ORIGINAL ARTICLE

A Bronchial Genomic Classifier for the Diagnostic Evaluation of Lung Cancer

Gerard A. Silvestri, M.D., Anil Vachani, M.D., Duncan Whitney, Ph.D., Michael Elashoff, Ph.D., Kate Porta Smith, M.P.H., J. Scott Ferguson, M.D., Ed Parsons, Ph.D., Nandita Mitra, Ph.D., Jerome Brody, M.D., Marc E. Lenburg, Ph.D., and Avrum Spira, M.D., for the AEGIS Study Team*

ABSTRACT

BACKGROUND

Bronchoscopy is frequently nondiagnostic in patients with pulmonary lesions suspected to be lung cancer. This often results in additional invasive testing, although many lesions are benign. We sought to validate a bronchial-airway gene-expression classifier that could improve the diagnostic performance of bronchoscopy.

METHODS

Current or former smokers undergoing bronchoscopy for suspected lung cancer were enrolled at 28 centers in two multicenter prospective studies (AEGIS-1 and AEGIS-2). A gene-expression classifier was measured in epithelial cells collected from the normal-appearing mainstem bronchus to assess the probability of lung cancer.

RESULTS

A total of 639 patients in AEGIS-1 (298 patients) and AEGIS-2 (341 patients) met the criteria for inclusion. A total of 43% of bronchoscopic examinations were non-diagnostic for lung cancer, and invasive procedures were performed after bronchoscopy in 35% of patients with benign lesions. In AEGIS-1, the classifier had an area under the receiver-operating-characteristic curve (AUC) of 0.78 (95% confidence interval [CI], 0.73 to 0.83), a sensitivity of 88% (95% CI, 83 to 92), and a specificity of 47% (95% CI, 37 to 58). In AEGIS-2, the classifier had an AUC of 0.74 (95% CI, 0.68 to 0.80), a sensitivity of 89% (95% CI, 84 to 92), and a specificity of 47% (95% CI, 36 to 59). The combination of the classifier plus bronchoscopy had a sensitivity of 96% (95% CI, 93 to 98) in AEGIS-1 and 98% (95% CI, 96 to 99) in AEGIS-2, independent of lesion size and location. In 101 patients with an intermediate pretest probability of cancer, the negative predictive value of the classifier was 91% (95% CI, 75 to 98) among patients with a nondiagnostic bronchoscopic examination.

CONCLUSIONS

The gene-expression classifier improved the diagnostic performance of bronchoscopy for the detection of lung cancer. In intermediate-risk patients with a nondiagnostic bronchoscopic examination, a negative classifier score provides support for a more conservative diagnostic approach. (Funded by Allegro Diagnostics and others; AEGIS-1 and AEGIS-2 ClinicalTrials.gov numbers, NCT01309087 and NCT00746759.)

From the Medical University of South Carolina, Charleston (G.A.S.); the University of Pennsylvania School of Medicine, Philadelphia (A.V., N.M.); Allegro Diagnostics (D.W., K.P.S., E.P.) and the Boston University School of Medicine (J.B., M.E.L., A.S.) — both in Boston; Veracyte, San Francisco (D.W., K.P.S.) and Elashoff Consulting, Redwood City (M.E.) — both in California; and the University of Wisconsin School of Medicine and Public Health, Madison (J.S.F.). Address reprint requests to Dr. Spira at the Division of Computational Biomedicine, Department of Medicine, Boston University Medical Center, 715 Albany St., Boston, MA 02118, or at aspira@bu.edu; or to Dr. Silvestri at the Division of Pulmonary and Critical Care Medicine, Medical University of South Carolina, 96 Jonathan Lucas St., Charleston, SC 29425, or at silvestr@musc.edu.

LESIONS THAT ARE SUSPICIOUS FOR LUNG cancer are frequently identified on chest imaging. The decision to pursue surveillance imaging or an invasive evaluation requires an assessment of the likelihood of cancer, the ability to biopsy, the surgical risk, and the patient's preferences.[1] When biopsy is required, the approach can include bronchoscopy, transthoracic needle biopsy, or surgical lung biopsy. The choice among these procedures is determined on the basis of considerations such as lesion size and location, the presence of adenopathy, the risk associated with the procedure, and local expertise. Bronchoscopy is relatively safe, with less than 1% of procedures complicated by pneumothorax.[2] Approximately 500,000 bronchoscopic examinations are performed each year in the United States.[3] Of these procedures, approximately half are for the diagnostic evaluation of suspected lung cancer. However, bronchoscopy is limited by its sensitivity, which ranges from 34 to 88%, depending on the location and size of the lesion.[4] Even with newer bronchoscopic guidance techniques, the sensitivity for the detection of lung cancer is only approximately 70% for peripheral lesions.[5]

