Successful Turing Test of a Physician-Specific Knowledge Based Planning Model for Prostate Radiotherapy

James M. Lamb[1], John S. Ginn[2], Christopher R. King[1], Nzhde Agazaryan[1], Michael L. Steinberg[1], Amar U. Kishan[1], Shun Tsai[1]

(1) Department of Radiation Oncology, University of California, Los Angeles (2) Department of Radiation Oncology, Washington University in St. Louis

JULY 12–16 2020 VIRTUAL JOINT AAPM | COMP MEETING
Improving Health QUALITY, Increasing Global IMPACT
EASTERN TIME [GMT-4]

OBJECTIVES

In recent years substantial efforts have been dedicated to the development of automated intensity modulation radiotherapy (IMRT) treatment planning using knowledge based planning (KBP). Several groups have reported that KBP automated planning produces plans that are clinically acceptable based on quantitative dose-volume metrics. However, very few reports describe blinded physician evaluation of KBP automated plans. Here we report such a study using a commercial implementation of KBP, helping to address a key hurdle to clinical acceptance of KBP automated planning.

METHODS

Patient Cohort

Our institution's clinical plan database was searched for prostate stereotactic body radiation therapy (SBRT) patients treated from 2014-2019. All included patients were treated by a single physician, using a Varian TrueBeam linear accelerator with Millenium 120 multileaf collimator. All plans were generated by a single dosimetrist. Prescriptions were for a total dose of 40 Gy delivered over 5 fractions, normalized to 95% coverage of the planning target volume (PTV). The planning target volumes included the prostate alone, or the prostate plus proximal seminal vesicles. All plans used volumetric modulated arc therapy (VMAT) with two full arcs. Exclusion criteria included the presence of hydrogel spacer (SpaceOAR, Augmentix, Inc. Bedford MA), as well as patients whose treatment fields include pelvic lymph nodes, who underwent prostatectomy surgery prior to treatment, or who previously had a hip replacement. A total of 153 patients meeting all criteria were identified. Data partitioning was performed by randomly selecting patients for model training, validation and testing. A total of 100 patients were used to build the KBP model. A total of 10 and 20 patients were selected for model validation and testing respectively

Model Training and Validation

KBP modeling was performed using RapidPlan in Eclipse (version 15.6.05). Outlier removal was not performed during the training process. Optimization objectives within the model were iteratively tuned to generate plans that closely matched the set of 10 model validation clinical plans. Model validation plans were not included in the model itself. A total of 17 model iterations were performed before model accuracy was deemed acceptable according to quantitative DVH metrics as well as visual assessment of the DVHs and 3-dimensional dose distributions by an experienced planner. Normal organ contours and tuning structures used in the final model, as well as the nature of the constraint generated by the model, are listed in **Table 1**. Five out of the 100 plans did not include all three of the tuning structures used by the model. All 100 plans included the same set of normal organ contours.

Physician Blinded Review

Twenty patient cases were used to test the model performance. These test cases were not used for model building or model validation. A fully automated plan was created from each test case. Two separate evaluations were made. First, the 20 automated plans and the corresponding clinical plans were presented in a blinded fashion to the same physician who originally treated all patients physician in a blinded fashion. The 40 plans were presented to the physician in a randomized order. The plans were graded on a scale of 1-5 where a score ≥ 3 indicated a plan that was clinically acceptable and 5 indicating the best quality. Subsequently, the matched pairs of clinical and automated plans were presented to the physician head-to-head, but still blinded with regard to planning method. The physician then identified the plans as equivalent, one better than the other but the differences were not clinically significant, or one plan is preferable to the other.

Statistical analysis

At our institution both PTV coverage and organs at risk (OAR) dose metrics are used to evaluate plan quality and clinical acceptability. In this work we evaluate the same dosimetric endpoints for clinical acceptability, and statistically to determine whether or not KBP yields significantly different dosimetric values. Specifically, we use a paired t-test to evaluate these differences.

RESULTS

Complete, un-blinded results of the blinded physician evaluate are listed in **Table 2**. For the plan grading under randomized presentation, the mean clinical plan score was 4.5 (min: 4, max: 5, mode: 4.5, rms: 0.28). The mean automated plan score was 4.7 (min: 4.5, max: 5, mode: 4.5, rms: 0.26). In the blinded head-to-head comparison 5 automated plans were graded as better than the clinical plans in a clinically significant way, 5 automated plans were graded as better than the clinical plans but in a clinically non-significant way, and 10 plans were graded as no difference.

DISCUSSION

The key result of this study is that a RapidPlan knowledge-based planning model could be used to produce fully automated prostate-only radiotherapy plans that are essentially indistinguishable from clinical plans according to a blinded physician review. We believe blinded plan comparison is a crucial hurdle to clinical acceptable of automated KBP planning, and thus our study addresses an important gap in the KBP literature because, surprisingly, few such studies have been published. This work was undertaken as a result of our institution's initial experience with clinical KBP. At our institution, KBP for prostate radiotherapy was initially investigated using a publically available RapidPlan model. Physicians specializing in genitourinary radiotherapy determined that the plans produced by the publically available model were clinically acceptable, but were less preferred compared to the manual plans.

Contour	Description	Constraint Type
PTV	Typically 5 mm isotropic margin about CTV, except 4 mm posteriorly.	Fixed dose, volume and priority
CTV	Prostate gland	Fixed dose, volume and priority
Ring	Subtraction of PTV expanded by 45 mm and PTV expanded by 15 mm.	Generated line
Rectum	Rectum	Fixed dose, generated volume, fixed priority (multiple)
Rectum_05	Rectum cropped by a 5 mm expansion of the PTV	Fixed dose, generated volume and priority (multiple)
Rectum_10	Rectum cropped by a 10 mm expansion of the PTV	Fixed dose, generated volume and priority (multiple), and generated line.
Bladder	Bladder	Generated line (preferring OAR)
Femur_L, Femur R	Left and right femoral heads	Fixed volume and dose, generated priority

Table 1: Targets, normal organs, and tuning structures utilized by the KBP model.

Patient	Physician Plan score		Physician Head-to-Head Comparison
	Clinical	Automated	
1	4.5	4.5	No difference
2	4.5	4.5	No difference
3	4.5	5	No difference
4	4.5	4.5	No difference
5	5	4.5	No difference
6	4.5	4.5	Auto-plan better, clinically significant
7	4.5	5	No difference
8	4.5	5	No difference
9	4.5	4.5	Auto-plan better, clinically non-significant
10	4.5	5	No difference
11	4	4.5	Auto-plan better, clinically non-significant
12	4	4.5	Auto-plan better, clinically non-significant
13	4.5	5	Auto-plan better, clinically non-significant
14	4	5	Auto-plan better, clinically non-significant
15	4.5	4.5	Auto-plan better, clinically significant
16	4.5	4.5	Auto-plan better, clinically significant
17	4	5	Auto-plan better, clinically significant
18	4.5	4.5	Auto-plan better, clinically significant
19	5	5	No difference
20	4.5	5	No difference

Table 2: Plan evaluation for each clinical and knowledge-based treatment plan pair.

UCLA RADIATION ONCOLOGY
Division of Medical Physics

Email: jlamb@mednet.ucla.edu