# Multimodal Late Fusion Meta-Learning for Pre-treatment Radiation Pneumonitis Prediction

MEDICAL PHYSICS COMMUNITY · INSPIRATION INNOVATION IMPACT
2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

Yong Hum Na[1], David J. Carlson[1], Ming Chao[2]

1. Yale University School of Medicine, New Haven, CT, 2. Icahn School of Medicine at Mount Sinai, New York, NY

Yale University School of Medicine

Icahn School of Medicine at Mount Sinai

## PURPOSE

Radiation pneumonitis (RP) remains a clinically significant dose-limiting toxicity in thoracic radiotherapy. Accurate RP prediction is challenging due to its multifactorial etiology and complex interactions among contributing factors. Although multimodal data including radiomic, dosimetric, and clinical variables are essential for improved prediction, effective learning is hindered by their high dimensionality and heterogeneity. This study proposes a robust multimodal learning framework to improve RP prediction.

## METHODS

Pre-treatment CT images, dosimetric data, and clinical variables from 421 lung cancer patients enrolled in NRG Oncology RTOG 0617 were analyzed. Features were partitioned into five predefined blocks: multilobe radiomics, ipsilateral lung dose metrics, contralateral lung dose metrics, treatment arm–specific variables, and general clinical factors. For each feature block, base learners including logistic regression (LR), support vector machine (SVM), random forest (RF), extra trees (ET), eXtreme Gradient Boosting (XGB), k-nearest neighbors (KNN), and Tabular Prior-data Fitted Network (TabPFN) were trained using stratified 5-fold cross-validation. Dimensionality reduction (none, neighborhood component analysis, or principal component analysis) was optimized per block and classifier by maximizing fold-level AUROC.

Optimized block-level predicted probabilities were combined using a logistic regression meta-learner in a late fusion framework. Performance was evaluated using mean AUROC ± standard deviation and bootstrap-derived 95% confidence intervals (CI). Figure 1 illustrates general flowchart of this study