Patients with a nondiagnostic bronchoscopic examination often undergo further invasive testing. Surgical lung biopsy is one approach, but it has a complication rate of approximately of 5% and a 30-day mortality of approximately 1%.[6] Furthermore, 20 to 25% of surgical biopsies are performed in patients who are ultimately found to have benign lesions.[7,8] Transthoracic needle biopsy is also associated with substantial morbidity, including a 15% rate of pneumothorax[9] and a 6% rate of pneumothorax necessitating chest tube drainage.[10] Given the pitfalls of invasive procedures, alternative approaches are needed to identify patients with a reduced likelihood of cancer who are appropriate candidates for imaging surveillance.

The use of gene expression in the classification of biologic disease states in clinical specimens is well established.[11] Cancer-associated gene-expression patterns are found in cytologically normal epithelium collected from the proximal airways of current and former smokers with lung cancer.[12] Recently, we developed a gene-expression classifier in bronchial epithelial cells collected from the mainstem bronchus by means of bronchoscopy that distinguishes patients with lung cancer from those without lung cancer among current and former smokers.[13] We undertook the present studies to prospectively validate this classifier in patients undergoing bronchoscopy for suspected lung cancer and to assess how this classifier alters the diagnostic performance of bronchoscopy.

METHODS

STUDY DESIGN, POPULATION, AND PROTOCOL

Current and former smokers who were undergoing bronchoscopy for suspected lung cancer at 28 sites in the United States, Canada, and Ireland (Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org) were enrolled in the Airway Epithelial Gene Expression in the Diagnosis of Lung Cancer (AEGIS) trials (AEGIS-1 and AEGIS-2), two independent, prospective, multicenter, observational studies. Cytology brushes were used to collect epithelial cells from the normal-appearing mainstem bronchus during bronchoscopy. Results of the classifier analysis were not reported to physicians or patients. Exclusion criteria included an age less than 21 years, no history of smoking (defined as having ever smoked <100 cigarettes), and a concurrent cancer or history of lung cancer. Patients were followed until a diagnosis was established or until 12 months after bronchoscopy. A diagnosis of lung cancer was established at the time of bronchoscopy or subsequently by means of biopsy with the use of a transthoracic needle, a surgical biopsy, a second bronchoscopic examination, or another invasive procedure. The specific bronchoscopic method used (which is detailed in Table S2 in the Supplementary Appendix) and any subsequent testing were at the discretion of the treating physician.

Patients who were defined as cancer-free had a specific diagnosis of a benign condition or radiographic stability or resolution at 12 months. Patients without a definitive diagnosis of cancer, a specific diagnosis of a benign condition, or stability or resolution at the 12-month follow-up were not included in further analyses. The treating physician assessed each patient's pretest probability of having cancer before bronchoscopy with the use of a five-level scale (probabilities of <10%, 10 to 39%, 40 to 60%, 61 to 85%, and

>85%). The study protocols (available at NEJM.org) were approved by the institutional review board at each center, and all patients provided written informed consent before enrollment. The study population is described in more detail in the Supplementary Appendix.

LABORATORY METHODS

All bronchial epithelial-cell specimens were processed to isolate and analyze RNA for quality and yield before gene-expression analysis; only specimens with an RNA yield of at least 1 μg and sufficient RNA integrity were run on Gene 1.0 ST microarrays. Further details are provided in the Supplementary Appendix. All microarray data have been deposited in the Gene Expression Omnibus (GEO) database under accession number GSE66499.

GENE-EXPRESSION ANALYSIS

The normalization and preprocessing of microarray data are described in the Supplementary Appendix. Patients enrolled in AEGIS-1 were randomly assigned to independent training and validation sets (Fig. S1 in the Supplementary Appendix). The classifier algorithm was derived only within the AEGIS-1 training set and locked, as described previously.[13] Scores for each sample in the AEGIS-1 validation set and for all AEGIS-2 samples were generated with the use of this prespecified classifier, which was based on the expression of 23 genes and patient age. These scores were dichotomized as test-positive and test-negative with the use of a prespecified threshold value.

STATISTICAL ANALYSIS

The performance of the classifier was evaluated with the use of receiver-operating-characteristic curves, calculation of area under the curve (AUC),[14] and estimates of sensitivity, specificity, negative predictive value, positive predictive value, and the negative likelihood ratio, defined as $(1-\text{sensitivity})\div\text{specificity}$. A Mann–Whitney nonparametric test was used for the analysis of continuous variables, and Fisher's exact test was used for categorical variables. All confidence intervals are reported as two-sided binomial 95% confidence intervals. Statistical analysis was performed with R software, version 3.01 (R Project for Statistical Computing).

