan Hospital were included. All enrolled patients were divided chronologically into the training (67 patients) and validation (39 patients) groups. Higher than or equal to grade 2 RP was defined as the end point. Logistic regression and RF models were established and compared using the receiver operating characteristic (ROC) and a confusion matrix in the training and validation groups.

Results: The cumulative incidence of the end point was 25.4% and 17.9% in the training and validation groups, respectively. Logistic regression models were constructed by dosage parameters of total lungs, ipsilateral or contralateral lungs, respectively. ROC analysis revealed that the dosimetric factors of total lungs yielded a superior classification performance than did that of the ipsilateral or contralateral lungs [area under the curve (AUC) =0.920, AUC =0.701, and AUC =0.661, respectively]. Furthermore, the RF model yielded a better prediction capacity than did the traditional logistic model based on the dosimetric factors of the total lungs (accuracy: 88.06%; precision: 84.62%; sensitivity: 64.71%; specificity: 96.00%). Moreover, the RF identified mean lung dose [MLD; mean decrease gini (MDG) =5.74], V20 (MDG =4.62), and V35 (MDG =3.08) of total lungs as the most common primary differentiators of RP.

Conclusions: Our RF model established based on the dosimetric parameters of the total lungs could accurately predict the RP risk in patients with NSCLC treated with moderately hypofractionated tomotherapy.

Keywords: Non-small cell lung cancer (NSCLC); hypofractionated radiotherapy; random forest (RF); radiation pneumonitis (RP)

Introduction

Radiotherapy remains the mainstay of treatment for locally advanced unresectable non-small cell lung cancer (NSCLC). Theoretically, tumor control is dependent on the dose of radiation delivered. However, the classic lung cancer study by the Radiation Therapy Oncology Group (RTOG) 0617 revealed that 74 Gy radiation given in 2 Gy fractions was not superior to 60 Gy for patients with stage III NSCLC (1). The unfavorable results may be attributed to the repopulation of tumor cells due to a prolonged treatment time and the excess toxicity associated with conventional radiation techniques (2).

Moderately hypofractionated radiotherapy (hypo-RT), which is the delivery of RT with a dose per fraction ranging from 2.3 to 5.5 Gy, is an alternative means to achieving a higher biological equivalent dose within a short treatment time (3). It is currently recognized as a potentially powerful treatment for reducing the effect of the accelerated repopulation of tumor cells (4). Moreover, with the advent of image-guided radiation therapy (IGRT), RT has entered a new era of precision. Our previous studies demonstrated that moderately hypo-RT delivered by IGRT provided a superior prognosis and tolerable toxicity than did conventional RT in Chinese patients with locally advanced NSCLC (5,6). In line with our results, moderately hypo-RT is also being increasingly used to treat unresectable NSCLC worldwide (7,8), particularly in the United Kingdom (9).

However, concerns prevail regarding the severe side effects associated with this treatment modality, such as radiation pneumonitis (RP). Over the years, parameters predicting RP following thoracic RT with conventional dose fractionation (1.8–2.2 Gy/fraction) or stereotactic body radiation therapy (>5.5 Gy/fraction) have been extensively analyzed (10-16). However, the findings regarding these variables cannot be directly extrapolated to moderately hypo-RT due to the different dose fractions and the potential difference in radiobiological response. Currently, there is insufficient data concerning moderately hypo-RT. Thus, the present study was designed to examine the incidence of RP and to establish and validate a prediction model for RP in patients with unresectable NSCLC treated with moderately hypo-RT. We present the following article in accordance with the TRIPOD reporting checklist (available at https://atm.amegroups.com/article/view/10.21037/atm-22-3049/rc).

Methods

Patients

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institutional Review Board of Zhongshan Hospital, Fudan University (No. B2021-136), and informed consent was obtained from patients before enrollment. We retrospectively collected information on all patients with unresectable NSCLC who received moderately hypofractionated helical tomotherapy at Zhongshan Hospital, Fudan University, between August 2011 and June 2020. Patient inclusion criteria were the following: (I) newly histologically diagnosed unresectable NSCLC; (II) undergone treatment with definitive hypofractionated helical tomotherapy (54–70.4 Gy, 2.3–3.0 Gy/fraction); (III) stage II–IVa [according to the eighth edition of the American Joint Committee on Cancer (AJCC) staging system]; and (IV) available clinical and dosimetric data. Patients were excluded if they (I) had undergone thoracic RT or surgery previously; (II) had received consolidation immune therapy; or (III) were lost to follow-up. All enrolled patients were divided chronologically into the training (67 patients) and validation (39 patients) groups.

Interstitial lung abnormality score (ILAS)

ILAS was determined according to a previously published paper (17). ILASs were rated as follows: 0, none; 1, abnormality without honeycombing (ground-glass attenuation, fine reticular opacity, and microcysts); and 2, honeycombing.

