Biomarkers for Lung Cancer Screening and Detection

Edwin J. Ostrin, David Sidransky, Avrum Spira and Samir M. Hanash

ABSTRACT

Lung cancer is the leading worldwide cause of cancer mortality, as it is often detected at an advanced stage. Since 2011, low-dose CT scan–based screening has promised a 20% reduction in lung cancer mortality. However, effectiveness of screening has been limited by eligibility only for a high-risk population of heavy smokers and a large number of false positives generated by CT. Biomarkers have tremendous potential to improve early detection of lung cancer by refining lung cancer risk, stratifying positive CT scans, and categorizing intermediate-risk pulmonary nodules. Three biomarker tests (Early CDT-Lung, Nodify XL2, Percepta) have undergone extensive validation and are available to the clinician. The authors discuss these tests, with their clinical applicability and limitations, current ongoing evaluation, and future directions for biomarkers in lung cancer screening and detection.

See all articles in this CEBP Focus section, "NCI Early Detection Research Network: Making Cancer Detection Possible."

Introduction

Lung cancer is the leading worldwide cause of cancer mortality (1). A major contributor to the lethality of lung cancer is that most lung cancers are diagnosed at an already advanced stage (2). Screening for lung cancer was generally not considered successful until 2011, when the National Lung Screening Trial (NLST) demonstrated a reduction in lung cancer mortality of 20% through low-dose CT screening of a selected pool of high-risk patients (3). Although the trial was a success, the screening protocol produced a high number of false-positive results, with 24% of scans read as positive and only 6% of these confirmed as true positive. Through the 3 years of screening in the trial, 39% of the CT-screened subjects had at least one false-positive scan. In the United States, the accepted entry criteria (>55 years old, >30 pack-year history of smoking, currently smoking or <15 years since quit) are somewhat restrictive, with only 27% of lung cancer diagnoses estimated to occur in screening-eligible patients. The expense and logistics of implementing such a screening trial are thus complex, and the false-positive rate is a tangible downside for patients considering screening (4, 5). Thus, in the United States, only 4% of eligible high-risk patients undergo screening yearly for lung cancer (Fig. 1; ref. 6). As a compounded limitation, while CT screening is recommended by the U.S. Preventive Services Task Force, adoption by other health services worldwide has been piecemeal at best (7, 8). Even with full implementation, only around 10% of U.S. lung cancer deaths could be averted (9).

Outside of true cancer screening, approximately 2% of the U.S. population undergoes CT of the chest yearly for various reasons (10). For both low-dose CT screening and incidentally discovered nodules, there is a substantial need for biomarkers that could accurately discriminate benign lesions from early cancers at the time of imaging. Several diagnostic biomarkers have come to market, with several other biomarkers in active development. Here, we will explore the uses and limitations of existing biomarkers and the path to development of additional, clinically useful biomarkers for early detection of lung cancer.

Discussion

Clinical utility

There are many potential roles for lung cancer biomarkers, both within the setting of low-dose CT screening and independent of combined testing. In patients being considered for screening, a biomarker may help to stratify risk, optimally performing to refine existing lung cancer risk models based on age, smoking, and family history (13–15). Such a risk model would then help to identify subjects at a high risk for lung cancer who may benefit from CT screening but are excluded by standard NLST or National Comprehensive Cancer Network criteria. It also may encourage a subset of patients who meet the established criteria but who are reluctant to undergo screening. These patients may test negative with a biomarker and thus could be clinically followed with less frequent CT scans or serial biomarker testing.

After scanning, when nodules are detected, there is an obvious role for biomarkers in classifying nodule risk. Again, an effective biomarker should perform in combination with a clinical nodule risk score, for instance, the Lung-RADS criteria or a model such as proposed by McWilliams and colleagues or Swensen and colleagues (14, 15). Here, a biomarker may be of utility both in the setting of CT screening and in the workup of incidentally discovered lung nodules. It is in this diagnostic setting that this field has seen most progress, with three approved tests available. These tests, two of which are blood based and the other performed on airway epithelial brushings collected during bronchoscopy, are reviewed below.