RESULTS

CHARACTERISTICS OF THE STUDY PARTICIPANTS

A total of 298 patients from AEGIS-1 were included in the first validation set, and 341 patients from AEGIS-2 were included in a second validation set (Table 1, and Fig. S1 in the Supplementary Appendix). The prevalence of lung cancer was 74% and 78% in the AEGIS-1 and AEGIS-2 cohorts, respectively. Patients with lung cancer were older and had higher cumulative tobacco exposure than patients without cancer (P<0.001 for both comparisons) (Table S3 in the Supplementary Appendix). A summary of cancer stages and diagnosed benign conditions are shown in Tables S4 and S5, respectively, in the Supplementary Appendix. Comparisons of the clinical characteristics of the 639 patients who were included in the combined studies with the characteristics of those not included (either because of poor RNA quality, loss to follow-up, or absence of a final diagnosis at 12 months) are shown in Tables S6, S7, and S8 in the Supplementary Appendix.

PERFORMANCE OF BRONCHOSCOPY

A total of 639 patients who were in the combined study cohort underwent bronchoscopy for suspected lung cancer. Of those bronchoscopic examinations, 272 (43%; 95% confidence interval [CI], 39 to 46) were nondiagnostic, including in 120 of 487 patients (25%; 95% CI, 21 to 29) in whom lung cancer was ultimately diagnosed. The sensitivity of bronchoscopy for the detection of lung cancer was 74% (95% CI, 68 to 79) in AEGIS-1 and 76% (95% CI, 71 to 81) in AEGIS-2. Data on follow-up procedures were available for 267 of 272 of the patients with a nondiagnostic bronchoscopic examination (98%); 170 of these patients (64%; 95% CI, 58 to 69) underwent an invasive procedure, including 52 of 147 patients (35%; 95% CI, 28 to 44) with benign lesions and 118 of 120 patients (98%; 95% CI, 94 to 99) with cancer. Surgical lung biopsy was performed in 76 patients, 27 of whom had benign lesions (36%; 95% CI, 25 to 47).

PERFORMANCE OF GENE-EXPRESSION CLASSIFIER

In AEGIS-1, the classifier had an AUC of 0.78 (95% CI, 0.73 to 0.83) and accurately identified 194 of 220 patients with cancer (sensitivity, 88%; 95% CI, 83 to 92) and 37 of 78 patients without

Table 1. Demographic and Clinical Characteristics of the Study Participants.*

Characteristic	AEGIS-1 (N=298)	AEGIS-2 (N=341)
Sex — no.†		
Female	125	106
Male	173	235
Median age (IQR) — yr†	62 (55–70)	64 (57–71)
Race — no.‡		
White	226	267
Black	55	66
Other	15	4
Unknown	2	4
Smoking status — no.		
Current	146	169
Former	152	172
Median cumulative tobacco use (IQR) — pack-yr	40 (24–60)	45 (25–63)
Lesion size — no.		
<2 cm	48	83
2 to 3 cm	41	39
>3 cm	155	188
Infiltrate§	32	28
Unknown	22	3
Lesion location — no.		
Central	98	127
Peripheral	86	108
Central and peripheral	90	102
Unknown	24	4
Lung-cancer histologic type — no.¶	220	267
Small-cell	38	42
Non–small-cell	175	222
Adenocarcinoma	69	100
Squamous	72	81
Large-cell	8	8
Non–small-cell not otherwise specified	26	33
Unknown	7	3
Diagnosis of a benign condition — no.	78	74
Infection	18	14
Sarcoidosis	16	15
Resolution or stability	27	24
Other	17	21

* IQR denotes interquartile range.
† The difference between the groups was significant (P<0.05).
‡ Race was self-reported. The P value for race was calculated for white versus nonwhite.
§ Infiltrates are pulmonary lesions with ill-defined margins and a diameter that cannot be accurately quantified.
¶ The P value for lung-cancer histologic type was calculated for non–small-cell lung cancer versus small-cell lung cancer.

cancer (specificity, 47%; 95% CI, 37 to 58) (Fig. 1). In AEGIS-2, the classifier had an AUC of 0.74 (95% CI, 0.68 to 0.80) and correctly identified 237 of 267 patients with cancer (sensitivity, 89%; 95% CI, 84 to 92) and 35 of 74 patients without cancer (specificity, 47%; 95%, 36 to 59) (Fig. 1). The combination of the classifier plus bronchoscopy increased the sensitivity to 96% (95% CI, 93 to 98) and 98% (95% CI, 96 to 99) in AEGIS-1 and AEGIS-2, respectively, as compared with 74% and 76% for bronchoscopy alone (P<0.001 for both comparisons).

