#6865

IMAGING AI PROGNOSIS OF EARLY-STAGE LUNG CANCER USING CT RADIOMICS

Ann Valter[1], Tanel Kordemets[2], Aydan Gasimova[3], Brennen Heames[3], Noah Waterfield Price[3], Gerald Hodgkinson[4], Tõnu Vanakesa[5], Ingemar Almre[5], Lutz Freitag[3], David Paul Carbone[6], Kersti Oselin[1]

[1]Department of Chemotherapy, Clinic of Oncology and Hematology, North Estonia Medical Centre, Tallinn, Estonia; [2]Department of Radiology, Clinic of Diagnostics, North Estonia Medical Centre, Tallinn, Estonia;
[3]Research Department, Optellum, Ltd., Oxford, United Kingdom; [4]Surgical Robotics, Medtronic - Operational Headquarters, Minneapolis, United States of America;
[5]Cardio-Thoracic Surgery, North Estonia Medical Centre Foundation, Tallinn, Estonia; [6]Internal Medicine department, Ohio State University Medical Center, Columbus, OH, United States of America

Background

Accurate prediction of lung cancer recurrence risk is crucial for treatment decisions and follow-up, particularly for Stage I patients who are not eligible for (neo-)adjuvant therapy but approximately one-third of whom still experience recurrence after surgical resection [1,2]. We present a machine learning model that uses patient computed tomography (CT) images and clinical features to predict lung cancer recurrence.

Methods: Data

A dataset of 968 clinical stage I-III lung cancer patients who underwent surgical resection was gathered from the US National Lung Screening Trial [3], the North Estonia Medical Centre, the Stanford University School of Medicine and Palo Alto Veterans Affairs Healthcare System [4]. Of these patients, 32.3% (313/968) had lung cancer recurrence, with Stage I recurrence rate at 28.5% (221/776).

Table 1. Patient characteristics.

	Recurred	No recurrence
N	313 (32.3%)	655 (67.7%)
Sex		
Male	209 (66.8%)	376 (57.4%)
Female	104 (33.2%)	279 (42.6%)
Age (mean, std)	66.6, 7.3	66.5, 8.4
Nodule size (mean, std)	27.7, 14.6	23.8, 11.9
Clinical Stage		
Stage I	221 (70.6%)	555 (84.7%)
Stage II	51 (16.3%)	72 (11.0%)
Stage III	41 (13.1%)	28 (4.3%)
Lobe		
Upper	182 (58.1%)	422 (64.4%)
Lower	131 (41.9%)	233 (35.6%)
Attenuation		
Solid	241 (77.0%)	455 (69.5%)
Part-solid	62 (19.8%)	143 (21.8%)
GGO	4 (1.3%)	34 (5.2%)
Other	6 (1.9%)	23 (3.5%)

Methods: Training and validation

The pre-operative survival model was trained to predict the likelihood of recurrence at each time-point using radiomic features extracted from CT images and relevant clinical variables. An 8-fold cross-validation strategy was used, and performance evaluated using the time-dependent Area-Under-the-ROC-Curve (AUC), disease-free survival (DFS), hazard ratio (HR) and log-rank test against clinical staging.

Results: Risk Stratification

The ML survival model was better able to stratify patients into high and low-risk (HR=2.7, p<0.005) compared with Stage I vs II-III (HR=2.2, p<0.005). The same was observed for the Stage I sub-group (HR=2.4, p<0.005) when compared with using Stage IA vs IB (HR=1.1, p=0.79). The gaps between the high and low-risk DFS at 1, 2, and 5 years are larger for the ML model than separation by staging. ML model thresholds were set to match on high/low-risk patient counts.

Machine learning survival model better able to stratify patients by risk of lung cancer recurrence than clinical staging alone

Table 2. Disease-free survival (DFS) and hazard ratios (HR) of high vs low-risk patient populations.

