Advances in Sputum Analysis for Screening and Early Detection of Lung Cancer

Melvyn S. Tockman, MD, PhD

Background: Screening for lung cancer using currently available techniques is not effective in reducing mortality from the disease.

Methods: Archived sputum specimens and clinical data linking specimens to lung cancer outcomes from prior screening programs have been reexamined to evaluate altered gene expressing, including specific oncogene activation and tumor suppressor gene deletion, as well as genomic instability and abnormal methylation.

Results: Several of these tests allow determination of a molecular diagnosis of cancer years before clinical presentation.

Conclusions: These sputum tests provide an impetus to reconsider screening for lung cancer. Prospective trials are required to confirm test performance characteristics, and management and intervention strategies must be developed that are appropriate to the stage at which lung cancer is diagnosed.

Abnormal Protein (hnRNP A2/B1) Overexpression

Sputum cells can now be probed for altered gene expression to reveal more about the preneoplastic state of the airway. Morphologic changes in sputum epithelial cells and standard chest radiography, long the only preclinical lung cancer diagnostic evaluations available, had been carefully studied 20 years ago in the Collaborative Early Lung Cancer Detection studies of the National Cancer Institute (NCI) at Memorial Sloan-Kettering, Mayo Clinic, and Johns Hopkins. These specific diagnostic tests, while useful for indi-
individual case finding, were found to be not sufficiently sensitive for lung cancer screening. Less than half (49%) of the lung cancer cases that arose during screening were detected by either standard sputum cytology or chest radiograph, and only 11% were detected by cytology alone.

During the Johns Hopkins Lung Project (JHLP), we developed an archive of sputum specimens and associated clinical data linking specimens to lung cancer outcome. To identify cancer-associated protein overexpression, we tested promising clinically available antibodies plus a series of murine monoclonal antibodies (MoAbs) raised by colleagues at the NCI. Differential display of two of these monoclonal antibodies (MoAbs 703D4 and 624H12) identified biomarkers of lung cancer in archived sputum specimens two years prior to clinical detection of lung cancer. For the JHLP archived specimens, preserved for their cytologic atypia (moderate or grave atypical metaplasia), these antibodies together showed a sensitivity of 91% and a specificity of 88% for the diagnosis of lung cancer within two years.

We have found that MoAb 703D4 recognizes an epitope of hnRNP A2 and its splice variant, hnRNP B1. This target antigen for 703D4 was purified using Western blot detection after SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) separation. Purification steps included anion exchange chromatography, preparative isoelectric focusing, and polymer-based reverse-phase high-performance liquid chromatography (HPLC). After 25,000-fold to 50,000-fold purification, the principal immunostaining protein was more than 95% pure. Three sequences, including one across a site of alternate exon splicing, all identified a single protein, heterogeneous nuclear ribonucleoprotein-A2 (hnRNP-A2). The splice variant hnRNP-B1 was a minor copurifying immunoreactive protein. The hnRNPs are members of a family of ribonucleic acid proteins that are generally thought to regulate the shuttling of nascent RNA transcripts between the nucleus and cytoplasm. Interactions of these molecules are also thought to regulate mRNA splicing, capping, and polyadenylation. Recent information suggests other fundamental roles for this important nucleoprotein family. The hnRNP A2/B1 family of antigens is frequently observed in transformed bronchial epithelium, and its increased expression is associated with a critical phase of fetal lung development for three mammalian systems, suggesting an oncofetal role for this protein.

We began a clinical trial to evaluate the performance of the hnRNP A2/B1 protein as a biomarker for the early detection of second primary lung cancer (SPLC). The patients at risk for SPLC have the highest annual incidence of lung cancer (2% to 5%) among asymptomatic populations. The Lung Cancer Early Detection Working Group (LCEDWG), which is composed of thoracic surgeons and medical oncologists from leading medical centers throughout the United States and Canada, collaborate in this ongoing trial. Accrual to this trial is now supported by a Collaborative Research and Development Agreement (CRADA) among the NCI, Moffitt Cancer Center, and Chiron/Bayer Diagnostics.

