CHEST

Assessment of Plasma Proteomics Biomarker's Ability to Distinguish Benign From Malignant Lung Nodules

Results of the PANOPTIC (Pulmonary Nodule Plasma Proteomic Classifier) Trial

*Gerard A. Silvestri, MD; Nichole T. Tanner, MD; Paul Kearney, PhD; Anil Vachani, MD; Pierre P. Massion, MD; Alexander Porter, MD; Steven C. Springmeyer, MD; Kenneth C. Fang, MD; David Midthun, MD; Peter J. Mazzone, MD, MPH; for the PANOPTIC Trial Team**

BACKGROUND: Lung nodules are a diagnostic challenge, with an estimated yearly incidence of 1.6 million in the United States. This study evaluated the accuracy of an integrated proteomic classifier in identifying benign nodules in patients with a pretest probability of cancer (pCA) ≤ 50%.

METHODS: A prospective, multicenter observational trial of 685 patients with 8- to 30-mm lung nodules was conducted. Multiple reaction monitoring mass spectrometry was used to measure the relative abundance of two plasma proteins, LG3BP and C163A. Results were integrated with a clinical risk prediction model to identify likely benign nodules. Sensitivity, specificity, and negative predictive value were calculated. Estimates of potential changes in invasive testing had the integrated classifier results been available and acted on were made.

RESULTS: A subgroup of 178 patients with a clinician-assessed pCA ≤ 50% had a 16% prevalence of lung cancer. The integrated classifier demonstrated a sensitivity of 97% (CI, 82-100), a specificity of 44% (CI, 36-52), and a negative predictive value of 98% (CI, 92-100) in distinguishing benign from malignant nodules. The classifier performed better than PET, validated lung nodule risk models, and physician cancer probability estimates ($P < .001$). If the integrated classifier results were used to direct care, 40% fewer procedures would be performed on benign nodules, and 3% of malignant nodules would be misclassified.

CONCLUSIONS: When used in patients with lung nodules with a pCA ≤ 50%, the integrated classifier accurately identifies benign lung nodules with good performance characteristics. If used in clinical practice, invasive procedures could be reduced by diverting benign nodules to surveillance.

TRIAL REGISTRY: ClinicalTrials.gov; No.: NCT01752114; URL: www.clinicaltrials.gov).

CHEST 2018; 154(3):491-500

KEY WORDS: biomarker; diagnosis; lung cancer; proteomics; pulmonary nodules; risk models

FOR EDITORIAL COMMENT, SEE PAGE 467

ABBREVIATIONS: AUC = area under the receiver-operating characteristic curve; NPV = negative predictive value; pCA = probability of cancer; TTNA = transthoracic needle biopsy; VA = Veterans Affairs

AFFILIATIONS: From the Thoracic Oncology Research Group Division of Pulmonary and Critical Care Medicine (Drs Silvestri and Tanner), Medical University of South Carolina, Charleston, SC; Health Equity and Rural Outreach Innovation Center (Dr Tanner), Ralph H. Johnson Veterans Affairs Hospital, Charleston, SC; Integrated Diagnostics (Drs Kearney, Porter, Springmeyer, and Fang), Seattle, WA; Division of Pulmonary, Allergy, and Critical Care Medicine (Dr Vachani), Penn Lung Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA; Thoracic Program (Dr Massion),

Chest CT imaging has increasingly been used for a myriad of medical presentations such that the estimated incidence of pulmonary nodules in the United States is 1.6 million per year.[1] Add to that the findings of the National Lung Screening Trial, which showed low-dose chest CT imaging led to a nodule detection rate of 25% on the prevalence screen, the evaluation and management of pulmonary nodules are certain to be an increasingly common issue needing to be addressed.[2]

The approach to a pulmonary nodule begins with estimating the probability of malignancy.[3] Depending on probability of cancer (pCA), patients are managed with either CT surveillance, additional diagnostic testing (including PET, CT-guided transthoracic needle biopsy [TTNA], and/or bronchoscopy), or definitive therapy with surgical excision. All approaches should account for patient health status and preferences. Although physicians are skilled at predicting malignancy in pulmonary nodules, they may not follow guideline-based care when managing patients with differing risks.[4,5]

We previously developed a 13-protein blood test to differentiate benign from malignant lung nodules using multiple reaction monitoring mass spectrometry.[6] The proteomic classifier was more accurate than a four-parameter clinical model, independent of patient age, tobacco use, nodule size, and the presence of COPD.[7] The test was further refined to improve performance for lower risk nodules in which the diagnostic dilemma between surveillance and invasive procedures is most challenging. It was discovered that the accuracy of two of the proteins, LG3BP and C163A (independently linked to lung cancer and the inflammatory response to cancer),[8-10] could be optimized for evaluating lower risk nodules by integrating them with five clinical risk factors. LG3BP, in particular, has been previously studied in a network analysis implicating its overexpression to lung cancer pathways and associated transcription factors.[6] This integrated classifier was trained on an independent set of 222 prospectively collected plasma samples from patients with 8- to 20-mm lung nodules.[11] The current article is a second prospective, multicenter, observational clinical validation study with retrospective evaluation assessing the performance characteristics of this integrated classifier in patients presenting with undiagnosed low to moderate risk lung nodules.

Patients and Methods

Pulmonary Nodule Plasma Proteomic Classifier Study Design and Enrollment

The Pulmonary Nodule Plasma Proteomic Classifier (PANOPTIC) study is a prospective, multicenter, observational study with retrospective evaluation of the performance of the integrated classifier test comprising two proteins and five clinical risk factors. Physicians, study subjects, and laboratory and statistical personnel were blinded to the results of the test and clinical information. The blinding protocol was strictly followed, and the results of the test did not direct or influence patient care. Thirty-three sites (31 US and two Canadian) were included. All sites had institutional review board approval, and informed written consent was obtained from all eligible participants (e-Table 1).

