Accuracy of the Next Generation Percepta Genomic Sequencing Classifier (GSC) for the Diagnosis of Suspicious Intermediate Pulmonary Nodules

Sangeeta Bhorade,[1] Michael Bernstein,[2] Travis L. Dotson,[3] David Feller-Kopman,[4] Hans Lee,[4] Yoonha Choi,[1] Lori Lofaro,[1] Jing Huang,[1] Duncan Whitney,[5] Christopher Stevenson,[5] Avrum Spira,[5] Giulia Kennedy,[1] Momen Wahidi,[6] and Peter Mazzone[7]

1. Veracyte, Inc. 2. Stamford Hospital. 3. Wake Forest Baptist Health. 4. Johns Hopkins University. 5. Johnson & Johnson Innovation LLC. 6. Duke University. 7. Cleveland Clinic

BACKGROUND

- The goal of lung nodule evaluation is to expedite treatment of malignant nodules and minimize procedures for benign nodules.
- While guidelines recommend bronchoscopic workup for patients with intermediate risk nodules, up to 60% of these bronchoscopies are nondiagnostic.
- The Percepta Genomic Sequencing Classifier (GSC) development is based on the concept of "field of injury", which shows altered cancer associated gene expression in bronchial epithelial cells in current/ prior smokers with malignancy.
- Percepta GSC utilizes RNA sequencing of transcripts from 1232 genes, as well as clinical factors, to achieve both down-classification and up-classification of pre-test ROM.

OBJECTIVE AND ENROLLMENT CRITERIA

Objective:

We aimed to prospectively validate the accuracy of next-generation Percepta, the Genomic Sequencing Classifier (GSC) using a combined cohort from AEGIS I and II & Percepta Registry

Patient eligibility criteria:

INCLUSION:

- 21 years or older with a pulmonary nodule on Chest CT scan
- Current or former smoker (>100 cigarettes in a lifetime)
- An inconclusive bronchoscopy

EXCLUSION:

- Patients with concurrent or prior cancers at the time of enrollment

METHODS

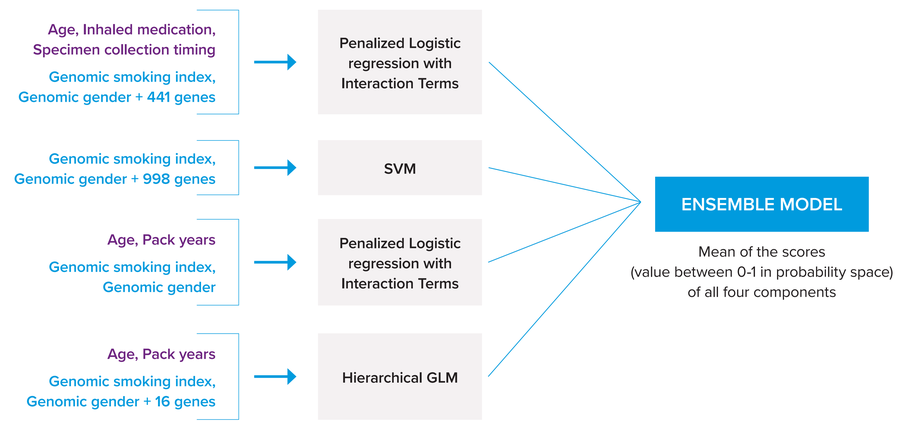
Study Design and Patient Population:

- The training set included 1611 patients from four cohorts and the validation set included 412 patients from three independent cohorts.
- Two bronchial brushings were obtained during the initial bronchoscopy from cytologically normal appearing bronchial epithelial cells in the mainstem bronchus.
- Patients in each cohort were followed until a diagnosis (adjudicated by panel of 3 physicians) was established or for 12 months after bronchoscopy.
- The treating physician determined the pre-test probability of risk of malignancy (low (<10%), intermediate (10-60%), high (>60%)) prior to the bronchoscopy.

