driven risk stratification can enhance screening effectiveness, reduce disparities, and improve personalized screening strategies. **Methods:** This ambidirectional, prospective-retrospective study analyzed patients who underwent lung cancer screening (LCS) at University of Illinois Health system (UIH). Multimodal AI inference (Zakkar et al. MedRxiv, 2024) was used to predict lung cancer risk from clinical and radiographic features. Outcomes of lung cancer diagnosis was determined by review of Electronic Health Record. **Results:** At the time of this analysis, a total of 1,584 individuals who underwent LDCT procedures were included, with a median follow-up period of approximately six months. The predictive model demonstrated strong discriminative ability, achieving a Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) of 0.82.Precision-recall analysis resulted in an Area Under the Curve (PR AUC) of 0.41. At a sensitivity threshold of 0.79, the model achieved a specificity of 0.96. Additionally, the positive predictive value (PPV) was 0.26 while the negative predictive value (NPV) was 0.97. **Conclusions:** In this interim analysis of diverse lung cancer screening patients, preliminary findings show multimodal AI has promising predictive ability of diagnosed lung cancer. These results highlight the model's effectiveness in distinguishing cases with a high degree of specificity and negative predictive value, making it a potentially valuable tool in clinical decision-making for LDCT-based assessments. Long-term follow up to ensure accuracy of complete diagnostic workups are required to definitively assess performance. **Keywords:** early detection, multimodal AI, LDCT

MA05.05

Analysis of Longitudinal Low-Dose CT Sybil Scores in Two Canadian Cohorts

C.H. Durney,[1] **S. Atkar-Khattra,**[2] **S. Lam,**[2,3] **R. Meza**[2,3] [1]*BC Cancer Research Institute, Vancouver/BC/CA,* [2]*British Columbia Cancer Research Institute, Vancouver/BC/CA,* [3]*University of British Columbia, Vancouver/BC/CA*

Introduction: Sybil is a deep learning model developed to predict lung cancer risk using a single low-dose computed tomography (LDCT) scan, without requiring clinical data. While prior studies have validated its predictive accuracy, time-patterns in Sybil scores from single-patient longitudinal LDCT scans, and variability of scores across risk groups remain under-explored. Understanding these is critical to assess the potential of Sybil for clinical use. We thus conducted an analysis of longitudinal Sybil risk scores using data from two Canadian lung cancer screening cohorts. **Methods:** We analyzed LDCT imaging and clinical data from 2537 participants in the Pan-Canadian Early Detection of Lung Cancer Screening Study and 2110 participants in the British Columbia Lung Screening Study. Screening eligibility criteria were: age 50-75 with PLCOm2008 risk score 2%, and 50-80 years with a PLCOm2012 risk score of $>1.51\%$ over 6 years and/or USPSTF 2013 criteria, respectively. On average, cancer cases had 3.3 scans and non-cases had 2.96 scans per patient. A total of 270 (179,91 respectively) lung cancer cases were detected. Sybil scores were computed for each LDCT scan and trajectories were examined for individuals who developed lung cancer (cases) and those who did not (non-cases). We investigated how Sybil scores evolved relative to the time of diagnosis, age at baseline, and smoking status. **Results:** Sybil scores were consistently higher for individuals who developed lung cancer compared to those who remained cancer-free (Figure 1A, $p < 0.001$). Among cases, mean/median Sybil scores increased as time to diagnosis decreased (Figure 1B). However, there was high heterogeneity in trajectories, with some cases having decreasing scores, and some non-cases having increasing scores (Figure 1A). Notably, we observed an age effect among cases: Sybil scores rose with increasing age at scan, a pattern not observed among controls, whose scores remained relatively flat across age (Figure 1C). When stratified by smoking status, both current and former smokers who developed lung cancer exhibited similar Sybil score trajectories. **Conclusions:** Our analysis suggests that Sybil scores differentiate in aggregate individuals at risk of lung cancer with average scores increasing as diagnosis approaches. The model appears robust to smoking status and reflects subtle age-related risk dynamics in cancer cases, suggesting it detects radiographic features of malignancy independent of risk factors. These findings support a potential role of Sybil in risk stratification. However, the large heterogeneity in individual trajectories indicates that further refinements might be needed before clinical use.