MoAb 703D4 binds hnRNP A2/B1 within selected epithelial cells exfoliated in the sputum. In all cells correctly diagnosed by immunocytochemistry, we recognized at least a proplastic morphology. Proplasia consists of minimal cytologic changes that are usually regarded as normal epithelial responses to proliferative stimuli. To assure consistency in the selection of proplastic cells and reduce the possibility of false-positive diagnoses, we have agreed on a set of morphologic criteria based on the original description of these cells by Frost. These morphologic criteria reflect proliferative changes in nuclear morphology and a level of cytoplasmic immaturity. When such cells bind MoAb, we consider them to be sentinel cells for preclinical lung cancer.

Sentinel cells expressing upregulated levels of hnRNP A2/B1 are found infrequently (1 in 5,000) among cells that express normal low levels of this protein. By developing a cell-based diagnostic approach rather than a traditional mass extraction assay, we preserve in these isolated cells the natural upregulated signal compared with background noise.

Cellular distributions and concentrations of hnRNP A2/B1 indicative of lung cancer status were measured using semiautomated, quantitative image cytometry. We programmed a commercially available workstation to perform feature extraction of digitally recorded transmission optical microscope video images of immunostained sentinel cells at 510 and 600 nm. Multidimensional cluster analysis led to the selection of features for a discriminant function evaluation. Differential light transmission of immunostained cells was scored against the gold standard of known histologic lung cancer, resulting in a diagnostic accuracy of 87%. Development of hnRNP A2/B1 and its assay on a high-throughput platform are objectives in our CRADA with Chiron/Bayer Diagnostics.

Detection of hnRNP A2/B1 upregulation in morphologically “normal-appearing” sentinel cells permits greater generalizability of results compared with the results from the earlier JHLP specimens with moderately or gravely atypical metaplastic appearance. After the first
year of the LCEDWG trial, 13 SPLCs were identified. The sensitivity and specificity of the hnRNP A2/B1 biomarker for later SPLC were 77% to 82% and 65% to 81% respectively. Among the cases identified as positive by immunocytochemistry and image cytometry, 67% developed SPLC within one year. This diagnostic accuracy exceeds that commonly found in prostate-specific antigen (PSA) cancer screening tests. Working independently, Sueoka and colleagues recently published confirmation of this epitope to detect preclinical lung cancer in Japan. They propose to initiate lung cancer screening in that country. Detection of hnRNP A2/B1 overexpression in sputum epithelial cells with proplastic morphology appears to be the basis of a cytotest that could initiate a strategy of preclinical lung cancer diagnosis.

We have discussed an example of successful translation of a lung cancer biomarker (hnRNP A2/B1) from the laboratory through platform development and to application in a clinical trial. More than 6,000 individuals have been screened with this biomarker in ongoing clinical trials in North America, the United Kingdom, China, and Japan. Other less well-studied biomarkers have also been applied to exfoliated sputum epithelial cells.

Specific Oncogene Activation (ras) and Tumor Suppressor Gene Deletion (p53, 3p, 9p)

Tumor development progresses through a series of specific genetic changes in protooncogenes and tumor suppressor genes. Changes that cannot be repaired and do not trigger a program of cell death (apoptosis) may lead to a cellular growth advantage. Many of these genetic changes are acquired prior to and during the earliest stages of clonal expansion and are retained by daughter cells through the course of carcinogenesis and malignancy. If detected during the premalignant period, these specific genetic changes could serve as cancer markers.

Three closely related genes, H-ras, N-ras and K-ras, comprise the ras family of oncogenes. The highly conserved 21 kDa protein products of these genes are important signal transduction elements that participate in cell cycle regulation by controlling proliferation. Mutation of the K-ras2 oncogene is one of the most commonly occurring genetic lesion in colorectal cancer and is frequently mutated in lung cancer. The JHLP archive of preclinical sputum linked to tumor outcome allowed us to demonstrate that specific mutations could be detected in nonmalignant sputum specimens prior to clinical lung cancer. In this pilot study, we selected 15 participants in the JHLP with no malignancy in sputum cytology who went on to develop adenocarcinoma or large-cell carcinoma of the lung. These histologic cell types were selected because they have a higher incidence of K-ras mutations (30%) than other lung tumors. We also looked for p53 gene mutations because these are among the most common genetic alterations found in lung cancers (and other cancers). The first exon of K-ras or exons 5-8 of the p53 gene were amplified by polymerase chain reaction (PCR) from DNA extracted from the paraffin-embedded primary lung tumor. After cloning, the K-ras gene was sequenced to detect mutations. Tumors not containing K-ras mutations were sequenced for p53 mutations to detect tumor-specific markers. Once mutations specific for each tumor were identified, oligonucleotide probes were prepared, specific for the wild-type sequence or individual mutant K-ras and p53. These probes were hybridized to sputum DNA that had been amplified by PCR, cloned into a phage vector, and transferred to nylon membranes. Ten of the 15 patients had primary tumors that contained either a K-ras or p53 gene mutation. Identical mutations were detected in nonmalignant sputum cells from 8 of 10 patients who had tumors containing oncogene mutations. Patients whose tumors did not contain mutations as well as control patients without cancer were negative for sputum mutations by this assay.