In patients with a nondiagnostic bronchoscopic examination, the classifier accurately identified cancer in 49 of 57 patients in AEGIS-1 (sensitivity, 86%; 95% CI, 74 to 94) and in 58 of 63 patients in AEGIS-2 (sensitivity, 92%; 95% CI, 82 to 97). Because there was no significant difference between the two cohorts with regard to the classifier AUC either among all patients (P=0.32) or among those with a nondiagnostic bronchoscopic examination (P=0.61) (Fig. 1), we combined the two cohorts for subsequent analyses of subgroups. The sensitivity of bronchoscopy was lower for lesions that were smaller than 3 cm in diameter (P<0.001) or peripherally located (P<0.001) (Table 2), as well as in patients without hilar or mediastinal adenopathy (P<0.001). In contrast, the sensitivity of the classifier and of the classifier combined with bronchoscopy were consistently high and not significantly associated with the size or location of the lesion (Table 2), cancer stage (Table S9 in the Supplementary Appendix), histologic type of the cancer (Table S10 in the Supplementary Appendix), or presence of adenopathy (Table S11 in the Supplementary Appendix). The combination of the classifier plus bronchoscopy had a sensitivity of 96% among patients without hilar or mediastinal adenopathy (285 patients).

ACCURACY OF THE CLASSIFIER IN PATIENTS WITH AN INTERMEDIATE PROBABILITY OF CANCER

We binned the physician-assessed probability of cancer into categories of low (<10%), intermediate (10 to 60%), and high (>60%) probability (Table 3, and Table S12 in the Supplementary Appendix), to align with guideline recommendations for assessing lung-cancer risk.[1] Bronchoscopy was nondiagnostic for cancer in 83% of patients with an intermediate pretest probability (101 patients), despite a cancer prevalence rate of
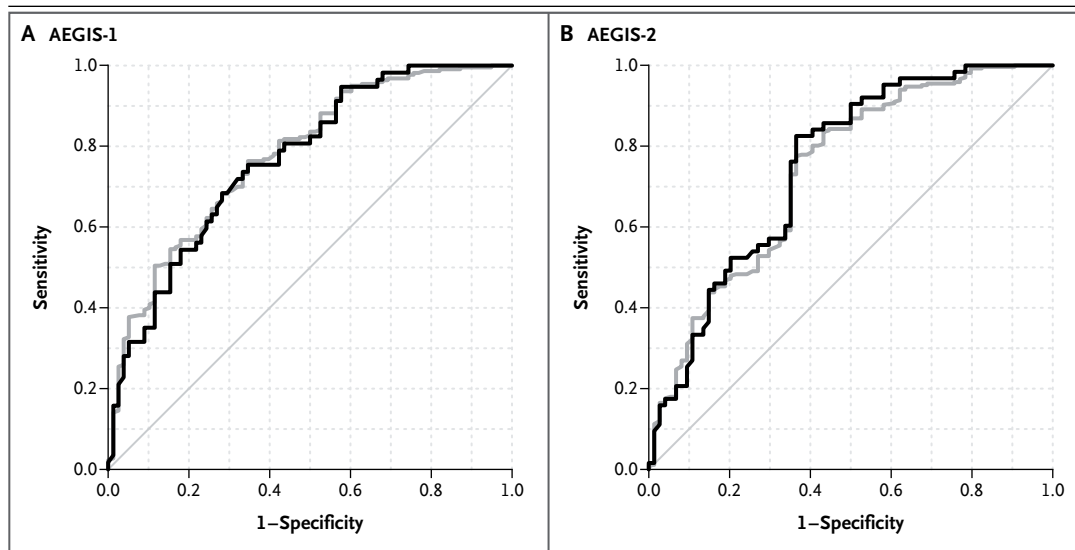
Figure 1. Classifier Performance in the AEGIS-1 and AEGIS-2 Studies.

