Toward an AI-Automated Radiation Therapy Workflow: A Dual-Model Approach to Target Volume Contouring and Dose Optimization

Nayoon Justina Lee, Brook Byrd, Maksym Sharma, Bolin Li, Nicholas Molina, William Green, and Rafe McBeth
[1]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, [2]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA

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

ABSTRACT

Purpose: Prostate cancer male reproductive system malignancy and the most common male cancer treated by radiation. Given the high incidence rate and complexities of treatment planning, artificial intellige nce (AI) solutions in medical physics are essential to reducing healthcare system burden and improving access to patient-specific care. This study developed and validated two AI models to address resource-intensive bottlenecks in the care pathway: target volume (TV) contouring and dose optimization (DO), both of which are highly manual and vulnerable to inter-observer and intra-observer variations.

Methods: Two AI models were developed: 1) a TV segmentation model utilizing U-Net, a convolutional neural network architecture, and 2) a RapidPlan DO model with knowledge-based planning architecture. Both models were trained on 21 intermediate prostate cancer cases for risk-specific optimization. The TV model was validated by comparing its contours to human-expert contoured clinical target volumes using the DICE similarity coefficient. Performance of the DO model was evaluated on external beam radiation therapy protocols and cross-evaluated against clinical and outsourced models' plans.

Results: The TV model achieved a DICE score of 0.842 in a validation case, exceeding the benchmark average of 0.81 observed with TheraPanacea across 90 prostate cancer patients, despite facing significant challenges from extreme class imbalance (average of 0.0562% foreground voxels) and wide variation in prostate volume. The RapidPlan DO model demonstrated performance comparable to the clinical plan, even in a case where the outsourced model failed. However, model limitations were observed in patients with anatomical features outside the training dataset thresholds, highlighting the impact of data scarcity (n=21 compared to a typical n>60).

Conclusion: AI-driven automation for the medical physics workflow in the treatment of prostate cancer demonstrates high potential for decreasing time spent on manual, iterative tasks and reducing risks of plan quality variabilities while meeting the growing demand for efficient, patient-specific care

INTRODUCTION

Radiation therapy is one of the most precise treatment methods in oncology. However, the treatment planning process remains time-consuming, highly manual, and subject to variability. Going from computed tomography simulation to the administering of treatment ("sim to start") often requires several hours to days.[1]

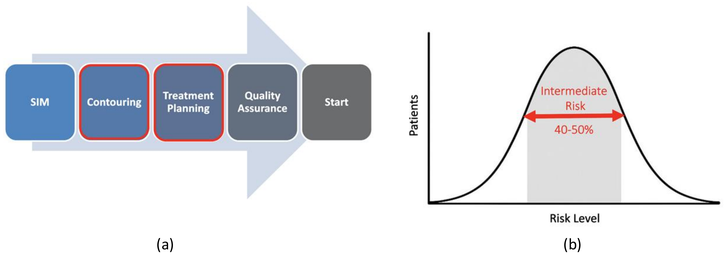
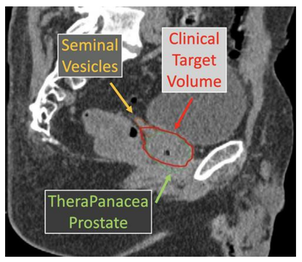


Figure 1: AI-automation will target (a) the treatment workflow in radiation oncology and (b) the high volume intermediate risk group.

In the the care path, target contouring and treatment planning are particularly susceptible to inter- and intra-observer variabilities. These steps are also the most resource-intensive components of the workflow. Artificial intelligence (AI) methods, trained on large repositories of historical clinical data, can accelerate these areas of the care path while improving efficiency and the quality of care.[2]

KEY CLINICAL MISMATCH



Target and Critical Normal Tissue Constraints

Structure Name	DVH Objective	Evaluator	Variation Acceptable	Priority	Evaluate BED/EQD2
PTV_7000	D97%[cGy]	>= 6930	6650	3	☐
PTV_7000	D99%[cGy]	>= 6300	5985	3	☐
CTV_7000	D99%[cGy]	>= 7000	6650	1	☐
Rectum	V3600cGy[%]	<= 50	<=55	2	☐
Rectum	V4000cGy[%]	<= 40	<=50	2	☐
Rectum	V5300cGy[%]	<= 20	<=25	2	☐
Rectum	V6100cGy[%]	<= 10	<=15	2	☐
Rectum	D1cc[cGy]	<= 7210	<=7350	2	☐
Bladder	V4000cGy[%]	<= 50	<=55	2	☐
Bladder	V4400cGy[%]	<= 40	<=45	2	☐
Bladder	V5700cGy[%]	<= 20	<=25	2	☐
Bladder	V6100cGy[%]	<= 15	<=20	2	☐
Bowel_Small	V3200cGy[%]	<= 50	<=55	2	☐
Bowel_Small	V3600cGy[%]	<= 25	<=30	2	☐
Bowel_Small	D1cc[cGy]	<= 4830	<=4950	2	☐
Bowel_Large	D1cc[cGy]	<= 5756	<=5928	2	☐
Femur_Head_L	V4000cGy[%]	<= 5	<=10	3	☐
Femur_Head_R	V4000cGy[%]	<= 5	<=10	3	☐
Body	Max[cGy]	<= 7490		3	☐
External_Pat	Max[cGy]	Choose operator and enter numeric value		3	☐

(b)

Figure 2: Key shortcomings of currently employed clinical solutions including (a) TheraPanacea's exclusion of the seminal vesicles from the clinical target volume contour and (b) the Washington University in St. Louis (WUSTL) model's exclusion of the small and large bowel structures from its training parameters.

