Development and validation of a prospective radiomics-clinical signature for locoregional recurrence in patients with locally advanced head and neck cancer

Balu Krishna S[1], Amal Joseph Varghese[1], Hasan Shaikh[1], Rajendra Benny Kuchipudi[1], Ezhil Sindhanai[1], Julia Priyadarshini Rao[1], Vijayshree C[1], Jino Wilson Victor[1], Simon Pavamani[1], Rajesh I[1], Manu Mathew[1], Leonard Wee[2], Andre Dekker[2], and Hannah Mary Thomas T[1]

[1]Quantitative Imaging Research and Artificial Intelligence Lab, Department of Radiation Oncology, Unit 2, Christian Medical College, Vellore, Tamil Nadu, India

[2]Department of Radiation Oncology (Maastro), GROW Research Institute for Oncology and Reproduction, Maastricht University Medical Centre, Maastricht, The Netherlands

Abstract

Purpose:
The lack of standardized protocols and prospective data is a frequently cited limitation in radiomics. This study aimed to develop and internally validate a CT-based clinical-radiomics signature for predicting locoregional tumour recurrence (LRR) in patients with locally advanced head and neck squamous cell carcinoma (LA-HNSCC), using prospective, standardized planning CT and clinical data from a large tertiary care center in India.

Methods:
We prospectively enrolled 1,467 patients with LA-HNSCC treated between 2020 and 2024. Patients who received primary radiochemotherapy and had a minimum follow-up of 12 months after completion of radiotherapy were eligible for this analysis. All patients were imaged using a standard acquisition protocol, with images harmonized across two CT simulators (Siemens and GE Discovery). LRR was defined as local or regional recurrence identified on CT imaging and clinical examination.

A total of 107 quantitative imaging features were extracted from pre-treatment CT scans using PyRadiomics to characterize tumour shape and texture heterogeneity[1]. First, a radiomics-only model was trained using CT imaging features, after which demographic and clinical parameters were added [2]. A final clinical-radiomics signature was derived using repeated five-fold cross-validation on the discovery cohort. Seven feature selection methods and five machine learning classifiers were evaluated. All models used an 80/20 train-test split while maintaining the LRR event rate across training and hold-out sets. Performance was evaluated on the internal hold-out dataset using the area under the receiver operating characteristic curve (AUC) over 1000 bootstrap iterations with 95% confidence intervals [3].

Results:
Of 367 patients who received primary chemoradiation, 176 met all inclusion criteria and were included in the final analysis. Primary tumour sites comprised larynx (n = 79, 44%), hypopharynx (n = 36, 20%), oropharynx (n = 20, 11%), tonsil (n = 16, 9%), oral cavity subsites (n = 15, 8%), base of tongue (n = 6, 3%), and other locations (n = 4, 2%). Fifty-six patients had LRR, and 120 remained disease-free at 12 months post-treatment completion.

The training cohort included 141 patients, and the hold-out test cohort included 35 patients. The clinical-radiomics signature achieved a test AUC of 0.82 (95% CI: 0.66–0.94), exceeding the performance of the radiomics-only model (AUC 0.78; 95% CI: 0.61–0.91). The final signature included two clinical and eight radiomics features.

Conclusion:
We identified and internally validated a clinical-radiomics signature for LRR in locally advanced HNSCC using a prospective dataset from an Indian population. Prospective external validation is planned to evaluate the generalizability of this signature.

Keywords: Head and neck cancer; Radiomics, Loco regional recurrence.

*Corresponding Author: Dr. Hannah Mary Thomas T, Email: hannah.thomas@cmcvellore.ac.in

