Response Prediction in Lung SBRT with Artificial Intelligence

Durmuş ETİZ,1 Melek YAKAR,1 Güntülü AK,2 Deniz KÜTRİ,1 Özer ÇELİK,3 Muzaffer METİNTAŞ2

1Department of Radiation Oncology, Eskişehir Osmangazi University Faculty of Medicine, Eskişehir-Türkiye
2Department of Chest Diseases, Eskişehir Osmangazi University Faculty of Medicine, Eskişehir-Türkiye
3Department of Mathematics-Computer, Eskişehir Osmangazi University, Eskişehir-Türkiye

OBJECTIVE
Lung cancer is the leading cause of cancer-related death worldwide. Although the majority of patients have locally advanced or metastatic disease at diagnosis, the incidence of early-stage non-small-cell lung cancer is expected to increase due to the wider use of thoracic CT scans. Primary tumor control and distant metastasis rates in early-stage lung cancer are similar for stereotactic body radiation therapy (SBRT) and surgery. Overall survival (OS) is lower for SBRT compared to surgery. Although some studies provide guidance on which cases will have a good response to SBRT, there is still no standard guideline. SBRT results are not the same in cases at the same stage or with the same metastatic burden. It is thought that there may be other parameters other than stage or tumor burden that affect the response. It is aimed to predict the response to SBRT with artificial intelligence in early-stage lung cancer, recurrent lung cancer, and lung metastases.

METHODS
Between September 2016 and April 2021, 137 cases and 148 lesions in which SBRT was applied by Eskişehir Osmangazi University Faculty of Medicine Radiation Oncology Department were evaluated. To create a balanced data set, Synthetic Minority Oversampling Technique technique was used and 200 lesions were evaluated. Logistic Regression (LR), multilayer perceptron Classifier, Extreme Gradient Boosting, Support Vector Classifier, Random Forest Classifier, and Gaussian Naive Bayes algorithms were used. The data sets are divided into 85% training and 15% prediction sets. Models were created using the training set and validated using the prediction set.

RESULTS
Complete response was obtained in 41 tumors out of 148 tumors. The median OS after SBRT is 18 (2–61) months, and progression-free survival is 16 (0–61) months. Important variables are tumor diameter, NLR, presence of biopsy at diagnosis, tumor location and type, diagnosis, and histopathology. LR algorithm was determined as the best estimating algorithm with 80% accuracy (Confidence Interval, CI: 0.65–0.94, ROC AUC: 0.60), 66% sensitive and 90% specificity.

CONCLUSION
In order to use the current algorithm in clinical practice, it is necessary to increase the diversity of data and the number of patients by sharing data between centers.

Keywords: Artificial intelligence; lung cancer; response prediction; stereotactic body radiotherapy.

Copyright © 2023, Turkish Society for Radiation Oncology
INTRODUCTION

Lung cancer is the leading cause of cancer-related death worldwide.[1] Although the majority of patients have locally advanced or metastatic disease at diagnosis, the incidence of early-stage non-small-cell lung cancer (NSCLC) is expected to increase due to the wider use of thoracic CT scans.[2] While early-stage NSCLC is treated with lobectomy and mediastinal/hilar lymph node sampling, sublobar resection may be an option for patients with low pulmonary reserve.[3] Alternative effective treatment options should be preferred in cases where surgery is at high risk or surgery is rejected by the patient, depending on the comorbidities of the patients. Stereotactic body radiation therapy (SBRT), also referred to as stereotactic ablative radiotherapy, has emerged as a standard treatment option for early-stage NSCLC in the last decade. SBRT is a conformal technique that can deliver a very high dose (ablative dose) to the target in 1–5 fractions.[4]

Primary tumor control and distant metastasis rates in early-stage lung cancer are similar for SBRT and surgery. Overall survival (OS) is lower for SBRT compared to surgery. This is largely attributable to inadequate cardiopulmonary function as the primary selection criterion for SBRT and the reduced survival of these patients due to their comorbid disease, independent of lung cancer. The lung is the second-most common site of metastatic focus. It is estimated that 20–54% of malignant tumors that develop in other parts of the body will metastasize to the lungs.[5] Lungs are the sole site of metastasis in 80% of patients with sarcoma and in 2–10% of patients with carcinoma.[6] In 1995, the paradigm of treating patients with limited metastases was defined as the “oligometastatic state” by Hellman and Weichselbaum.[7] Patients diagnosed with metastatic disease at any point in the disease course have a wide range of total metastatic burden, from a single lesion to extensive disease. Conventionally, systemic therapy has been the mainstay of therapy for these patients, and radiotherapy for palliation has been used if necessary.[8,9] However, recently, this treatment paradigm has been changing, especially in oligometastatic cases. Randomized studies in NSCLC show that radical local treatments increase progression-free survival and OS in oligometastatic disease.[10–12] Although some studies provide guidance on which cases will have a good response to SBRT, there is still no standard guideline. SBRT results are not the same in cases at the same stage or with the same metastatic burden. It is thought that there may be other parameters other than stage or tumor burden that affect the response. In this study, it is aimed to predict the response to SBRT with artificial intelligence techniques.