Eligible patients were ≥ 40 years old with pulmonary nodules 8 to 30 mm in diameter presenting within 60 days of the baseline CT scan to a pulmonologist and/or a thoracic surgeon. Mediastinal lymphadenopathy was not an exclusion. Patients were ineligible if they had any of the following: undergone any attempt at a previous biopsy of the nodule in question; undergone a prior CT scan or PET/CT scan that had identified the lung nodule under consideration; a current or previous diagnosis of any cancer within 2 years of lung nodule detection (except for nonmelanoma skin cancer); or received any blood products within 30 days of study enrollment.

A power analysis was performed to assess the number of benign and malignant nodules required to perform a statistical test of significance for area under the receiver-operating characteristic curve (AUC) assuming an α of 0.05, power of 0.80, and an assumed performance with AUC of 0.70. The number necessary to have adequate power was 21 benign and 21 malignant nodules.

Data Collection

At enrollment and prior to subsequent testing, physicians (pulmonologists or thoracic surgeons to whom the patients had been referred) assessed nodule pretest pCA as 0% to 5%, 6% to 10%, 11% to 20%, and so forth. Physicians estimated the pretest probability at their own discretion. Patient demographic and nodule characteristics were collected at baseline and from subsequent CT imaging studies. Imaging studies were interpreted by site radiologists as part of usual care, and nodule characteristics (including location,

Vanderbilt-Ingram Comprehensive Cancer Center, Nashville, and Veterans Affairs, Tennessee Valley Healthcare System, Nashville Campus, Nashville, TN; Mayo Clinic (Dr Midthun), Rochester MN; and the Respiratory Institute (Dr Mazzone), Cleveland Clinic, Cleveland, OH.

Drs Silvestri and Tanner contributed equally to this article.

FUNDING/SUPPORT: Integrated Diagnostics provided funding for the study as well as laboratory and biostatistical support.

*A list of investigators and coordinators in the PANOPTIC (Pulmonary Nodule Plasma Proteomic Classifier) trial is provided in the supplementary material.

CORRESPONDENCE TO: Gerard A. Silvestri, MD, Medical University of South Carolina, 96 Jonathan Lucas St, Room 816 CSB, Charleston, SC 29425; e-mail: silvestri@musc.edu

DOI: https://doi.org/10.1016/j.chest.2018.02.012

size, and edge characteristics) were collected from the reports by site study personnel. Blood samples were obtained and processed for storage and shipment. Data on subsequent procedures, including bronchoscopy, transthoracic needle biopsy, and surgery, were obtained and entered by the enrolling site until definitive diagnosis, nodule resolution, or at least 1 year of radiographic stability according to chest CT imaging. PET results were recorded as positive for standard uptake values $\geq$ 2.5. Pathology reports were collected from biopsy samples or surgical resection.

Plasma Proteomic Analysis

Proteomic analysis of two plasma proteins, LG3BP and C163A, was performed by using multiple reaction monitoring mass spectroscopy as previously described.[6,7] These results were integrated with the five clinical risk factors (age [in years], smoking status [never, current, or former], nodule diameter [largest diameter], edge characteristics [smooth, spiculated, or lobulated], and location]) to yield a posttest probability of a lung nodule being benign as previously defined in a separate prospective study (e-Fig 1, e-Table 2).[11]

Analysis of Integrated Classifier Performance

The performance of the integrated classifier was assessed by using standard metrics, including AUC, sensitivity, specificity, and negative predictive value (NPV). These metrics were also used to compare the performance of the integrated classifier vs other lung nodule cancer risk stratification methods, including physician assessment (pCA), clinical prediction models (Mayo and Veterans Affairs [VA]), and PET. Because the intended use of the integrated classifier is to identify benign lung nodules and reduce the number of invasive procedures, the performance of the test, and its comparators, was assessed by using the NPV as the primary metric. The threshold at which maximum NPV was observed in the discovery study was validated in the current study.[11] Furthermore, the NPV at this threshold in the current study is used to report the performance of the test.

Nodules were defined as benign for any of the following: (1) definitive pathologic diagnosis; (2) radiographic resolution; or (3) no evidence of growth according to CT scan over 1 year. Although the standard surveillance is 2 years of radiologic stability on chest CT imaging, for this analysis, 1 year of CT imaging stability was chosen based on a study observing lung nodules/masses by CT scan demonstrating no growth in nodules stable at 1 year on subsequent 2-year follow-up.[12] A malignant diagnosis was based on histopathologic findings. Those patients who were treated empirically with radiation and/or chemotherapy for a suspected malignancy without a confirmatory diagnosis were excluded from the analysis.

Assessment of Potential Impact of the Integrated Classifier

In the intended use cohort, those nodules with a "likely benign" integrated classifier result were stratified according to final diagnosis. Percentages of benign nodules correctly classified by using the integrated classifier and malignant nodules incorrectly classified by using the integrated classifier were calculated. Invasive procedure utilization directly following initial nodule detection was tabulated and stratified according to final diagnosis to assess the number and percentages of invasive testing that could have been avoided had the integrated classifier test been available and used for nodule management.