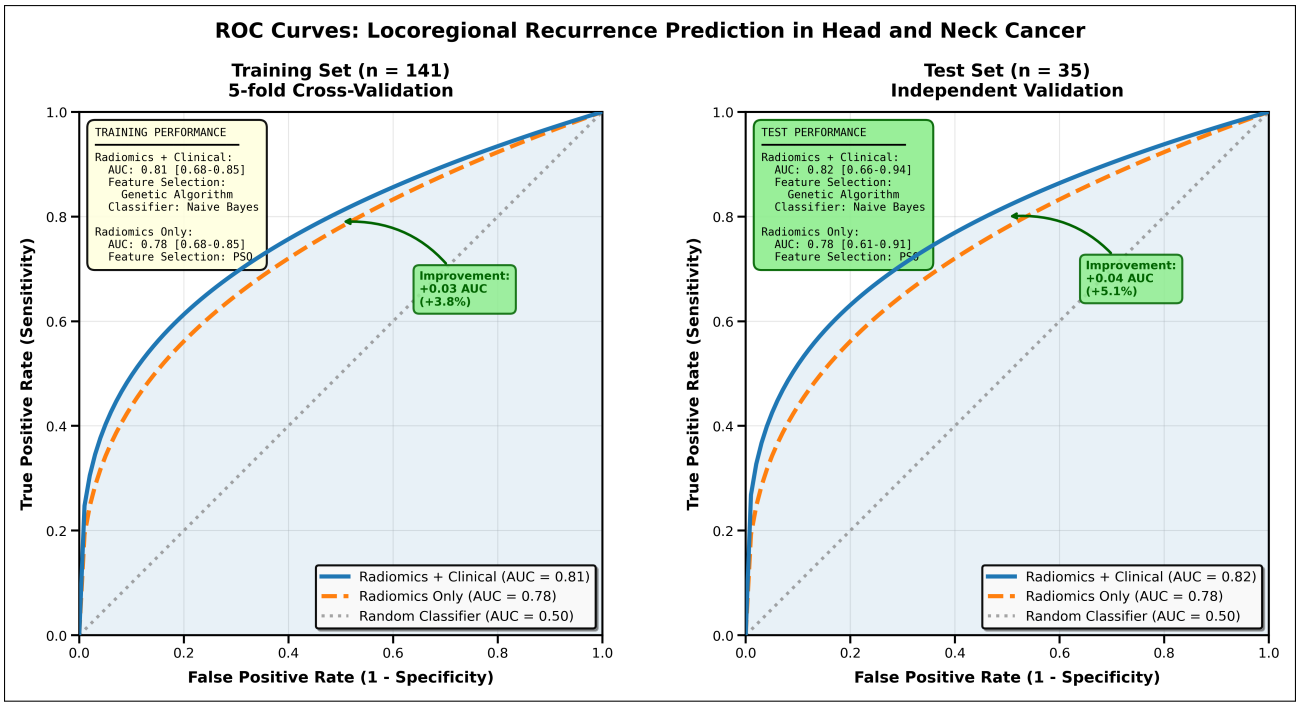

Figure 1: **ROC curves comparing radiomics-only and clinical-radiomics models for locoregional recurrence prediction.** Left panel shows training set performance using 5-fold cross-validation (n = 141). Right panel shows independent test set validation (n = 35). The clinical-radiomics signature (solid blue line) demonstrates superior discrimination compared to the radiomics-only model (dashed orange line), with an improvement of +0.04 AUC (+5.1%) on the test set. Dotted grey line represents random classifier (AUC = 0.50). Performance metrics were calculated using 1,000 bootstrap iterations with 95% confidence intervals.

Table 1: Clinical characteristics and association with locoregional recurrence

Characteristic	**Total Cohort** (n = 176)	**LRR** (n = 56, 31.8%)	**Non-LRR** (n = 120, 68.2%)	**p-value**
Age (years)				0.234
Mean ± SD	58.3 ± 12.4	59.1 ± 11.8	57.9 ± 12.7	
Range	21–85	28–82	21–85	
Gender, n (%)				0.421
Male	157 (89.2)	51 (91.2)	106 (88.3)	
Female	19 (10.8)	5 (8.9)	14 (11.7)	
Smoking history, n (%)				0.156
Non-smoker	69 (39.2)	19 (33.9)	50 (41.7)	
Former/Current smoker	107 (60.8)	37 (66.1)	70 (58.3)	
Chewable tobacco use[a], n (%)				**0.009**
No	106 (62.4)	27 (50.0)	79 (68.1)	
Yes	64 (37.6)	27 (50.0)	37 (31.9)	
Tumour location, n (%)				**<0.0001**
Larynx	77 (43.8)	15 (26.8)	64 (52.5)	
Hypopharynx	36 (20.5)	18 (32.1)	18 (14.8)	
Oropharynx	20 (11.4)	8 (14.3)	12 (10.0)	
Tonsil	16 (9.1)	7 (12.5)	9 (7.5)	
Other sites[b]	27 (15.3)	8 (14.3)	17 (14.2)	
T-stage[a], n (%)				**0.0001**
T1/T1a/T1b	14 (8.2)	2 (3.7)	12 (10.2)	
T2	41 (24.1)	8 (14.5)	34 (28.8)	
T3	71 (41.8)	23 (42.6)	49 (41.5)	
T4a/T4b	44 (26.0)	21 (40.0)	23 (19.8)	
N-stage[a], n (%)				0.089
N0	82 (48.2)	23 (42.6)	59 (50.9)	
N1	27 (15.9)	9 (16.7)	18 (15.5)	
N2 (a/b/c)	37 (21.8)	14 (25.9)	23 (19.8)	
N3	24 (14.1)	8 (14.8)	16 (13.8)	
AJCC stage[a] (8th ed.), n (%)				**0.018**
I	8 (4.8)	1 (2.0)	7 (6.1)	
II	25 (15.0)	5 (9.8)	20 (17.4)	
III	62 (37.7)	19 (37.3)	43 (37.4)	
IVA	38 (23.4)	14 (27.5)	24 (20.9)	
IVB	30 (18.6)	12 (23.5)	18 (15.7)	
HPV/p16 status, n (%)				0.387
Positive	14 (9.2)	3 (6.7)	11 (10.2)	
Negative	39 (25.0)	11 (24.9)	27 (25.0)	
Not available/tested	100 (65.8)	31 (68.9)	69 (63.9)	

