# Ensemble Model-Driven Disease Recurrence Prediction for NSCLC Patients Treated with SBRT

Kaushik Halder, PhD[1] Michael D. Mix, MD[1] Tithi Biswas, MD[2] Jeffrey A. Bogart, MD[1] and Tarun K. Podder, PhD[1]

[1]SUNY Upstate Medical University, Syracuse, NY, USA [2]University of Florida, Gainesville, FL, USA

2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

## ABSTRACT

***Purpose:*** Early prediction of distant recurrence in early-stage non-small cell lung cancer (NSCLC) patients may assist clinical decision making. Recent studies demonstrated hardly any benefit while adding systemic therapy to SBRT, and identification of patients at highest risk for developing distant relapses may be critical. Current study aims to develop a distant recurrence prediction model by combining radiomic, dosiomic, and clinical features for patients treated with SBRT.

***Methods:*** Total 141 node-negative early-stage NSCLC patients treated with SBRT (48-60 Gy in 3-5 fractions) were included in this study. Radiomic features from planning CT images were gleaned from the gross tumor volume (GTV) zone and evaluated based on the histogram, gray level co-occurrence matrix, gray level run length matrix, and gray level size zone matrix. A novel ensemble model was developed based on:(1) Support Vector Machine, (2) Gradient Boosting, and (3) Random Forest algorithms with soft voting technique, while considering a combination of 47 radiomic features and 11 most influential dosiomic and clinical features as inputs. Output layer of the designed model utilized an optimal thresholding layer to predict distant recurrence after two-years of SBRT.

***Results:*** Patients' median age was 73years (range: 52-91); T-stage was T0-T2bN0;cohort had 46% female. Designed distant recurrence prediction model demonstrated superior performance with mean AUC of 0.75 (0.68-0.83), F1-score of 0.85 (0.80-0.92), sensitivity of 0.80 (0.67-0.96), and specificity of 0.70 (0.50-1). In contrast, model with the combination of radiomic features, clinical, and dosiomic variables exhibited an improvement of 8.7% in AUC compared to the model relies solely on radiomic features.

***Conclusion:*** Leveraging radiomic, dosiomic, and clinical features has the potential for developing a distant recurrence prediction model for NSCLC patients treated with SBRT. A future study with a larger patient cohort from multiple institutes is desirable to verify the robustness of the developed ML model.

## CONTACT

**Tarun K. Podder, PhD**
SUNY Upstate Medical University,
Syracuse, NY 13210 USA
poddert@upstate.edu
215-464-5999

## INTRODUCTION

- Stereotactic Body Radiation Therapy (SBRT) is a highly effective modality of treating early-stage non-small lung cancer (NSCLC) patients.
- Despite advances in curative treatments, the mortality rate associated with distant recurrence still exceeds 25% after SBRT [1, 2].
- Predicting distant recurrence in patients with NSCLC is important for
  - Optimizing treatment regimens
  - Improving patient outcomes
  - Advancing the overall disease management

## METHODS AND MATERIALS

- **Data cohort →** Study includes 141 early-stage (stage I-II) NSCLC patients' planning CT images during SBRT (as shown in Table 1). SBRT characteristics are presented in Table 2.
- **Input Features →**
  - Eleven influential clinical and dosiomic features.
  - Radiomic features (as shown in Fig. 2) extracted from planning CT-images within GTV ROI.
- **Influential Clinical & Dosiomic Features →**
  - **Clinical features:**
    **T-stage, Performance status, Smoking pack per year.**
  - **Dosiomic features:**
    **Conformity index, PTV size, Prescription dose, Min dose to PTV, Min Dose to GTV, Mean dose to PTV, Mean dose GTV, Mean lung dose.**