Data Analysis

Statistical analyses were performed by using MATLAB, version 8.3.0.532 (MathWorks), and MedCalc, version 16.4 (MedCalc Software bvba). A χ^2 of ANOVA testing was used to compare groups, and a P value $\leq$ .05 was considered significant. For comparison of cancer risk predictors, the McNemar statistical test was used at a fixed sensitivity or specificity.[13]

Results

Of 685 patients enrolled prospectively in the PANOPTIC study from November 2012 to December 2015, a total of 293 were excluded for the reasons listed in Figure 1, yielding 392 eligible for analysis. The demographic characteristics for the 392 patients are presented in e-Table 1. This study of the integrated classifier's performance focused on the subgroup of 178 patients having a lung nodule with a pCA $\leq$ 50%. No differences were observed in demographic or nodule characteristics between included and excluded (n = 234) patients (e-Table 3). Table 1 provides baseline patient demographic characteristics and radiographic parameters for nodules having a pCA $\leq$ 50%, stratified according to final histologic diagnosis. The prevalence of malignancy was 16%, with the majority being adenocarcinoma histology (59%). Both groups had significant smoking histories. Malignant nodules were significantly larger (16.5 vs 13.5 mm; $P < .01$).

The demographic characteristics, outcomes, and procedure utilization across groups of physician-assessed risk of malignancy are presented in e-Table 4. The risk of cancer increases with nodule size ($P < .001$). The use of serial imaging decreases, while the use of PET, biopsy, and surgery all increase, as the risk of malignancy increases ($P < .001$).

For all 392 eligible subjects, clinicians assigned a PCA $\leq$ 50% to the majority with an eventual benign diagnosis, with only 6% of subjects assigned to the very lowest risk group with a pCA of 0% to 5% (Fig 2). The performance of clinician assessments (pCA) was compared with that of the VA and Mayo Clinic lung nodule risk models.[14,15] The AUC for physician assignment of benign nodules into low to moderate risk ($\leq$ 50%) or malignant nodules into moderate to high risk ($>$ 50%) groups was 0.85, which was greater than the AUCs of 0.75 and 0.78 observed for the VA and Mayo models, respectively.[5]

Further analyses focused on the pCA $\leq$ 50% subgroup (n = 178). The integrated classifier revealed a sensitivity of 97% (CI, 82-100), a specificity of 44% (CI, 36-52), and a posttest probability of 98% (CI, 92-100) in distinguishing benign from malignant nodules. The integrated classifier demonstrated an AUC of 0.76,

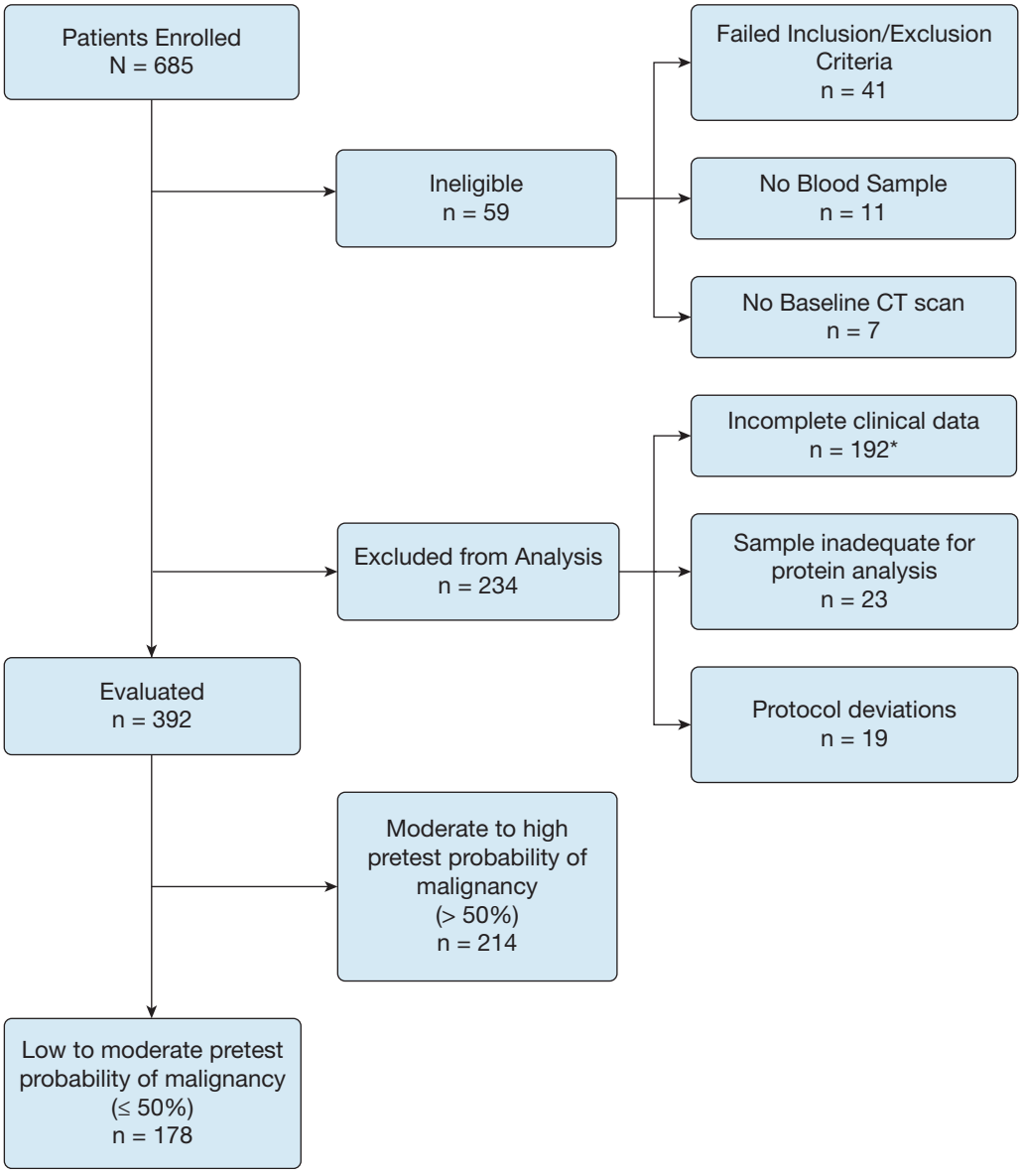

Figure 1 – *Eligibility of the Pulmonary Nodule Plasma Proteomic Classifier study patients for integrated classifier performance analysis in lung nodules according to the probability of malignancy.* *Incomplete clinical data are broken down as follows: n = 9, no pretest probability provided; n = 48, no follow-up procedure documented; n = 88, no 1-year follow-up CT scan; n = 39, no follow-up after PET scan; n = 5, time between interval scans did not reach 1 year; and n = 3, biopsy performed without documentation of results.

