

Supplementary

Kyungwon Kim¹[0009-0005-4760-8886], Yongmoon Lee², Doohyun Park¹[0000-0001-6237-468X], Taejoon Eo¹[0000-0002-3546-0184], Daemyung Youn³, Hyesang Lee³, and Dosik Hwang^{1,4,5,6}[0000-0002-2217-2837]

- ¹ School of Electrical and Electronic Engineering, Yonsei University, Seoul, Republic of Korea
`{yskgw93, dosik.hwang}@yonsei.ac.kr`
- ² Department of Pathology, Dankook University School of Medicine, Cheonan, Republic of Korea
- ³ DoctorWorks Co., Ltd., Seoul, Republic of Korea
- ⁴ Department of Radioogy and Center for Clinical Imaging DataScience, Yonsei University, Seoul, Republic of Korea
- ⁵ Department of Oral and Maxillofacial Radiology, College of Dentistry, Yonsei University, Seoul, Republic of Korea
- ⁶ Artificial Intelligence and Robotics Institute, Korea Institute of Science and Technology, Seoul, Republic of Korea

Table 1: Demographic and pathological characteristics per hospital. Characteristic with † is written in mean \pm standard deviation form. If standard deviation is not specified, data represents the number of patients. Among all 6 university hospitals, AJMC, CNUH, and PNUH were used for training, and HUMC, SCHMC, and EUMC were designated as external validation set.

Characteristic	AJMC (N=296)	PNUH (N=267)	CNUH (N=13)	HUMC (N=111)	SCHMC (N=68)	EUMC (N=41)
Mean age (year)†	72.9 \pm 9.7	72.3 \pm 9.2	72.6 \pm 9.4	73.7 \pm 11.5	74.4 \pm 8.7	77.9 \pm 11.0
Sex						
Men	215 (73%)	189 (71%)	82 (62%)	79 (71%)	48 (71%)	27 (66%)
Women	81 (27%)	98 (37%)	51 (38%)	32 (29%)	20 (29%)	14 (34%)
Smoking history						
No	208 (70%)	141 (53%)	49 (37%)	51 (46%)	28 (41%)	24 (59%)
Yes	88 (30%)	126 (47%)	84 (63%)	60 (54%)	40 (59%)	17 (41%)
TNM stage						
1	106 (36%)	79 (30%)	86 (65%)	26 (23%)	8 (12%)	10 (24%)
2	36 (12%)	33 (12%)	19 (14%)	4 (4%)	3 (4%)	16 (39%)
3	59 (20%)	63 (24%)	19 (14%)	25 (22%)	16 (24%)	3 (7%)
4	95 (32%)	92 (34%)	9 (7%)	26 (23%)	41 (60%)	12 (29%)
T stage						
1	91 (31%)	65 (24%)	65 (49%)	31 (28%)	8 (12%)	11 (27%)
2	112 (38%)	103 (39%)	45 (34%)	26 (23%)	20 (29%)	15 (37%)
3	42 (14%)	43 (16%)	15 (11%)	16 (14%)	18 (26%)	3 (7%)
4	51 (17%)	56 (21%)	8 (6%)	38 (34%)	22 (32%)	12 (29%)
N stage						
0	146 (49%)	111 (42%)	98 (74%)	37 (33%)	16 (24%)	17 (41%)
1	35 (12%)	32 (12%)	12 (9%)	4 (4%)	12 (18%)	7 (17%)
2	51 (17%)	58 (22%)	15 (11%)	27 (24%)	26 (38%)	9 (22%)
3	64 (22%)	66 (25%)	8 (6%)	43 (39%)	14 (21%)	8 (20%)
M stage						
0	198 (67%)	178 (67%)	124 (93%)	53 (48%)	26 (38%)	31 (76%)
1	98 (33%)	89 (33%)	9 (7%)	58 (52%)	42 (62%)	10 (24%)
Location of tumor						
Right superior lobe	78 (26%)	70 (26%)	39 (29%)	30 (27%)	17 (25%)	13 (32%)
Right middle lobe	19 (6%)	31 (12%)	10 (8%)	7 (6%)	1 (1%)	3 (7%)
Right inferior lobe	74 (25%)	59 (22%)	35 (26%)	32 (29%)	21 (31%)	8 (20%)
Left superior lobe	74 (25%)	58 (22%)	26 (20%)	25 (23%)	18 (26%)	11 (27%)
Left inferior lobe	51 (17%)	49 (18%)	23 (17%)	14 (15%)	11 (16%)	6 (15%)
Type of tumor						
Adenocarcinoma	190 (64%)	185 (69%)	93 (70%)	75 (68%)	31 (46%)	26 (63%)
Squamous cell carcinoma	106 (36%)	82 (31%)	40 (30%)	36 (32%)	37 (54%)	15 (37%)
Survival						
5-year Overall Survival	98 (33%)	100 (37%)	91 (27%)	24 (13%)	13 (19%)	8 (20%)
5-year Death	198 (67%)	167 (63%)	42 (84%)	76 (76%)	55 (81%)	3 (80%)

Table 2: Demographic and pathological characteristics per validation set. Characteristic with † is written in mean \pm standard deviation form. If standard deviation is not specified, data represents the number of patients. p -values were calculated between train dataset and external validation set by the chi-squared test and t test.

Characteristic	Train (N=516)	Validation		p -value
		Internal (N=178)	External (N=214)	
Mean age (year)†	72.5 \pm 9.5	72.9 \pm 9.3	74.7 \pm 10.63	0.01
Sex				0.94
Men	361 (70%)	124 (70%)	151 (71%)	
Women	155 (30%)	54 (30%)	63 (29%)	
Smoking history				0.02
No	292 (57%)	105 (59%)	100 (47%)	
Yes	224 (43%)	73 (41%)	114 (53%)	
TNM stage				<0.01
1	203 (39%)	68 (38%)	44 (21%)	
2	60 (12%)	28 (28%)	21 (10%)	
3	103 (20%)	37 (37%)	43 (20%)	
4	150 (29%)	45 (45%)	106 (50%)	
T stage				<0.01
1	168 (33%)	53 (30%)	50 (23%)	
2	188 (36%)	72 (40%)	60 (28%)	
3	75 (15%)	25 (14%)	35 (16%)	
4	85 (16%)	28 (16%)	69 (32%)	
N stage				<0.01
0	266 (52%)	88 (49%)	69 (32%)	
1	60 (12%)	19 (11%)	20 (9%)	
2	90 (17%)	33 (19%)	62 (29%)	
3	100 (19%)	38 (21%)	63 (29%)	
M stage				<0.01
0	366 (71%)	133 (75%)	107 (50%)	
1	150 (29%)	45 (25%)	107 (50%)	
Location of tumor				0.59
Right superior lobe	136 (26%)	50 (28%)	59 (28%)	
Right middle lobe	38 (7%)	22 (12%)	11 (5%)	
Right inferior lobe	130 (25%)	37 (21%)	59 (28%)	
Left superior lobe	115 (22%)	43 (24%)	52 (24%)	
Left inferior lobe	97 (19%)	26 (15%)	33 (15%)	
Type of tumor				0.05
Adenocarcinoma	349 (68%)	117 (66%)	128 (60%)	
Squamous cell carcinoma	167 (32%)	61 (34%)	86 (40%)	
Survival				<0.01
5-year Overall Survival	212 (41%)	77 (43%)	48 (22%)	
5-year Death	304 (59%)	101 (57%)	166 (78%)	