Dosimetric Scorecards to Create and Validate a Head and Neck Knowledge Based Planning Model from a Foundation Model for Subsite Specific Application on Cases with Bilateral Parotid Involvement to Create High Quality Treatment Plans without Revision

JULY 21–25 | LOS ANGELES, CA
AAPM 2024
66TH ANNUAL MEETING & EXHIBITION

R. Clark[1], A. Magliari[1], L. Rosa[1], J. Paisley[2] and S. Beriwal[1]
1 Office of Medical Affairs, Varian, A Siemens Healthineers Company, Palo Alto, CA
2 Novant Health New Hanover Regional Medical Center, Radiation Oncology, Wilmington, NC

NOVANT HEALTH

varian
A Siemens Healthineers Company

INTRODUCTION

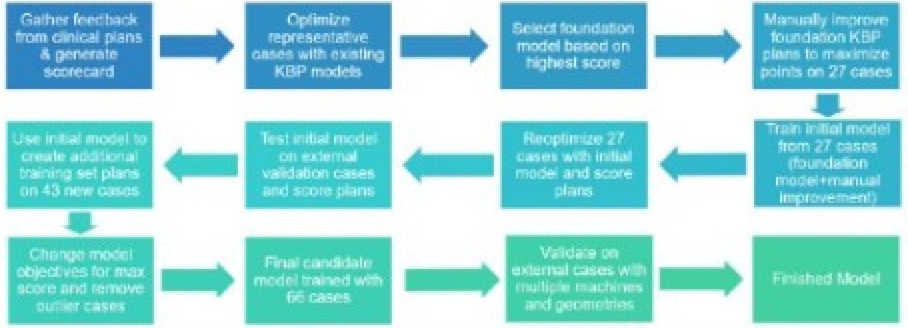
Several institutions have created both sequential boost and SIB Head and Neck Knowledge Based Planning (KBP) RapidPlan models over the last decade, however, few are publicly shared and those that are, often lack sufficient documentation to assess their performance without time consuming manual evaluation [1].

AIM

The aim is to develop and validate a KBP model specifically for head-and-neck cases with Bilateral Parotid Involvement (BPI), to improve overall dosimetric plan quality and parotid sparing. A dosimetric scorecard guided model creation and validation, ensuring comprehensive articulation of clinical intent and trade-off preferences between competing metrics [2].

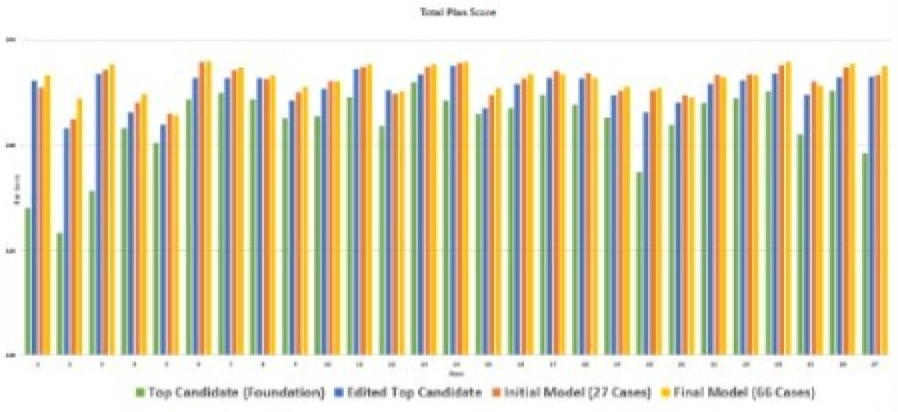
METHOD

The HN-SIB-BPI model was developed using an initial dataset of 27 BPI cases identified through AI clustering. A dosimetric scorecard was then designed to reflect clinical priorities and trade-offs beyond the NRG HN004 protocol. A foundation model was selected from four existing RapidPlan models based on scorecard performance. The training set was created using the foundation model as a starting point and improved upon with new structures and further OAR sparing through manual adjustments to optimization objectives with scorecard feedback. The initial model was then expanded with 39 additional BPI cases, all replanned with the initial model and manual optimization adjustments, resulting in a final model of 66 cases. DVH predictions improved, while optimization and normal tissue objective priorities were further tuned. Validation was performed on 8 external cases (4 with 3 targets, 4 with 2 targets), tested on both Halcyon and TrueBeam delivery systems. The model's performance was evaluated for both Eclipse (versions 15.6 & 18.0) and assessed using the dosimetric scorecard. Results were compared against the foundation model's performance and at each iterative step in the model's development process until completion.


Gather feedback from clinical plans & generate scorecard
Optimize representative cases with existing KBP models
Select foundation model based on highest score
Manually improve foundation KBP plans to maximize points on 27 cases
Train initial model from 27 cases (foundation model+manual improvement)
Reoptimize 27 cases with initial model and score plans
Test initial model on external validation cases and score plans
Use initial model to create additional training set plans on 43 new cases
Change model objectives for max score and remove outlier cases
Final candidate model trained with 66 cases
Validate on external cases with multiple machines and geometries
Finished Model

RESULTS

HN-SIB-BPI demonstrated significant improvements over the foundation model in a 27-case subset of the training set. The average scorecard score increased from 210.5 to 231.7, with mean parotid dose reductions of 1.05 Gy for ipsilateral and 1.58 Gy for contralateral glands. External validation using 4-arc VMAT showed excellent performance on both Halcyon and TrueBeam delivery systems. For Halcyon, 3PTV cases scored 221.38/260 and 2PTV cases scored 196.1/228.5. TrueBeam results were similar, with 3PTV cases scoring 222.01/260 and 2PTV cases scoring 202.24/228.5. The model also performed well in 19-field IMRT validation on Halcyon, with 3PTV cases scoring 218.5/260 and 2PTV cases scoring 200.1/228.5. These results demonstrate the model's excellent performance and compatibility across different delivery systems and techniques for both VMAT and IMRT.