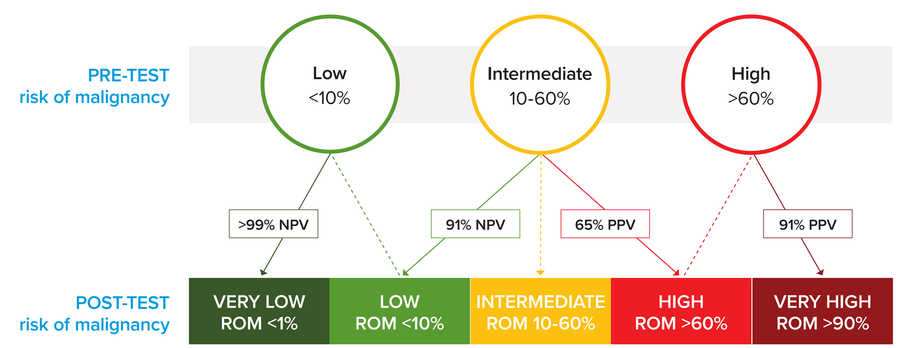
Algorithm Development and Clinical Validation:

- Machine-learning algorithms inputting RNA sequencing results and clinical factors were used from >1600 patients to train the new Percepta GSC.
- The final classifier is an ensemble of four models which combine genomic and clinical features, as well as their interactions. The algorithm was locked, and analytical verification studies performed to ensure robustness in the laboratory.
- We then assessed the diagnostic accuracy of the locked algorithm on an independent validation test set of patients (n=412).
- We measured positive predictive value (PPV), negative predictive value (NPV), specificity, and sensitivity in each pre-test risk category.

Percepta GSC Combines Clinical and Genomic Features



Percepta GSC Performs Seven Types of Risk Classifications



REFERENCES

1. Silvestri G, Vachani A, Whitney D, et al. A bronchial genomic classifier for the diagnostic evaluation of lung cancer. *N Engl J Med*. 2015.
2. Rivera P, Mehta A, and Wahidi, M. Establishing the diagnosis of lung cancer. Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines *Chest* 2013.
3. Ost DE, Ernst, A, Lei X, et al. Diagnostic yield and complications of bronchoscopy for peripheral lung lesions. Results of the AQuIRE registry. *Am J Resp Crit Care*. 2016.

Clinical Validation Dataset: Patient Demographics

Characteristic		AEGIS I and II (N=246)	Percepta Registry (N=166)	Total (N=412)
Sex	Female	83	84	167
	Male	163	82	245
Race	White	192	132	324
	Black	42	29	71
	Other	12	4	16
	Unknown	0	1	1
Smoking status	Current	107	73	180
	Former	139	93	232

Clinical Validation Dataset: Breakdown by Pre-Test Risk of Malignancy and Cancer Prevalence

Cohort	LOW RISK		INTERMEDIATE RISK		HIGH RISK		Total
	Benign	Malignant	Benign	Malignant	Benign	Malignant	
AEGIS I &II	56	2	58	24	21	85	246
Registry	20	2	77	29	17	21	166
Total	80 (19%)		188 (46%)		144 (35%)		412
Cancer Prevalence in each group	4/80 = 5%		53/188 = 28%		106/144 = 74%		

CONCLUSION

1. Percepta GSC classifier utilizes RNA sequencing of transcripts from 1232 genes, as well as clinical factors, in four different algorithms that are combined in a final ensemble model.
2. Percepta GSC was trained on >1600 samples from four cohorts and prospectively validated on three independent cohorts (412 patients).
3. Percepta GSC provides increased functionality in the intermediate-risk patients:
   - Down-classifies intermediate-risk patients to low-risk with 91% NPV
   - Up-classifies intermediate-risk patients to high risk with PPV >65%
4. Percepta GSC provides additional functionalities:
   - Down-classifies low-risk patients to very low-risk with NPV >99%
   - Up-classifies high-risk patients to very high with PPV >90%

© 2019 Veracyte, Inc. All rights reserved. The Veracyte and Percepta names and logos are trademarks of Veracyte. PC221.1.1907