

Metabolomic Profiling for the Early Detection of Lung Cancer

Philippe Joubert ⁽¹⁾, David Wishart ⁽²⁾, Jean-François Haince ⁽³⁾, Horacio Bach ⁽⁴⁾, Rashid Ahmed Bux ⁽⁵⁾, Paramjit S. Tappia ^(6 †), and Bram Ramjiawan ^(6 †)

⁽¹⁾ Department of Pathology, Institut Universitaire de Cardiologie et de Pneumologie de Québec, Université Laval, Québec, QC, Canada, ⁽²⁾ Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada, ⁽³⁾ BioMark Diagnostic Solutions Inc., Québec, QC, Canada,

⁽⁴⁾ Department of Medicine, Division of Infectious Diseases, University of British Columbia, Vancouver, BC, Canada, ⁽⁵⁾ BioMark Diagnostic Inc., Richmond, BC, Canada, ⁽⁶⁾ Asper Clinical Research Institute, St. Boniface Hospital, Winnipeg, MB, Canada

[†] Corresponding Authors: PTappia@sbrc.ca and BRamjiawan@sbrc.ca

OVERVIEW

Currently, the five-year survival rate of lung cancer patients is very low, largely attributed to newly diagnosed patients presenting with locally advanced or metastatic disease. The lung cancer five-year survival rate (18.6%) is lower than many other leading cancer sites, such as colorectal (64.5%), breast (89.6%) and prostate (98.2%). The five-year survival rate for lung cancer is 56% for cases detected when the disease is still localized (within the lungs). However, less than 30% of lung cancer cases are diagnosed at an early stage. (Stage I & II). For distant tumors (spread to other organs) the five-year survival rate is only 5%. More than 50% of lung cancer cases die within one year of being diagnosed. Accordingly, early diagnosis is key to the successful treatment, management and care of lung cancer.

INTRODUCTION

Cancer can be regarded as a metabolic disease. Since metabolism reflects the biochemical state of a healthy or unhealthy cell, then the unique metabolic characteristics (fingerprints) may assist in the determination of the stage and location of a carcinoma and allow for the distinction between metabolic processes in healthy cells as well as cancer cells. Here we describe a quantitative metabolomic follow on study that has succeeded in revalidating the metabolites discovered and validated in a previous clinical study (Cancers, 2020 Mar 7;12(3):622). The large sample size (total of 813 plasma samples) allows for the incorporation of other clinical parameters such as smoking history and identifies additional metabolites that can be included to further enhance the robustness of the assay. This large validation study also included individuals with other pulmonary diseases (pneumonia, tuberculosis, chronic obstructive pulmonary disease (COPD), asthma) in the control group to help determine whether previously discovered set of blood metabolites were specific to lung cancer alone or whether they are also markers for general lung distress.

METHODS

Targeted metabolomic techniques were used to discover and validate plasma biomarkers for the diagnosis of early-stage non-small cell lung cancer (NSCLC). A total of 142 different metabolites were tested by our quantitative LC-MS/MS method. In quantitative metabolomic studies, missing values normally indicate that the metabolite fell below the assay's limit of detection (LOD). Therefore, metabolites with more than 80% of missing values (in all groups) were removed from further analysis. Logistic regression with a Lasso feature selection algorithm was used to develop predictive models of lung cancer using both metabolite and clinical variables. The area under the receiver-operator characteristic curves (AUC), sensitivities/specificities at selected cut-off points and the 95% confidence intervals were calculated for each cancer stage. Cut-off points were selected by calculating the Youden Index ($J = \max \{ \text{Sensitivity} + \text{Specificity} - 1 \}$). Predictive models with $AUC > 0.8$ were developed and validated using selected metabolites and other clinical data for detecting different stages of NSCLC. Analysis and predictive modeling was performed by Dr. Lun Zhang at The Metabolomics Innovation Centre (TMIC).

