Reconstructing Delivered Dose In Real Time: A Beam Physics-Embedded, Language-Model-Driven Approach

Yuli Wang[1*], Jing Wang[1*], Jiahan Zhang[1], Yuhang Sun[2], Ming Chao[1], Junyi Xia[1], Ruirui Liu[2], Shuo Wang[2], Yu Lei[2], Karyn A Goodman[1], Tian Liu[1] and Yang Lei[1]
1 Icahn School of Medicine at Mount Sinai, New York, NY;
2 University of Nebraska Medical Center, Omaha;

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

ABSTRACT

Accurate dose prediction accelerates radiotherapy planning and QA, but image-only networks ignore beam-physics priors, leaving residual errors along beam paths, field edges, and the high-dose core. Conventional treatment planning systems store optimized fluence but do not provide a direct mapping between delivered control points and physically meaningful beam fluence, limiting real-time dose estimation. We introduce a two-stage framework that (1) uses an LLM agent to retrieve AAPM beam-profile guidance and tokenize the 6 MV beam profile into text, and (2) conditions a BERT-based U-Net on those tokens to refine beam-eye-view (BEV) dose. Model training employs a composite loss combining mean absolute error, mean squared error, edge-aware loss, and high-dose region-weighted loss, with weighting coefficients optimized using 3D gamma analysis.

On an internal lung cohort (train 135 / hold-out test 45) and an external prostate cohort (5 cases) from MSH, the model yields the smallest residual difference maps among CNN and CLIP baselines and transfers across anatomy and site.

INNOVATIONS

The key innovation of this work is the explicit integration of an LLM as a clinical knowledge encoder within a dose prediction pipeline.

In addition, we introduce a customized physics-aware loss function for model training.

By decoupling clinical knowledge representation from numerical model training, the framework enables scalable updates and safer translation to real-time or adaptive radiotherapy workflows.

KEY CONTRIBUTIONS

- LLM tokenizes beam-physics knowledge into text.
- Text-conditioned BERT-based U-Net refines BEV dose.
- Physics-aware loss: consistency + edge + high-dose.
- Cross-site eval: internal lung + external prostate.

EVALUATION & RESULT

- Beam-profile validation
- Image similarity (NPE & MAE)
- Gamma analysis
- A mean 3D gamma passing rate (3%, 2 mm/10%) of 0.975 ± 0.056, outperforming 3D U-Net (0.931 ± 0.113), Dose-Net (0.939 ± 0.120), and CLIP-Unet (0.945 ± 0.087) ($p < 0.05$).

CONTACT

Yuli Wang or Yang Lei
yuliwang@mountsinai.org
Yang.Lei@mountsinai.org
Dept. of Radiation Oncology
Mount Sinai (MSH)

INTRODUCTION

Dose prediction is a workhorse for treatment planning, plan QA, and automation in external-beam radiotherapy. Convolutional models (e.g. 3D-U-Net) operate purely on images and cannot consume the beam characteristics, such as the 6 MV beam profile in AAPM guidance, that govern dose deposition, leaving systematic residuals at beam paths, field edges, and the high-dose core.

We ask: How can beam-physics knowledge be injected into a dose-prediction network in a usable form, without hand-engineering physics features?

Our answer: let an LLM convert beam knowledge into text tokens and condition the network on them.

METHODS & MATERIALS

Fig. 1: Network framework

Stage 1 — LLM tokenization.

An LLM agent retrieves AAPM beam-profile and characteristics and encodes the 6 MV beam profile into text tokens.

Stage 2 — Beam-wise refinement.

A beam-eye-view (BEV) generator projects CT, dose and fluence into BEV stacks; a BERT-based U-Net, conditioned by a text encoder on the tokenized profile, predicts refined BEV dose.

Loss Function.

L_{Final} = L1/MSE consistency + λ_2·edge-enhancement + λ_3·high-dose-enhancement.

Data & evaluation.

Internal lung (train 135 / test 45) and external prostate (5) cohorts from MSH; beam-profile validation, image-similarity (NPE & MAE) and gamma analysis.

135	45	5
Internal train · lung	Internal test · lung	External test · prostate

RESULTS

Loss ablation (Fig 2).

Residual |prediction − GT| difference maps shrink monotonically: baseline → +MSE → +edge → full model; beam-path and edge errors dominating the baseline are largely removed.

Comparison with State-of-art Method (Fig 3).

Against 3D-U-Net, Dose-Net and CLIP-U-Net, the proposed model's line profile tracks ground truth most closely through the high-dose peak and yields the smallest residual on both a lung and a prostate case.