Shown are receiver-operating-characteristic curves for all patients (gray) and the subset of patients with a nondiagnostic bronchoscopic examination (black) in the AEGIS-1 and AEGIS-2 cohorts. In AEGIS-1, the area under the curve (AUC) was 0.78 (95% CI, 0.73 to 0.83) for all patients and 0.76 (95% CI, 0.68 to 0.83) for patients with a nondiagnostic examination (P=0.31). In AEGIS-2, the AUC was 0.74 (95% CI, 0.68 to 0.80) and 0.75 (95% CI, 0.68 to 0.82), respectively (P=0.85). The AUC was also not significantly different for patients with a nondiagnostic examination in the comparison between AEGIS-1 and AEGIS-2 (P=0.61).

41% (95% CI, 31 to 51). In this subgroup, the classifier achieved a negative predictive value of 91% (95% CI, 75 to 98) and a positive predictive value of 40% (95% CI, 27 to 55) (Table 3). In patients with indeterminate nodules measuring less than 3 cm and a low or intermediate probability of cancer (73 patients), the classifier had a sensitivity of 88% and a negative predictive value of 94% (Table S13 in the Supplementary Appendix). In the subset of patients who had a pulmonary nodule measuring less than 3 cm without adenopathy (36 patients), the sensitivity of the classifier was 92%, with a negative predictive value of 92%.

Although the classifier had a high negative predictive value in patients with a nondiagnostic bronchoscopic examination, 13 such patients had lung cancer and a negative classifier score (i.e., false negatives). The majority of these patients (10 of 13) had a high (>60%) probability of cancer, and only 3 patients in the group had an intermediate (10 to 60%) pretest probability of cancer. The classifier performance, stratified according to the number of patients in each pretest probability category, is shown in Table S14 in the Supplementary Appendix.

The negative likelihood ratio of the classifier in combination with bronchoscopy was calculated to determine the range of pretest probabilities of cancer in which the posttest probability would be less than 10%. When bronchoscopy was combined with the classifier, the negative likelihood ratio improved from 0.24 (95% CI, 0.21 to 0.29) to 0.06 (95% CI, 0.03 to 0.10). As a result, when both bronchoscopy and the classifier were negative, the posttest probability of cancer was reduced to less than 10% for patients with a pretest probability of up to 66% (Fig. 2).

DISCUSSION

In this study, we describe the evaluation of a bronchial genomic classifier for lung-cancer diagnosis among patients undergoing bronchoscopy in two independent prospective cohorts. We found that the gene-expression classifier had high sensitivity across different lesion sizes, locations, stages, and cell types of lung cancer. The combination of the classifier plus bronchoscopy had a sensitivity of 96% and 98% in the AEGIS-1 and AEGIS-2 validation cohorts, respectively.

Table 2. Sensitivity of the Classifier, Bronchoscopy, and the Combined Approaches, According to Radiologic Imaging Characteristics.

Group	All Patients	Patients with Cancer	Sensitivity*		
			Bronchoscopy	Classifier	Classifier plus Bronchoscopy
	no. of patients		*percent (95% confidence interval)*		
All patients	639	487	75 (71–79)	89 (82–94)	97 (95–99)
Lesion size					
<2 cm	131	73	55 (43–66)	91 (76–98)	96 (88–99)
2 to 3 cm	80	60	58 (45–71)	92 (74–99)	97 (88–100)
>3 cm	343	313	82 (78–86)	85 (74–93)	97 (95–99)
Infiltrate	60	25	84 (64–95)	100 (40–100)	100 (86–100)
Unknown	25	16	80 (54–96)	100 (29–100)	100 (79–100)
Lesion location					
Central	225	174	84 (78–89)	81 (62–94)	97 (93–99)
Peripheral	194	133	55 (46–63)	90 (79–96)	95 (90–98)
Central and peripheral	192	164	82 (75–87)	97 (82–100)	99 (96–100)
Unknown	28	16	81 (54–96)	67 (9–99)	94 (70–100)

* The sensitivity of bronchoscopy was determined for patients with lung cancer in each category. The sensitivity of the classifier was determined for the patients with lung cancer whose cancer was not diagnosed during bronchoscopy. The sensitivity of the classifier combined with bronchoscopy was calculated for all patients with lung cancer in each category.

We also report several additional findings that support the clinical value of this type of classifier. First, our studies confirm that nondiagnostic bronchoscopic examinations are common (particularly in patients with an intermediate pretest probability of cancer) and lead to further invasive testing (including surgical biopsy) in patients who are ultimately found to have benign lesions. Second, in contrast to the classifier, bronchoscopy had poor sensitivity for the detection of lung cancer in patients with small, peripheral, or early-stage lesions, as well as in those without lymphadenopathy. Third, the classifier had a high negative predictive value in patients with an intermediate probability of cancer and a nondiagnostic bronchoscopic examination. These findings suggest that this classifier has the potential to assist in clinical decision making for patients with an intermediate probability of cancer, in whom the prevalence of lung cancer is 41% but the sensitivity of bronchoscopy is only 41%.