**Table 1:** Demographic and clinical data cohort (n=141).

| Feature Groups | Feature Variables | Median (range) or (%) |
|---|---|---|
| Clinical numerical features | Age | Median (73), Range (52-91) |
| | PET Max SUV | Median (5.1), Range (1-37.4) |
| | aCCI | Median (8), Range (4-14) |
| | Smoking pack per year | Median (50), Range (0-150) |
| Clinical categorical features | Sex | Male (53.9 %), Female (46.1 %) |
| | Location of Tumor | Upper (67.4 %), Middle (7.8 %), Lower (24.8 %) |
| | Performance status | Zero (24 %), One (44 %), Two (26.2 %), Three (5.8 %) |
| | History of COPD | Yes (71.6 %), No (28.4 %) |
| | T-stage | 1-stage (2.1 %), 1a-stage (15.6 %), 1b-stage (50.3 %), 1c-stage (21.3 %), 2a-stage (9.9 %), 2b-stage (0.7 %) |

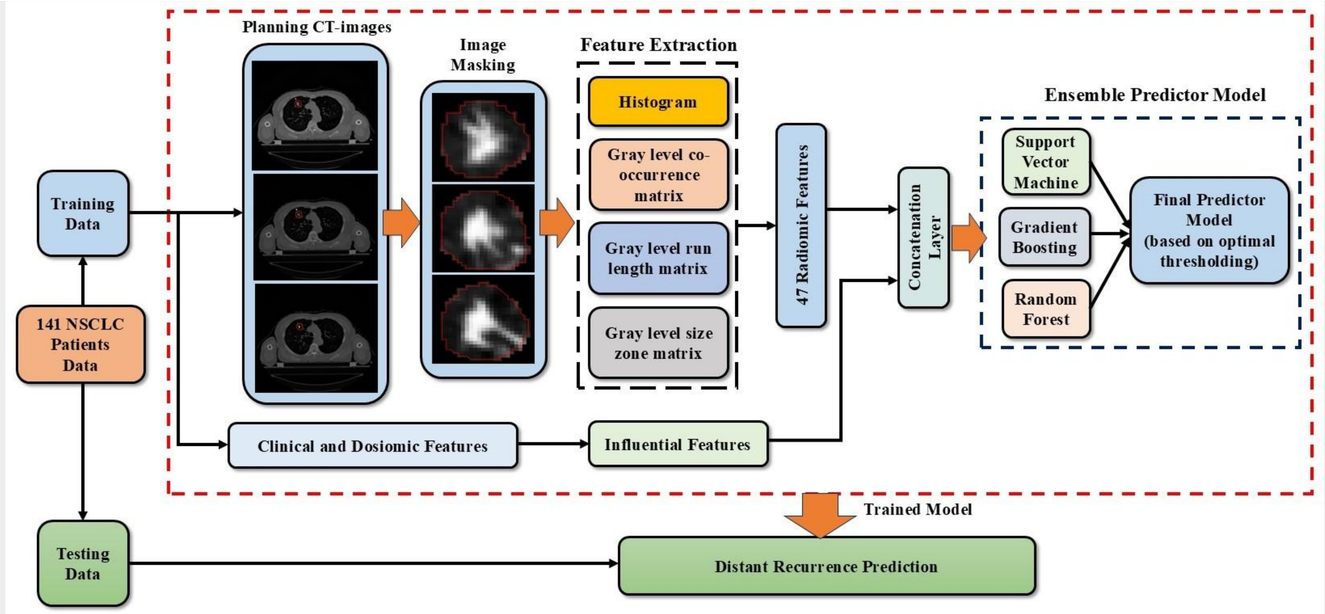


**Figure 1:** The proposed ensemble machine learning-based disease recurrence prediction framework.

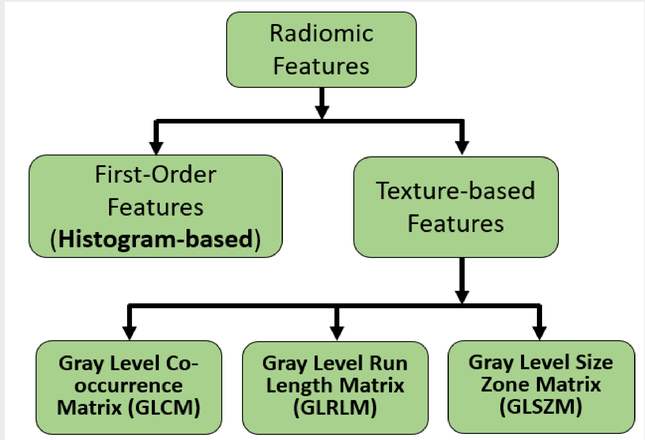


**Figure 2:** Extracted 47 radiomic features, including histogram, GLCM, GLRLM, and GLSZM-based.

- **Planning CT-scans preprocessing steps →**
  - Step-1: Image Masking based on GTV ROI
  - Step-2: Dimensional reshaping
  - Step-3: Intensity normalization
- **Designed Prediction Model →** Ensemble model involving Support Vector Machine, Gradient Boosting, and Random Forest algorithms, as given in Fig. 1.

**Table 2:** Treatment parameters of the patients (n=141).

| SBRT Dose/ # of fraction | Number of patients (%) | Dose in BED | Study Endpoint | Number of patients (%) |
|---|---|---|---|---|
| 50 Gy/ 4fx | 53 (37.6) | 112.5 | Distant Recurrence | Recurrence: 23 (16.3) |