whereas the AUCs for physician pCA, the VA and Mayo models, and PET were 0.69, 0.60, 0.69, and 0.58, respectively (Fig 3). In e-Figure 2 and e-Table 5, the relative contribution of the component elements of the integrated classifier are presented. Using the McNemar test in which sensitivity was fixed at 90%, the integrated classifier performed significantly better than physician pCA, the VA and Mayo models, and PET ($P \leq .001$). Comparison to PET was restricted to the 75 subjects who had PET performed. Similar performance of the integrated classifier when stratified according to lung cancer subtype is shown in e-Table 6. The effects of the integrated classifier on adjusting pCA in the intended-use population stratified according to benign and malignant diagnoses are shown in e-Figure 3.

The potential impact of the integrated classifier was evaluated in two populations. For intended-use subjects, 66 of 178 were "likely benign" according to the integrated classifier. Of these, 65 had a benign nodule and 1 had a malignant lung nodule. Because there are 149 benign lung nodules and 29 malignant lung nodules in the study, 44% of benign lung nodules (65 of 149) were correctly labeled "likely benign," and 3% of malignant nodules

TABLE 1] Demographic Characteristics of the PANTOPIC Cohort With Lung Nodule pCA ≤ 50%

Characteristic	All Patients	Cancer	Benign	*P* Value
No. of patients	178	29	149	
Age, y	65.52 ± 1.55	66.05 ± 3.05	65.42 ± 1.76	.772
Sex				
Male	95 (53.37%)	12 (41.38%)	83 (55.70%)	.157
Female	83 (46.63%)	17 (58.62%)	66 (44.30%)	
Smoking history				
Status				.855
Never	42 (23.60%)	6 (20.69%)	36 (24.16%)	.725
Former	99 (55.62%)	16 (55.17%)	83 (55.70%)	.972
Current	37 (20.79%)	7 (24.14%)	30 (20.13%)	.665
Pack-year mean	43.56 ± 6.17	43.66 ± 11.73	43.54 ± 7.06	.989
Lung nodule				
Size	13.95 ± 0.76	16.48 ± 2.18	13.46 ± 0.78	.006
Nodule location				
Upper lobe		20 (68.97%)	70 (46.98%)	.128
Lower lobes		9 (31.03%)	79 (53.02%)	.123
Histology				
Benign nodule diagnosis				
Granuloma			9 (6.04%)	
Hamartoma			6 (4.03%)	
CT scan stable/resolution			116 (77.85%)	
Other			15 (10.07%)	
NA			3 (2.01%)	
Malignant nodule type				
Adenocarcinoma		17 (58.62%)		
Squamous cell		4 (13.79%)		
Large cell		0		
Mixed/nonspecified NSCLC		1 (3.45%)		
Small cell		2 (6.90%)		
Carcinoid		3 (10.34%)		
Other		2 (6.90%)		

Data are presented as mean ± SD unless otherwise indicated. NA = not available; NSCLC = non-small cell cancer lung cancer; PANOPTIC = Pulmonary Nodule Plasma Proteomic Classifier; pCA = probability of cancer.

(1 of 29) were incorrectly labeled "likely benign." In the 58 subjects who underwent an invasive procedure directly following the initial nodule detection, 35 nodules were benign and 23 were malignant. Of the 35 benign lung nodules, 14 (40%) were identified as "likely benign" by the integrated classifier, whereas of the 23 malignant lung nodules, 1 (4%) was incorrectly identified as "likely benign." The actual number of malignant nodules routed initially to surveillance was 13 of 29 (45%) in the intended-use subjects.

Table 2 details invasive diagnostic procedure utilization for nodules with pCA ≤ 50% stratified according to histologic diagnosis, as well as potential reduction of procedure had the integrated classifier been used in decision-making. Among surgeries performed as the sole diagnostic procedure, 7 of 20 (35%) were benign. Among biopsies performed as the only diagnostic procedure, 28 of 37 (76%) were benign. Fifty percent of patients who underwent both biopsy and surgical procedures received a benign diagnosis.

Discussion

This trial is the first clinical validation study evaluating the performance characteristics of an integrated blood

