Self-Supervised Bayesian Multimodal Learning for Uncertainty-aware Prediction of Radiation Pneumonitis

Jin Yang, PhD[1]; Jing Wang, PhD[1]; Tian Liu, PhD[1]; Ming Chao, PhD[1]
[1]Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York, NY, 10029

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

ABSTRACT

This study aims to develop a self-supervised multimodal framework for radiation pneumonitis (RP) risk prediction while employing Bayesian inference to explicitly quantify uncertainty, thereby enhancing interpretability and supporting clinical decision-making in thoracic radiation therapy.

INTRODUCTION

- Radiation pneumonitis (RP) is a major dose-limiting toxicity in thoracic cancer radiation therapy (RT) and can substantially impair survivors' quality of life.
- Traditional lung dose-volume histograms (DVHs) enforce dose constraints to reduce RP risk during treatment planning.
- Existing machine learning and deep learning methods rely on supervised learning and employ deterministic models to produce point estimates without uncertainty estimation.
- However, uncertainty quantification is necessary for interpretability and reliability of DL models in clinical practice.

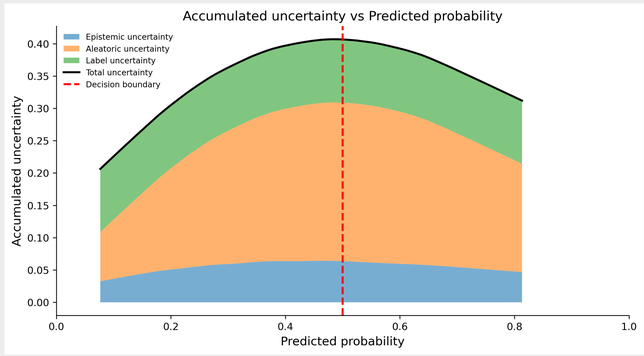
METHODS AND MATERIALS

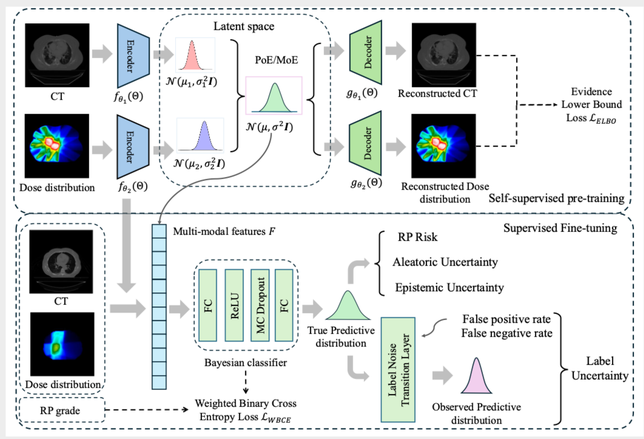
- Pretrained a multimodal variational autoencoder using self-supervised learning from planning CTs and spatial dose distributions
- Extracted modality-specific features using encoders and fused them in the latent space to form unified multimodal representations
- Trained a Bayesian classifier for supervised RP risk prediction
- Employed Monte Carlo dropout to approximate Bayesian inference through repeated stochastic forward passes, yielding a predictive distribution for each patient.
- Calculated the mean of the predictive distribution as the final prediction
- Quantified multiple sources of uncertainty:
  - **Epistemic uncertainty** was estimated as the variance across MC samples.
  - **Aleatoric uncertainty** was derived from the expected data noise captured in the predictive distribution.
  - **Label uncertainty** was explicitly modeled using a label noise transition layer to estimate ambiguity of labels from stochastic predictions.
- Evaluation materials:
  - 416 patients from RTOG 0617
  - RP using a grade ≥ 2 threshold

RESULTS

- CNN-based encoders with PoE achieved the best predictive performance, outperforming MoE.
- For ViT-based encoders, PoE slightly exceeded MoE, though overall performance remained lower than CNN models.
- Using the optimal CNN+PoE strategy, uncertainty are quantified as:
  - Epistemic ([0.0272, 0.0743])
  - Aleatoric ([0.0702, 0.2473])
  - Label ([0.0976, 0.0977])

Model architectures		Latent fusion		Stochastic process		AUC	Specificity	Sensitivity
CNN	ViT	PoE	MoE	100	200			
✓		✓		✓		83.86	80.85	78.69
✓		✓			✓	83.77	81.13	80.33
✓			✓	✓		82.32	81.13	72.13
✓			✓		✓	82.18	81.13	73.77
	✓	✓		✓		80.38	80.56	63.93
	✓	✓			✓	81.21	81.69	67.21
	✓		✓	✓		80.08	79.72	70.49
	✓		✓		✓	79.83	79.15	67.21





DISCUSSION

Three representative cases illustrate their clinical relevance

1. A false-positive case (664275) (RP grade 0; prediction 0.569249) exhibited highest epistemic uncertainty, indicating limited model knowledge and the need for expert review.
2. Another misclassified case (547163) (RP grade 0; 0.510142) showed highest aleatoric but low epistemic uncertainty, suggesting reliable modeling but intrinsically ambiguous input data.
3. Case 440507 (RP grade ≥ 2) had the highest label uncertainty despite low epistemic and aleatoric uncertainties, indicating potential annotation ambiguity, such as borderline toxicity or inter-observer variability.

Case #	RP grade	RP Risk	Epistemic Uncertainty	Aleatoric Uncertainty	Label Uncertainty
0617-664275	0	0.569249	0.074250	0.239691	0.097648
Percentile			100.00	75.48	80.53
0617-547163	0	0.510142	0.051010	0.247295	0.097640
Percentile			15.14	100.00	74.04
0617-440507	2	0.813245	0.053263	0.149041	0.097680
Percentile			22.12	8.41	100

CONCLUSIONS

- In this work, we introduced a Bayesian multi-modal framework that integrates CT imaging features with dosimetric data to predict risk of RP.
- Beyond providing risk stratification, our model provides quantitative measures of epistemic uncertainty, aleatoric uncertainty, and label uncertainty.
- These output on risk predictions and corresponding uncertainty enables clinicians to distinguish between high-risk cases and high-uncertainty cases, facilitating more proactive and individualized toxicity management.
- Following independent external validation, this framework holds the potential to serve as a robust decision-support tool in the radiotherapy clinic and significantly enhance personalized treatment planning and toxicity management.

CONTACT

Jin Yang
Ichan School of Medicine at Mount Sinai
New, NY, 10029
jin.yang@mountsinai.org
3142038098