An Automated Radiomics Based Approach for Predicting HPV Status in Oropharyngeal Squamous Cell Carcinoma

[1]Rikhil D. Makwana, BA; [1]Fahmida Haque, PhD; [1]Stephanie Harmon, PhD; [2]Evrim Turkbey, MD; [3]James L. Gulley, MD, PhD; [4]Clint. T. Allen, MD; [3]Charalampos S. Floudas*, MD; [1]Baris Turkbey*†, MD

[1]Artificial Intelligence Resource, National Cancer Institute, National Institute of Health, Bethesda, MD, 20814, USA; [2]Molecular Imaging Branch, National Cancer Institute, National Institute of Health, Bethesda, MD, 20892 USA ; [3]Center for Immuno-Oncology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA; [4]Surgical Oncology Program, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

Background

HPV-associated (HPV+) oropharyngeal SCC (OPSCC) has unique pathophysiology, **better prognosis**, and is **less likely to recur** than HPV(-) OPSCC

Radiomic based feature extraction from tumor and lymph node (LN) contrast enhanced CT (CECT) coupled with machine learning represents a noninvasive method of determining HPV status in patients with OPSCC

Methods

Retrospective analysis of 5 cohorts contoured with an in-house segmentation model

Features extracted with PyRadiomics

- Preprocessing: Z-score normalization + clipping of image intensity

Cohorts split into 80%/20% Training and Test sets

- HPV: Cohorts 1, 2, HPV (-) from 3

Model Selection: 5-fold cross validation applied to the Training set

- Cases per fold balanced by HPV status/recurrence

Performance is reported as the ensembled results across all folds applied to the:

- Internal Test Set
- External Validation Cohorts
  - HPV: 4, 5, 3 (HPV+)

Dataset	Total	HPV +	HPV -
1	755	564	191
2	18	14	4
3	150	128	22
4	60	54	6
5	80	23	57
Total	1063	783	280

Table 1: Cohort Demographics. Cohorts 1, 2, 3, and 5 are open access datasets. Cohort 4 is in-house.

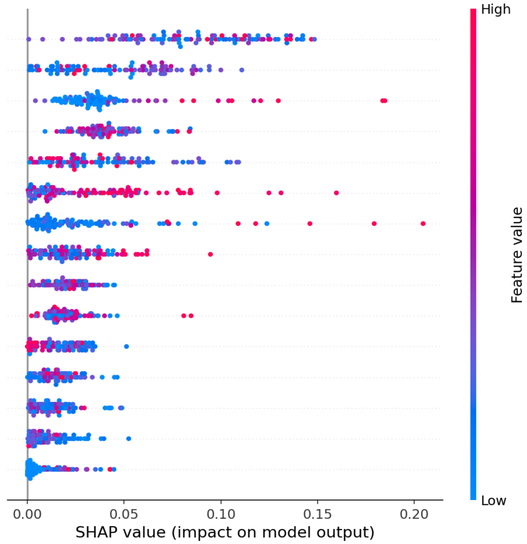

Figure 1: Bee Swarm Plot of the Top 15 Most Predictive Features by SHAP value. SHAP = SHapley Additive exPlanations

y-axis: Feature names
x-axis: SHAP contribution averaged across folds per feature

Each point is a patient and overlap of points shows the density of impact or the relative impact of that feature in the output.

Red Points: SHAP value when feature values are high
Blue Points: SHAP value when feature values are low

Results

	Model Configuration
Inputs	Lesion + LN Mask
Cohort	
Train (+ %)	507 (76%)
Test (+ %)	126 (76%)
External (+ %)	234 (79%)
Normalization	Z-score
Correlated Features	Removed features with > 0.85 Spearman rank correlation
Selection	Recursive feature elimination
Models	Random Forrest, XGBoost
Outcome	HPV Status

Table 2: Model Configuration and Set Composition

Test Set	Internal (n = 96/126)		External (n = 221/287)	
Model	**Random Forrest**	**XGBoost**	**Random Forrest**	**XGBoost**
TP	78	96	155	172
TN	20	7	17	11
FP	10	23	31	37
FN	18	0	31	14
Sensitivity [CI]	0.81 [0.72, 0.88]	1.0 [0.96, 1.0]	0.83 [0.77, 0.88]	0.92 [0.88, 0.95]
Specificity [CI]	0.67 [0.49, 0.81]	0.23 [0.12, 0.41]	0.35 [0.23, 0.5]	0.23 [0.13, 0.37]
Accuracy [CI]	0.78 [0.7, 0.84]	0.82 [0.74, 0.88]	0.74 [0.68, 0.79]	0.78 [0.72, 0.83]
PPV [CI]	0.89 [0.8, 0.94]	0.80 [0.73, 0.88]	0.83 [0.77, 0.88]	0.82 [0.77, 0.87]
NPV [CI]	0.53 [0.37, 0.68]	1.0 [0.65, 1.0]	0.35 [0.23, 0.5]	0.44 [0.27, 0.63]
AUC	0.78	0.80	0.63	0.61

Table 3: Performance Metrics
TP = true positive
TN = true negative
FP = false positive
FN = false negative
CI = confidence intervals
PPV = positive predictive value
NPV = negative predictive value
AUC = area under the receiver operating characteristic curve

Conclusion

- Both the Random Forrest and XGBoost based models achieve similar performance at predicting HPV status from radiomic features extracted from tumor and LN regions of interest from CECT in patients with OPSCC
- Our models shows similar performance across internal & external testing suggesting the generalizability of our results compared to alternative radiomic based classification models that forego external validation
- LN signals have a strong association with HPV status—this matches the clinical gestalt of HPV+ OPSCC being strongly associated with large cystic lymph nodes

Future Directions

- Train tumor only and LN only models to further identify important HPV associated features radiomic features in patients with OPSCC
- Use both models to run inference on expert radiologist's contours to determine model generalizability outside of automated segmentation methods

Artificial Intelligence Resource

Presented by: Rikhil Makwana
Email: rikhil.makwana@my.rfums.org
LinkedIn: https://www.linkedin.com/in/rikhil-makwana/