Given the multiple roles of biomarkers and different approaches to screening, assessment of a biomarker’s performance clearly necessitates reporting its sensitivity, specificity, negative (NPV) and positive predictive value (PPV), as well as likelihood ratio (LR). For instance,
Like an iceberg, the bulk of cases of lung cancer are not only not detected by screening, but occur in patients not eligible for screening. Approximately 4% of NLST-eligible patients undergo screening. If all NLST-eligible patients underwent screening, only 27% of lung cancers would be detected, similar to the portion of an iceberg seen above water. The remaining 73% of lung cancers occur in those ineligible for screening, for instance those with only a light smoking history, those who have quit >15 years ago, or those who have never smoked. Biomarkers in high-risk individuals can decrease the rate of false positives after CT-based screening. In lower-risk individuals, biomarkers can be used to identify patients at higher risk who may benefit from screening.

Figure 1.

The clinical utility of a biomarker intended for use in combination with a risk score or cut-off can be easily assessed using a diagnostic LR. As an example, one can point to the Lung-RADS criteria used in CT-based screening (12). In this population, the pretest probability for lung cancer in the setting of a positive CT scan is 4%. With classification of these positive scans using the Lung-RADS guidelines, it is recommended that patients undergo more intensive follow-up if they are in category 4A or higher, whereby the pre-test probability of malignancy is >5%. In the Lung-RADS guidelines, nodules in class 2 have a pretest probability of <1% and thus return to standard annual screening is advised. Lung-RADS category 4B has a pretest probability of >15% and further workup by PET-CT or biopsy is indicated. These should then be our goalsposts, such that if we could move the posttest probabilities to >15% or <1%, our clinical decision making would be enhanced. Given the pretest probabilities, we can use Bayes theorem to calculate the needed positive and negative LRs (LR+ and LR−) to deliver these posttest probabilities. We can thus estimate that a diagnostic test, given a pretest probability of 5%, should have an LR+ of 3.38 to deliver a posttest probability of 15%, and should have an LR− of 0.191 to deliver a posttest probability of <1%.

Another outlook is to consider a test’s NPV or PPV. This allows a rapid assessment of a value of a test: if negative, a test with an NPV of 95% would only miss 5% of cases. However, the NPV and PPV depend on disease prevalence, which varies both in the populations on which these tests were calibrated and in their intended clinical use. For example, a test with 95% specificity would still yield a low PPV in most at-risk smoking populations for lung cancer bringing in hundreds of patients without disease for further assessment.

Blood-based biomarkers

Detection of biomarkers in the blood holds numerous advantages, including the relative noninvasive nature of blood draws and well-established laboratory pipelines for isolation and analyses of various assays from plasma, exosomes, circulating nucleic acids, and circulating cells. Two tests, mostly intended for classification of indeterminate pulmonary nodules (IPN), are currently available.

EarlyCDT-Lung (OnclImmun), is a seven-autoantibody panel, first developed in 2010 and extensively validated in seven different cohorts (16–20). This panel, consisting of autoantibodies against p53, CAGE, NY-ESO-1 (CTAG1B), SOX2, GBU4−5, HuD, and MAGE-A4, has shown good performance in classifying indeterminate pulmonary nodules, with important information produced both when the panel is positive or negative. Work to assemble this panel began in 2006, through several different assessments of lung cancer autoantibodies. These utilized forward discovery tools such as phage display and serological analysis of tumor antigens by recombinant cDNA (SEREX), as well as work analyzing humoral response to known cancer antigens. From these, a six-autoantibody panel was selected, with clinical validation in three different matched cohorts of patients with newly diagnosed lung cancer, showing a specificity of approximately 90% at sensitivities of around 40%. The panel was technically validated in three separate cohorts and four additional postvalidation cohorts of newly diagnosed patients with lung cancer versus control (21, 22).