PATIENT DEMOGRAPHICS

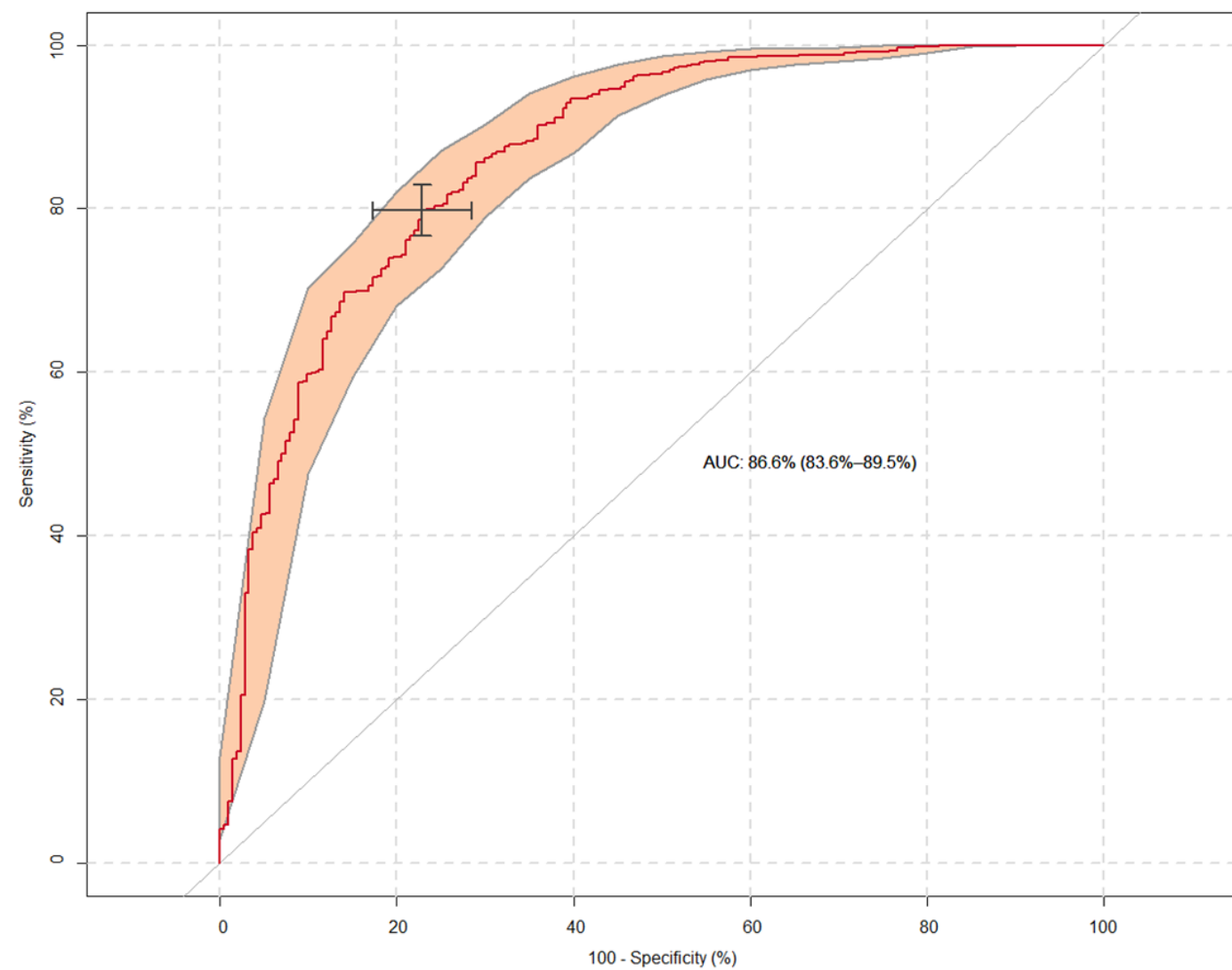
Table 1 Patient Demographics by Study Groups

		Case	Control
Age (Median)		65	60
Sex	Male	284 (47%)	116 (54%)
	Female	315 (53%)	98 (46%)
Smoking Status	Current	137 (23%)	28 (13%)
	Former	416 (69%)	97 (45%)
	Never	46 (8%)	89 (42%)
Cancer Stages	Stage I Adenocarcinoma	200 (33%)	
	Stage I Squamous	75 (13%)	
	Stage II Adenocarcinoma	98 (16%)	
	Stage II Squamous	43 (7%)	
	Advanced NSCLC	50 (8%)	
	NETs	120 (20%)	
	Mesothelioma	13 (2%)	
Total		599	214

