Lung Cancer Detection via Whole-Transcriptome RNA Sequencing of Nasal Epithelium

Carla Lamb,[1] Jie Ding,[2] Saeed Saberi,[2] Dan Pankratz,[2] Joshua Babiarz,[2] Janna Chamberlin,[2] Sangeeta Bhorade,[2] Jonathan Wilde,[2] P. Sean Walsh,[2] Lori Lofaro,[2] Jing Huang,[2] Chris Stevenson,[3] Marc Lenburg,[4] Duncan Whitney,[3] Giulia C. Kennedy[2] and Avrum Spira[3]

1. Lahey Hospital & Medical Center. 2. Veracyte Inc. 3. Johnson & Johnson Innovation LLC. 4. Boston University.

PURPOSE

- The goal of lung nodule evaluation is to expedite treatment of malignant nodules and minimize procedures for benign nodules.
- The Percepta Genomic Sequencing Classifier (GSC) is based on the concept of “field of injury”, where altered gene expression is seen in bronchial epithelial cells in current/ prior smokers with malignancy.[1,2]
- This “field of injury” principle has been extended to the nose with evidence that gene expression changes associated with cancer can be detected in the nasal epithelium.[3]
- We report the feasibility of a molecular test developed from non-invasive nasal swab samples for lung cancer detection in current and former smokers.

METHODS

- Swab samples of the nasal epithelium were prospectively collected during the AEGIS I and II clinical trials from current and former smokers with suspected lung cancer lesions found on chest CT.
- Patients were followed for up to one year until a final diagnosis of lung cancer or benign disease was made. Adjudicated Benign and Malignant diagnoses were obtained for each patient.
- A total of 675 subjects were divided into a training set of 411 nasal samples and an independent test set of 261 nasal samples.
- Extracted RNA was analyzed using whole-transcriptome RNA sequencing.[4]
- Sequencing data was analyzed by the Veracyte feature extraction pipeline.
- Machine learning models were developed using gene expression as well as clinical factors such as age, gender, smoking status and nodule characteristics.
- The nasal classifier performance was compared to a clinical risk model (Gould) for risk prediction of malignancy.

CONCLUSIONS

- For Low risk classification, at fixed high sensitivity (96.6%)
  - Clinical genomic classifier achieved overall specificity of ~45%, almost 20 absolute points higher than the Gould model (specificity ~26%)
  - In a population with 25% cancer prevalence, the Nasal Classifier results in over 70% more Benign patients being classified as low risk compared to the Gould model
- For high risk classification, at fixed high specificity (~94%)
  - Clinical genomic classifier achieved overall sensitivity of 50% compared to 42% for Gould model
  - In a population with 25% cancer prevalence, the Nasal Classifier results in over 18% more malignant patients being classified as high risk than the Gould model
- A nasal genomic-clinical classifier developed using whole transcriptome RNA sequencing shows promise in detecting lung-cancer related gene expression changes in nasal swabs from high-risk smokers. Further research is warranted to understand its impact on patient care in a real-world clinical setting.

CLINICAL IMPLICATIONS

A nasal genomic-clinical classifier has the potential to serve as a non-invasive tool for lung cancer risk-stratification among patients with pulmonary nodules.

Training/Independent test set demographic information

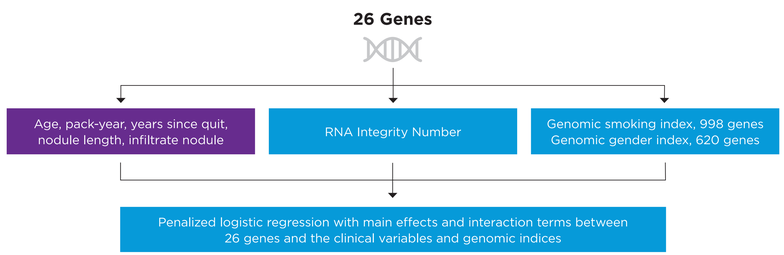
Training Set: 411 patients

Category N=411	Sub-category	Benign 85	Malignant 326
Sex	Male	51	210
	Female	34	116
Age	Median	58	65
Smoking status	Current	25	146
	Former	60	180
Pack-year	Median	30	47
Nodule size (cm)	< 1	9	8
	1 to 2	16	52
	>2 to <3	9	45
	>=3	27	195
	Ill defined infiltrate	20	14
	Unknown	4	12
Nodule location	Central	22	120
	Peripheral	33	90
	Both	26	105
	Unknown	4	11
Histology	SCLC	—	42
	NSCLC	—	250
	Other	—	34
NSCLC type	Adenocarcinoma	—	106
	Squamous	—	100
	Large Cell	—	11
	Other	—	33