| 50 Gy/ 5fx | 68 (48.2) | 100 | | |
| 54 Gy/ 3fx | 17 (12.1) | 151.2 | | |
| 55 Gy/ 5fx | 1 (0.7) | 115.5 | | Non-recurrence: 118 (83.7) |
| 57.5 Gy/ 5fx | 1 (0.7) | 123.6 | | |
| 60 Gy/ 3fx | 1 (0.7) | 180 | | |

## RESULTS

- Model's effectiveness is assessed using 100 bootstrap iterations, with 70% data for training and rest 30% for testing in each iteration.
- Only radiomic features based model achieved mean AUC of 0.69 (range: 0.58-0.79).
- All combined features-based model yielded mean AUC of 0.75 (range: 0.68-0.83).
- Integrating clinical and dosiomic features to radiomics improved AUC by 8.7%, as summarized in Fig. 3.

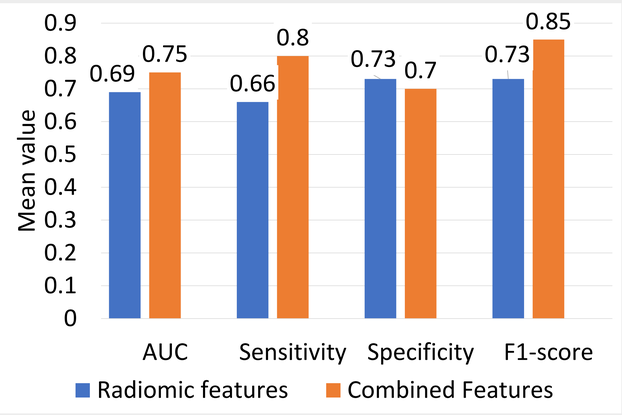


**Figure 3:** Performance of the ensemble model while incorporating radiomic, dosiomic, and clinical features.

## CONCLUSIONS

- Demonstrated the effectiveness of the planning CT images during the SBRT in developing disease recurrence prediction model.
- Proposed Deep Neural Network architecture may potentially assist in clinical decision-making process with improved clinical responses.
- **Future study:** A large, multi-institutional patient cohort is essential to develop a more robust and generalizable model.

## REFERENCES

1. Chi A, et al. Systemic review of the patterns of failure following stereotactic body radiation therapy in early-stage non-small-cell lung cancer: clinical implications. *Radiother Oncol,* 2010;94(1):1-11.
2. Timmerman RD, et al. Long-term results of stereotactic body radiation therapy in medically inoperable stage I non-small cell lung cancer. *JAMA Oncol,* 2018;4(9):1287-1288.