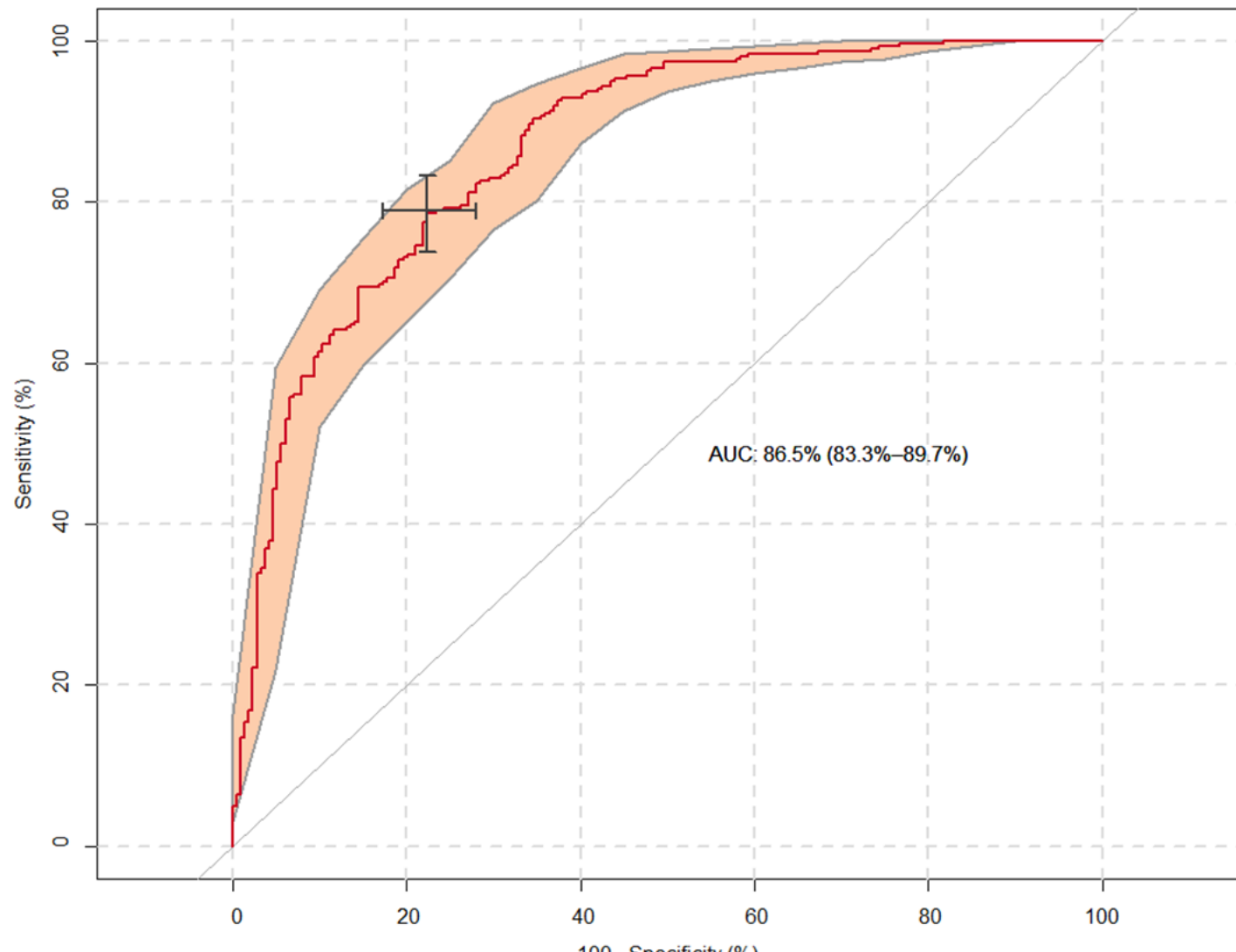
Plasma samples from **599 patients** with biopsy-confirmed lung cancer along with age and sex-matched plasma samples from **214 controls** were analyzed. The control group consists of 90 healthy individuals and 124 with other lung diseases including asthma, COPD, bronchiectasis and COVID. The lung cancer sub-groups include early-stage (Stage I & II) lung adenocarcinoma and squamous cells carcinoma, advanced stages (Stage III & IV) NSCLC, Lung neuroendocrine tumors (NETs) and malignant mesothelioma.