RT and dosimetric parameters

Definitive moderately hypo-RT was delivered by helical tomotherapy with a total dose of 54 to 70.4 Gy (2.3 to 3.0 Gy/fraction) 5 times per week. The gross tumor volume

(GTV) included all detectable carcinomas and lymph nodes with a short-axis diameter >1 cm as shown by computed tomography (CT) or positive features demonstrated by positron emission tomography-CT scans. The clinical target volume (CTV) of the primary tumor was determined using an isotropic 6 to 8 mm margin around the GTV. The CTV of the lymph nodes comprised the entire involved lymph node region detected by CT or positron emission tomography-CT. The planning target volume was delineated using a margin of 0.5–1.0 cm around the CTV to compensate for the breathing movement and setup errors. The objective of all RT plans was to deliver the prescription dose to ≥95% of the planning target volume. The institutional dosimetric constraints for normal organs at risk (lungs, heart, esophagus, spinal cord, etc.) have been described in detail previously (6).

The normal lung volume in this study was defined as the volume of the lungs minus the GTV, with Vx denoting the percentage of normal lungs receiving a dose of ≥ x Gy. The following dosimetric parameters of the ipsilateral, contralateral, and total lungs were generated: mean lung dose (MLD), V5, V10, V15, V20, V25, V30, V35, V40, V45, and V50. The dosimetric parameters of normal lungs were not converted into an equivalent biological dose of 2 Gy/fraction in the present study. Dosimetric and clinical parameters were collected in the absence of information on RP.

Follow-up and assessment of RP

All patients included in the study were examined 1 month after the initial therapy and followed up every 3 months, provided that the patient's condition remained stable. The end point of this study was symptomatic RP (grade ≥2), defined in accordance with the Common Terminology Criteria for Adverse Events 4.03. The diagnosis of symptomatic RP was reached through a consensus of at least 2 radiation oncologists blinded to other clinical and dosimetric parameters. All patients diagnosed with RP underwent a CT scan of their chest.

Statistical analysis

Statistical analysis in this study was performed using R v. 4.1.3 (The R Foundation for Statistical Computing). The Pearson chi-squared or Fisher exact test was conducted to compare categorical variables between the 2 groups. Continuous variables are presented as mean ± standard deviation (normally distributed) or median (interquartile range; nonparametric). Continuous variables were analyzed with an independent t-test or a nonparametric test. Significant parameters identified during univariate analysis (P<0.05) were subjected to multivariate analysis. Both binary logistic regression and random forest (RF) models for the prediction of RP were established. The discriminative capacity of the prediction models was evaluated by comparing the receiver operating characteristic (ROC) or the confusion matrix results. All statistical tests were 2-tailed.

Results

Baseline characteristics of patients in the training and validation groups

From August 2011 to June 2020, 121 patients with unresectable NSCLC received moderately hypo-RT with definitive intention in Zhongshan Hospital, Fudan University. Patients were excluded due to being lost to follow-up (n=6) or for having incomplete data on the RT plan (n=9). A total of 106 cases were finally enrolled and divided chronologically into the training (67 patients) and validation (39 patients) groups. A total of 17 (25.40%) and 7 (17.90%) patients in the training and validation groups, respectively, developed RP with a grade higher than or equal to 2. The median occurrence time of RP calculated from the end of RT was 3.0 months (range, 0.4–10.9 months). Detailed information about patients in the 2 groups is shown in *Table 1*.

Univariate analysis of clinical and dosimetric parameters in the training group

The univariate analysis showed that none of the clinical variables were significantly associated with symptomatic RP (Table S1). The following dosimetric parameters were closely related to the development of symptomatic RP: V5 of the contralateral lung; MLD, V5, V10, V15, V25, V30, and V40 of the ipsilateral lung; and MLD, V5, V10, V20, V25, V30, V35, and V40 of the total lungs (Table S2).

Evaluation of multivariate logistic RP models established based on dosimetric parameters from ipsilateral, contralateral, and total lungs

Three binary logistic models were established based on the significant dosimetric parameters in the univariate analysis of the ipsilateral, contralateral, and total lungs. The