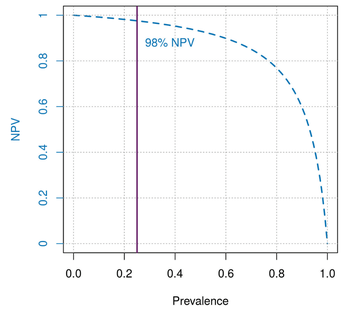
Independent Test Set: 261 patients

Category N=261	Sub-category	Benign 57	Malignant 204
Sex	Male	38	131
	Female	19	73
Age	Median	57	66
Smoking status	Current	23	104
	Former	34	100
Pack-year	Median	20	48
Nodule size (cm)	< 1	7	6
	1 to 2	17	33
	>2 to <3	6	25
	>=3	14	125
	Ill defined infiltrate	10	10
	Unknown	3	5
Nodule location	Central	24	72
	Peripheral	21	61
	Both	8	65
	Unknown	4	6
Histology	SCLC	—	26
	NSCLC	—	161
	Other	—	17
NSCLC type	Adenocarcinoma	—	69
	Squamous	—	62
	Large Cell	—	10
	Other	—	20

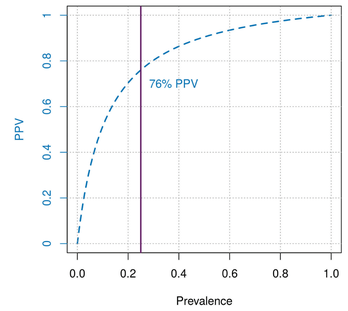
Preliminary Nasal Classifier Combines CLINICAL and GENOMIC Features



The Nasal Classifier performance modeled on a population with 25% cancer prevalence

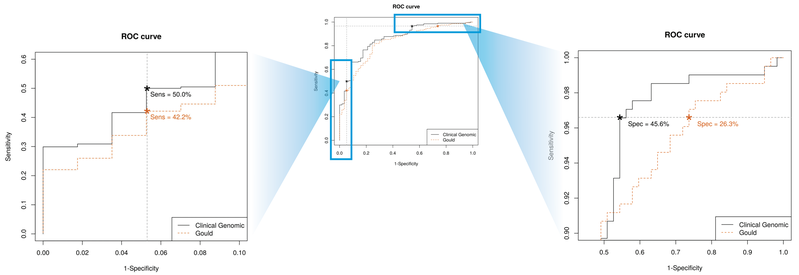


Classification as Low Risk with a specified cut-off
Sensitivity = ~96 %, Specificity = ~45%



Classification as High Risk with a specified cut-off
Sensitivity = ~50%, Specificity = ~94%

Comparison of the Nasal Classifier with a Clinical Model (Gould)



NPV/PPV at 25% prevalence						
	Sensitivity	Specificity	NPV	Sensitivity	Specificity	PPV
Nasal Classifier	96.6%	45.6%	97.6% [94.8 – 98.9]	50.0%	94.7%	76.0% [51.1 – 90.6]
Gould	96.6%	26.3%	95.8% [90.8 – 98.2]	42.2%	94.7%	72.8% [46.7 – 89.0]

Initial Independent Test Set Performance using the Nasal Classifier with one cut-off

Incoming Suspicious Nodules	Sensitivity	Specificity
All (N=261)	**96.6%** [93.1 – 98.6]	**45.6%** [32.4 – 59.3]

Subsequent Independent Test Set Performance using Classifier with two cut-offs

Sensitivity > 96% for Classification as Low Risk

Incoming Suspicious Nodules	Sensitivity	Specificity
All (N=261)	**96.6%** [93.1 – 98.6]	45.6% [32.4 – 59.3]

Specificity > 94% for Classification as High Risk

Incoming Suspicious Nodules	Sensitivity	Specificity
All (N=261)	50.0% [42.9 – 57.1]	**94.7%** [85.4 – 98.9]

REFERENCES

1. Spira A, Beane JE, Shah V, Steiling K, Liu G, Schembri F, Gilman S, Dumas YM, Calner P, Sebastiani P, Sridhar S, Beamis J, Lamb C, Anderson T, Gerry N, Keane J, Lenburg ME, Brody JS. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. Nat Med 2007; Mar;13(3):361-6. Epub 2007 Mar 4.
2. Silvestri G, Vachani A, Whitney D, et al A bronchial genomic classifier for the diagnostic evaluation of lung cancer. N Engl J Med 2015; Jul 16;373(3):243-51.
3. Billatos E, Vick JL, Lenburg ME, Spira AE. The Airway Transcriptome as a Biomarker for Early Lung Cancer Detection. Clin Cancer Res. 2018 Jul 1;24(13):2984-2992.
4. Hao Y, Choi Y, Babiarz J et al. Analytical validation performance of Afirma Genomic Sequencing classifier in the diagnosis of cytologically indeterminate thyroid nodules. Front Endocrinol (Lausanne). 2019; 10: 438.
5. Gould, MK et al. Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013;143(5 Suppl):e93S-e120S.