RESULTS

All Lung Cancers Using Previously Identified 5-Biomarker Panel



Stage I & II NSCLC Using Previously Identified 5-Biomarker Panel



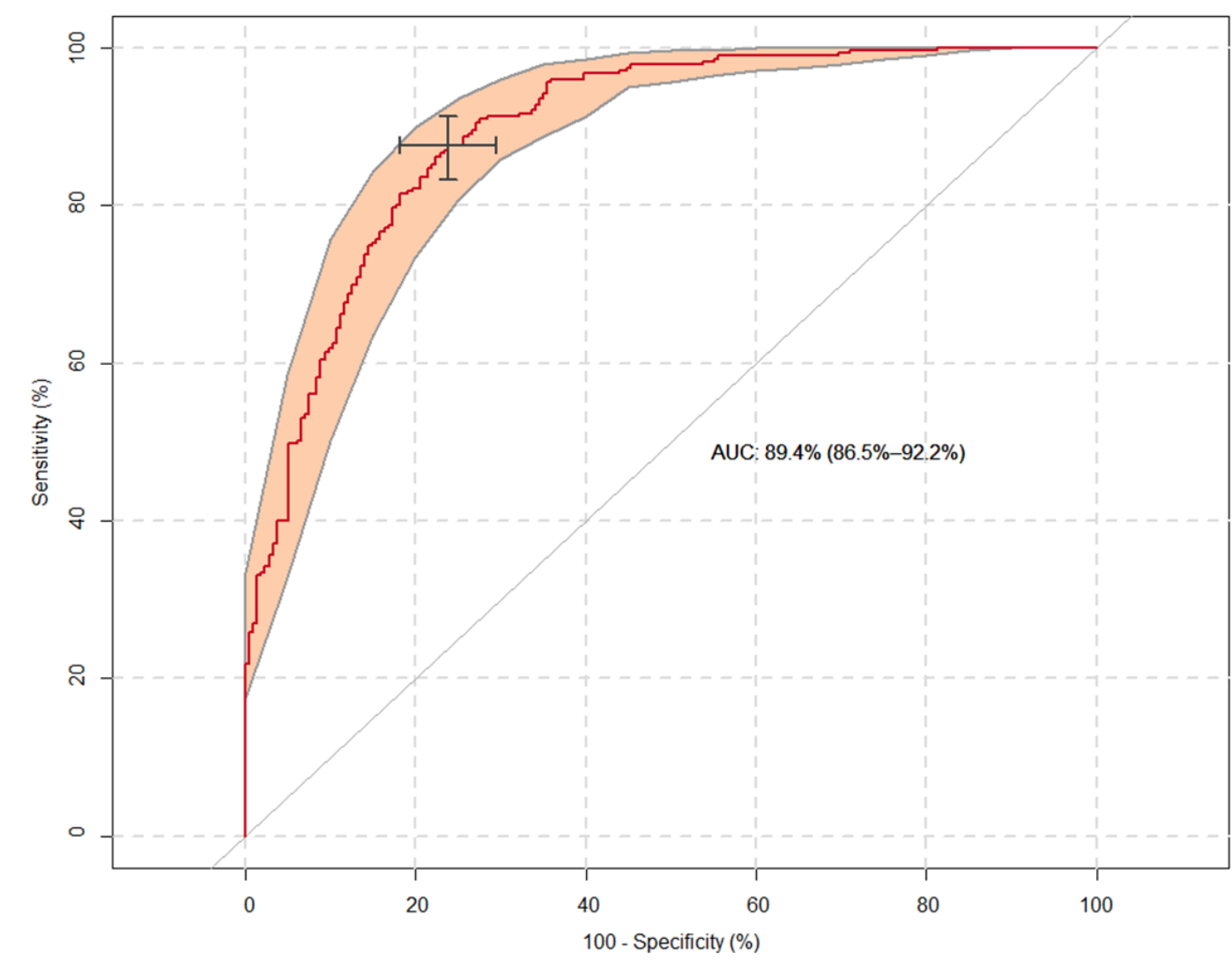
Univariate and multivariate statistical analysis previously identified 5 biomarkers that included β -hydroxybutyric acid, LysoPC 20:3, PC ae C40:6, citric acid, and fumaric acid as being significantly different between healthy controls and patients with early stage (I/II) lung cancer.