[a] Missing data excluded from percentage calculations. [b] Other sites include base of tongue, anterior tongue, buccal mucosa, and alveolus. Bold p-values indicate statistical significance ($p < 0.05$).

Table 2: Machine learning methodology and optimal model characteristics

Parameter	Details
DATASET CONFIGURATION	
Total radiomics features	107 features extracted from pre-treatment CT scans
Total clinical features	4 features (tumour site, T-stage, chewable tobacco use, AJCC stage)
Total input features	111 features
Training set	n = 141 (80%)
Test set	n = 35 (20%)
Cross-validation	5-fold repeated cross-validation
MODEL DEVELOPMENT	
Machine learning classifiers (n=5)	Logistic Regression, Naive Bayes, SVM, Decision Tree, Random Forest
Feature selection techniques (n=7)	LASSO, SelectKBest, Particle Swarm Optimization, Whale Optimization Algorithm, Grey Wolf Optimizer, Genetic Algorithm, Simulated Annealing
Bootstrap iterations	1,000 iterations
Performance metric	Area Under ROC Curve (AUC) with 95% confidence intervals
OPTIMAL MODEL	
Best classifier + feature selector	Naive Bayes + Genetic Algorithm
Test set AUC [95% CI]	**0.82 [0.66–0.94]** (clinical-radiomics model)
Radiomics-only AUC [95% CI]	0.78 [0.61–0.91]
Final selected features	10 features (8 radiomics + 2 clinical)
SELECTED CLINICAL FEATURES (n=2)	
1. Tumor location	$p < 0.0001$
2. T-stage	$p = 0.0001$
SELECTED RADIOMICS FEATURES (n=8)	
3. GLDM Dependence Non-Uniformity	Texture heterogeneity
4. First-order Kurtosis	Intensity distribution
5. GLRLM High Gray Level Run Emphasis	Texture coarseness
6. GLCM Inverse Difference Normalized (IDN)	Texture homogeneity
7. GLRLM Short Run High Gray Level Emphasis	Fine texture with high intensity
8. GLSZM Size Zone Non-Uniformity	Regional texture heterogeneity
9. Shape Least Axis Length	Tumor geometry (minimum dimension)
10. GLRLM Short Run Emphasis	Fine texture pattern

GLDM = Gray Level Dependence Matrix; GLRLM = Gray Level Run Length Matrix; GLCM = Gray Level Co-occurrence Matrix; GLSZM = Gray Level Size Zone Matrix.

References

[1] Hugo J. W. L. Aerts, Emmanuel Rios Velazquez, Ralph T. H. Leijenaar, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5:4006, 2014.

[2] Tanmay Gangil, Kundan Sharan, B. Dinesh Rao, Karthick Palanisamy, Bidyut Chakrabarti, and Rajagopal Kadavigere. Utility of adding radiomics to clinical features in predicting the outcomes of radiotherapy for head and neck cancer using machine learning. *PLoS ONE*, 17(12):e0277168, 2022.

[3] Abby J. Varghese et al. Multi-centre radiomics for prediction of recurrence following radical radiotherapy for head and neck cancers: Consequences of feature selection, machine learning classifiers and batch-effect harmonization. *Physics and Imaging in Radiation Oncology*, 26:100450, 2023.