Predictors	1-year DFS % low-risk	1-year DFS % high-risk	*p*-value	2-year DFS % low-risk	2-year DFS % high-risk	*p*-value	5-year DFS % low-risk	5-year DFS % high-risk	*p*-value	HR (5% CI)
Stage I-III										
cTNM Stage I (n=776) vs Stage II-III (n=192)	92.0	77.1	0.718	85.6	66.7	0.202	74.0	50.8	<0.005	2.2 (1.7, 2.8) *p*<0.005
ML model Low-risk (n=776) vs High-risk (n=192)	93.2	72.6	0.767	86.8	61.9	0.003	75.3	46.2	<0.005	2.7 (2.1, 3.4) *p*<0.005
Stage I subgroup										
cTNM Stage IA (n=116) vs Stage IB (n=660)	92.2	91.0	0.778	86.1	82.9	0.646	74.4	74.0	0.318	1.1 (0.7, 1.5) *p*=0.79
ML model Low-risk (n=116) vs High-risk (n=660)	94.0	80.9	0.600	88.0	72.0	0.017	77.4	55.1	<0.005	2.4 (1.8, 3.3) *p*<0.005

Results: Classification Performance

The ML survival model had better prediction accuracy, with the time-dependent AUCs being significantly better than staging alone at 1, 2, and 5-year marks.

Table 3. Time-dependent Area under Receiver Operating Characteristic (AUC) curve of staging vs ML model.

Predictor	1-year AUC	2-year AUC	5-year AUC
Stage I-III			
cTNM	0.664	0.643	0.608
ML model	0.742	0.696	0.680
p-value	0.025	0.065	0.004
Stage I subgroup			
cTNM	0.515	0.525	0.502
ML model	0.670	0.637	0.635
p-value	0.006	0.011	<0.005

Results: AUC and CIR

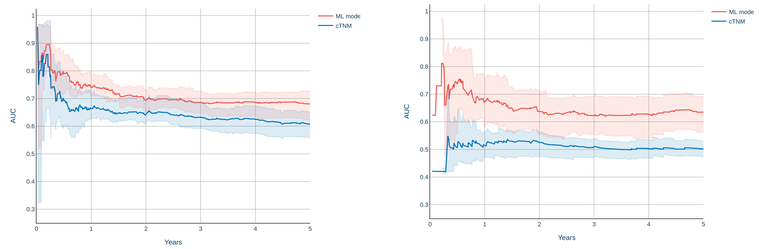

Figure 1. Time-dependent AUC of a) Stage I-III patients and b) Stage I sub-group.

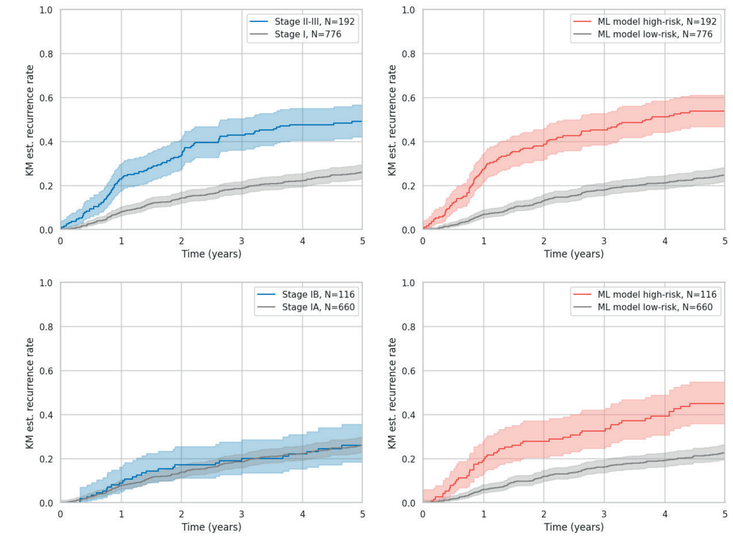

Figure 2. Cumulative incidence of recurrence (CIR) with 5% confidence intervals of all Stage I-III patients risk-stratified by a) stage and b) the ML model, and Stage I patients by c) stage and d) the ML model. ML model thresholds were set to match on high/low-risk patient counts.

Conclusions

The ML survival model outperforms clinical staging in patient risk-stratification and time-dependent lung cancer recurrence prediction. With further development, this algorithm could prove a valuable, non-invasive tool to aid the management of lung cancer patients.

REFERENCES
1) Sugimura H et al. Survival after recurrent NSCLC after complete pulmonary resection. Annals of Thoracic Surgery, 2007.
2) Goldstraw P et al. The IASLC lung cancer staging project: proposals for the revision of the TNM classification for lung cancer. Journal of Thoracic Oncology, 2016.
3) The National Lung Screening Trial Research Team. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic screening. NEJM, 2011.
4) Bakr, S. et al. Data for NSCLC Radiogenomics (Version 4) [Data set]. The Cancer Imaging Archive, 2017

Ann Valter declares no conflict of interest with this work.
This study was sponsored by Optellum.

Corresponding author e-mail: ann.valter@regionaalhaigla.ee