Table 1 Clinical characteristics of the training and validation groups

Characteristics	Training group (n, %)	Validation group (n, %)
n	67	39
Age (years, mean ± SD)	63.06±10.60	66.56±12.70
Gender		
Female	6 (9.00)	6 (15.40)
Male	61 (91.00)	33 (84.60)
Smoke		
No	37 (55.20)	21 (53.80)
Yes	30 (44.80)	18 (46.20)
Emphysema		
No	41 (61.20)	16 (41.00)
Yes	26 (38.80)	23 (59.00)
ILAS		
0	50 (74.63)	33 (84.62)
1	10 (14.93)	6 (15.38)
2	7 (10.45)	0 (0.00)
Diabetes		
No	61 (91.00)	38 (97.40)
Yes	6 (9.00)	1 (2.60)
Pathology		
Adenocarcinoma	26 (38.80)	16 (41.00)
Squamous	36 (53.70)	23 (59.00)
Others	5 (7.50)	0 (0.00)
T		
1–2	38 (56.70)	21 (53.80)
3–4	29 (43.30)	18 (46.20)
N		
0–1	15 (22.40)	15 (38.50)
2–3	52 (77.60)	24 (61.50)
Stage		
II–III	58 (86.60)	37 (94.90)
IV	9 (13.40)	2 (5.10)
Modality		
conCRT	18 (26.90)	9 (23.10)
RT alone	9 (13.40)	13 (33.30)
seqCRT	40 (59.70)	17 (43.60)
GTV (cm³), median [IQR]	75.00 [37.50, 102.80]	54.00 [27.00, 146.50]
RP	17 (25.40)	7 (17.90)

ILAS, interstitial lung abnormality score; conCRT, concurrent chemoradiotherapy; RT, radiotherapy; seqCRT, sequential chemoradiotherapy; GTV, gross tumor volume; IQR, interquartile range; RP, radiation pneumonitis.

efficacies of these models were compared by ROC analysis. *Figure 1* shows that the area under the curve (AUC) for the models of the ipsilateral, contralateral, and total lungs was 0.701, 0.661, and 0.920, respectively. Interestingly, the logistic model constructed by the total lung radiation dosage possessed superior classification performance. MLD [odds ratio (OR) =2.176, 95% CI: 1.306–3.625; P=0.003] and V35 (OR =1.635, 95% CI: 1.057–2.528; P=0.027) of the total lungs were found to be significant factors impacting RP occurrence (*Table 2*). The risk of RP of a grade greater than or equal to 2 could be predicted by the following formula: $Y = \dfrac{e^x}{1+e^x}$, where x = − 15.056 + 0.778 × MLD (Gy) + 0.492 × V35 (%).

RF model for RP prediction

The discriminative performance of the RF model achieved the highest capacity at ntree =2,000 with an out-of-bag (OOB) value of 14.93% (*Figure 2A*). The variables affecting RP were ranked using the permutation feature importance method: MLD of total lungs (TMLD) [mean decrease gini (MDG): 5.74], TV20 (MDG: 4.62), TV35 (MDG: 3.08), TV5 (MDG: 2.04), TV25 (MDG: 1.36), GTV (MDG: 1.35), TV15 (MDG: 1.33), TV10 (MDG: 1.32), TV30 (MDG: 1.17), TV45 (MDG: 1.09), TV40 (MDG: 0.99), and TV50 (MDG: 0.84; *Figure 2B*).

Efficacy contrast of logistic and RF models in the training and validation groups

In the tests conducted with these models, the model success rates were determined by confusion matrix metrics based on the accuracy, precision, sensitivity, and specificity. *Figure 3* and *Table 3* display the confusion matrices for the logistic and RF classifiers. It is worth noting that the RF model performed better did than logistic model in the training and validation groups. In the training group, the accuracy, precision, sensitivity, and specificity of the RF model were 88.06%, 84.62%, 64.71%, and 96.00%, respectively, which were superior to those of the logistic regression model (accuracy: 86.57%; precision: 78.57%; sensitivity: 64.71%; specificity: 94.00%). As shown in *Table 3*, similar results were observed in the validation group.

Discussion
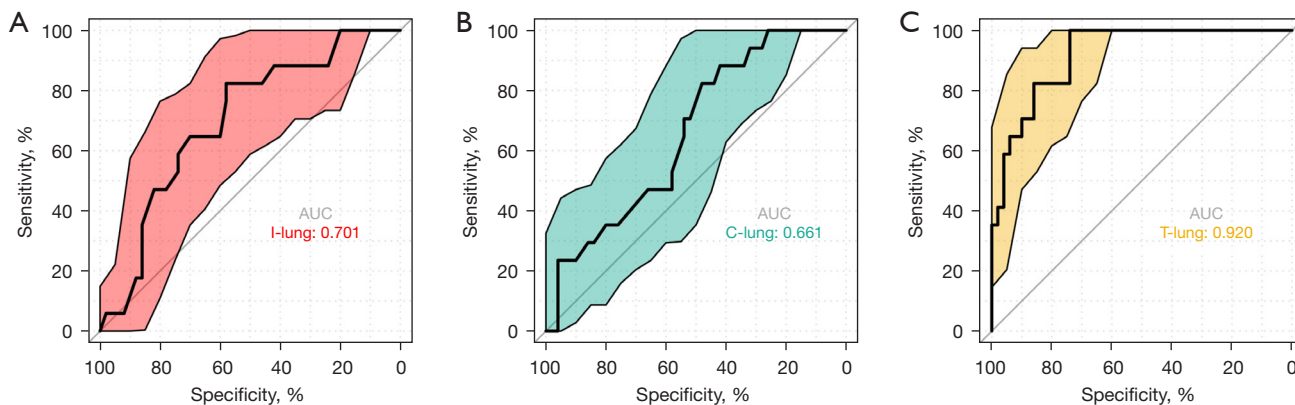
The hypofractionated dose escalation associated with