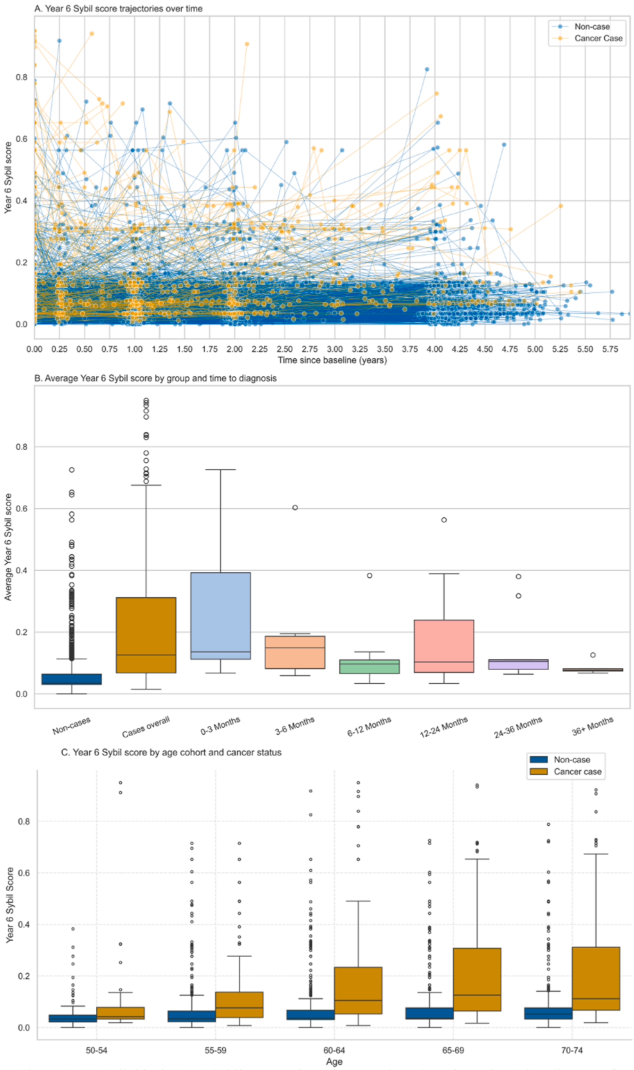

Figure 1: (A) Individual Year 6 Sybil score trajectories over time since the patient's baseline scan for cancer cases and non-cases. Each line represents a single participant. (B) Distribution of Year 6 Sybil scores by group. Cancer cases are further stratified by time from scan to diagnosis. (C) Distribution of Year 6 Sybil score stratified by age at scan and cancer status. Cancer Case refers to individuals diagnosed with lung cancer during follow-up; Non-case refers to those not diagnosed.

Keywords: Lung cancer screening, Deep learning, Longitudinal analysis

---

MA05.06

Integrating Circulating Protein Biomarkers in Nodule Malignancy Prediction Model Based on Five International Lung Cancer Screening Cohorts

E. Khodayari Moez,[1] **K. Murison,**[1] **M.T. Warkentin,**[1,2] **B. Diergaarde,**[3,4] **D.O. Wilson,**[3] **J.-M. Yuan,**[3,4] **S. Lam,**[5] **J. Field,**[6] **M.P. Davies,**[6] **G. Liu,**[7] **L.M. Montuenga,**[8,9] **Y. Brhane,**[1] **R. Phellan Aro,**[1] **C.I. Amos,**[10] **H.A. Robbins,**[11]