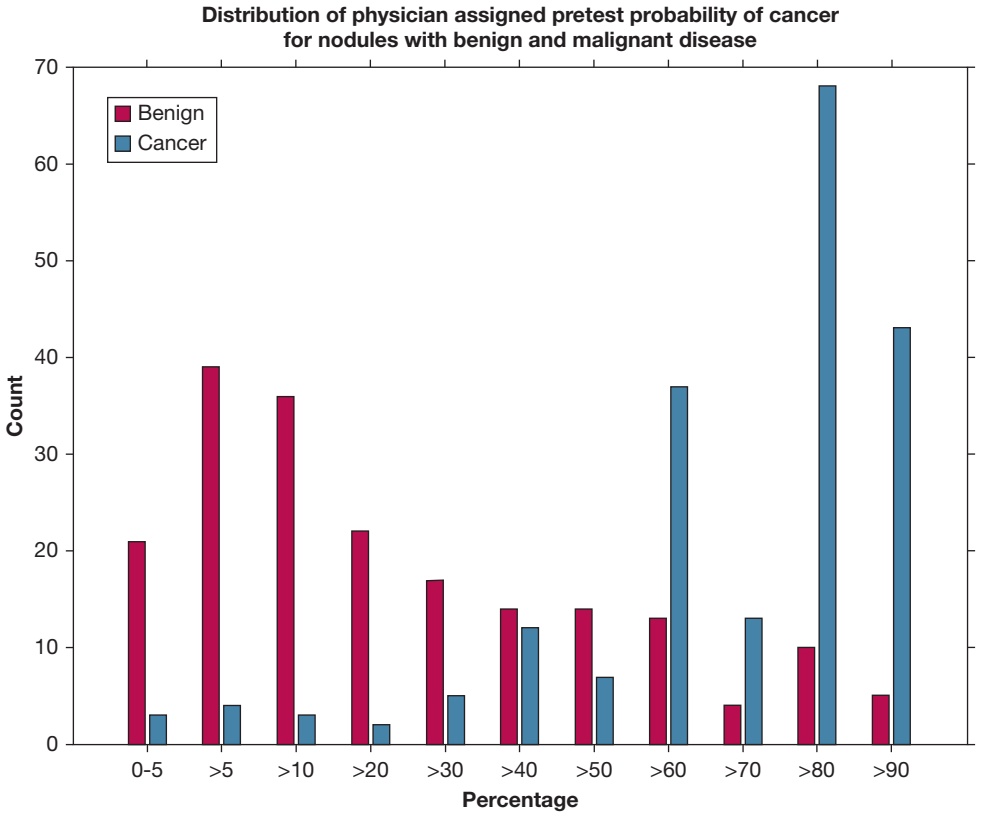

Figure 2 – *Distribution of physician-assigned pretest pCA for eligible Pulmonary Nodule Plasma Proteomic Classifier study patients (n = 392) by deciles. Shown are the physician-assigned pCA percentages for nodules with either a malignant (n = 197) or benign (n = 195) diagnosis. Note: the first two columns represent 5% pCA increments. pCA = probability of cancer.*

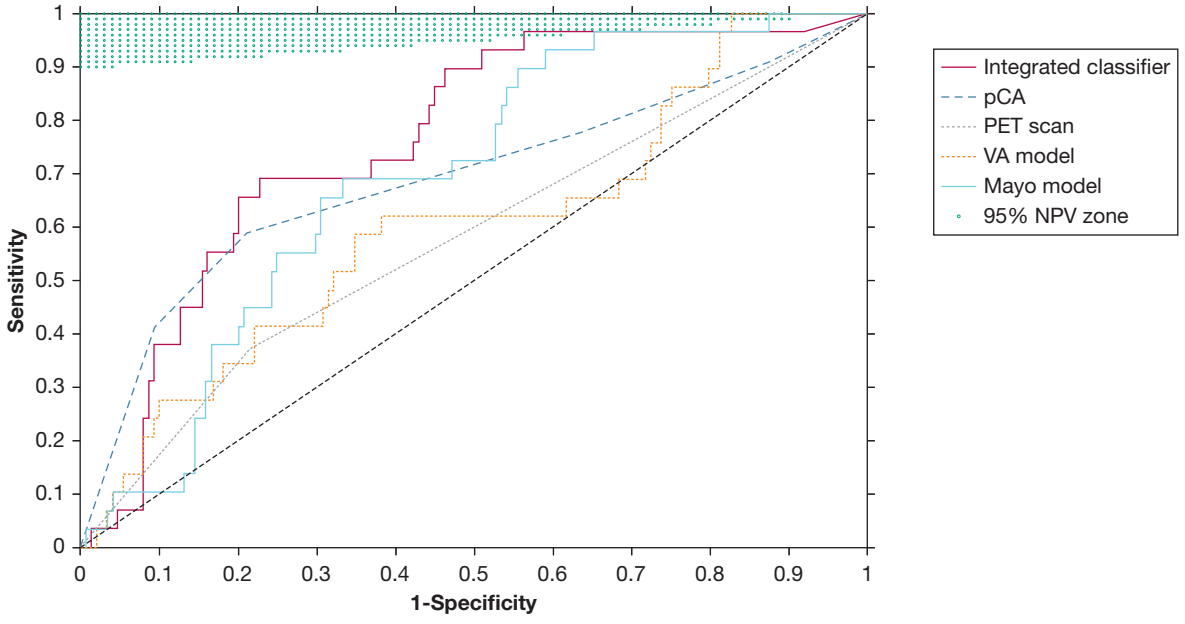

Figure 3 – *Comparison of the area under the receiver-operating characteristic curves of lung nodule malignancy risk assessment tools relative to the 95% NPV zone. Shown are the receiver-operating characteristic curves for subjects with lung nodules assigned a pCA ≤ 50% (n = 178) comparing the integrated classifier vs the physician-assigned pCA, PET, and the VA and Mayo cancer risk models. The shaded area indicates the ≥ 95% NPV diagnostic performance zone, which corresponds to the 5% cancer risk threshold specified in the CHEST guidelines for lung management. NPV = negative predictive value; VA = Veterans Affairs. See Figure 2 legend for expansion of other abbreviation.*

TABLE 2] Use of Invasive Diagnostic Procedures in the PANOPTIC Study Subgroup With Lung Nodules Having pCA ≤ 50%

