

Flagging high-risk individuals with a ML model improves NSCLC early detection in a USPSTF-eligible population

Eran N. Choman¹, Alon Lanyado¹, Prof. Michael K. Gould, MD, MS²

¹ Medial EarlySign, Hod Hasharon, Israel.

² Department of Health Systems Science, KAISER PERMANENTE BERNARD J. TYSON SCHOOL OF MEDICINE, CA, USA.

Background

The USPSTF recommends annual lung cancer screening with LDCT in adults aged 50 to 80 years who have a ≥20 pack-year smoking history and currently smoke or have quit within the past 15 years. Risk prediction models are an alternative approach to identify high-risk individuals for screening that may have advantages compared to age and smoking history-based selection. We compared the performance of two risk prediction models, LungFlag and adapted to EHR data PLCom2012 (mPLCom2012).

Methods

Data from a large US health system including 6,505 case patients with non-small cell lung cancer (NSCLC) and 189,597 contemporaneous NSCLC-free controls were used to evaluate the performance of an optimized version of a previously published machine-learning model (LungFlag) to detect NSCLC among individuals who meet the USPSTF criteria compared to the performance of mPLCom2012. The model used existing routine out-patient lab measurements, smoking history, comorbidities, and demographic data.

Results

Data were analyzed using the area under the receiver operating characteristic curve (AUC), and diagnostic sensitivity on the USPSTF screen-eligible population (Tables 1-3) and Ever Smokers ages 50-80 population (Table 4). The risk predictor was calculated for a 3-12-month window prior to the diagnosis date (Dx) using cut-offs yielding specificity levels of 97%, 95% or 90%.

Conclusion

By using available information existed in the EHR, the LungFlag model was more accurate for early diagnosis of NSCLC than mPLCom2012, demonstrating the potential to help prevent lung cancer deaths through early detection among the subgroup of USPSTF as well as the Ever Smokers population.

Contact

Eran N Choman
eran@earlysign.com
+972 52 5611154
www.earlysign.com



Table 1: All NSCLC USPSTF Eligible

NSCLC Stage	AUC		Sensitivity @ Specificity Level					
			Sens @ 97% Spec		Sens @ 95% Spec		Sens @ 90% Spec	
	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012
Stage 0 (in situ)	0.840* [0.770,0.890]	0.770 [0.710,0.850]	26.5 [11.5,46.3]	23.5 [10.0,38.7]	44.1* [26.7,61.3]	23.5 [10.3,41.2]	55.9* [37.9,69.7]	35.3 [18.8,53.1]
Stage I	0.793* [0.770,0.822]	0.747 [0.722,0.775]	21.4* [16.8,26.6]	11.0 [7.4,15.2]	28.5* [24.1,35.2]	16.7 [12.6,21.9]	42.7* [38.3,49.7]	31.7 [25.5,37.2]
Stage II	0.824* [0.785,0.852]	0.764 [0.729,0.808]	26.6 [18.1,34.7]	18.8 [12.7,26.5]	32.0 [24.0,41.3]	25.0 [17.6,33.8]	44.5* [35.7,54.6]	33.6 [25.7,42.4]
Stage III	0.777* [0.746,0.805]	0.739 [0.709,0.770]	16.7 [11.2,21.2]	15.9 [11.2,20.6]	24.7 [18.5,29.6]	18.8 [13.9,24.0]	40.2 [32.6,45.8]	33.1 [26.5,39.0]
Stage IV	0.753* [0.733,0.775]	0.726 [0.704,0.748]	15.5* [12.2,19.1]	11.7 [9.0,14.9]	21.7 [17.3,25.4]	17.4 [12.8,21.3]	33.1 [28.1,38.0]	29.1 [24.1,33.0]
Total	0.772* [0.757,0.788]	0.735 [0.718,0.749]	17.7* [15.2,20.5]	12.1 [9.6,14.1]	23.7* [21.2,26.8]	18.0 [14.7,20.3]	37.5* [33.6,40.5]	29.5 [26.5,33.2]

* Statistically significant difference (P-value<0.05)