M. Johansson,[11] **R.J. Hung**[1,2] [1]*Lunenfeld-Tanenbaum Research Institute, Toronto/ON/CA,* [2]*University of Toronto, Toronto/ON/CA,* [3]*University of Pittsburgh, Pittsburgh/PA/USA,* [4]*UPMC Hillman Cancer Center, Pittsburgh/PA/USA,* [5]*British Columbia Cancer Agency, Vancouver/BC/CA,* [6]*University of Liverpool, Liverpool/GB,* [7]*Princess Margaret Cancer Center, Toronto/ON/CA,* [8]*University of Navarra, Pamplona/ES,* [9]*Pamplona and CIBERONC, Madrid/ES,* [10]*Baylor College of Medicine, Houston/TX/USA,* [11]*International Agency for Research on Cancer, Lyon/FR*

Introduction: Low-Dose Computed Tomography (LDCT) screening can significantly improve early lung cancer detection and reduce its mortality. However, it remains a challenge to optimize the clinical management of pulmonary nodules detected within screening programs. Our previous work identified a set of informative circulating protein markers for predicting pulmonary nodule malignancy. In this study, we evaluated whether incorporating these circulating protein biomarkers along with clinical, epidemiological and image data could enhance the predictive accuracy of risk models for pulmonary nodules. **Methods:** This study is part of the Integrative Analysis of Lung Cancer Risk and Etiology (INTEGRAL) research program and is based on data from 5 LDCT screening programs in US, Canada and Europe with a total of 2,039 participants (683 cancer patients, 687 frequency matched controls with benign nodules and 669 healthy controls without nodules, nested in the same cohort). From a proteomic screening of over 1,100 circulating protein markers, 21 top-performing markers were selected in collaboration with the Lung Cancer Cohort Consortium, including six population-based cohorts. These 21 selected proteins were then quantified using absolute values. Model development was conducted using data from four LDCT screening studies: Pan-Canadian Early Detection of Lung Cancer Study (PanCan), International Early Lung Cancer Action Program (IELCAP)-Toronto, UK Lung Cancer Screening Trial (UKLS), and Pamplona-IELCAP. The model was externally evaluated in an independent cohort, the Pittsburgh Lung Screening Study-PLuSS (258 cancer patients, 257 frequency-matched controls with benign nodules).We identified the top-performing sets of protein markers using a support vector machine-recursive feature elimination (SVM-RFE) approach. For each set, we trained models to predict the risk of a benign nodule becoming malignant within five years post screen using machine learning (ML) approaches. ML models included logistic regression, random forest, and support vector machine (SVM), with logistic regression achieving the highest area under the receiver operating characteristic curve (AUC) when combined with epidemiological and imaging features through an ensemble approach. The final model was recalibrated using a logistic recalibration method. **Results:** In the independent validation set, the INTEGRAL-Proteins for Nodules (INTEGRAL-PEN) model demonstrated strong predictive performance, with an overall AUC of 0.87 (95%CI=0.83-0.90) and 0.80 (95%CI=0.77-0.83) for predicting malignancy within the first- and 5-years post screen, respectively. Our model significantly outperformed both the Brock/PanCan model (AUC=0.74) and Lung Imaging Reporting and Data System (LungRADS) (AUC=0.65) in predicting malignancy up to 5 years post screen. Compared to the other risk assessment approaches, the INTEGRAL-PEN model performed better at identifying cancers among small nodules (<10mm), among those with longer follow-up (>1 year), and at identifying early-stage lung cancers. Those without nodules have similar protein profiles to those with benign nodules. This finding suggests that the protein profiles we identified are specific to the presence of malignant nodules, not just to the presence of any type of nodules. **Conclusions:** The INTEGRAL-PEN model, validated in an external cohort, demonstrated strong performance in distinguishing malignant from benign pulmonary nodules. These findings highlight the added value of circulating protein biomarkers in improving nodule stratification in lung cancer screening. **Keywords:** Lung Cancer, Nodule malignancy, Protein biomarkers