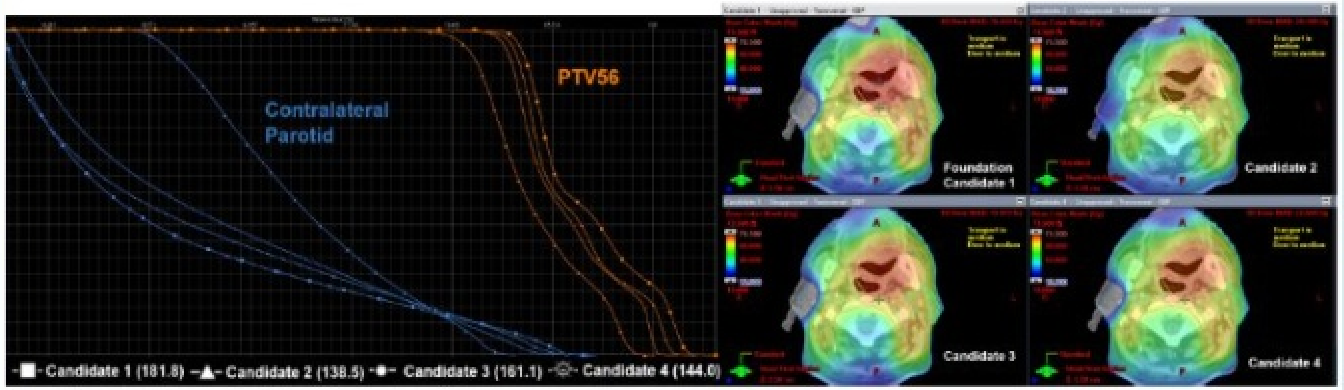


Dosimetric scorecard performance comparison between top candidate, foundation model and iterations of the HN-SIB-BPI model for 27 cases. Average scores: Top candidate-210.15; Edited candidate-226.96; Draft model-230.1; Final model-231.7.

Key Target Metrics	PTV70 Volume at 70Gy [%]		PTV70 Dose at 0.03CC [Gy]		V70 Dose at 0.03CC [Gy]		PTV56 Volume at 56Gy [%]		PTV56-PTV63 Volume at 58.8Gy [%]		RingPTV56 Dose at 0.03CC [Gy]	
	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1
Average	98.74	97.32	74.24	75.09	73.67	72.67	98.02	97.93	33.81	33.77	57.07	58.09

Key OAR Metrics	SpinalCord_05 Dose at 0.03CC [Gy]		ParotidIps-PTV Mean Dose [Gy]		ParotidCon-PTV Mean Dose [Gy]		SubmandL-PTV Mean Dose [Gy]		SubmandR-PTV Mean Dose [Gy]		OCavity-PTV Mean Dose [Gy]	
	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1	HN-SIB-BPI	Candidate1
Average	16.51	42.86	16.75	17.88	12.58	13.72	33.88	35.94	37.43	40.48	25.67	22.75

Key Target/OAR Metrics between Foundation Model and final HN-SIB-BPI model on 27 case training subset



Four candidate models evaluated, Candidate 1 selected as the foundation model in this work after Dosimetric Scorecard evaluation, qualitative isodose and select PTV + OAR DVH analysis all indicated the best performance of the group.

CONCLUSIONS

This model demonstrates excellent performance for head and neck cases with BPI, offering a single-click solution for high-quality treatment plans. It is compatible with different delivery platforms, and has the potential to improve plan quality and efficiency in clinical practice. The dosimetric scorecard approach proved valuable for model development, assessment, and validation, enabling precise articulation of clinical intent and trade-off preferences. Further validation and investigation of clinical impact are recommended to fully assess its potential benefits in routine clinical practice. To facilitate wider use and evaluation, the model, scorecard, tools, and an example case have been shared publicly online [3].

REFERENCES

1. Luca K, Roper J, Wolf J, et al. Evaluating the plan quality of a general head-and-neck knowledge-based planning model versus separate unilateral/bilateral models. Medical Dosimetry. 2023;48(1):44-50. doi:10.1016/j.meddos.2022.10.002

2. Nelms BE, Robinson G, Markham J, et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. Practical Radiation Oncology. 2012;2(4):296-305. doi:10.1016/j.prro.2011.11.012

3. Varian Medical Affairs - Bilateral Head&Neck 70/63/56Gy (NRG HN004) [RapidPlan]. Varian Medical Affairs. Accessed March 24, 2024. https://medicalaffairs.varian.com/hn-sib-bpi-rapidplan-vmat2

ACKNOWLEDGEMENTS

The authors would like to sincerely thank the following for their contributions to this work: Vanessa Magliari from Varian for providing the Foundation RapidPlan model used in this study. Kenny Guida of the University of Kansas Cancer Center for preliminary testing of the final HN-SIB-BPI model.

CONTACT INFORMATION

Ryan Clark – Ryan.rc.clark@Varian.com
Anthony Magliari – Anthony.Magliari@Varian.com