METHODS

TARGET VOLUME CONTOURING MODEL

- A 3D U-Net was implemented with MONAI framework and trained on 21 cases with full bladder, spacer, fiducials, and seminal vesicles included in the target volume.
- DiceFocalLoss was used to resolve the high average class imbalance of 0.056%.

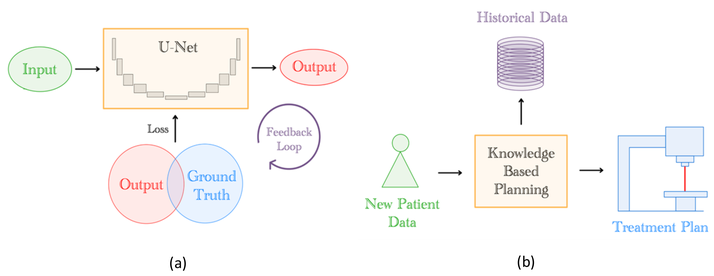


Figure 3: Framework for AI-automated (a) target volume contouring and (b) dose optimization.

DOSE PLANNING OPTIMIZATION MODEL

- Varian's RapidPlan was used to train a knowledge-based planning model to predict achievable dose-volume histograms and objectives.
- Training was done on 21 patients of the defined cohort with 19 patients treated with 7000 cGy / 28 fx and 2 patients with 7920 cGy / 44 fx.

RESULTS

TARGET VOLUME CONTOURING MODEL

- Validation case achieved a Dice score of 0.842 which exceeds the benchmark average of 0.81 observed with TheraPanacea across 90 prostate cancer patients.

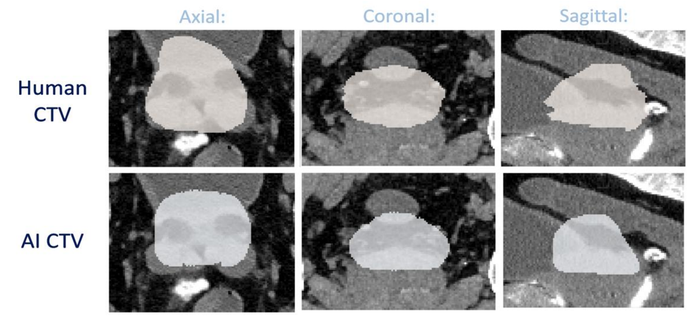


Figure 4: Axial, coronal, and sagittal view comparison of the CTV created by a human expert versus the CTV generated by the U-Net model.

DOSE PLANNING OPTIMIZATION MODEL

- In the first validation case, the in-house model's plan passes EBRT protocols while the WUSTL plan fails which demonstrates that the Penn-made model achieves clinical acceptability in cases where the WUSTL model fails.
- In the second validation case, both the in-house model's plan and the WUSTL plan do not pass EBRT protocols, but the in-house plan has a more favorable dose distribution for rectum, bladder, and body.

StructureName	DVHObjective	Evaluator	Variation	Priority	Pt 1 Prostate: Clinical	Pt 1 Prostate: Validation	Pt 1 Prostate: WUSTL
PTV_7000	D97%[cGy]	>=6930	6650	3	6985.0 cGy	6985.0 cGy	7020.2 cGy
PTV_7000	D99%[cGy]	>=6300	5985	3	6880.9 cGy	6890.9 cGy	6906.7 cGy
CTV_7000	D99%[cGy]	>=7000	6650	1	7074.7 cGy	7023.0 cGy	7044.4 cGy
Rectum	V3600cGy[%]	<=50	55	2	25.05 %	25.00 %	26.44 %
Rectum	V4000cGy[%]	<=40	50	2	21.22 %	21.12 %	22.87 %
Rectum	V5300cGy[%]	<=20	25	2	11.40 %	11.93 %	13.83 %
Rectum	V6100cGy[%]	<=10	15	2	6.71 %	7.29 %	8.83 %
Rectum	D1cc[cGy]	<=7210	7350	2	7060.8 cGy	7123.9 cGy	7180.3 cGy
Bladder	V4000cGy[%]	<=50	55	2	15.63 %	15.66 %	15.42 %
Bladder	V4400cGy[%]	<=40	45	2	13.81 %	13.91 %	13.77 %
Bladder	V5700cGy[%]	<=20	25	2	8.92 %	9.09 %	9.02 %
Bladder	V6100cGy[%]	<=15	20	2	7.73 %	7.91 %	7.85 %
Bowel_Small	V3200cGy[%]	<=50	55	2	0.00 %	0.00 %	0.00 %
Bowel_Small	V3600cGy[%]	<=25	30	2	0.00 %	0.00 %	0.00 %
Bowel_Small	D1cc[cGy]	<=4830	4950	2	97.4 cGy	96.4 cGy	100.1 cGy
Bowel_Large	D1cc[cGy]	<=5756	5928	2	458.9 cGy	446.0 cGy	464.8 cGy
Femur_Head_L	V4000cGy[%]	<=5	10	3	0.00 %	0.00 %	0.00 %
Femur_Head_R	V4000cGy[%]	<=5	10	3	0.00 %	0.00 %	0.00 %
Body	Max[cGy]	<=7490		3	7391.5 cGy	7477.5 cGy	7536.5 cGy
External_Pat	Max[cGy]			3	7391.5 cGy	7477.5 cGy	7536.5 cGy