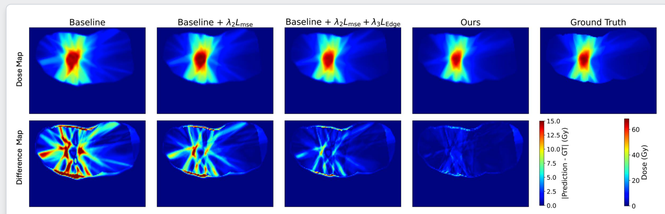

Fig 2. Loss ablation — dose maps (top) and difference maps (bottom): Baseline → +λ_2L_MSE → +λ_3L_Edge → full model → ground truth.

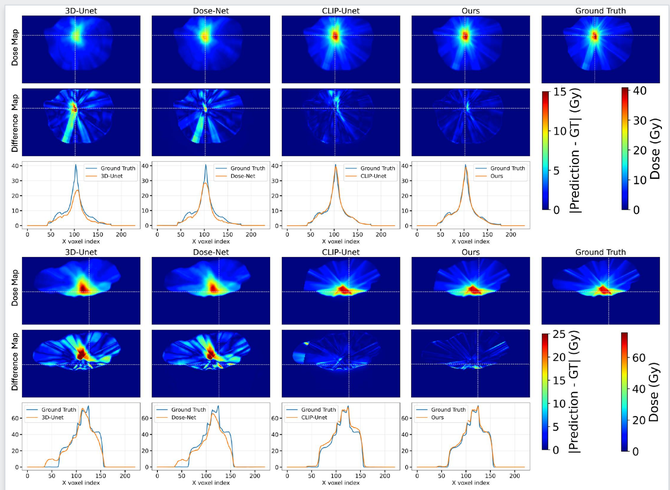

Fig 3. Comparison vs. baselines — dose maps, difference maps and line profiles for 3D-U-Net, Dose-Net, CLIP-U-Net, the proposed model and ground truth

KEY FINDINGS

Smallest residual

Lowest |prediction − GT| among 3D-U-Net, Dose-Net and CLIP-U-Net.

Method	3D Gamma	2D Gamma	MAE (Gy)	NPE
3D-U-Net	0.931 ± 0.113	0.925 ± 0.135	2.862 ± 1.495	0.067 ± 0.021
Dose-Net	0.939 ± 0.120	0.933 ± 0.135	2.701 ± 1.131	0.062 ± 0.018
CLIP-U-Net	0.945 ± 0.087	0.939 ± 0.099	2.676 ± 1.023	0.058 ± 0.016
Ours	**0.975 ± 0.056**	**0.962 ± 0.073**	**2.243 ± 0.703**	**0.047 ± 0.012**

Monotonic gain

Each loss term (MSE → edge → high-dose) lowers residual error.

Method	3D Gamma	MAE (Gy)
Baseline (L_{MAE})	0.905 ± 0.102*	2.801 ± 1.033*
Baseline + L_{MSE}	0.934 ± 0.071*	2.676 ± 1.012*
Baseline + L_{MSE} + L_{Edge}	0.965 ± 0.058	2.356 ± 0.947*
Baseline + L_{MSE} + L_{Edge} + L_{High} (Ours)	**0.975 ± 0.056**	**2.243 ± 0.703**

Cross-site transfer

Lung-trained model reproduces external prostate dose.

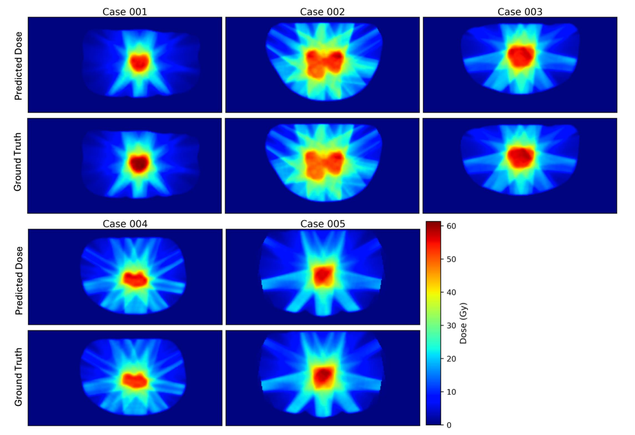

Fig 4. External prostate cases — predicted vs. ground-truth dose (cases 001–005); trained on lung, the model reproduces prostate distributions.