Table 2 Logistic Regression Based Optimal Model For Each Cancer Stage

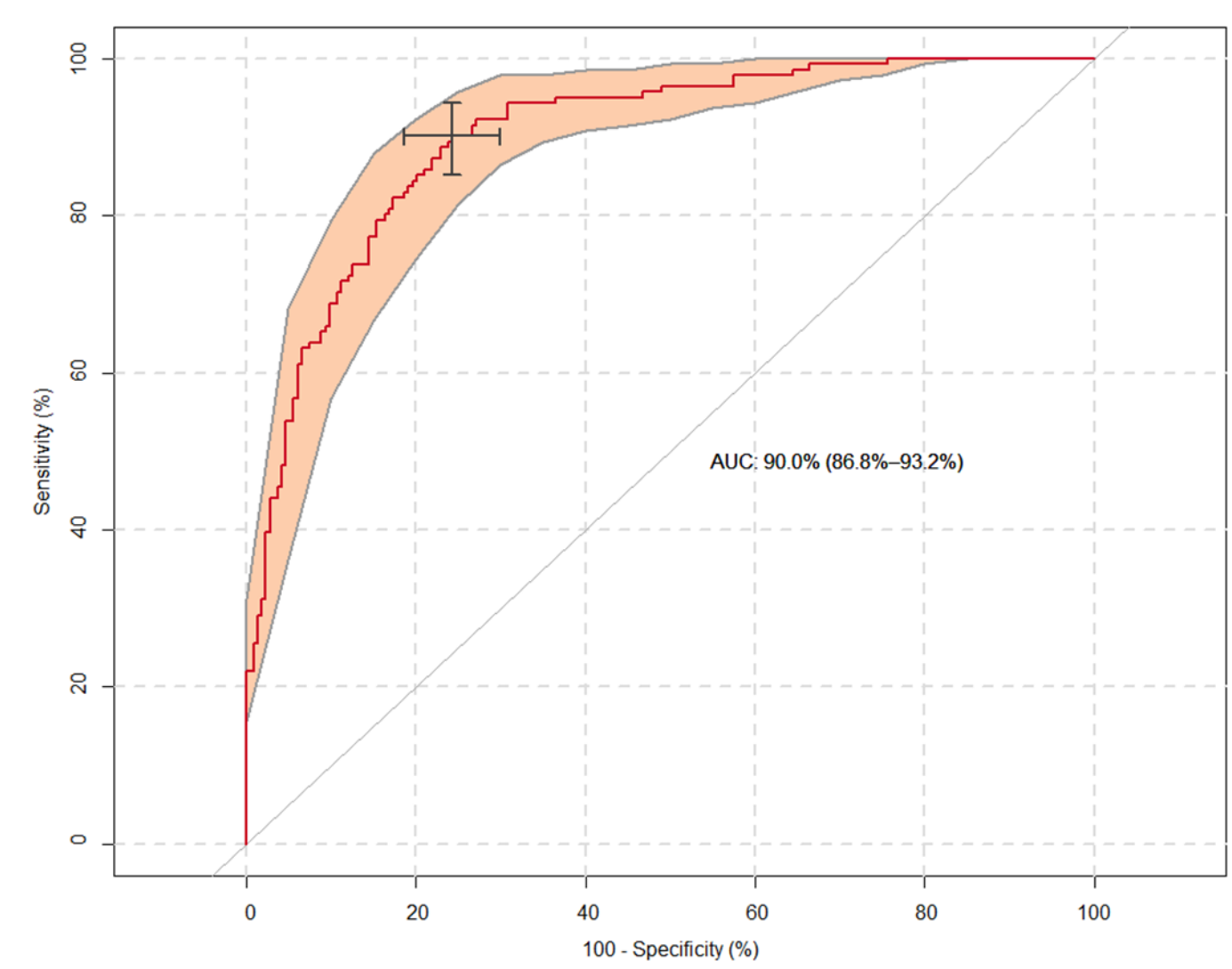
Performance of Logistic Regression Model		AUC (95% CI)	Sensitivity	Specificity
All Lung Cancers Cases vs Control	5-Biomarker Panel	86.6% (83.6% - 89.5%)	80.0%	75.7%
Stage I & II Lung Cancers Cases vs Control	5-Biomarker Panel	86.5% (83.3% - 89.7%)	78.5%	77.6%
Stage I NSLC Cases vs Control	9-Biomarker Panel	89.4% (86.5% - 92.2%)	87.6%	76.2%
Stage II NSLC Cases vs Control	9-Biomarker Panel	90.0% (86.8% - 93.2%)	90.0%	75.7%
Stage III & IV NSLC Cases vs Control	9-Biomarkers Panel	92.7% (89.0% - 96.3%)	82.0%	91.1%
Stage I & II NSLC Cases vs Control	9-Biomarkers Panel	89.5% (86.9% - 92.2%)	89.9%	75.2%
Stage I & II NSLC Cases vs Control	9-Biomarker Panel + Smoking	91.4% (89.0% - 93.8%)	91.3%	77.6%

