

## Supplementary Material

**Table S1** Survival probabilities via Kaplan-Meier method at 3-year overall survival (OS), 3-year disease specific survival (DSS) and 2-year locoregional free survival

Characteristics	Overall Survival, OS		Disease-specific Survival, DSS		Locoregional Recurrence-free Survival, LRFS	
	3-year OS (%) <sup>1</sup>	p-value <sup>2</sup>	3-year DSS (%) <sup>1</sup>	p-value <sup>2</sup>	2-year LRFS (%) <sup>1</sup>	p-value <sup>2</sup>
<b>All patients</b> <i>N</i> =148	62.6 (54.1, 72.4)	-	68.2 (59.7, 77.9)	-	70.3 (62.5, 78.9)	-
<b>AJCC 8th Stage I/II</b> <i>N</i> =59	82.5 (72.6, 93.8)	0.002	83.9 (74.2, 95.0)	0.02	71.8 (60.5, 85.3)	0.88
<b>AJCC 8th Stage III/IV</b> <i>N</i> =89	52.0 (41.5, 65.1)		59.5 (48.5, 72.9)		69.1 (59.1, 80.9)	
<b>p16-positive</b> <i>N</i> =64	88.5 (79.9, 97.9)	<0.001	89.9 (81.7, 99.0)	<0.001	79.9 (69.7, 91.5)	0.001
<b>p16-negative</b> <i>N</i> =65	34.6 (23.4, 51.2)		41.6 (29.0, 59.8)		52.0 (39.8, 68.1)	

Denominators, *N* that do not equal sample sizes are due to missing data

<sup>1</sup>Survival probability (%) with the 95% confidence interval (CI)

<sup>2</sup>Log-rank test

**Table S2** Determination of cutpoint for neutrophil to lymphocyte ratio (NLR) for prediction of poorer survival (*N*=139)

Survival Outcome	Method	<i>N</i> <sup>1</sup>	Cutpoint	High Risk NLR ≥ Cut-point <sup>2</sup>	HR (95% CI) <sup>3</sup>	p-value <sup>4</sup>	Performance Metrics <sup>5</sup>
<b>Overall survival (OS)</b>	Maximally selected statistics using Hothorn and Lausen method	139	1.61	123 (88%)	4.58 (1.11, 18.9)	0.02	C=0.56
	Maximally selected statistics using Contal and O'Quigley method	139	3.63	48 (35%)	2.03 (1.17, 3.54)	0.01	C=0.60
	ROC method using outcome at 2-year	105	2.51	85 (61%)	1.76 (0.97, 3.19)	0.06	AUROC=0.61
<b>Disease-specific survival (DSS)</b>	Maximally selected statistics using Hothorn and Lausen method	139	3.53	51 (37%)	2.51 (1.31, 4.80)	0.004	C=0.60
	Maximally selected statistics using Contal and O'Quigley method	139	3.56	50 (36%)	2.59 (1.36, 4.95)	0.003	C=0.60
	ROC method using outcome at 2-year	96	2.51	85 (61%)	1.95 (0.96, 3.96)	0.06	AUROC=0.61
<b>Locoregional recurrence-free survival (LRFS)</b>	Maximally selected statistics using Hothorn and Lausen method	139	1.61	123 (88%)	2.00 (0.71, 5.61)	0.18	C=0.56
	Maximally selected statistics using Contal and O'Quigley method	139	2.09	105 (76%)	1.73 (0.83, 3.60)	0.14	C=0.56
	ROC method using outcome at 2-year	91	2.67	75 (54%)	1.45 (0.79, 2.63)	0.23	AUROC=0.58

<sup>1</sup>*N* that do not equal sample sizes are due to missing data. *N* is the number of patients used in determining the optimal cutpoints.

<sup>2</sup>*n* (%); Denominator *N* = 139

<sup>3</sup>HR = Hazard Ratio, CI = Confidence Interval based on univariate Cox proportional hazards model

<sup>4</sup>p-value is based on log-rank test

<sup>5</sup>C = Concordance statistic, AUROC = Area under the receiver operating curve

**Table S3** Determination of cutpoint for neutrophil to lymphocyte ratio (NLR) for prediction of poorer survival in p16-positive oropharyngeal squamous cell carcinoma group (N=60)