Further development included testing and validation of a related seven-autoantibody panel (dropping annexin I and adding HuD and MAGE A4), which showed an improvement in performance, raising specificity from 82% to 90% in a prospective validation in an at-risk population from the previous six-autoantibody panel (23). On the basis of this, the seven-marker panel has replaced the original six-marker panel as the commercially available EarlyCDT-Lung test, and in a postmarketing audit of over 1,600 patients presenting with a nodule, showed a similar sensitivity of 41% at a specificity of 87% (24). A cost effectiveness study indicated that the use of EarlyCDT-Lung in patients presenting with nodules of approximately 8–30 mm is around $24,000 per quality-of-life adjusted life year gained (25).

This test, while extensively validated in newly diagnosed lung cancer and nodule cohorts, has not been evaluated extensively as a part of a traditional low-dose CT-based screening program. However, if one were to assume that it would perform similarly in a screening population after the detection of nodules, with a sensitivity of 40% and specificity of 90%, this would generate a positive diagnostic LR (LR+) of 4, calculated as sensitivity/(1−specificity), which is enough to reclassify some low-probability nodules into a category where they should have radiographic follow-up sooner. Interestingly, mathematical modeling of a high- and low-specificity version of EarlyCDT-Lung has shown the ability to both reclassify nodules into a higher- and a lower-risk group (26). Recently, a double-blinded randomized trial was published describing the use of EarlyCDT-Lung followed by CT scan in a population at higher risk for lung cancer (27). Participants were randomized into an intervention arm that began with the EarlyCDT-Lung test then CT scan if the biomarker test was positive versus control, or standard of care that did not include CT scans and relied on symptomatic presentation. Over 2 years, EarlyCDT-Lung did not increase the frequency of detection of lung cancer, but lung cancers detected in the intervention arm were at earlier stage. They detected 56 cancers in the intervention arm, of which 23 were early stage (41.1%), and 71 in the control arm, of which 19 were early stage (26.8%). Because most subjects did not under CT scanning, sensitivity and specificity were estimated using cancer registry data. Sensitivity was lower than in previous trials, at 32.1%, with a preservation of specificity at 90.3%. Of note, sensitivity was particularly low for stage III or IV cancers, at 18.2%. Continued longitudinal monitoring, to better measure the true cancer incidence in each arm, is warranted.
A second test, Nodify XL2 (Biodexis), is also available for the classification of indeterminate pulmonary nodules (28). This test measures a panel of blood proteins using a mass spectrometry-based assay. This test originated in a 2013 study that reported a 13-protein proteomic classifier, with scoring by a logistic regression model that gave a 90% NPV for benign nodules (29). The model was developed in a discovery set of patients with IPNs between 4 and 20 mm, and with a lung cancer incidence of 20%. Markers were measured using a mass spectrometry multiple reaction monitoring technique. The panel was then validated in two independent sets, in which it gave similar performance. Later work by the same group used a five-marker subset of the original 13 proteins plus 6 normalization markers, which they validated in a similar population (30). A subsequent study claimed clinical utility based on the test’s NPV potentially sparing invasive procedures for 31.8% of subjects (31).

The panel underwent continued refinement on a cohort of 222 patients presenting with nodules 8–20 mm in size who had undergone invasive workup (32). The panel was integrated with clinical risk factors, with a specific focus on the marker performance in this high-prevalence population (81% diagnosed with lung cancer). Using decision tree analyses, an integrated model was found to have mildly improved performance versus the Mayo clinical risk model or the proteomic classifier individually (ROC from 58% for the clinical risk score, 60% for the proteomic classifier alone, to 63% for the integrated model). This integrated model, termed Xpresys Lung (XL2; Integrated Diagnostics), was then prospectively validated in the PANOPTIC study of 685 patients presenting with 6–30 mm nodules (33). While XL2 was tuned for a high-incidence population, PANOPTIC revealed its best performance in a subgroup with clinicians’ estimated pretest probability of cancer of <50%. In this subgroup of 178 patients, who had a lung cancer prevalence of 16%, the classifier showed a sensitivity of 97% and a specificity of 44%, with an NPV of 98% (and a LR− of 0.07). These findings supported the use of this panel for identifying low-risk pulmonary nodules, outperforming PET/CT, physician estimates, and lung nodule risk scores.