RESULTS

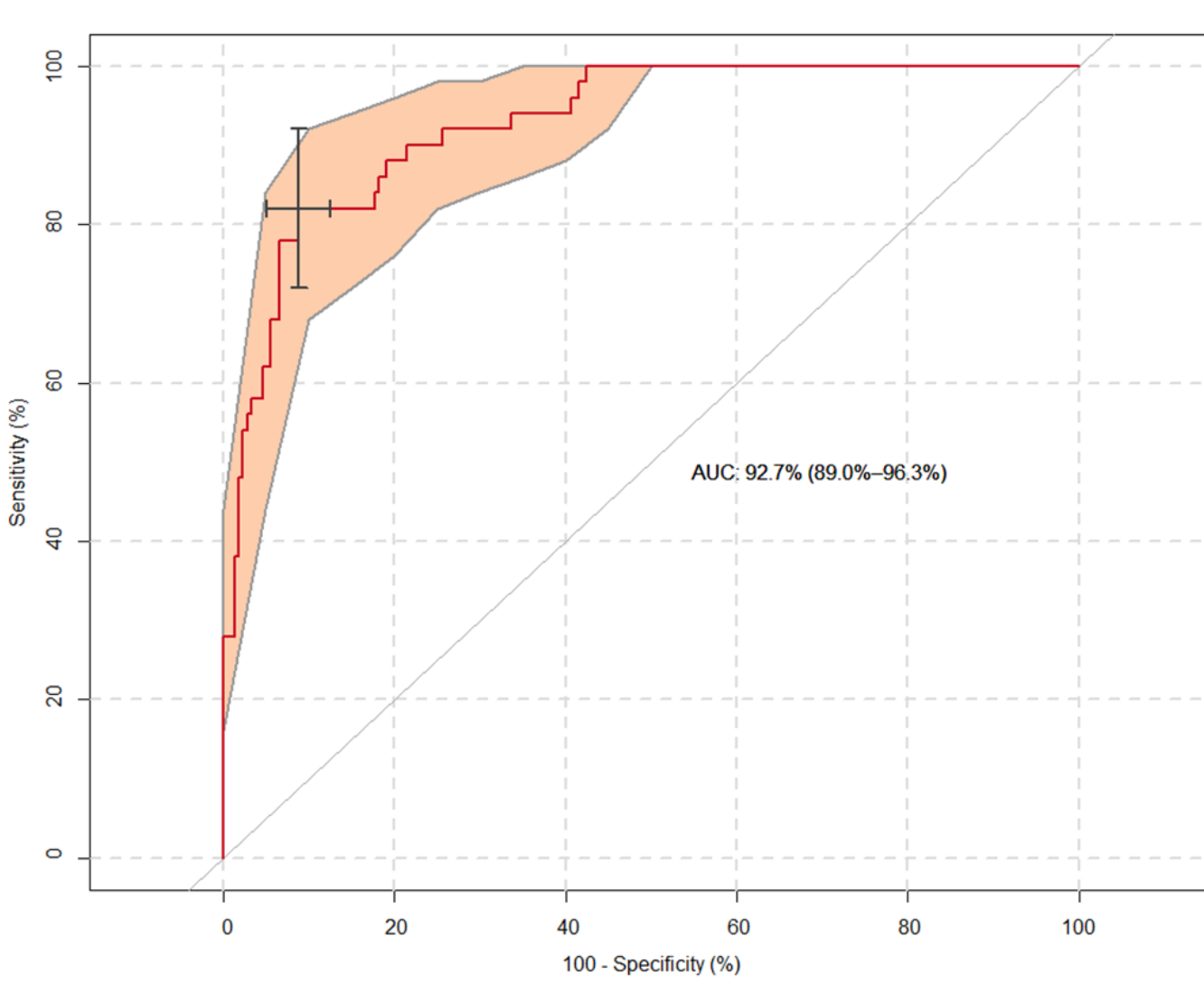
Stage I NSCLC 9-Biomarker Panel



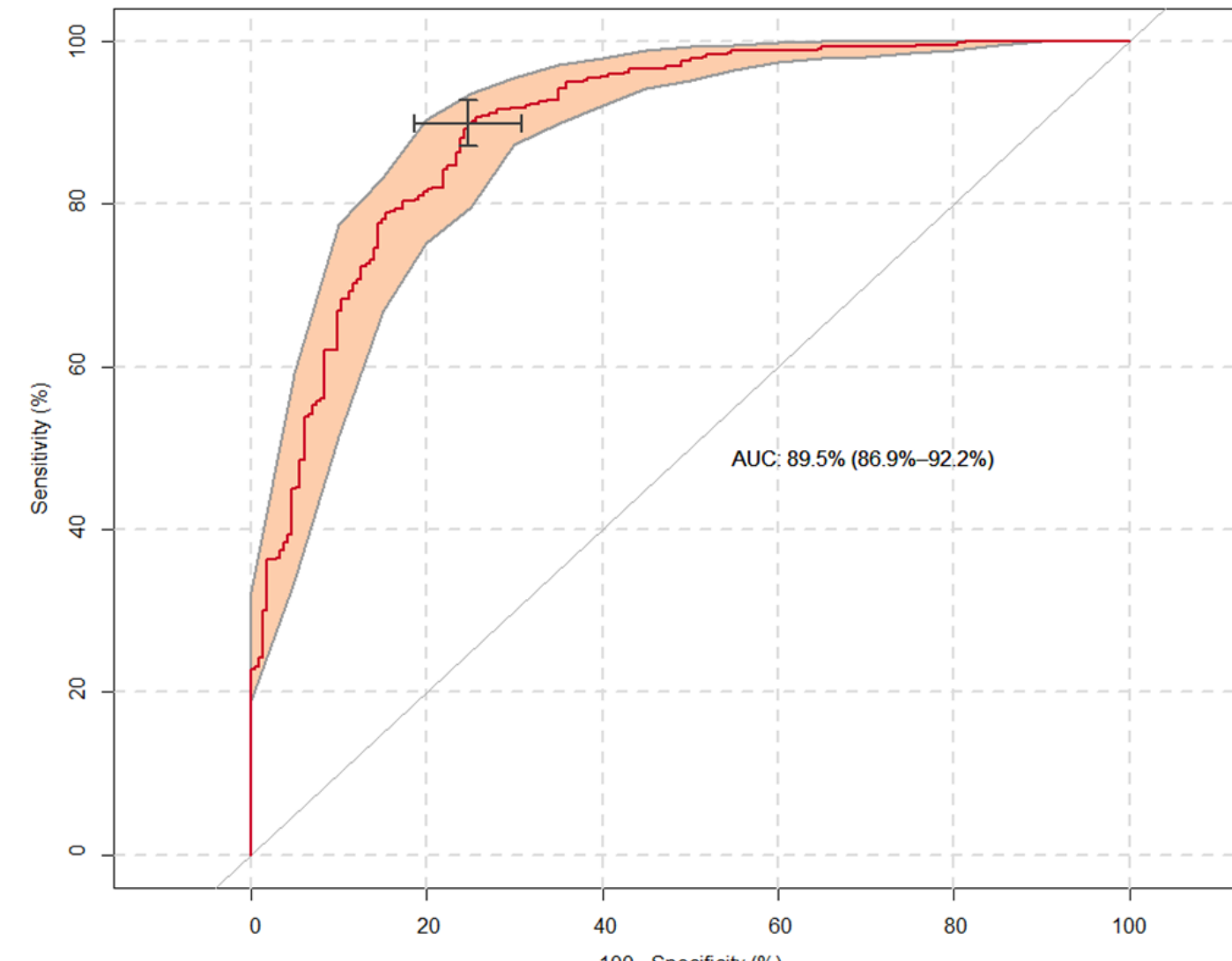
Stage II NSCLC 9-Biomarker Panel



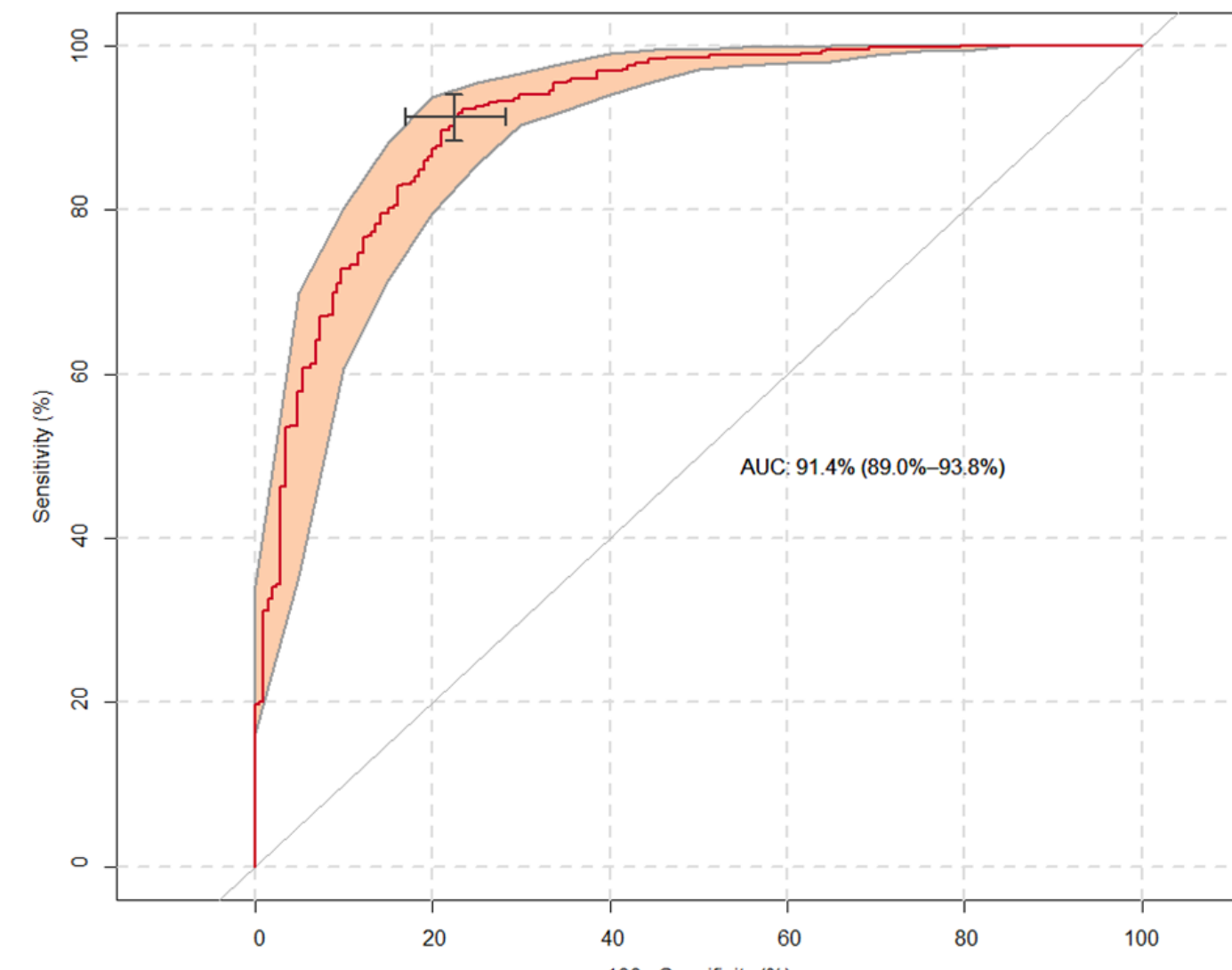
Stage III & IV NSCLC 9-Biomarker Panel



Stage I & II NSCLC 9-Biomarker Panel



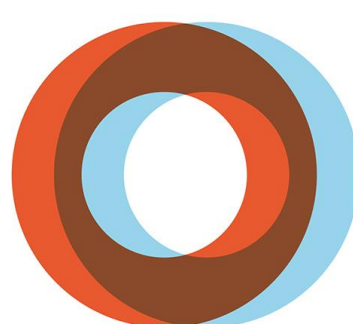
Stage I & II NSCLC 9-Biomarker Panel + Smoking History



- The metabolites discovered were robust and continue to demonstrate reproducibility for early-stage NSCLC cancer detection.
- Confounding lung diseases did not impact the sensitivity of the assay. The markers discovered can differentiate lung cancer from lung diseases quite reliably.