Survival Outcome	Method	N <sup>1</sup>	Cutpoint	High Risk NLR ≥ Cut-point <sup>2</sup>	HR (95% CI) <sup>3</sup>	p-value <sup>4</sup>	Performance Metrics <sup>5</sup>
<b>Overall survival (OS)</b>	Maximally selected statistics using Hothorn and Lausen method	60	4.16	11 (18%)	2.73 (0.49, 15.2)	0.23	C=0.61
	Maximally selected statistics using Contal and O'Quigley method	60	4.17	10 (17%)	3.28 (0.58, 18.6)	0.16	C=0.62
	ROC method using outcome at 2-year	37	1.92	43 (72%)	0.85 (0.16, 4.66)	0.85	AUROC=0.47
<b>Disease-specific survival (DSS)</b>	Maximally selected statistics using Hothorn and Lausen method	60	4.16	11 (18%)	3.87 (0.63, 23.8)	0.12	C=0.61
	Maximally selected statistics using Contal and O'Quigley method	60	4.17	10 (17%)	4.77 (0.76, 30.0)	0.07	C=0.62
	ROC method using outcome at 2-year	36	4.17	10 (17%)	4.77 (0.76, 30.0)	0.07	AUROC=0.59
<b>Locoregional recurrence-free survival (LRFS)</b>	Maximally selected statistics using Hothorn and Lausen method	60	3.09	22 (37%)	0.72 (0.22, 2.34)	0.58	C=0.50
	Maximally selected statistics using Contal and O'Quigley method	60	3.25	21 (35%)	0.49 (0.13, 1.78)	0.27	C=0.49
	ROC method using outcome at 2-year	38	3.09	22 (37%)	0.72 (0.22, 2.34)	0.58	AUROC=0.52

<sup>1</sup>N that do not equal sample sizes are due to missing data. N is the number of patients used in determining the optimal cutpoints.

<sup>2</sup>n (%); Denominator N = 60

<sup>3</sup>HR = Hazard Ratio, CI = Confidence Interval based on univariate Cox proportional hazards model

<sup>4</sup>p-value is based on log-rank test

<sup>5</sup>C = Concordance statistic, AUROC = Area under the receiver operating curve

**Table S4** Determination of cutpoint for neutrophil to lymphocyte ratio (NLR) for prediction of poorer survival in p16-negative oropharyngeal squamous cell carcinoma group (N=63)

Survival Outcome	Method	N	Cutpoint	High Risk NLR ≥ Cut-point <sup>2</sup>	HR (95% CI) <sup>3</sup>	p-value <sup>4</sup>	Performance Metrics <sup>5</sup>
<b>Overall survival (OS)</b>	Maximally selected statistics using Hothorn and Lausen method	63	3.35	28 (44%)	1.79 (0.93, 3.43)	0.08	C=0.60
	Maximally selected statistics using Contal and O'Quigley method	63	3.40	26 (41%)	1.71 (0.89, 3.27)	0.10	C=0.58
	ROC method using outcome at 2-year	53	3.39	27 (43%)	1.90 (0.99, 3.64)	0.05	AUROC=0.53
<b>Disease-specific survival (DSS)</b>	Maximally selected statistics using Hothorn and Lausen method	63	3.39	27 (43%)	1.95 (0.94, 4.07)	0.07	C=0.60
	Maximally selected statistics using Contal and O'Quigley method	63	3.56	26 (41%)	1.96 (0.94, 4.08)	0.07	C=0.58
	ROC method using outcome at 2-year	46	2.63	39 (62%)	1.28 (0.60, 2.76)	0.52	AUROC=0.53
<b>Locoregional recurrence-free survival (LRFS)</b>	Maximally selected statistics using Hothorn and Lausen method	63	4.66	16 (25%)	0.54 (0.19, 1.57)	0.25	C=0.55
	Maximally selected statistics using Contal and O'Quigley method	63	4.76	15 (24%)	0.54 (0.19, 1.57)	0.25	C=0.53
	ROC method using outcome at 2-year	40	2.80	32 (51%)	1.09 (0.51, 2.33)	0.82	AUROC=0.54

<sup>1</sup>N that do not equal sample sizes are due to missing data. N is the number of patients used in determining the optimal cutpoints.

<sup>2</sup>n (%); Denominator N = 63

<sup>3</sup>HR = Hazard Ratio, CI = Confidence Interval based on univariate Cox proportional hazards model

<sup>4</sup>p-value is based on log-rank test

<sup>5</sup>C = Concordance statistic, AUROC = Area under the receiver operating curve