MATERIALS AND METHODS

Patients Characteristics

Between September 2016 and April 2021, 137 cases and 148 lesions in which SBRT was applied were evaluated by the Radiation Oncology Department of Eskişehir Osmangazi University. Early-stage medically inoperable cases (T1-T2N0M0, ≤5 cm), cases with single isolated lung metastasis or cases matching the definition of oligometastatic with lung metastases,[13] and lung cancer cases with recurrence in follow-up were included in the study. Ultracentral, central and peripheral tumors were included in the study. For staging purposes, FDG PET CT and brain MRI were seen in each case. The cases were evaluated in the Eskişehir Osmangazi University Oncology Council after staging and treatment decisions were taken multidisciplinary. The study was initiated after the approval of the Non-Invasive Clinical Research Ethics Committee.

Treatment Characteristics

The patients were immobilized in the supine position with their hands on the head and with a T-bar/Wingboard. For all patients, a 4DCT was acquired on a Somatom Definition AS® comprising ten respiratory phases, where phases are indicated in percent of the breathing cycle. Computed tomography (CT) images of all cases were taken with a cross-sectional interval of 2 mm. Average CT (avgCT) and maximum intensity projection (MIP) datasets were generated from the phases. The ITV was contoured on the 4DCT MIP (4DCT ITV), and evaluated on the selected breathing phase. A 3-mm isotropic margin was placed around the ITV to create the planning target volume (PTV). Treatment plans were generated for a Varian TrueBeam® linear accelerator with a 6 MV flattening filter-free beam. Planning criteria were based on RTOG 0915[14] such that at least 95% of the PTV is covered by the prescription dose. For each fraction, a free-breathing 4DCBCT was acquired. Varian Real-time Position Management was used to track patient breathing.

Different fractionation schemes were used according to the location of the tumor. While more hypofractionated regimens are preferred in ultracentral tumors, 10–12 Gy × 5 fraction regimens are preferred in peripheral tumors. 3D CBCT images were taken before and after beam was on. During SBRT, no simultaneous chemotherapy was applied. At least 48 h were inter-
rupted between fractions. Cases were evaluated in the outpatient clinic twice a week for possible toxicity.

**Selected Variables**

A total of 18 variables were evaluated to predict the response to lung SBRT, which are: tumor location type (central, ultracentral, peripheral, chest wall), diagnosis (early stage/relapse/metastasis), age, gender, Karnofsky Performance Score, smoking history, history of chronic disease, presence of biopsy at diagnosis, histopathology, tumor location (right/upper left/middle/lower), tumor largest diameter, GTV, PTV, fraction dose, fraction number, BED10, neutrophil/lymphocyte (NLR) and platelet/lymphocyte ratio.

**Follow-up after Radiotherapy**

Patients were called for control 6 and 12 weeks after SBRT, then every 2 years, every 3 months, every 6 months for the next 2 years, and annually thereafter. Physical examinations were performed at each follow-up visit.

Thoracic CT imaging was performed first at 6 weeks after RT and then at each follow-up to assess response and toxicity. All patients underwent FDG PET CT examination 3 months and 1 year after SBRT to evaluate the response. Cases with suspected recurrence were evaluated in multidisciplinary councils. Tumors were measured at each follow-up visit using the Response Evaluation Criteria in Solid Tumors, and the response was graded according to the international criteria proposed in the Response Evaluation Criteria in Solid Tumors Guideline version 1.1.5.[15]

**Artificial Intelligence and Statistics**

Patients with missing data were excluded from the study. In case the sample sizes of the subgroups of the dependent variable to be estimated in machine learning are not equal, biased estimations are made as a result of overfitting. To get rid of this situation, it is necessary to create a balanced data set.[16] Synthetic Minority Oversampling Technique (SMOTE) technique was used to create a balanced data set. In SMOTE, each minority class sample is taken, and synthetic samples are created by looking at any or all of the k neighbors of this sample. Thus, the minority class becomes oversampled. The main difference from other sampling methods is that synthetic samples are produced by looking at their close neighbors instead of duplicating the samples in the minority class.[17]