Variable	No.	% of Category	No. of Procedures if Integrated Classifier Used	RRR	Misclassified Patient
All patients with invasive procedures	71				
Benign nodule patients	42	59	27	36%	1
Malignant nodule patients	29	41	28		
All patients with direct[a] invasive procedures	58				
Benign nodule patients	35	60	21	40%	1
Malignant nodule patients	23	40	22		
Patients with surgical procedures alone	20				
Benign nodule patient receiving surgery	7	35	5	29%	1
Malignant nodule patient receiving surgery	13	65	12		
Patients with biopsies alone	37				
Benign nodule patients receiving biopsy	28	76	16	43%	0
Malignant nodule patients receiving biopsy	9	24	9		
Patients with biopsies and surgical procedures	14				
Benign biopsy and surgery patients	7	50	6	14%	0
Malignant biopsy and surgery patients	7	50	7		

RRR = relative risk reduction. See Table 1 legend for expansion of other abbreviations.
[a]Direct denotes that the patient was not placed into CT surveillance but entered into a pathway that led directly to an invasive procedure.

proteomics classifier, comprising both proteins and clinical parameters, in distinguishing benign from malignant nodules. This large, multicenter observational study has several important findings. First, over the entire risk space, physicians perform very well; however, they are less accurate when the pCA is ≤ 50%. Second, when the probability of malignancy in a nodule is ≤ 50%, a "likely benign" integrated proteomic classifier result accurately identifies patients with benign nodules. This approach may enable physicians to re-categorize nodules from a strategy in which further testing is indicated to one in which CT surveillance is advisable. Thus, if incorporated into the current algorithm for managing nodules, the test may reduce patients' exposure to the morbidity and cost of avoidable invasive procedures. In addition, use of the test would lead to only a small percentage of patients with malignancy being misclassified as benign, potentially leading to an inappropriate surveillance strategy. These findings substantiate the need for a future clinical utility study to assess how clinical decision-making and use of invasive procedures are altered with knowledge of the results of this test in the right clinical setting.

When choosing a strategy for evaluating patients with lung nodules, clinicians should consider both the probability that the nodule is malignant and the advantages and disadvantages of management strategies. Serial surveillance has the advantage of being noninvasive and is recommend if the pCA is < 5%. However, in this study and another, < 10% of nodules ≥ 8 mm fell into this risk category.[4] In the latter study, all patients with a pCA < 5% were ultimately found to have benign disease. Although this approach avoids the potential harm of procedures, there is potential for interval growth of malignant nodules. The penalty for that growth is unknown and based on the biology of the tumor. One study of lung cancer experts found that there was little agreement in the penalty of missing growth between scans with estimates of stage shift from a lower stage to higher stage ranging from 1 to 50% with a resultant reduction in life expectancy.[16] On the opposite end of the spectrum, nodules with a pCA > 65% should be promptly resected in those healthy enough to tolerate surgery, providing both a diagnosis and treatment. The harms associated with this strategy include a morbidity of 5% and a mortality of 0.5%.[3]

Perhaps the most challenging group to manage are those with intermediate risk nodules (pCA 5%-65%), in which a substantial proportion of patients (45%) in this study are grouped and where 81% of the nodules were benign. Guidelines suggest further evaluation with PET scan, TTNA, or bronchoscopy.[3] Integrated PET/CT imaging has good sensitivity (86%- 91%) and specificity (71%- 81).[17] However, false-positive findings

(eg, granulomatous disease) and false-negative findings (eg, carcinoid) can mislead the clinician in either direction. TTNA has a yield between 70% (nodules < 15 mm) and 90% (nodules > 15 mm) but a 1% risk of hemorrhage and a pneumothorax rate of 16%, in which 6% require a chest tube.[3,18] The side-effect profile for bronchoscopy is acceptably low, but the yield (51%-70%) for pulmonary nodules even when using navigation is lower than TTNA.[19,20] In addition, a negative score from a bronchial-airway gene expression classifier of epithelial cells brushed from normal mucosa at the time of bronchoscopy improves the NPV in patients with an intermediate pretest probability for cancer and nondiagnostic bronchoscopy; however, not all patients with nodules routinely undergo bronchoscopy.[12] Clinicians can be left uncertain in management decisions, leading to potentially avoidable testing in patients with benign disease. One analysis found that of the total diagnostic costs of evaluating nodules, 43.1% was due to biopsy of patients without lung cancer.[21]

This intermediate-risk group may be those individuals in whom an integrated proteomics classifier would be useful. In patients with a pCA ≤ 50%, a likely benign test result could reduce the number of invasive procedures in those with benign nodules by 36% (an absolute risk reduction of 10.1% for all patients with benign nodules). This scenario assumes that all patients with likely benign results would be shifted to a surveillance strategy. It is important to consider that 3% of the patients with malignant nodules would also be placed in surveillance. Thus, it is critical for clinicians to continue to follow up until resolution or an acceptable period of CT radiographic stability has occurred.

What is unclear from these data is the extent to which clinicians would change their management based on the results of this test, although a survey of experienced pulmonologists found that a hypothetical blood test resulted in significant alterations in a decision to pursue invasive testing.[22] What is clear is that multiple studies reveal significant deviation from guideline-based care.[4,5,23] This trial is a clinical validation study that will need to undergo the clinical utility phase of biomarker development as outlined in the American Thoracic Society policy statement.[24]

The clinical validation phase of biomarker development establishes the accuracy of the biomarker in the intended-use population. Biomarker developers can use this accuracy, as well as an understanding of the potential benefit of a true result and the harm from a false result, to estimate whether the biomarker performs well enough to justify further evaluation in a clinical utility study. More than one method is available to help with this estimate.[24] One formula that has been proposed for a rule-out test states the specificity/(1 – sensitivity) ≥ (prevalence/1 – prevalence) × harm/benefit, where harm/benefit is the number of true negative results required to justify one false-negative result. If we use the prevalence of malignancy in the intended-use population in this study (16.3%) and the specificity (44%) and sensitivity (97%) reported here, the formula would produce a harm/benefit of 75.2 (ie, 75 true-negative results for each false-negative result). If the prevalence in the intended-use population was 40%, this number would be 22. These results strongly suggest that assessing the biomarker in a clinical utility study would be worthwhile.