Table 1: The first validation case's EBRT scorecard comparing the human expert created plan labelled as "Clinical," the Penn-made model's generated plan labelled as "Validation," and the WUSTL model's generated plan labelled as "WUSTL."

StructureName	DVHObjective	Evaluator	Variation	Priority	Pt 5 Prostate: Clinical	Pt 5 Prostate: Validation	Pt 5 Prostate: WUSTL
PTV_7000	D97%[cGy]	>=6930	6650	3	6952.3 cGy	6979.0 cGy	6965.0 cGy
PTV_7000	D99%[cGy]	>=6300	5985	3	6790.1 cGy	6884.7 cGy	6877.4 cGy
CTV_7000	D99%[cGy]	>=7000	6650	1	7084.4 cGy	6993.8 cGy	6967.4 cGy
Rectum	V3600cGy[%]	<=50	55	2	29.29 %	26.88 %	27.10 %
Rectum	V4000cGy[%]	<=40	50	2	24.99 %	23.74 %	24.06 %
Rectum	V5300cGy[%]	<=20	25	2	14.83 %	15.62 %	15.95 %
Rectum	V6100cGy[%]	<=10	15	2	9.91 %	11.21 %	11.35 %
Rectum	D1cc[cGy]	<=7210	7350	2	7150.6 cGy	7161.9 cGy	7174.7 cGy
Bladder	V4000cGy[%]	<=50	55	2	17.35 %	15.84 %	15.84 %
Bladder	V4400cGy[%]	<=40	45	2	15.07 %	13.81 %	13.95 %
Bladder	V5700cGy[%]	<=20	25	2	9.47 %	8.92 %	9.06 %
Bladder	V6100cGy[%]	<=15	20	2	8.15 %	7.76 %	7.83 %
Bowel_Small	V3200cGy[%]	<=50	55	2	0.00 %	0.00 %	0.00 %
Bowel_Small	V3600cGy[%]	<=25	30	2	0.00 %	0.00 %	0.00 %
Bowel_Small	D1cc[cGy]	<=4830	4950	2	117.4 cGy	117.8 cGy	116.8 cGy
Bowel_Large	D1cc[cGy]	<=5756	5928	2	447.5 cGy	453.0 cGy	452.3 cGy
Femur_Head_L	V4000cGy[%]	<=5	10	3	0.00 %	0.00 %	0.00 %
Femur_Head_R	V4000cGy[%]	<=5	10	3	0.00 %	0.00 %	0.00 %
Body	Max[cGy]	<=7490		3	7463.2 cGy	7561.6 cGy	7617.6 cGy
External_Pat	Max[cGy]			3	7463.2 cGy	7561.6 cGy	7617.6 cGy

Table 2: The second validation case's EBRT scorecard comparing the human expert created plan labelled as "Clinical," the Penn-made model's generated plan labelled as "Validation," and the WUSTL model's generated plan labelled as "WUSTL."

CONCLUSIONS

AI-automated target volume contouring and dose optimization is a dual-model approach that has the potential to reduce traditional trade offs between time and quality in radiation therapy planning. Even on a limited training set of 21 cases, both models show strong promise, and a larger dataset, scaled up by 10-20 times, will enhance generalizability for real clinical deployment.

REFERENCES

1. Chowdhry, V. K., Daniel, Z., & Hackett, R. (2021). Departmental quality initiative to establish turnaround times from simulation to treatment. *Technical Innovations & Patient Support in Radiation Oncology, 19,* 37–40.
2. Huynh, E., Hosny, A., Guthier, C., Bitterman, D. S., Petit, S. F., Haas-Kogan, D. A., Kann, B., Aerts, H. J. W. L., & Mak, R. H. (2020). Artificial intelligence in radiation oncology. *Nature Reviews Clinical Oncology, 17*(12), 771–781.

ACKNOWLEDGEMENTS

This work was supported by the Summer Undergraduate Program Educating Radiation Scientists, an NIH/NCI funded program.

CONTACT

Nayoon Justina Lee
University of Pennsylvania
nayoon.lee@pennmedicine.upenn.edu