Figure 1 The AUC of the logistic models established based on dosimetric parameters from ipsilateral (A), contralateral (B), or total lungs (C) were 0.701, 0.661, and 0.920, respectively, which indicated that dose constraints of total lungs played a prominent role. AUC, area under the curve; I, ipsilateral; C, contralateral; T, total.
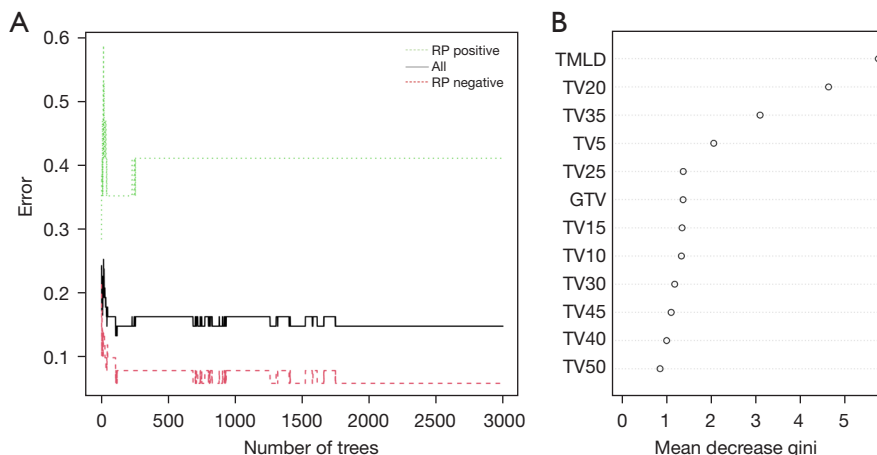

Figure 2 RF model for RP. (A) The change in the error rates of the RF model according to the number of trees. (B) Permutation feature importance plot. RF, random forest; RP, radiation pneumonitis; T, total lungs; MLD, mean lung dose; Vx, the percentage of normal lungs receiving a dose of $\geq$ x Gy; GTV, gross tumor volume.

Table 2 Multivariate analysis of the dosimetric parameters of total lungs

Characteristics	B	Exp(B)	Exp(B) 95% CI	P
TMLD	0.778	2.176	1.306–3.625	0.003
TV35	0.492	1.635	1.057–2.528	0.027
Constant	−15.056			

TMLD, mean lung dose of total lungs; TV35, V35 of total lungs.

fewer RT fractions represents a good strategy to achieving better tumor control. However, for a constant tumor control probability, hypo-fractionation is expected to be

more toxic than standard or hyper-fractionation under the linear-quadratic model when the traditional values of α/β are applied and repopulation effects are neglected (18). Increased lung toxicity was observed in a previous hypofractionated dose escalation trial using 3-dimensional conformal RT (19). It has been established that hypo-fractionation increases the accuracy requirements for radiation delivery (20). Compared with conventional plans, tomotherapy can potentially decrease the RT dose delivered to the surrounding normal organs, including the esophagus, lungs, and heart (21,22). Tomotherapy involves an integrated megavoltage CT for the application of IGRT.
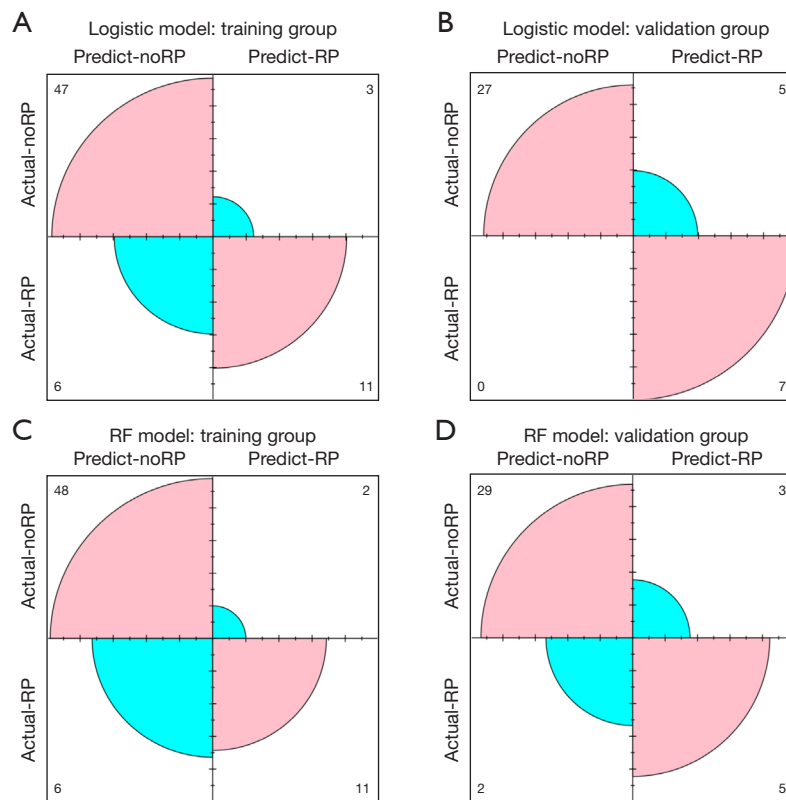
Page 6 of 9