---

MA05.07

Breathomics for Early Non-Invasive Detection of Lung Cancer: A Propectively Multicenter Study Using TD-GC-MS and Portable Micro-GC

R. Wang,[1] **H. Liang,**[1] **M. Peng,**[2] **B. Tian,**[3] **P. Fan,**[2] **H. Wang,**[2] **S. Chen,**[2] **Z. Chen,**[4] **R. Zhong,**[1] **X. Fan,**[1] **A. Ailifeire,**[1] **S. Yang,**[4] **Q. Li,**[3] **J. Wang,**[2] **J. He**[1] [1]*the First Affiliated Hospital of Guangzhou Medical University, Guangzhou/CN,* [2]*ChromX Health Co., Ltd, Guangzhou/CN,* [3]*Sichuan Clinical Research Center for Cancer, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, Affiliated Cancer Hospital of University of Electronic Science and Technology of China, Chengdu/CN,* [4]*the First People's Hospital of Foshan, Foshan/CN*

Introduction: This study evaluates the diagnostic accuracy of breathomics, using TD-GC-MS for model development and a portable micro gas chromatography- micro photoionisation detector (μGC-μPID) system for validation, to establish a non-invasive, accessible lung cancer screening method aimed at reducing mortality rates and healthcare costs of late-stage lung cancer. **Methods:** A prospective, multicenter study enrolled 5,292 participants across four hospitals, including 1,676 lung cancer patients and 4,616 controls (472 healthy individuals, 2,993 with chronic respiratory diseases). Thermal desorption-gas chromatography-mass spectrometry (TD-GC-MS) analysis of exhaled breath samples, with feature selection via orthogonal partial least squares discriminant analysis, was used to train (n=2,132) and internally validate (n=719) a logistic regression model. External validation (2,441 samples) used data from the portable μGC-μPID system, which integrates three microfabricated chips for volatile organic compound capture, separation, and detection. **Results:** Fourteen breathomic biomarkers were identified for distinguishing lung cancer patients from controls. In internal validation, the model achieved an AUC of 0.91 (95% CI: 0.88-0.94), accuracy of 78%, sensitivity of 95%, specificity of 60%, and a Negative Predictive Value (NPV) of 91%. External validation using the μGC-μPID system confirmed these results, with an AUC of 0.87, accuracy of 51%, sensitivity of 91%, specificity of 45%, and NPV of 96%. The model demonstrated high sensitivity in detecting lung cancers of various sizes, particularly excelling in identifying lesions smaller than 10 mm (AUC 0.92, sensitivity 98.2% internally; AUC 0.90, sensitivity 92.8% externally) and AJCC stage IA1 lung cancer (sensitivity 95.5% internally and 94.8% externally). In a real-world screening subset of the external validation group (n=2,055), the model correctly identified 80% of lung cancer cases and excluded 45.7% of non-cancer individuals. **Conclusions:** This largest-scale study to date demonstrates the potential of machine learning-enhanced breathomics as a highly sensitive and non-invasive approach for early lung cancer detection. Utilizing portable μGCs, this cost-effective screening method offers feasible deployment across diverse clinical settings, particularly benefiting high-risk populations by minimizing radiation exposure and improving access to care. By facilitating earlier diagnosis, breathomics has the potential to significantly reduce lung cancer mortality and enhance quality of life, supporting its integration into public health strategies. **Keywords:** Breathomics, Lung cancer, Cancer screening

---

MA05.08

Cell-Free DNA Methylation Markers in Serum and Bronchial Washing for Lung Cancer Detection: A Prospective Observational Study

Y. Song,[1] **H.S. Kwak,**[1] **Y. Seo,**[2] **S. Jang,**[2] **S. Kim,**[2] **C.-M. Choi,**[1] **T.J. Oh,**[2] **S. An,**[2] **W. Ji**[1] [1]*Asan Medical Center, University of Ulsan College of Medicine, Seoul/KR,* [2]*Genomictree Inc, Daejeon/KR*