- Adding the smoking history to early-stage NSCLC patients slightly increased both the sensitivity and specificity of the assay.

CONCLUSION

- A view of cancer as primarily a metabolic disease will impact approaches to cancer management and prevention.
- We identified and re-validated a simple, high-performing, metabolite-based test for detecting early stage (I/II) NSCLC patients in plasma.
- This validation of previously identified biomarkers in a larger cohort illustrate how metabolomics fingerprinting has the potential to map out early biochemical changes in cancer cells and hence provides an opportunity for faster and more sensitive early diagnosis where treatment can be more effective.
- The 9-biomarker panel could enable the establishment of a blood-based routine screening test for smokers at high risk of lung cancer in a cost-effective, accurate and reliable manner.



INSTITUT UNIVERSITAIRE
DE CARDIOLOGIE
ET DE PNEUMOLOGIE
DE QUÉBEC
UNIVERSITÉ LAVAL



TMIC The Metabolomics
Innovation Centre



Hôpital St-Boniface Hospital
RECHERCHE • RESEARCH



BioMark™
DIAGNOSTIC SOLUTIONS INC



NRC-CIRC
Canada
NRC Industrial Research
Assistance Program



CIHR IRSC
Canadian Institutes of
Health Research
Instituts de recherche
en santé du Canada



MEDTEQ
L'INNOVATION POUR LA SANTÉ
INNOVATION FOR HEALTH