Zhang et al. Random forest model for RP in NSCLC receiving moderately hypo-RT

Figure 3 Classification results are expressed as confusion matrices in the logistic regression and random forest models in both the training and validation groups. (A) The logistic model of RP for the training group. (B) The logistic model of RP for the validation group. (C) The RF model of RP for the training group. (D) The RF model of RP for the validation group. RF, random forest; RP, radiation pneumonitis.

Table 3 Performance of logistic and RF models in the training and validation groups

Parameters	Training group (%)		Validation group (%)	
	Logistic model	RF model	Logistic model	RF model
Accuracy	86.57	88.06	87.18	87.18
Precision	78.57	84.62	58.33	62.50
Sensitivity	64.71	64.71	100	71.43
Specificity	94.00	96.00	84.38	90.63

RF, random forest.

Treatment with IGRT has been established to result in a high local control rate and reduced toxicity (23-25). Using the MLD model of normal tissue complication probability, Vogelius *et al.* (18) showed that a hypofractionated schedule delivered with modern tomotherapy theoretically resulted in a very limited change in the predicted RP risk compared with standard fractionation. In our hypofractionation clinical study, the incidence of symptomatic RP was 22.64%, similar to previously reported rates with conventional RT (26).

In the present study, we revealed that the RP prediction model built on dosage parameters of total lungs had the highest capacity, while the model constructed based on the contralateral lung was the worst, which is consistent with the literature (16,27,28). This finding suggests that the radiation received by the total and ipsilateral lung tissue

should be strictly restrained during clinical practice.

RF is an increasingly popular approach for dealing with high-dimensional data. An RF model is an ensemble-based decision tree method for classification and feature selection (29). Even though it has satisfactory prediction potential, RF is rarely used in RP risk evaluation. To the best of our knowledge, this is the first study to predict RP among patients with NSCLC receiving moderately hypo-RT based on the RF algorithm. The accuracy, precision, sensitivity, and specificity of the RF model were satisfactory. The RF model yielded a higher discriminative performance in RP than did the traditional logistic regression model.

In the present study, the dosimetric parameters of normal lungs were not converted into an equivalent biological dose of 2 Gy/fraction for the following reasons. First, the exact α/β ratio of the remaining normal lungs was unclear. Therefore, the biological equivalent doses estimated based on an uncertain α/β ratio may be characterized by large deviation compared to the real situation. Second, the range of fraction dose in the present study was very small (2.3–3.0) compared to stereotactic ablative RT. Finally, no matter how accurate the dose conversion based on the L-Q model is, it is not equivalent to the actual situation.

According to a recent study by Kashihara *et al.* (17), the original simple ILAS is an easy-to-use tool and is a significant predictive factor of steroid-RP in patients. with NSCLC treated with definitive RT. Nevertheless, our present study indicated that ILAS was not associated with RP. Several differences prevailed between these 2 studies that might have affected the data interpretation. First, the primary outcome of the research of Kashihara *et al.* (17) was RP requiring steroidal treatment, while in our study, it was defined as symptomatic RP. Second, participants in their study were subjected to conventional radiation modality using 3-dimensional conformal technology, while all cases in the present study were treated with moderately hypo-RT conducted by tomotherapy. Finally, the small sample size of ILAS higher than or equal to 1, might have also contributed to this discrepancy.

The present study excluded patients with NSCLC who received maintenance durvalumab. Importantly, our real-world data clarified that RP incidence rates were higher than those in the previous report without durvalumab (30). Moreover, the predictive factors of RP associated with this treatment modality may be different from the conventional chemoradiotherapy mode (31). In addition, clinical data on the combination of moderately hypo-RT and consolidation immune therapy are not yet available. Accordingly, only a few patients who received moderately hypo-RT were subjected to durvalumab maintenance. Under such scenarios, we are more inclined to conduct conventional RT. Therefore, cases with adjuvant durvalumab were excluded to reduce the statistical bias.