Although the high negative predictive value of the classifier could help avoid unnecessary invasive procedures in patients with an intermediate probability of cancer who are classifier-negative, a small number of patients in this subgroup had lung cancer. A negative classifier score may delay further invasive testing in such patients. However, patients with a nondiagnostic bronchoscopic examination and a negative classifier score would probably undergo active surveillance with the use of imaging, which is the standard practice when an immediate invasive strategy is not used.[1,15] This would identify lesion growth and trigger additional invasive testing to establish a definitive diagnosis. In contrast to the high negative predictive value of the classifier that we observed for patients with an intermediate probability of cancer, the positive predictive value of 40% was modest for these patients. Thus, a positive classifier result does not warrant alteration of the decision between an invasive strategy and an imaging-surveillance strategy. Future registry studies will further determine the effect of the classifier on the evaluation of patients with suspected lung cancer after a nondiagnostic bronchoscopic examination.

This gene-expression classifier is measured in proximal bronchial epithelial cells and not in

Table 3. Performance of Bronchoscopy and the Classifier, Stratified According to the Pretest Probability of Cancer.*

Variable	Low Pretest Probability of Cancer (N = 62)	Intermediate Pretest Probability of Cancer (N = 101)	High Pretest Probability of Cancer (N = 426)	Unknown Pretest Probability of Cancer (N = 50)
Patients with lung cancer — no. (%)	3 (5)	41 (41)	405 (95)	38 (76)
Patients with benign lesions — no. (%)	59 (95)	60 (59)	21 (5)	12 (24)
Bronchoscopy performance				
Sensitivity — % (95% CI)	33 (1–91)	41 (26–58)	79 (74–82)	82 (66–92)
Patients with nondiagnostic bronchoscopic examination — no. (%)†	61 (98)	84 (83)	108 (25)	19 (38)
Classifier performance				
Sensitivity — % (95% CI)‡	100 (16–100)	88 (68–97)	89 (80–94)	100 (59–100)
Specificity — % (95% CI)§	56 (42–69)	48 (35–62)	29 (11–52)	33 (10–65)
Negative predictive value — % (95% CI)¶	100 (89–100)	91 (75–98)	38 (15–65)	100 (40–100)
Positive predictive value — % (95% CI)¶	7 (1–24)	40 (27–55)	84 (75–91)	47 (21–73)
Combined classifier and bronchoscopy sensitivity — % (95% CI)	100 (29–100)	93 (80–98)	98 (96–99)	97 (91–100)

* Before the bronchoscopic examination, the treating physician assessed each patient's pretest probability of having cancer; the results were divided into categories of low (<10%), intermediate (10 to 60%), and high (>60%) probability.
† The bronchoscopic examination was nondiagnostic for 272 of 639 patients (43%; 95% CI, 39 to 46), including 120 of 487 patients (25%; 95% CI, 21 to 29) who had lung cancer.
‡ The classifier accurately predicted the presence of cancer in 107 of 120 patients overall (89%; 95% CI, 82 to 94). Sensitivity is reported in each probability category for patients with nondiagnostic bronchoscopic examinations.
§ The classifier accurately predicted the absence of cancer in 72 of 152 patients overall (47%; 95% CI, 40 to 55). Specificity is reported in each probability category for patients with nondiagnostic bronchoscopic examinations.
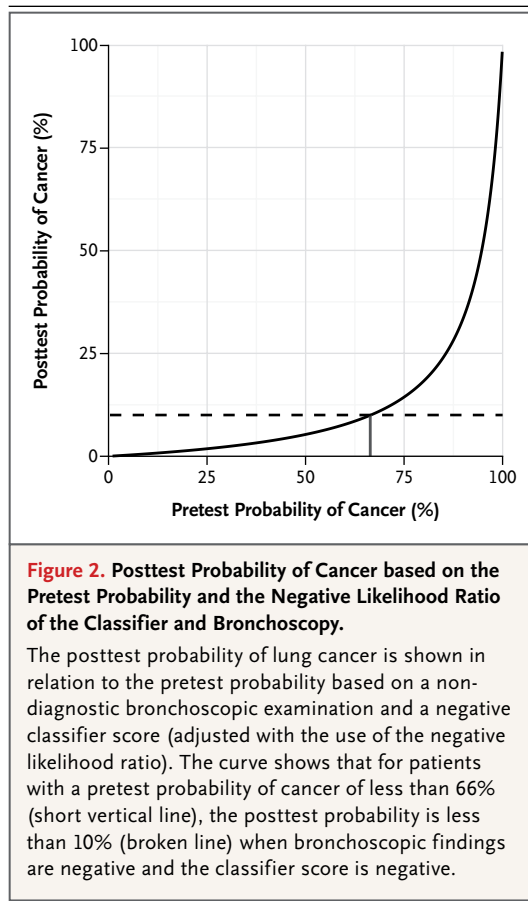¶ Negative and positive predictive values are reported for patients with nondiagnostic bronchoscopic examinations.