**Airway gene expression classifiers**

Cigarette smoking produces gene expression alterations throughout the epithelial cells that line the respiratory tract, leading to an airway “field of injury” (34–36). Furthermore, cancer-associated gene expression patterns are found in cytologically normal epithelium collected from the bronchial airways of current and former smokers with lung cancer (37). As an initial proof of concept in 2007, investigators identified an 80-gene expression set, measured from brushings of histologically normal bronchial airways that could distinguish smokers with and without lung cancer with 80% sensitivity and 84% specificity in a validation set of samples (37). Accuracy was maintained in early-stage cancers, with 90% sensitivity in stage I tumors. Interestingly, the accuracy of the panel was not affected by the location of the airway brushing relative to the cancer, implying a broad field of injury throughout the bronchial epithelium. This panel was additive to bronchoscopy cytopathology obtained during the same procedure. Cytology alone only yielded a diagnosis of cancer in 32 of 60 subjects with cancer and ruled out cancer in 5 of 69 subjects without cancer. Among the cytopathologically nondiagnostic bronchoscopies, the classifier performance was consistent (89% sensitivity, 83% specificity).

On the basis of these findings, two large prospective multicenter trials (AEGIS-1 and AEGIS-2) were conducted, enrolling patients undergoing bronchoscopy for suspicion of lung cancer (38). During these bronchoscopies, airway brushings were performed on the normal-appearing mainstem bronchi and underwent RNA expression profiling by microarray. Both cohorts had high prevalence of lung cancer (74% and 78%), with the cancer subjects being older, heavier smokers, and more likely to have lymph node involvement of lymph nodes. To define a clinical utility, the authors further examined the performance of the classifier based on physician-assessed prebronchoscopy probability of cancer. In patients with an intermediate pretest probability and a negative bronchoscopy (who did not have cancer prevalence of 41%), the classifier had a 91% NPV. In a combined group of low and intermediate probability patients with nodules <3 cm, the reported sensitivity was 88% with an NPV of 94%.

Combining the classifier with bronchoscopy increased sensitivity from 74% to 76% for bronchoscopy alone to 96%–98%. In patients with nondiagnostic bronchoscopies, the sensitivity was maintained, at 86%–92%, with no impact from size or location of the lesion, cancer stage, or involvement of lymph nodes. To define a clinical utility, the authors further examined the performance of the classifier based on physician-assessed prebronchoscopy probability of cancer. In patients with an intermediate pretest probability and a negative bronchoscopy (who had cancer prevalence of 41%), the classifier had a 91% NPV. In a combined group of low and intermediate probability patients with nodules <3 cm, the reported sensitivity was 88% with an NPV of 94%.

These findings spurred additional research on the classifier, including work defining possible clinical utility and cost-effectiveness within the AEGIS collection (39, 40). Use of Percepta was projected to reduce unnecessary invasive procedures by as much as 50%, with a false-negative rate of 11%. Use of the classifier reduced costs for invasive procedures, with a small benefit in quality adjusted life years, giving a modest gain in incremental cost-effectiveness ratio. The test has been successfully deployed and in clinical use since 2015, with continued postmarketing data collection that facilitated refinement of the test. In 2019, an updated test based on next-generation RNA transcriptome sequencing, termed Percepta Genomic Sequencing Classifier, was released, which allows for both up-classification and down-classification among those at intermediate pretest risk of lung cancer.