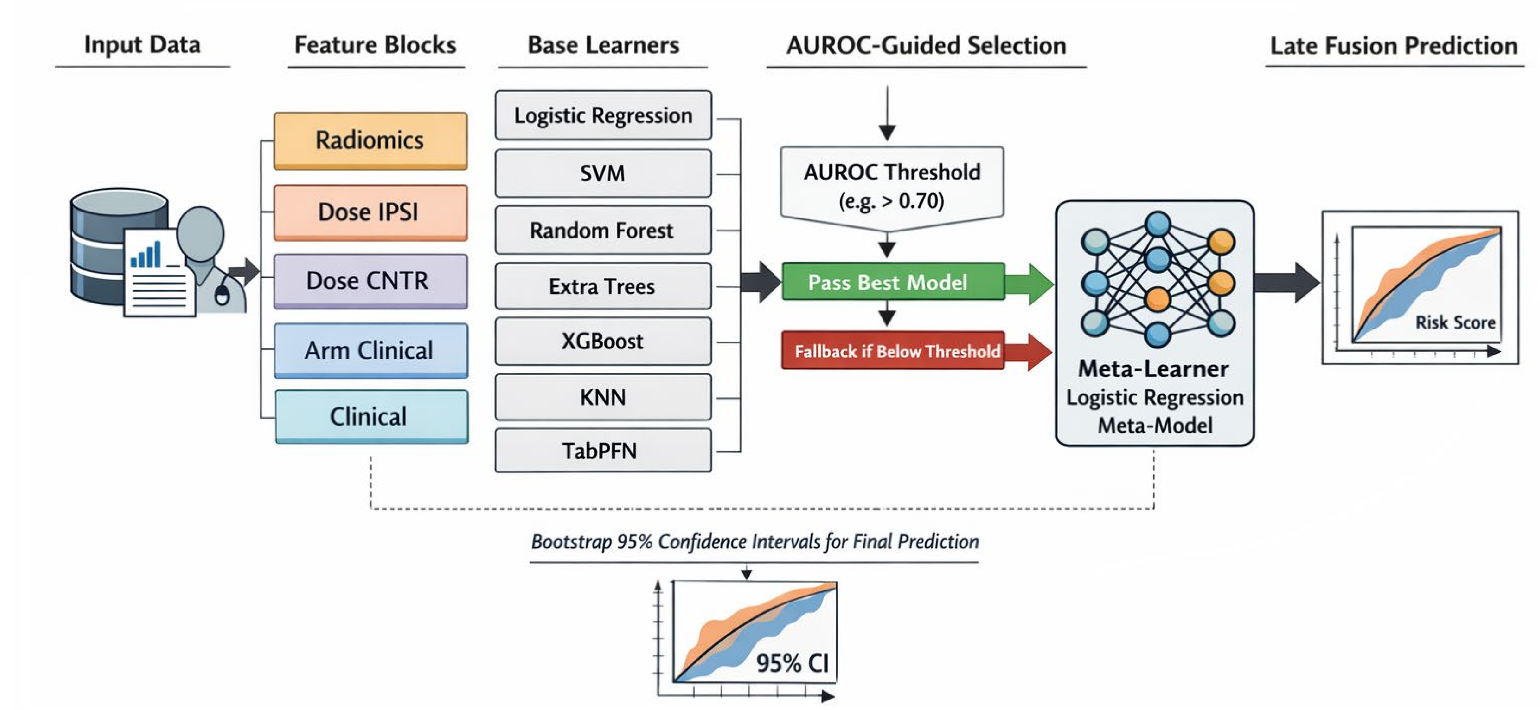


Fig 1. General flowchart

## RESULTS

Predictive performance was evaluated using stratified 5-fold cross-validation on multimodal features derived from planning CTs. Figure 2 and 3 illustrate mean ROC curves and performance, highlighting the improved discriminative performance of late fusion relative to individual feature blocks. The RF-based fusion model achieved the highest performance (AUROC = 0.687 ± 0.067; 95% CI: 0.610–0.766), followed by KNN (AUROC = 0.678 ± 0.063; 95% CI: 0.610–0.755), and SVM (AUROC = 0.654 ± 0.078; 95% CI: 0.519–0.672). Extra Trees achieved AUROC = 0.653 ± 0.052 (95% CI: 0.569–0.729), while logistic regression yielded AUROC = 0.633 ± 0.061 (95% CI: 0.569–0.717).

Transformer-based TabPFN and gradient boosting (XGB) demonstrated lower performance and increased sensitivity to feature sparsity. As shown in Table 1, the bootstrap analysis confirmed that late fusion improved robustness and reduced variability compared with single-modality models.

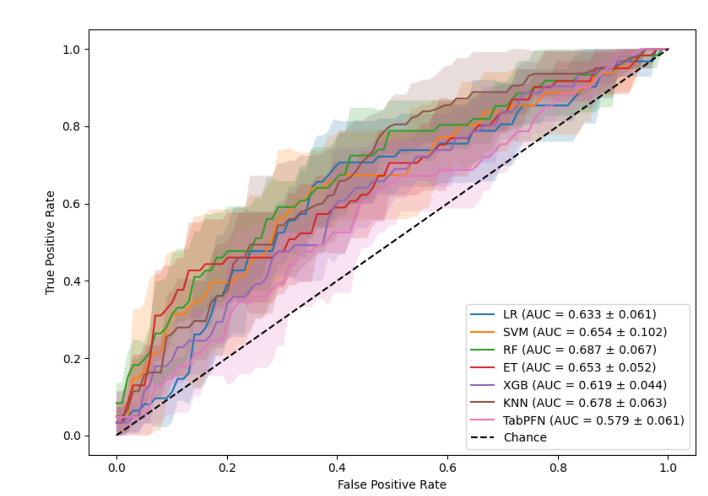

Fig 2. AUROC-Guided Late Fusion Curves

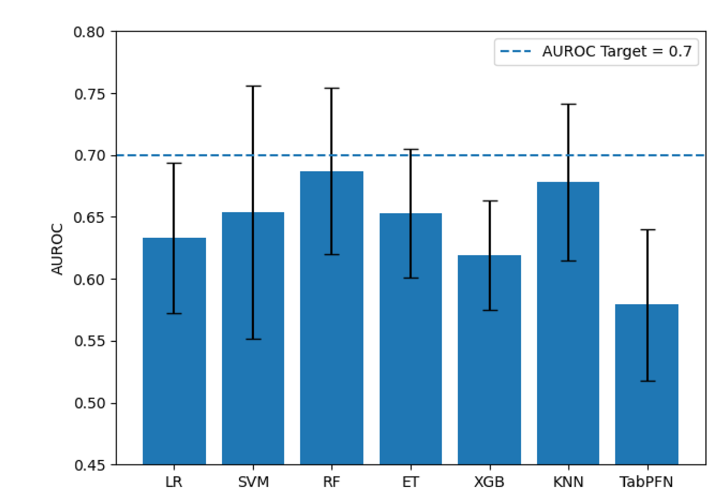


Fig 3. AUROC-Guided Late Fusion Performance

| Base Learner | Mean AUROC ± SD | 95% CI |
|---|---|---|
| RF | 0.687 ± 0.067 | 0.610–0.766 |
| KNN | 0.678 ± 0.063 | 0.610–0.755 |
| ET | 0.653 ± 0.052 | 0.569–0.729 |
| SVM | 0.654 ± 0.078 | 0.519–0.672 |
| LR | 0.633 ± 0.061 | 0.569–0.717 |
| XGBoost | 0.619 ± 0.044 | 0.538–0.691 |
| TabPFN | 0.579 ± 0.061 | 0.503–0.655 |

Table 1 Bootstrap analysis with 95% Confidence Interval

## CONCLUSION

The proposed late fusion meta-learning enables robust and interpretable integration of radiomic, dosimetric, and clinical data for RP prediction. This approach improves pre-treatment risk stratification and shows promise for supporting personalized treatment planning in lung cancer radiotherapy.