cells within the pulmonary lesion. The ability of alterations in gene expression in the cytologically normal proximal airway to enable the detection of lung cancer within the lung parenchyma stems directly from the "field of injury" paradigm.[12] We have previously shown that there is a distinct pattern of alterations in gene expression in cytologically normal bronchial epithelial cells among current and previous smokers with lung cancer.[12,16] In addition, oncogenic signaling pathways are activated in the proximal-airway epithelium of smokers with lung cancer and smokers with premalignant airway lesions.[17] More recently, Kadara et al.[18] showed that many of the genes that have altered expression in lung-cancer lesions and in the adjacent small-airway epithelium are also altered in the proximal-airway epithelium, which suggests that the changes in gene expression within the proximal airway reflect, in part, the altered transcriptome observed in lung tumors.

There are several important limitations to this study. First, specimens from 155 patients (11%) yielded insufficient or poor-quality RNA,

precluding measurement of the classifier. However, similar rates of insufficient RNA quality or quantity have been observed with other gene-expression tests that have been integrated into clinical practice,[19] and it may be possible to improve sample quality by decreasing the time between sample collection and RNA isolation. Patients who were not included in the study for this reason do not appear to differ in terms of cancer prevalence or other clinical features in comparison with the overall study population (Table S6 in the Supplementary Appendix); however, it cannot be determined whether the classifier has similar performance in this group.

Second, 9% of patients were lost to follow-up, and 5% did not have a definitive diagnosis established at 12 months after bronchoscopic examination. This rate of loss to follow-up is not unexpected in an observational trial in which the subsequent evaluation after bronchoscopic examination was not mandated to occur at the study center. Although our follow-up period was limited to 12 months, it is unlikely that we missed a substantial number of cancers that

Figure 2. Posttest Probability of Cancer based on the Pretest Probability and the Negative Likelihood Ratio of the Classifier and Bronchoscopy.

The posttest probability of lung cancer is shown in relation to the pretest probability based on a non-diagnostic bronchoscopic examination and a negative classifier score (adjusted with the use of the negative likelihood ratio). The curve shows that for patients with a pretest probability of cancer of less than 66% (short vertical line), the posttest probability is less than 10% (broken line) when bronchoscopic findings are negative and the classifier score is negative.

would have been found with an additional year of follow-up. Although guidelines suggest 24 months of surveillance, these recommendations are based on older studies regarding solitary pulmonary nodules discovered on chest radiography (not computed tomography [CT]).[20,21] The high sensitivity of CT makes it unlikely that solid nodules that are stable in the first year will have subsequent growth; this is supported by studies of lung-cancer screening in which nodules that were stable for 1 year had a conversion rate to cancer of only 1 per 1000 during year 2.[22]

Third, the exclusion criteria in this study limit the generalizability of these findings among lifetime nonsmokers and smokers with a history of lung cancer. It is unclear whether a similar field of injury exists in people who have never smoked or in very light smokers who have lung cancer and whether the field of injury persists after tumor resection; further studies are needed to evaluate these questions.

Fourth, we considered bronchoscopy to be

"diagnostic" only when the procedure yielded a lung-cancer diagnosis. There were 49 bronchoscopic examinations in which a specific benign cause was identified, but 31 of the patients received further invasive testing, including 4 patients who ultimately had lung cancer diagnosed on subsequent lung biopsy; this suggests that the concern for cancer remained elevated despite the initial benign finding on bronchoscopic examination.

Finally, we did not assess the accuracy of a model incorporating the classifier in combination with clinical variables. Although risk-prediction models have been developed for solitary pulmonary nodules,[1,23,24] there are no validated models for patients undergoing diagnostic bronchoscopic examination, which includes patients with a broad range of findings, including larger lesions (i.e., >3 cm), infiltrates, or lymphadenopathy. Thus, most patients are selected for bronchoscopy on the basis of the physician's qualitative assessment of the probability of lung cancer. We show that our classifier performs well in patients with an intermediate probability of cancer as assessed by a physician in a process that incorporates the available clinical risk factors.