A total of 200 lesions were evaluated in the present study. Logistic Regression (LR), multilayer perceptron Classifier (MLP), Extreme Gradient Boosting (XGB), Support Vector Classifier (SVC), Random Forest Classifier, and Gaussian Naïve Bayes algorithms are used. After the correlation analysis, the permutation-based variable selection method was used as the variable selection method. Permutation-based variable selection is defined as the reduction in the model score when a single variable value is randomly mixed. This process breaks the relationship between the variable and the target, so the decrease in the model score indicates how much the model is dependent on the feature. This technique can be calculated many times with different permutations of the variables in the model.[18] The data sets are divided into 85% training and 15% prediction sets. Models were created using the training set and validated using the prediction set. In the tests of these models, the accuracy (accuracy), sensitivity (sensitivity), and specificity values, which are the confusion matrix metrics, and the model success rates were determined with the receiver operating characteristic (ROC) curve area under the curve (AUC).[19]

In statistics, the ROC curve is a graphical plot showing the diagnostic ability of the dual classification system. AUC indicates the classification performance of the installed model and takes a value between 0 and 1. AUC value close to 1 means that the classification performance of the model is high.[20] Accuracy Rate (ACC), which is a widely used success evaluation method, was used in our study. The accuracy method is the ratio of the system’s class of facts (True Positive [GP] and True Negative [GN]) to the total number of samples. The error rate is the ratio of the number of incorrectly calculated samples (False Positive (F) and False Negative (F)) to the total number of samples.[21]

Statistical analyzes and machine learning algorithms were performed using Python software (Python Software Foundation. Python Language Reference, version 3.5. Available at http://www.python.org) and Scikit Learn library.[22] All analyses and operations were performed using a computer with Windows 10, 64-bit operating system, and Intel Core i7–9750 CPU with 2.6 GHz 12MB Cache and 16GB 2666MHz DDR4 Ram memory.

**RESULTS**

The median age was 68 (min:40-max:88). Male/female is 102/35. The median KPS was 80 (min: 70-max:100). Among the patients, smokers and non-smokers were 98 (71.5%) and 39 (28.5%). The rate of biopsy before SBRT in cases was 40%.
148 tumors of 137 cases were evaluated. The number of ultracentral/central/peripheral/chest wall tumors was 4 (2%), 12 (8%), 114 (77%), and 18 (12%), respectively. The median tumor diameter is 20 (5–50) mm. Tissue diagnosis was present in 48 (32%) of the tumors. When we look at the tumor diagnoses, the numbers of early-stage lung cancer, recurrent lung cancer, solid single lung metastasis, and oligometastatic lung cancer were 56 (37%), 43 (29%), 48 (32%), and 1 (0.6%), respectively. Median GTV and PTV volumes were 7.9 (min: 0.5, max: 72) cc and 23.5 (min: 1.94 max: 122) cc, respectively. The median BED10 value was 100 (min:72, max:132) Gy. The median NLR is 2.7 8min: 0.47, max: 19.7). Tumor and treatment characteristics are summarized in Table 1. The median OS after SBRT was 18 (2–61) months, and progression-free survival was 16 (0–61) months. According to FDG PET CT and 6th-month tomography findings performed for response evaluation at 3 months after SBRT, the rates of complete response/partial response/stable response and progressive response rates were 41 (27.7%), 74 (50%), 24 (16.2%), and 9 (6.1%).

Important variables are tumor diameter, NLR, presence of biopsy at diagnosis, tumor location and type, diagnosis, and histopathology. LR algorithm is determined as the best estimating algorithm with 80% accuracy (Confidence Interval, CI: 0.65–0.94, ROC AUC: 0.60), 66% sensitive and 90% specificity. The ACC for MLP, XGB, SVC, RF, and Gaussian NB, among other evaluated algorithms, are 76%, 43%, 46%, 60%, and 46%, respectively. ROC AUC graphs of the algorithms are given in Figure 1.

ROC AUC values for MLP, XGB, SVC, RF, and Gaussian NB are 0.50–0.36–0.54–0.53 and 0.64, respectively.

The confusion matrix of the LR algorithm is given in Table 2. The algorithm correctly predicted 20 of the 24 non-response cases, incorrectly predicted 4 of them, and correctly predicted 4 out of 6 complete response cases and incorrectly predicted 2 of them. The results of other algorithms are summarized in Table 3.