Table 2: Adenocarcinoma NSCLC USPSTF Eligible

NSCLC Stage	AUC		Sensitivity @ Specificity Level					
			Sens @ 97% Spec		Sens @ 95% Spec		Sens @ 90% Spec	
	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012
Stage 0 (in situ)	0.900* [0.780,0.980]	0.770 [0.710,0.850]	33.3 [0.0,83.3]	0.0 [0.0,0.0]	66.7* [16.7,100.0]	0.0 [0.0,0.0]	83.3* [37.5,100.0]	0.0 [0.0,0.0]
Stage I	0.783* [0.751,0.814]	0.747 [0.722,0.775]	17.9* [12.6,25.7]	6.9 [3.2,11.9]	25.5* [17.1,31.3]	12.4 [8.5,19.9]	41.4* [31.8,46.2]	24.8 [17.8,33.6]
Stage II	0.780 [0.720,0.830]	0.764 [0.729,0.808]	22.7* [7.3,31.2]	6.8 [0.0,17.2]	25.0 [11.5,35.1]	13.6 [4.3,22.6]	38.6* [19.5,46.5]	15.9 [6.5,26.8]
Stage III	0.749 [0.704,0.801]	0.739 [0.709,0.770]	11.8 [5.7,21.3]	10.5 [4.0,17.9]	18.4 [11.3,29.9]	11.8 [5.6,21.3]	34.2 [23.8,45.0]	25.0 [14.3,32.9]
Stage IV	0.737 [0.705,0.766]	0.726 [0.704,0.748]	14.4* [9.6,19.0]	8.5 [4.8,11.8]	19.9* [14.3,24.4]	13.1 [8.2,17.1]	30.5* [25.0,37.3]	25.0 [18.9,29.7]
Total	0.754* [0.733,0.775]	0.707 [0.685,0.731]	15.7* [12.0,18.8]	8.1 [5.8,11.0]	20.4* [17.1,25.5]	13.2 [9.7,16.5]	33.8* [28.4,37.9]	24.2 [20.2,28.5]

* Statistically significant difference (P-value<0.05)

Table 3: Squamous NSCLC USPSTF Eligible

NSCLC Stage	AUC		Sensitivity @ Specificity Level					
			Sens @ 97% Spec		Sens @ 95% Spec		Sens @ 90% Spec	
	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012
Stage 0 (in situ)	0.820 [0.720,0.900]	0.790 [0.680,0.880]	21.4 [0.0,41.7]	28.6 [7.1,55.6]	28.6 [6.2,53.8]	28.6 [10.0,60.0]	35.7 [12.5,66.7]	35.7 [12.5,64.3]
Stage I	0.838* [0.790,0.867]	0.784 [0.745,0.821]	25.8 [17.4,34.8]	13.4 [7.0,20.8]	34.0* [25.5,44.9]	21.6 [14.0,30.2]	50.5* [40.0,59.8]	39.2 [31.5,51.4]
Stage II	0.847* [0.808,0.887]	0.803 [0.755,0.851]	23.0 [15.4,39.0]	26.2 [15.4,37.1]	31.1 [20.8,43.1]	32.8 [22.1,47.1]	49.2 [35.2,60.9]	41.0 [30.9,53.6]
Stage III	0.794* [0.750,0.824]	0.749 [0.701,0.790]	18.3 [10.5,24.5]	17.4 [10.3,24.5]	24.3 [17.0,33.7]	21.7 [13.6,29.8]	41.7 [31.0,49.5]	34.8 [25.5,44.1]
Stage IV	0.778 [0.740,0.819]	0.763 [0.725,0.802]	18.7 [11.2,25.2]	13.8 [8.3,20.8]	24.4 [16.2,32.8]	22.8 [14.7,30.8]	36.6 [26.7,44.8]	31.7 [24.2,40.9]
Total	0.802* [0.783,0.829]	0.767 [0.742,0.792]	20.3* [16.7,26.2]	15.0 [11.0,19.1]	28.2* [23.5,32.8]	22.0 [17.8,27.2]	43.2* [37.4,48.4]	36.4 [31.6,41.6]

* Statistically significant difference (P-value<0.05)

Table 4: All NSCLC Ever Smokers Ages 50-80

NSCLC Stage	AUC		Sensitivity @ Specificity Level					
			Sens @ 97% Spec		Sens @ 95% Spec		Sens @ 90% Spec	
	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012	LungFlag	mPLCom2012
Stages 0-II	0.851* [0.839,0.863]	0.817 [0.805,0.830]	32.5* [29.3,35.8]	23.5 [20.9,26.3]	42.2* [39.2,45.5]	31.3 [28.4,34.2]	58.1* [55.2,61.1]	47.8 [44.7,51.3]
Stages III-IV	0.824* [0.815,0.833]	0.796 [0.786,0.806]	23.5* [21.5,25.6]	20.1 [18.2,22.2]	33.9* [31.5,36.4]	28.6 [26.3,30.8]	51.0* [48.3,53.6]	42.7 [40.4,45.0]
Total	0.831* [0.823,0.839]	0.802 [0.794,0.809]	26.1* [24.4,27.9]	21.0 [19.4,22.7]	36.2* [34.2,38.3]	29.4 [27.7,31.3]	52.8* [50.7,55.0]	44.1 [42.2,46.1]

* Statistically significant difference (P-value<0.05)