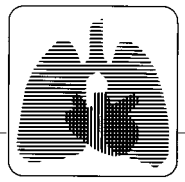
Detection of Lung Cancer With Volatile Markers in the Breath

Michael Phillips, Renee N. Cataneo, Andrew R.C. Cummin, Anthony J. Gagliardi,
Kevin Gleeson, Joel Greenberg, Roger A. Maxfield and William N. Rom

Chest 2003;123;2115-2123
DOI: 10.1378/chest.123.6.2115



preliminary report

Detection of Lung Cancer With Volatile Markers in the Breath*

Michael Phillips, MD; Renee N. Cataneo, MA; Andrew R.C. Cummin, DM; Anthony J. Gagliardi, MD, FCCP; Kevin Gleeson, MD; Joel Greenberg, BS; Roger A. Maxfield, MD, FCCP; and William N. Rom, MD, MPH, FCCP

Study objectives: **To evaluate volatile organic compounds (VOCs) in the breath as tumor markers in lung cancer. Alkanes and monomethylated alkanes are oxidative stress products that are excreted in the breath, the catabolism of which may be accelerated by polymorphic cytochrome p450-mixed oxidase enzymes that are induced in patients with lung cancer.**
Design: **Combined case-control and cross-sectional study.**
Setting: **Five academic pulmonary medicine services in the United States and the United Kingdom.**
Patients and participants: **One hundred seventy-eight bronchoscopy patients and 41 healthy volunteers.**
Intervention: **Breath samples were analyzed by gas chromatography and mass spectroscopy to determine alveolar gradients (*ie*, the abundance in breath minus the abundance in room air) of C4-C20 alkanes and monomethylated alkanes.**
Measurements: **Patients with primary lung cancer (PLC) were compared to healthy volunteers, and a predictive model was constructed using forward stepwise discriminant analysis of the alveolar gradients. This model was cross-validated with a leave-one-out jackknife technique and was tested in two additional groups of patients who had not been used to develop the model (*ie*, bronchoscopy patients in whom cancer was not detected, and patients with metastatic lung cancer [MLC]).**
Results: **Eighty-seven of 178 patients had lung cancer (PLC, 67 patients; MLC, 15 patients; undetermined, 5 patients). A predictive model employing nine VOCs identified PLC with a sensitivity of 89.6% (60 of 67 patients) and a specificity of 82.9% (34 of 41 patients). On cross-validation, the sensitivity was 85.1% (57 of 67 patients) and the specificity was 80.5% (33 of 41 patients). The stratification of patients by tobacco smoking status, histologic type of cancer, and TNM stage of cancer revealed no marked effects. In the two additional tests, the model predicted MLC with a sensitivity of 66.7% (10 of 15 patients), and it classified the cancer-negative bronchoscopy patients with a specificity of 37.4% (34 of 91 patients).**
Conclusions: **Compared to healthy volunteers, patients with PLC had abnormal breath test findings that were consistent with the accelerated catabolism of alkanes and monomethylated alkanes. A predictive model employing nine of these VOCs exhibited sufficient sensitivity and specificity to be considered as a screen for lung cancer in a high-risk population such as adult smokers.**

(CHEST 2003; 123:2115–2123)

Key words: breath test; detection; lung cancer; volatile organic compounds

Abbreviations: BMAC = breath-methylated alkane contour; CYP = cytochrome P450; MLC = metastatic lung cancer; NPV = negative predictive value; PLC = primary lung cancer; PPV = positive predictive value; VOC = volatile organic compound

Primary carcinoma of the lung is the leading cause of cancer death in both men and women in the United States.[1] A total of 99,000 men and 78,000 women are affected every year, and 86% of them are dead within 5 years of diagnosis. However, with early detection and treatment, the 5-year survival rate improves dramatically from 20% in patients with stage 3 lung cancer to 70% in patients with stage 1 disease. Researchers therefore have sought screening tests to detect lung cancer in its earliest stages,

For editorial comment see page 1788

and several promising new early markers have been proposed, including computer-assisted image analysis of chest radiographs, polymerase chain reaction-

based assays of sputum, fluorescence bronchoscopy, and spiral CT scanning.[2–4]