This study demonstrated that 8 (53%) of 15 patients with adenocarcinoma or large cell carcinoma of the lung could be detected by mutations in sputum cells from 1 to 13 months prior to clinical diagnosis. Less sensitive than the protein marker described above, the identification of specific gene abnormalities is further limited by the need to know the specific mutation sequence with which to probe the sputum specimens. In this pilot study, the mutation sequence was determined from the resected tumor. Presently, this approach is obviously not practical for screening undiagnosed individuals. Perhaps with future advances in gene chip technology, it might become feasible to probe for all possible mutations of common oncogenes and tumor suppressor genes in sputum specimens of asymptomatic individuals.

Genomic Instability (Loss of Heterozygosity, Microsatellite Alterations)

Chromosomal alterations have been extensively documented in lung cancer. In 63 non-small cell lung carcinomas, Tesa et al found loss of chromosomes 13 (71%) and 9 (65%) to be the most frequent changes,
while a gain on chromosome 7 was seen in 41%. The chromosomal arms with most frequent loss were 9p (79%), 3p, 6q, 8p, 9q, 13q, 17p, 18q, 19p, 21q, 22q, and the short arm of the acrocentric chromosomes. Regions of chromosomal loss are suspected to encode tumor suppressor genes, the loss of which confers a cellular growth advantage. For each chromosomal locus, individuals will have two alleles, one contributed by each parent. While occasionally parents may provide identical genetic contributions at a given locus (homozygosity), often slight differences are observed among alleles (heterozygosity). It has been suggested that allelic imbalance (generated by loss of heterozygosity, LOH) on the short arms of chromosomes 3, 9, and 17 indicates the location of tumor suppressor genes associated with the early stages of lung cancer development.25-28

Microsatellite markers are small repeating DNA sequences found in the introns (noncoding regions) of a gene. PCR amplification of these repeat sequences provides a rapid method for assessment of LOH and facilitates mapping of tumor suppressor genes.29,30 Yet microsatellites can provide additional information. Expansion and deletion of these repeating elements are called microsatellite alterations. These microsatellite alterations, acquired during division of a single transformed cell, are passed onto daughter cells during clonal expansion. Since they are not transcribed, microsatellite alterations provide no growth advantage to the cell. However, detection of microsatellite alterations in histologic specimens is equivalent to the detection of neoplastic (clonal) cell populations. Although the detection of microsatellite alterations does not indicate the specific genetic change in the tumor, detection of clonal cell populations might serve as a cancer screening marker.31

Widespread microsatellite instability was first reported in colorectal tumors.32 In hereditary nonpolyposis colorectal carcinoma (HNPCC), mutations of mismatch repair genes are probably responsible for microsatellite alterations at multiple locations in the genome.33 However, in non-HNPCC-associated tumors, including lung cancer, there is not a similar widespread loss of mismatch repair, indicating that another as yet unknown mechanism is responsible for somatic alterations of repeat sequences.34