**Future directions**

Given the high failure rate of cancer biomarkers, there needs to be substantial rigor in experimental design and extensive critical assessment of derived panels. There have been many guidelines published on best practices for validation of biomarkers, including the Institute of Medicine and REMARK guidelines (41, 42). However, we would add guidelines that mandate independent validation not only of the panel itself, but of the collection, the analysis, and the statistical techniques particularly when claims of clinical utility are made and the test is offered to patients. A comprehensive analysis of the missed costs of overdagnosis and underdiagnosis in a lung nodule cohort has not been performed. For instance, EarlyCDT-Lung is used to identify high-risk nodules, with false-positive tests leading to overdagnosis. Even small cumulative risks such as additional radiation from CT scans and PET-CTs, could overwhelm a small improvement in mortality from early detection of a subset of lung cancers. For Nodify XL2, a false-negative test could lead to delayed diagnosis. However, it is difficult to estimate whether any potential delay in diagnosis would be outweighed by the benefits of fewer unnecessary invasive procedures performed. Some of the analyses on Percepta have addressed these questions, and continued postmarketing data analysis is necessary to see whether these predictions correspond with real-world performance (39). These
are difficult questions to answer, and will require extensive, long-term investigation.

These biomarkers have demonstrated their performance through repeated validation. However, their best utility is in a limited clinical role, with EarlyCDT-Lung showing good specificity and Nodify XL2 showing good sensitivity in indeterminate pulmonary nodules, which thus precludes head-to-head comparison. However, modeling such a comparison, with both tests used in a potential “real world” scenario outside of their intended use, may still be valuable. Percepta, a high-sensitivity test done using bronchoscopic samples, likely will not find usage in a much higher-prevalence population than a blood-based biomarker. However, a comparison of Nodify XL2 and Percepta in patients with nodules undergoing bronchoscopy would be enlightening.

There are still several unmet needs in the lung cancer biomarker field. As mentioned, even with a full implementation of lung cancer screening under current guidelines, only a small fraction of lung cancer deaths would be averted (9). In a blinded validation study, a biomarker panel that combines a previously validated marker (43) with three additional markers has shown the ability to improve a clinical smoking-based risk model (44). A low-cost, easily implementable test has a high potential to alter lung cancer screening and diagnostics in the future. More accurate risk stratification, through a combined clinical and biomarker risk score, has the potential to reduce noneffective screening (and thus false positives) and also identify higher-risk subjects who do not meet NLST screening criteria yet may benefit from lung cancer screening. Future studies will need to assess the utility of this panel combined with screening.

An additional area for future exploration is the process of biomarker discovery itself. We believe that adopting rational evaluation of biomarker discovery would benefit the field, centering on biological plausibility of the marker panel. Biomarkers can be found in plasma, brushings, and sputum. For example, promoter methylation has shown promise in both plasma DNA, sputum, and effusions (45, 46). In an initial study in plasma, patients were divided into two groups, ground glass opacity (n = 23) and cancerous tumors (n = 70). Plasma DNA from age-matched nodule-free individuals were used as controls (n = 80). A total of 73% of patients with cancerous tumors showed methylation of at least one gene with a specificity of 71% and a methylation marker was found in only 22% of those with a ground glass opacity (46). Promoter methylation of a 6-gene panel was detected in 92.2% (83/90) of a training cohort with a specificity of 72.0% (18/25) and in 93.0% (40/43) of an independent cohort of stage IA primary non–small cell lung cancer (47). Tissue and cell-free DNA-based epigenomic approaches for cancer detection were recently reviewed elsewhere (48).