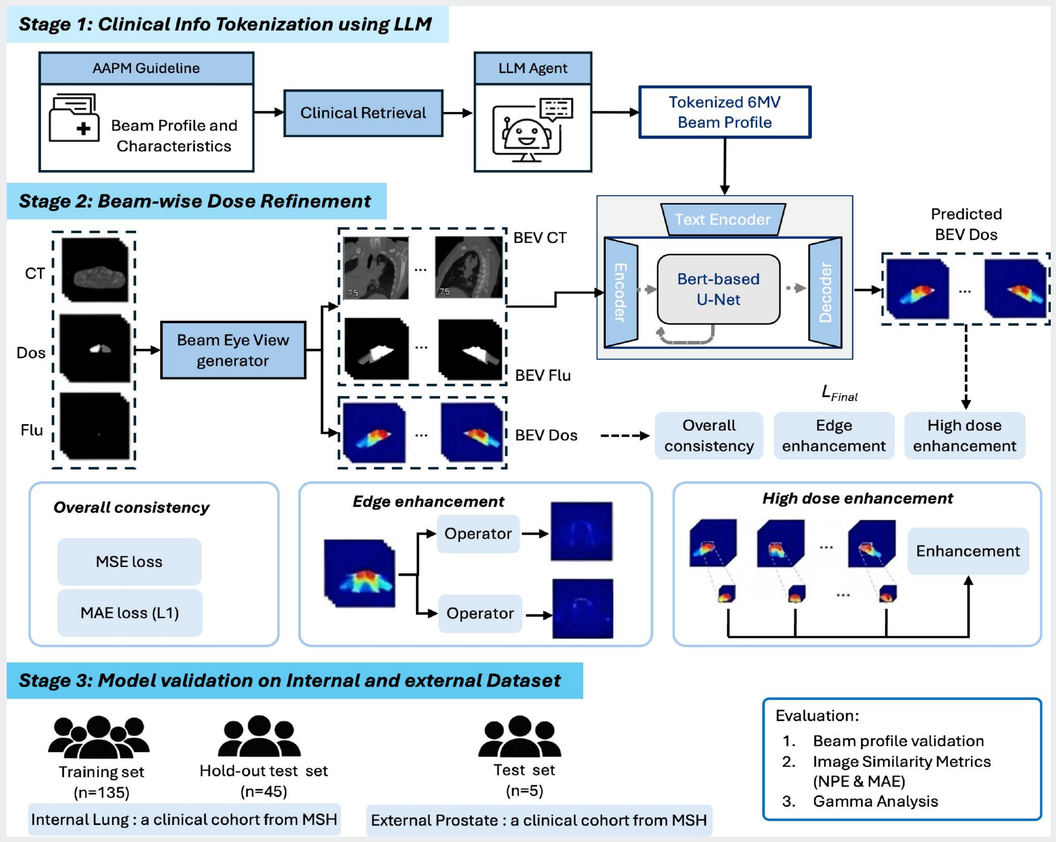

Fig 1. Two-stage framework: LLM tokenizes the 6 MV beam profile; a text-conditioned BERT-based U-Net refines beam-eye-view dose; validated on internal lung and external prostate cohorts.

DISCUSSION & CONCLUSIONS

We presented a physics-informed dose reconstruction framework that integrates LLM–encoded beam knowledge with a multimodal deep learning architecture.

By incorporating structured clinical knowledge and physics-aware loss functions into BEV dose modeling, the proposed approach improves dose prediction accuracy compared with existing learning-based methods.

Experimental results demonstrate strong agreement between predicted and reference dose distributions, with improved gamma passing rates and reduced voxel-wise dose errors.

The model also generalizes across anatomical sites, successfully reconstructing prostate dose distributions despite being trained only on lung plans.

These findings highlight the potential of combining knowledge-driven AI with physics-aware modeling for radiotherapy dose estimation.

REFERENCES

- G. Baiocco, S. et. al., A matter of space: how the spatial heterogeneity in energy deposition determines the biological outcome of radiation exposure, Radiation and Environmental Biophysics 61, 545–559 (2022).
- A. Ahnesj¨o, et. al., A pencil beam model for photon dose calculation, Medical physics 19, 263–273 (1992).
- A. Van Esch, et. al., Testing of the analytical anisotropic algorithm for photon dose calculation, Medical physics 33, 4130–4148 (2006).
- M. Ljungberg, et. al., A 3-dimensional absorbed dose calculation method based on quantitative SPECT for radionuclide therapy: evaluation for 131I using Monte Carlo simulation, Journal of Nuclear Medicine 43, 1101–1109 (2002).
- I. J. Chetty et al., Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning, Medical physics 34, 4818–4853 (2007).
- X. Jia, et. al., GPU-based fast Monte Carlo simulation for radiotherapy dose calculation, Physics in Medicine & Biology 56, 7017–7031 (2011)