The potential effect of these studies is bolstered by a number of key strengths. First, we report on two independent prospective studies in which the classifier was locked down before validation and then measured under conditions in which the test would be used clinically (i.e., before diagnosis). These are critical steps in moving molecular biomarkers from discovery studies to their ultimate clinical application. Second, the large multicenter design enabled the inclusion of patients undergoing bronchoscopic examination from different practice settings and geographic locations. Third, our data show that there is a high prevalence of lung cancer among patients with a nondiagnostic bronchoscopic examination and an intermediate probability of cancer, as well as a negative predictive value of 91% for the classifier in these patients, for whom there is the greatest uncertainty about cancer status. A negative classifier score in patients with a nondiagnostic bronchoscopic examination and an intermediate probability of cancer may warrant a more conservative diagnostic strategy that involves active surveillance with the use of imaging rather than immediate invasive procedures.

REFERENCES

1. 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-e120S.

2. Tukey MH, Wiener RS. Population-based estimates of transbronchial lung biopsy utilization and complications. Respir Med 2012;106:1559-65.

3. Ernst A, Silvestri GA, Johnstone D. Interventional pulmonary procedures: guidelines from the American College of Chest Physicians. Chest 2003;123:1693-717.

4. Rivera MP, Mehta AC, Wahidi MM. 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;143:Suppl 5: e142S-e165S.

5. Wang Memoli JS, Nietert PJ, Silvestri GA. Meta-analysis of guided bronchoscopy for the evaluation of the pulmonary nodule. Chest 2012;142:385-93.

6. Ost D, Fein AM, Feinsilver SH. The solitary pulmonary nodule. N Engl J Med 2003;348:2535-42.

7. Grogan EL, Weinstein JJ, Deppen SA, et al. Thoracic operations for pulmonary nodules are frequently not futile in patients with benign disease. J Thorac Oncol 2011;6:1720-5.

8. Detterbeck FC, Mazzone PJ, Naidich DP, Bach PB. Screening for 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: e78S-e92S.

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

10. Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. Ann Intern Med 2011;155:137-44.

11. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531-7.

12. Spira A, Beane JE, Shah V, et al. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. Nat Med 2007;13:361-6.

13. Whitney DH, Elashoff MR, Porta K, et al. Derivation of a bronchial genomic classifier for lung cancer in a prospective study of patients undergoing diagnostic bronchoscopy. BMC Med Genomics 2015; 8:18.

14. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.

15. MacMahon H, Austin JH, Gamsu G, et al. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. Radiology 2005;237:395-400.

16. Beane J, Sebastiani P, Whitfield TH, et al. A prediction model for lung cancer diagnosis that integrates genomic and clinical features. Cancer Prev Res (Phila) 2008;1:56-64.

17. Gustafson AM, Soldi R, Anderlind C, et al. Airway PI3K pathway activation is an early and reversible event in lung cancer development. Sci Transl Med 2010;2: 26ra25.

18. Kadara H, Fujimoto J, Yoo SY, et al. Transcriptomic architecture of the adjacent airway field cancerization in non-small cell lung cancer. J Natl Cancer Inst 2014;106:dju004.

19. Alexander EK, Kennedy GC, Baloch ZW, et al. Preoperative diagnosis of benign thyroid nodules with indeterminate cytology. N Engl J Med 2012;367:705-15.

20. Holin SM, Dwork RE, Glaser S, Rikli AE, Stocklen JB. Solitary pulmonary nodules found in a community-wide chest roentgenographic survey: a five-year follow-up study. Am Rev Tuberc 1959;79:427-39.

21. Lillington GA. Management of solitary pulmonary nodules. Dis Mon 1991; 37:271-318.

22. van Klaveren RJ, Oudkerk M, Prokop M, et al. Management of lung nodules detected by volume CT scanning. N Engl J Med 2009;361:2221-9.

23. McWilliams A, Tammemagi MC, Mayo JR, et al. Probability of cancer in pulmonary nodules detected on first screening CT. N Engl J Med 2013;369: 910-9.

24. Ost DE, Gould MK. Decision making in patients with pulmonary nodules. Am J Respir Crit Care Med 2012;185:363-72.
Copyright © 2015 Massachusetts Medical Society.

NEJM CLINICAL PRACTICE CENTER

Explore a new page designed specifically for practicing clinicians,
the NEJM Clinical Practice Center, at NEJM.org/clinical-practice-center.
Find practice-changing research, reviews from our Clinical Practice series,
a curated collection of clinical cases, and interactive features
designed to hone your diagnostic skills.