Another important aspect of determining if the results of the clinical validation study warrant pursuit of a clinical utility study is understanding what percentage of tests ordered will provide a meaningful result. For a rule-out test, where the negative test result may change management but the positive test result will not, a test with the characteristics presented here could lead to management change in a population with a prevalence of 16.3% approximately 37% of the time, and in a population with a prevalence of 40%, approximately 28% of the time. This method provides less convincing evidence of the potential for clinical utility, but the implications for those with a negative test result are still believed to be large enough to support clinical utility testing. It is only through the results of a well-designed clinical utility study that we can determine the true potential value of the test.

The present study has limitations. First, the effect the plasma protein test might have had on test ordering was retrospectively analyzed; thus, a prospective study to assess changes in practice is warranted. Second, community practices are underrepresented in this trial; however, based on a previous study,[4] the likely benefit in terms of reduction of invasive testing would be significantly greater than that reported here. Third, although there is precedent for reporting 1-year outcomes for stable nodules, traditionally, 2 years of nodule stability is what has been required to determine a nodule is benign. Two-year follow-up data will be reported in the future once finalized. Fourth, there were 88 patients without follow-up CT scan data at 1 year. This degree of missing data may not be random, as patients with a low risk of lung cancer may have been

less likely to be adherent with follow-up recommendations.

This study has several strengths. First, to the best of our knowledge, this study is the largest prospective, multicenter, geographically diverse observational clinical study to validate an integrated blood biomarker for the evaluation of pulmonary nodules. Second, test development was rigorous: the integrated classifier was developed on an independent set of prospectively collected samples; the integrated classifier, laboratory standard operating procedures, and quality control subjects were all predefined and documented; and physicians, patients, and laboratory and statistical personnel were compliant with a blinding protocol. Finally, by having physicians estimate the probability of cancer in a pulmonary guideline, concordance with management decisions could be assessed.

Conclusions

This study is the first we are aware of to assess the accuracy of an integrated plasma proteomics classifier in patients with pulmonary nodules in a geographically diverse population with varying risk of cancer. In those with low to moderate risk nodules (pCA $\leq$ 50%), a "likely benign" test result could safely allow patients to be followed up by using serial imaging. Further research is needed to assess the effect of incorporating this test into the diagnostic algorithm for nodule management in the hope of reducing unnecessary procedures in patients without cancer.

Acknowledgments

Author contributions: P. K. had full access to all of the data in the study and takes responsibility for the integrity of and accuracy of the data analysis; K. C. F. and P. K. contributed substantially to the study design, data analysis and interpretation, and the writing of the manuscript; and G. A. S., N. T. T., A. V., P. P. M., A. P., S. C. S., D. M., and P. J. M. contributed substantially to data analysis and interpretation, and the writing of the manuscript.

Financial/nonfinancial disclosures: The authors have reported to *CHEST* the following: G. A. S. received research grant funding from Integrated Diagnostics, Exact Sciences, and Olympus. N. T. T. has received other grant funding from the American Cancer Society and the CHEST Foundation; industry grant funding from Exact Sciences, Veracyte, Integrated Diagnostics, Oncimmune, Olympus, and Cook Medical; consulting monies from Integrated Diagnostics, Olympus, Cook Medical, Veran, and Oncocyte; and has participated in advisory board meetings for Veracyte. P. K. is an employee and stockholder of Integrated Diagnostics. A. V. has received research funding from Integrated Diagnostics. A. P. is the Vice President of Clinical Operations and Medical Affairs for Integrated Diagnostics. S. C. S. is an employee of Integrated Diagnostics. K. C. F. was the Chief Medical Officer and stockholder at Integrated Diagnostics and authored on issued patents. D. M. received royalties from UpToDate for chapter authorship and editorship. P. J. M. has served on clinical advisory boards for Integrated Diagnostics, Oncimmune, Exact Sciences, Grail, and Nucleix; and has participated in speaking activities for Oncocyte.

*Writing Committee Members for the PANOPTIC Trial Study Team consisted of the following investigators and coordinators (the city/state or province in which their institution is located and the number of patients enrolled are shown in parentheses): G.A. Silvestri, MD, L. Leake (Charleston, SC; 78), P. Mazzone, MD, M. Beukemann (Cleveland, OH; 63), D. Midthun, MD, P. McCarthy (Rochester, MN; 55), B. Sigal, MD, T. Deluca (Winston-Salem, NC; 51), F. Laberge, MD, B. Fortin (Quebec City, QC, Canada; 47), M. Balaan, MD, B. Dimitt (Pittsburgh, PA; 46), A. Pierre, MD, F. Allison (Toronto, ON, Canada; 41), L. Yarmus, MD, K. Oakjones-Burgess (Baltimore, MD; 27), N. Tanner, MD, L. Leake (VA Medical Center, Charleston, SC; 23), N. Ettinger, MD, T. Setchfield (Chesterfield, MO; 23), D. Madtes, MD, J. Hubbard (Seattle, WA; 20), W. McConnell, MD, K. Robinson (Louisville, KY; 20), A. Lackey, MD, L. Jacques (Royal Oak, MI; 16), E. Kuo, MD, V. Markland-Gentles (Phoenix, AZ; 16), P. Massion, MD, A. Muterspaug, (Nashville, TN; 14), J. Leach, MD, K. Rothe (Minneapolis, MN; 14), W. Rom, MD, H. Pass, MD, A. Sorenson (New York, NY; 13), A. Chesnutt, MD, A. Georgeson (Portland, OR; 13), A. Balekian, MD, J. Fisher (Los Angeles, CA; 13), R. Murali, MD, A. Overton (Greensboro, NC; 12), N. Desai, MD, A. Levesque (Elk Grove, IL; 11), W. Krimsky, MD, S. King (Baltimore, MD; 9), A. Vachani, MD, K. Maletteri (Philadelphia, PA; 9), K. Mileham, MD, L. Carter (Charlotte, NC; 8), G. Hong, MD, J. Ma (San Francisco, CA; 8), K. Voelker, MD, H. Barrentine (Sarasota, FL; 7), R. Aronson, MD, M. Henderson (Tucson, AZ; 7), J. Lamberti, MD, C. Krawiecki (Falls Church, VA; 7), A. Case, MD, L. Wilkins (Austell, GA; 4), J. M. Ayers, MD, K. Fangmann (Danville, PA; 4), J. Landis, MD, L. DeSouza (Springfield, MA; 3), Z. Hammoud, MD, D. Kah (Detroit, MI; 3), and J. Sanchez, MD, L. Murdoch, (Temple, TX; 0).