Some limitations were also present in this study. This was a retrospective study with uncertainty regarding the diagnosis and grading of RP. Moreover, patient homogeneity in this study was not optimal. In addition, given the relatively small sample size, we could not conduct a validation study based on a large cohort. Importantly, our ongoing multicenter, randomized, phase III clinical trial enrolled patients with unresectable stage III NSCLC receiving 60 Gy/20 fraction-concurrent chemoradiotherapy using tomotherapy (ChiCTR1800017367). After the completion of the trial, we will further analyze RP risk factors in this homogenous cohort.

Conclusions

An RF model was established based on the dosimetric parameters of total lungs. This RF model acted as an accurate machine learning method in identifying RP risk in patients with NSCLC treated with moderately hypofractionated tomotherapy. This model warrants further confirmation in large multicenter cohorts with high homogeneity.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://atm.amegroups.com/article/view/10.21037/atm-22-3049/rc

Data Sharing Statement: Available at https://atm.amegroups.com/article/view/10.21037/atm-22-3049/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://atm.amegroups.com/article/view/10.21037/atm-22-3049/coif).

The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institutional Review Board of Zhongshan Hospital, Fudan University (No. B2021-136), and informed consent was obtained from patients before enrollment.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Bradley JD, Hu C, Komaki RR, et al. Long-Term Results of NRG Oncology RTOG 0617: Standard- Versus High-Dose Chemoradiotherapy With or Without Cetuximab for Unresectable Stage III Non-Small-Cell Lung Cancer. J Clin Oncol 2020;38:706-14.
2. Fowler JF, Chappell R. Non-small cell lung tumors repopulate rapidly during radiation therapy. Int J Radiat Oncol Biol Phys 2000;46:516-7.
3. Bentzen SM. Quantitative clinical radiobiology. Acta Oncol 1993;32:259-75.
4. Beli I, Koukourakis G, Platoni K, et al. Hypofractionated radiotherapy in non small cell lung cancer: a review of the current literature. Rev Recent Clin Trials 2010;5:103-11.
5. Zhang Y, Li Z, Chen Y, et al. Outcomes of Image-Guided Moderately Hypofractionated Radiotherapy for Stage III Non-Small-Cell Lung Cancer. J Oncol 2021;2021:2721261.
6. He J, Huang Y, Chen Y, et al. Feasibility and efficacy of helical intensity-modulated radiotherapy for stage III non-small cell lung cancer in comparison with conventionally fractionated 3D-CRT. J Thorac Dis 2016;8:862-71.
7. Kaster TS, Yaremko B, Palma DA, et al. Radical-intent hypofractionated radiotherapy for locally advanced non-small-cell lung cancer: a systematic review of the literature.

8. Parisi G, Mazzola R, Ciammella P, et al. Hypofractionated radiation therapy in the management of locally advanced NSCLC: a narrative review of the literature on behalf of the Italian Association of Radiation Oncology (AIRO)-Lung Working Group. Radiol Med 2019;124:136-44.
9. Robinson SD, Tahir BA, Absalom KAR, et al. Radical accelerated radiotherapy for non-small cell lung cancer (NSCLC): A 5-year retrospective review of two dose fractionation schedules. Radiother Oncol 2020;143:37-43.
10. Bradley JD, Hope A, El Naqa I, et al. A nomogram to predict radiation pneumonitis, derived from a combined analysis of RTOG 9311 and institutional data. Int J Radiat Oncol Biol Phys 2007;69:985-92.
11. Hope AJ, Lindsay PE, El Naqa I, et al. Modeling radiation pneumonitis risk with clinical, dosimetric, and spatial parameters. Int J Radiat Oncol Biol Phys 2006;65:112-24.
12. Kong FM, Ao X, Wang L, et al. The use of blood biomarkers to predict radiation lung toxicity: a potential strategy to individualize thoracic radiation therapy. Cancer Control 2008;15:140-50.
13. Marks LB, Bentzen SM, Deasy JO, et al. Radiation dose-volume effects in the lung. Int J Radiat Oncol Biol Phys 2010;76:S70-6.
14. Wang C, Rimner A, Gelblum DY, et al. Analysis of pneumonitis and esophageal injury after stereotactic body radiation therapy for ultra-central lung tumors. Lung Cancer 2020;147:45-8.
15. Saha A, Beasley M, Hatton N, et al. Clinical and dosimetric predictors of radiation pneumonitis in early-stage lung cancer treated with Stereotactic Ablative radiotherapy (SABR) - An analysis of UK's largest cohort of lung SABR patients. Radiother Oncol 2021;156:153-9.
16. Liu Y, Wang W, Shiue K, et al. Risk factors for symptomatic radiation pneumonitis after stereotactic body radiation therapy (SBRT) in patients with non-small cell lung cancer. Radiother Oncol 2021;156:231-8.
17. Kashihara T, Nakayama Y, Ito K, et al. Usefulness of Simple Original Interstitial Lung Abnormality Scores for Predicting Radiation Pneumonitis Requiring Steroidal Treatment After Definitive Radiation Therapy for Patients With Locally Advanced Non-Small Cell Lung Cancer. Adv Radiat Oncol 2021;6:100606.
18. Vogelius IS, Westerly DC, Cannon GM, et al. Hypofractionation does not increase radiation pneumonitis risk with modern conformal radiation delivery techniques. Acta Oncol 2010;49:1052-7.
19. Oh D, Ahn YC, Park HC, et al. Prediction of radiation