The pattern of microsatellite alterations and LOH may be specific for different types of cancer. The high incidence of these changes on chromosomes 3, 5, 8, 9, 10, 11, 17, and 20 have been described in lung cancer specimens,34,35 although the role of these changes in carcinogenesis is not yet known. Perhaps it is the cumulative effect of these genetic injuries that is important. Wistuba et al.36 reported a progressive increase of overall LOH frequency within clones with increasing severity of histopathologic changes in lung squamous carcinoma. We have already shown that microsatellite alterations are clonal markers for the detection of human lung cancer. Also, we and others have shown microsatellite alterations at selected loci can be recognized in sputum cells prior to clinical lung cancer.31 More recently, we found that when microsatellite alterations occurred at more than one locus, there was a significant association with hnRNP A2/B1 overexpression.37 We tested 41 paired tumor and normal DNA specimens from non-small cell lung cancer patients surgically resected at Moffitt Cancer Center. Eleven dinucleotide and tetranucleotide repeat markers were selected for their high frequency of loss (LOH) or alteration in lung cancer. In 41 paired tumor/normal samples, we found 19 patients (46%) had more than two loci of microsatellite instability (loss or alteration), while 13 patients (32%) had only one locus of microsatellite instability. The total frequency for microsatellite alteration was 78%, which suggests that this panel of markers may have a sensitivity comparable to hnRNP protein markers for detection of lung cancer. We are now evaluating whether this performance is maintained in sputum specimens of high-risk individuals.

Abnormal Methylation

The p16 gene is located on the short arm of chromosome 9 (at 9p21) and is frequently mutated or inactivated in tumors and cell lines derived from lung cancer.38,39 This gene codes for a protein that binds to the cyclin-dependent kinases 4 or 6 (cdk4 or cdk6) and prevents the kinase from phosphorylating (activating) cyclin D1. When phosphorylated, the activated cyclin D1 phosphorylates the retinoblastoma protein to allow the release of E2F transactivator and progression through the cell cycle.40 Therefore, p16 acts as a tumor suppressor gene, inhibiting mitosis and cell proliferation. Downregulation or loss of p16 expression could contribute to the loss of cell cycle control and provide a cellular growth advantage. Mao and colleagues41 have shown that nearly two thirds of former smokers show some genetic alteration in their bronchial cells. These investigators show that the frequency of LOH at the tumor suppressor sites 9p21 (p16) and 17p13 (p53) is nearly the same for smokers and for former smokers. In contrast, lifetime nonsmokers show no LOH at 9p21, suggesting that loss of 9p21 sequences might be an early event for the development of lung cancer.

Established causes of loss of tumor suppressor function include the loss of p16 expression through
gene deletion and expression of an altered protein through gene mutation. More recently, Merlo and colleagues described inhibition of p16 gene transcription by promoter region hypermethylation, and Herman et al. described a novel PCR assay for detection of this condition. Briefly, the addition of methyl groups to a sequence motif (CpG islands) in the gene promoter region results in gene transcription failure. For p16, these sequence motifs start at the promoter and extend into exon (transcribed region) 1c. Myohanen and colleagues showed that CpG island methylation-induced transcriptional repression could be at least partly reversed by cell culture treatment with the demethylating agent 5-aza-2'-deoxycytidine.

Belinsky et al. recently measured hypermethylation of the CpG islands of the p16 gene in the sputum of lung cancer patients and demonstrated a high correlation with the early stages of non-small cell lung cancer. These investigators suggest that detection of p16 CpG island hypermethylation might be useful in the prediction of individuals who might develop lung cancer. As yet, however, no prospective studies have been conducted to assess the performance of the hypermethylation assay on samples from individuals at risk for developing lung cancer.

Conclusions

These sputum tests offer great promise in determining a molecular diagnosis of lung cancer far in advance of clinical presentation. Any or all of these tests could be incorporated into the routine management of individuals at risk of developing primary or secondary primary lung cancer. However, several issues must be considered before these tests are ready for clinical application. First, test performance characteristics must be confirmed in prospective trials. For several of these tests, those trials are currently underway. Second, we must develop a management and intervention strategy appropriate to the stage at which lung cancer is diagnosed. The ability to detect lung cancer at the stage of clonal expansion, well in advance of malignant invasion of the basement membrane, suggests that noninvasive chemoprevention might be appropriate in such cases. Preliminary studies of chemopreventive agents are now underway at the NCI. Several of these agents could be delivered by inhaler to place a maximum dose directly on the transformed epithelium. We must now begin to plan for clinical trials that evaluate combined diagnostic and therapeutic approaches to assess their impact on the incidence of clinical lung cancer. Finally, the larger public health issues of cost and accessibility of lung cancer screening must be considered before these advances in sputum screening can reach their potential.

References

Lung Tumors: Fundamental Biology and Clinical Management.