In addition, new biomarker panels should reflect the heterogeneity of the disease, including driver and passenger mutations, differences in immune response, and histologic subtypes. For biomarker discovery, the collection, isolation, and analysis protocols should be robust and reproducible, and based in biology. The members of a multicomponent panel need to be individually validated, preferably on a variety of platforms and in more than one center. Batch effect may be introduced by platform calibration, sample collection, sample preparation, and at numerous other steps. This could be addressed through reanalysis of replicates of the original biospecimens done at a later time. There should be a dose response between the detected abundances and the size of a lesion, to help to establish biologic plausibility behind a correlation.

Conclusions

Biomarkers are playing an emerging role in the early detection of lung cancer. There are many potential roles for biomarkers, from risk stratification to classification of nodules detected incidentally or through low-dose screening programs. Several multianalyte biomarker panels are available and have shown performance in classification of indeterminate pulmonary nodules. However, these panels must be used in an appropriate clinical context. In addition, there is still much work to be done on unfulfilled needs within and outside of CT-based screening.

Authors’ Disclosures

E.J. Ostrin reports grants from American Lung Association and The University of Texas MD Anderson outside the submitted work, as well as a patent for use of biomarker panel in nodule stratification pending to The University of Texas MD Anderson. A. Spira reports grants and personal fees from Johnson and Johnson during the conduct of the study; grants and personal fees from Johnson and Johnson outside the submitted work, and a patent for 10570454 Methods of Identifying Individuals at Increased Risk of Lung Cancer issued, a patent for 10314855 Methods Relating to Lung Cancer issued, a patent for 9920374 Diagnostic for Lung Disorders Using Class Prediction issued, a patent for 20200232045 Methods of Identifying Individuals at Increased Risk of Lung Cancer pending, a patent for 20200222446 Methods Relating to Lung Cancer pending, a patent for 20200115763 Detection Methods for Disorders of the Lung pending, a patent for 20200096513 Diagnostic and prognostic Methods for Lung Disorders Using Gene Expression Profiles from Nose epithelial Cells pending, a patent for 20200057053 Methods Related to Bronchial Premalignant Lesion Severity and Progression pending, a patent for 20190376148 Diagnostic for Lung Disorders Using Class Prediction pending, a patent for 20190629600 Nasal Epithelium Gene Expression Signature and Classifier for the Prediction of Lung Cancer pending, a patent for 20190247418 Methods Relating to Lung Cancer pending, a patent for 20190171418 Diagnostic for Lung Disorders Using Class Prediction pending, a patent for 20180171418 Diagnostic for Lung Disorders Using Class Prediction pending, a patent for 20180110197 Gene Expression-Based Biomarker for the Detection and Monitoring of Bronchial Premalignant Lesions pending, and a patent for 20170328908 Diagnostic and Prognostic Methods for Lung Disorders Using Gene Expression Profiles from Nose Epithelial Cells pending. S.M. Hanash reports a patent application filed for a four-marker panel for lung cancer pending to Cosmos, Dynex. No disclosures were reported by the other authors.

Acknowledgments

This work was supported by American Lung Association Lung Cancer Discovery Award (419882) (to E.J. Ostrin), NIH/NICHIU01CA206027 and NIH/NICHIU01CA208709 (to D. Sidransky), NIH/NICHIU20CA233280S1, NIH/NICHIU20CA233280-01, and NIH/NICHIU20CA233280-01 (to A. Spira); and NIH/NICHIU01CA194733-01A1, NIH/NICHIU01CA213285-01A1, and NIH/NICHIU01CA203654 (to S.M. Hanash). The authors would like to acknowledge Lila Ostrin for preparation of the iceberg illustration in Fig. 1.

Received June 10, 2020; revised October 1, 2020; accepted October 16, 2020; published first October 22, 2020.

References


Biomarkers for Lung Cancer Screening and Detection


Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-20-0865

Cited articles
This article cites 42 articles, 10 of which you can access for free at:
http://cebp.aacrjournals.org/content/29/12/2411.full#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/29/12/2411.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cebp.aacrjournals.org/content/29/12/2411.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.