Clin Lung Cancer 2015;16:71-9.

pneumonitis following high-dose thoracic radiation therapy by 3 Gy/fraction for non-small cell lung cancer: analysis of clinical and dosimetric factors. Jpn J Clin Oncol 2009;39:151-7.

20. Bentzen SM. High-tech in radiation oncology: should there be a ceiling? Int J Radiat Oncol Biol Phys 2004;58:320-30.

21. Scrimger RA, Tomé WA, Olivera GH, et al. Reduction in radiation dose to lung and other normal tissues using helical tomotherapy to treat lung cancer, in comparison to conventional field arrangements. Am J Clin Oncol 2003;26:70-8.

22. Cattaneo GM, Dell'oca I, Broggi S, et al. Treatment planning comparison between conformal radiotherapy and helical tomotherapy in the case of locally advanced-stage NSCLC. Radiother Oncol 2008;88:310-8.

23. Zelefsky MJ, Kollmeier M, Cox B, et al. Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. Int J Radiat Oncol Biol Phys 2012;84:125-9.

24. Bujold A, Craig T, Jaffray D, et al. Image-guided radiotherapy: has it influenced patient outcomes? Semin Radiat Oncol 2012;22:50-61.

25. Jaffray DA. Image-guided radiotherapy: from current concept to future perspectives. Nat Rev Clin Oncol 2012;9:688-99.

26. Palma DA, Senan S, Tsujino K, et al. Predicting radiation pneumonitis after chemoradiation therapy for lung cancer: an international individual patient data meta-analysis. Int J Radiat Oncol Biol Phys 2013;85:444-50.

27. Thompson MR, Dumane VA, Lazarev SA, et al. Dosimetric Correlates of Pulmonary Toxicity in Patients with Malignant Pleural Mesothelioma Receiving Radiation Therapy to the Intact Lungs. Pract Radiat Oncol 2019;9:e331-7.

28. Yorke ED, Jackson A, Rosenzweig KE, et al. Correlation of dosimetric factors and radiation pneumonitis for non-small-cell lung cancer patients in a recently completed dose escalation study. Int J Radiat Oncol Biol Phys 2005;63:672-82.

29. Patel H, Vock DM, Marai GE, et al. Oropharyngeal cancer patient stratification using random forest based-learning over high-dimensional radiomic features. Sci Rep 2021;11:14057.

30. Jung HA, Noh JM, Sun JM, et al. Real world data of durvalumab consolidation after chemoradiotherapy in stage III non-small-cell lung cancer. Lung Cancer 2020;146:23-9.

31. Mayahara H, Uehara K, Harada A, et al. Predicting factors of symptomatic radiation pneumonitis induced by durvalumab following concurrent chemoradiotherapy in locally advanced non-small cell lung cancer. Radiat Oncol 2022;17:7.

Table S1 Univariate analysis of clinical factors in the training group

Characteristics	No RP (n, %)	RP (n, %)	P
n	50.00	17.00	
Age (year, mean ± SD)	64.0±10.90	60.2±9.50	0.20
Gender			
Female	5 (83.33)	1 (16.67)	0.98
Male	45 (73.77)	16 (26.23)	
Smoke			
No	24 (64.86)	13 (35.14)	0.08
Yes	26 (86.67)	4 (13.33)	
Emphysema			
No	33 (80.49)	8 (19.51)	0.27
Yes	17 (65.38)	9 (34.62)	
ILAS			
0	37 (74.00)	13 (26.00)	0.91
1	8 (80.00)	2 (20.00)	
2	5 (71.43)	2 (28.57)	
Diabetes			
No	44 (72.13)	17 (27.87)	0.31
Yes	6 (100.00)	0 (0.00)	
Pathology			
Adenocarcinoma	16 (61.54)	10 (38.46)	0.07
Others	3 (60.00)	2 (40.00)	
Squamous	31 (86.11)	5 (13.89)	
T			
1–2	28 (73.68)	10 (26.32)	1.00
3–4	22 (75.86)	7 (24.14)	
N			
0–1	12 (80.00)	3 (20.00)	0.84
2–3	38 (73.08)	14 (26.92)	
Stage			
II–III	44 (75.86)	14 (24.14)	0.86
IV	6 (66.67)	3 (33.33)	
Modality			
conCRT	13 (72.22)	5 (27.78)	0.57
RT alone	8 (88.89)	1 (11.11)	
seqCRT	29 (72.50)	11 (27.50)	