Role of sponsors: Integrated Diagnostics designed the trial, collected the data, and assisted in interpretation of the results.

Additional information: The e-Figures and e-Tables can be found in the Supplemental Materials section of the online article.

References

1. Gould MK, Tang T, Liu IL, et al. Recent trends in the identification of incidental pulmonary nodules. *Am J Respir Crit Care Med*. 2015;192(10):1208-1214.
2. National Lung Screening Trial Research Trial, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011;365(5):395-409.
3. Gould MK, Donington J, Lynch WR, 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(suppl 5):e93S-120S.
4. Tanner NT, Aggarwal J, Gould MK, et al. Management of pulmonary nodules by community pulmonologists: a multicenter observational study. *Chest*. 2015;148(6):1405-1414.
5. Tanner NT, Porter A, Gould MK, Li XJ, Vachani A, Silvestri GA. Physician assessment of pretest probability of malignancy and adherence with guidelines for pulmonary nodule evaluation. *Chest*. 2017;152(2):263-270.
6. Li XJ, Hayward C, Fong PY, et al. A blood-based proteomic classifier for the molecular characterization of pulmonary nodules. *Sci Translation Med*. 2013;5(207):207ra142.
7. Vachani A, Hammoud Z, Springmeyer S, et al. Clinical utility of a plasma protein classifier for indeterminate lung nodules. *Lung*. 2015;193(6):1023-1027.
8. Grassadonia A, Tinari N, Iurisci I, et al. 90K (Mac-2 BP) and galectins in tumor progression and metastasis. *Glycoconj J*. 2002;19(7-9):551-556.

9. Yang L, Wang F, Wang L, et al. CD163+ tumor-associated macrophage is a prognostic biomarker and is associated with therapeutic effect on malignant pleural effusion of lung cancer patients. *Oncotarget*. 2015;6(12):10592-10603.

10. Moestrup SK, Moller HJ. CD163: a regulated hemoglobin scavenger receptor with a role in the anti-inflammatory response. *Ann Med*. 2004;36(5):347-354.

11. Kearney P, Hunsucker SW, Li XJ, Porter A, Springmeyer S, Mazzone P. An integrated risk predictor for pulmonary nodules. *PLoS One*. 2017;12(5):e0177635.

12. Silvestri GA, Vachani A, Whitney D, et al. A bronchial genomic classifier for the diagnostic evaluation of lung cancer. *N Engl J Med*. 2015;373(3):243-251.

13. Fagerland MW, Lydersen S, Laake P. The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Med Res Methodol*. 2013;13:91.

14. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. *Arch Intern Med*. 1997;157(8):849-855.

15. Gould MK, Ananth L, Barnett PG; Veterans Affairs SNAP Cooperative Study Group. A clinical model to estimate the pretest probability of lung cancer in patients with solitary pulmonary nodules. *Chest*. 2007;131(2):383-388.

16. Schultz EM, Silvestri GA, Gould MK. Variation in experts' beliefs about lung cancer growth, progression, and prognosis. *J Thorac Oncol*. 2008;3(4):422-426.

17. Deppen SA, Blume JD, Kensinger CD, et al. Accuracy of FDG-PET to diagnose lung cancer in areas with infectious lung disease: a meta-analysis. *JAMA*. 2014;312(12):1227-1236.

18. Wiener RS, Wiener DC, Gould MK. Risks of transthoracic needle biopsy: how high? *Clin Pulm Med*. 2013;20(1):29-35.

19. 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 Respir Critic Care Med*. 2016;193(1):68-77.

20. Wang Memoli JS, Nietert PJ, Silvestri GA. Meta-analysis of guided bronchoscopy for the evaluation of the pulmonary nodule. *Chest*. 2012;142(2):385-393.

21. Lokhandwala T, Bittoni MA, Dann RA, et al. Costs of diagnostic assessment for lung cancer: a Medicare claims analysis. *Clin Lung Cancer*. 2017;18(1):e27-e34.

22. Vachani A, Tanner NT, Aggarwal J, et al. Factors that influence physician decision making for indeterminate pulmonary nodules. *Ann Am Thorac Soc*. 2014;11(10):1586-1591.

23. Wiener RS, Gould MK, Slatore CG, Fincke BG, Schwartz LM, Woloshin S. Resource use and guideline concordance in evaluation of pulmonary nodules for cancer: too much and too little care. *JAMA Intern Med*. 2014;174(6):871-880.

24. Mazzone PJ, Sears CR, Arenberg DA, et al. Evaluating molecular biomarkers for the early detection of lung cancer: when is a biomarker ready for clinical use? An official American Thoracic Society policy statement. *Am J Respir Crit Care Med*. 2017;196(7):e15-e29.