**DISCUSSION**

Surgery is currently the standard of care for patients with stage I NSCLC. Radiation therapy, especially SBRT, is recommended for patients who are medically unsuitable for surgery. It is controversial whether SBRT is an appropriate treatment option for patients who are candidates for surgery. SBRT is an extremely well-tolerated procedure that does not require hospitalization and has been reported to provide local tumor control rates exceeding 90%. From this perspective, it is an attractive alternative to an invasive surgical procedure. Unfortunately, strong level 1 evidence comparing surgery and SBRT is lacking.[23] It is still unclear which cases might benefit more from SBRT. Despite major advances in therapeutic strategies over the past few decades, NSCLC is still the leading cause of death of cancer.
or which oligometastatic or oligorecurrent patient is suitable for SBRT. The importance of tumor size in the staging and prognosis of many tumors has been demonstrated by studies.[26,27] As the tumor size increases, the T stage progresses and the prognosis worsens.

Tumor microenvironment and especially inflammatory response and systemic inflammation play an important role in cancer development and progression. [28] Parameters that can be measured in blood and show systemic inflammation can be used as a biochemical marker to evaluate prognosis in cancer. Recently, elevated peripheral neutrophil-lymphocyte (NLR) ratio has been accepted as a poor prognostic indicator in various cancers.[29] In the current study, tumor diameter, NLR, presence of biopsy in diagnosis, tumor location and type, diagnosis, and histopathology were determined as important variables in estimating response to SBRT, and an ACC of up to 80% was obtained in the estimation algorithm established with these variables. Providing BED_{100} >100 Gy in ultracentral tumors results in serious morbidity and even mortality.[30] Therefore, the BED_{100} value was kept lower in these cases. This explains the effect of tumor location on dose and even prognosis.

Machine learning classification algorithms (classifiers) for the prediction of treatment response are becoming more popular in the radiotherapy literature. The general machine learning literature provides evidence in favor of some classifier families (random forest, support vector machine, gradient boosting) in terms of classification performance. Currently, there is no consensus on an optimal classification algorithm. Researchers choose algorithms for a variety of reasons: researcher’s experience, use in the literature, data characteristics and quality, default feature dependencies, availability of simple implementations, and model interpretability. One objective criterion for selecting a classifier is to maximize a chosen performance metric, for example, discrimination (expressed by the area under the ROC curve, AUC). In a study by Deist et al.,[31] different machine learning algorithms were evaluated in the evaluation of response to radiotherapy. As a result of the evaluation, random forest and LR provided higher discriminant performance in (chemo) radiotherapy outcomes and toxicity prediction than other classifiers studied. Therefore, one of these two classifiers is recommended as the first choice when creating classification models. In our current study, the LR algorithm is determined as the best estimating algorithm with 80% accuracy (Confidence Interval, CI: 0.65–0.94, ROC AUC: 0.60), 66% sensitive and 90% specificity. It was determined as the best estimating algorithm with specificity.

Table 2 Logistic regression algorithm confusion matrix

<table>
<thead>
<tr>
<th>Prediction</th>
<th>Complete response (-)</th>
<th>Complete response (+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Complete</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>response (-)</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3 The results of other algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy</th>
<th>ROC</th>
<th>Precision</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td>0.80</td>
<td>0.60</td>
<td>0.66</td>
<td>0.90</td>
</tr>
<tr>
<td>MLP</td>
<td>0.76</td>
<td>0.50</td>
<td>0.66</td>
<td>0.95</td>
</tr>
<tr>
<td>XGB</td>
<td>0.43</td>
<td>0.36</td>
<td>0.15</td>
<td>0.50</td>
</tr>
<tr>
<td>SVC</td>
<td>0.46</td>
<td>0.54</td>
<td>0.27</td>
<td>0.40</td>
</tr>
<tr>
<td>RF</td>
<td>0.60</td>
<td>0.53</td>
<td>0.16</td>
<td>0.77</td>
</tr>
<tr>
<td>GNB</td>
<td>0.46</td>
<td>0.64</td>
<td>0.33</td>
<td>0.27</td>
</tr>
</tbody>
</table>

ROC: Receiver operating characteristic; LR: Logistic regression; MLP: Multi-layer perceptron classifier; XGB: Extreme gradient boosting; SVC: Support vector classifier; RF: Random forest; GNB: Gaussian Naive Bayes
CONCLUSION

It is still unclear which patient would benefit more from SBRT. In order to predict this, machine learning algorithms, especially the LR algorithm is an algorithm recommended. However, studies still have not found standard modeling. SBRT response estimation can be made by creating models with higher accuracy with multicenter studies with more patients.

Peer-review: Externally peer-reviewed.

Conflict of Interest: All authors declared no conflict of interest.

Ethics Committee Approval: The study was approved by the Eskişehir Osmangazi University Non-Invasive Clinical Research Ethics Committee (no: 30, date: 21/02/2023).

Financial Support: None declared.


REFERENCES