RP, radiation pneumonitis; ILAS, interstitial lung abnormality score; conCRT, concurrent chemoradiotherapy; RT, radiotherapy; seqCRT, sequential chemoradiotherapy.

Table S2 Univariate analysis of dosimetric parameters of contralateral, ipsilateral, and total lungs

Dosimetric parameters	No RP	RP	P
n	50	17	
CMLD, median [IQR]	6.180 [4.168, 7.690]	6.590 [6.000, 7.410]	0.33
CV5, mean (SD)	34.059 (14.176)	42.235 (9.162)	0.03
CV10, median [IQR]	21.000 [12.250, 28.000]	24.000 [17.000, 28.000]	0.50
CV15, median [IQR]	10.000 [5.250, 18.000]	13.000 [9.000, 15.000]	0.38
CV20, median [IQR]	3.500 [1.000, 9.000]	8.000 [3.000, 12.000]	0.07
CV25, median [IQR]	1.500 [0.000, 5.000]	3.000 [0.000, 5.000]	0.63
CV30, median [IQR]	0.500 [0.000, 3.000]	2.000 [0.000, 3.000]	0.59
CV35, median [IQR]	0.000 [0.000, 2.000]	1.000 [0.000, 2.000]	0.44
CV40, median [IQR]	0.000 [0.000, 0.750]	0.000 [0.000, 1.000]	0.76
CV45, median [IQR]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]	0.88
CV50, median [IQR]	0.000 [0.000, 0.000]	0.000 [0.000, 0.000]	0.41
IMLD, median [IQR]	15.905 [13.037, 19.587]	18.120 [17.160, 19.790]	0.02
IV5, median [IQR]	61.500 [55.000, 70.500]	71.000 [65.000, 78.000]	0.014
IV10, median [IQR]	55.000 [45.000, 63.750]	61.000 [55.000, 69.000]	0.0245
IV15, median [IQR]	44.000 [34.250, 52.750]	50.000 [45.000, 55.000]	0.0406
IV20, mean (SD)	32.680 (12.286)	38.471 (6.746)	0.0696
IV25, mean (SD)	24.860 (10.150)	30.235 (5.750)	0.0427
IV30, mean (SD)	19.060 (8.515)	23.588 (4.611)	0.0411
IV35, mean (SD)	14.100 (6.855)	17.353 (3.534)	0.0664
IV40, median [IQR]	8.500 [7.000, 13.000]	12.000 [10.000, 14.000]	0.0137
IV45, median [IQR]	6.000 [4.000, 8.750]	7.000 [7.000, 9.000]	0.0703
IV50, median [IQR]	4.000 [3.000, 6.000]	5.000 [4.000, 6.000]	0.1031
TMLD, median [IQR]	10.950 [8.897, 11.677]	13.520 [12.530, 14.550]	<0.0001
TV5, median [IQR]	48.000 [39.500, 52.000]	54.310 [51.000, 58.000]	0.0036
TV10, median [IQR]	37.000 [28.250, 42.000]	41.000 [38.000, 43.000]	0.0377
TV15, median [IQR]	28.000 [18.750, 32.750]	31.000 [27.000, 33.000]	0.1028
TV20, median [IQR]	17.000 [13.000, 20.750]	25.000 [21.000, 26.000]	<0.0001
TV25, median [IQR]	14.000 [9.250, 17.000]	16.000 [14.000, 19.000]	0.0281
TV30, mean (SD)	9.701 (3.993)	12.347 (3.029)	0.0152
TV35, median [IQR]	6.500 [5.000, 8.000]	9.000 [7.000, 13.000]	<0.0001
TV40, mean (SD)	4.673 (2.095)	6.187 (1.697)	0.0091
TV45, median [IQR]	3.000 [2.000, 4.000]	3.050 [3.000, 4.000]	0.0511
TV50, median [IQR]	2.000 [1.000, 3.000]	2.000 [2.000, 3.000]	0.3309

RP, radiation pneumonitis; C, contralateral lung; I, ipsilateral lung; T, total lungs; MLD, mean lung dose; IQR; interquartile range; Vx, the percentage of normal lungs receiving a dose of ≥ x Gy.