In 1971, Pauling et al[5] reported a new method for the microanalysis of breath that revealed the presence of large numbers of previously undetected volatile organic compounds (VOCs) in normal human breath. It is now known that a sample of breath contains, on average, approximately 200 different VOCs, mostly in picomolar (*ie*, 10^{-12} mol/L) concentrations.[6] We and others[7–10] have identified apparent new markers of lung cancer among these VOCs, which are predominantly alkanes and methylated alkanes. Until recently, these were empirical observations that could not be readily explained by known pathophysiologic processes. However, a plausible explanation has now emerged from an improved understanding of the mechanisms and kinetics of VOC synthesis and clearance. Alkanes and methylated alkanes in the breath are apparent markers of oxidative stress, which are the toxic effects of reactive oxygen species comprising oxygen free radicals and hydrogen peroxide. Reactive oxygen species are constantly produced in the mitochondria and leak into the cytoplasm where they cause peroxidative damage to proteins, polyunsaturated fatty acids, and DNA.[11,12] Peroxidative changes to DNA bases may be carcinogenic,[13,14] and oxidative stress appears to be increased in some cancers,[15] although the evidence for a causal role is lacking. Lipid peroxidation of polyunsaturated fatty acids in cell membranes generates alkanes such as ethane and pentane, which are excreted in the breath,[16] and breath methylalkanes may be products of the same process.[17] Alkanes are metabolized to alkyl alcohols by cytochrome P450 (CYP)-mixed oxidase enzymes,[18] and a number of studies[19–22] have demonstrated that these enzymes are activated in lung cancer. Polyaromatic hydrocarbons in tobacco smoke induce the activity of CYP 1A1 in the lung and placenta, and CYP 1A2 in the liver, resulting in the accelerated metabolism of a number of drugs and the activation of some procarcinogens.[23] These findings provide a rational basis for a breath test for lung cancer. The activation of CYP enzymes in patients with lung cancer may accelerate the degradation of volatile alkanes and monomethylated alkanes that are produced by oxidative stress and result in measurable changes in the composition of the breath.

We have recently reported tests for the set of C4 to C20 alkanes and their monomethylated derivatives in the breath, which appear to vary with the amount of oxidative stress.[17,24] These breath VOCs were significantly more abundant in older than in younger healthy humans, a finding that is consistent with previous reports[25,26] that aging is accompanied by increased oxidative stress. We report here an evaluation of this breath test as a marker of disease in patients with lung cancer.

Materials and Methods

Human Subjects

This investigation involved the following four study groups: (1) patients with primary lung cancer (PLC); (2) patients with metastatic lung cancer (MLC); (3) patients with no histologic evidence of lung cancer; and (4) healthy volunteers. Patients in the first three study groups were classified based on bronchoscopy and biopsy findings following an abnormal chest radiograph finding. The healthy volunteers were recruited in Staten Island, NY, from members of the general population who had no history of cancer or any other chronic disease.[6] A subgroup of healthy volunteers who were $\geq$ 55 years of age was selected to serve as a control group for the patients with PLC, all of whom were $\geq$ 55 years of age (Fig 1). A tobacco smoking history was obtained from all subjects, and *ex-smokers* were defined as having been abstinent from the use of tobacco products for at least 2 years.

Patients were eligible to participate if they were $\geq$ 18 years of age, could understand the breath collection procedure, could give written informed consent, and did not have a history of previously diagnosed cancer at any site. One hundred seventy-eight patients who had been referred to a pulmonary physician for further evaluation of an abnormal chest radiograph were recruited from the pulmonary services of the following five academic medical centers: Charing Cross Hospital, Imperial College, London, UK (75 subjects); Columbia Presbyterian Medical Center, New York, NY (19 subjects); New York University Medical Center, New York, NY (5 subjects); Penn State Medical Center, Hershey, PA (69 subjects); and St. Vincent's Medical Center, New York, NY (10 subjects). The 41 healthy volunteers, who were recruited from Staten Island, NY, had no known history of cancer. The research was approved by the institutional review boards of all participating institutions.

Study Design

This study included the following two components: a case-control study and a cross-sectional study. The case-control study

*From Menssana Research Inc (Dr. Phillips, Ms. Cataneo, and Mr. Greenberg), Fort Lee, NJ; the Department of Medicine (Dr. Gagliardi), New York Medical College, Valhalla, NY; the National Heart & Lung Institute (Dr. Cummin), Imperial College School of Medicine, London, UK; Pulmonary/Critical Care Division (Dr. Gleeson), Department of Medicine, M.S. Hershey Medical Center of the Pennsylvania State University, School of Medicine, Hershey, PA; the Department of Medicine (Dr. Maxfield), Division of Pulmonary, Allergy, and Critical Care Medicine, Columbia University College of Physicians & Surgeons, New York, NY; and the Division of Pulmonary and Critical Care Medicine (Dr. Rom), New York University Medical Center, New York, NY.

This research was supported by SBIR grant 1R43 CA77098-01 from the National Institutes of Health. Dr. Rom was supported by National Institutes of Health grants M01 00096, EDRN U01 CA 86137, and UO1CA8617. Dr. Phillips is President and Chief Executive Officer of Menssana Research, Inc.

Manuscript received January 25, 2002; revision accepted October 7, 2002.

Correspondence to: Michael Phillips, MD, Menssana Research, Inc, 1 Horizon Rd, Suite 1415, Fort Lee, NJ 07024

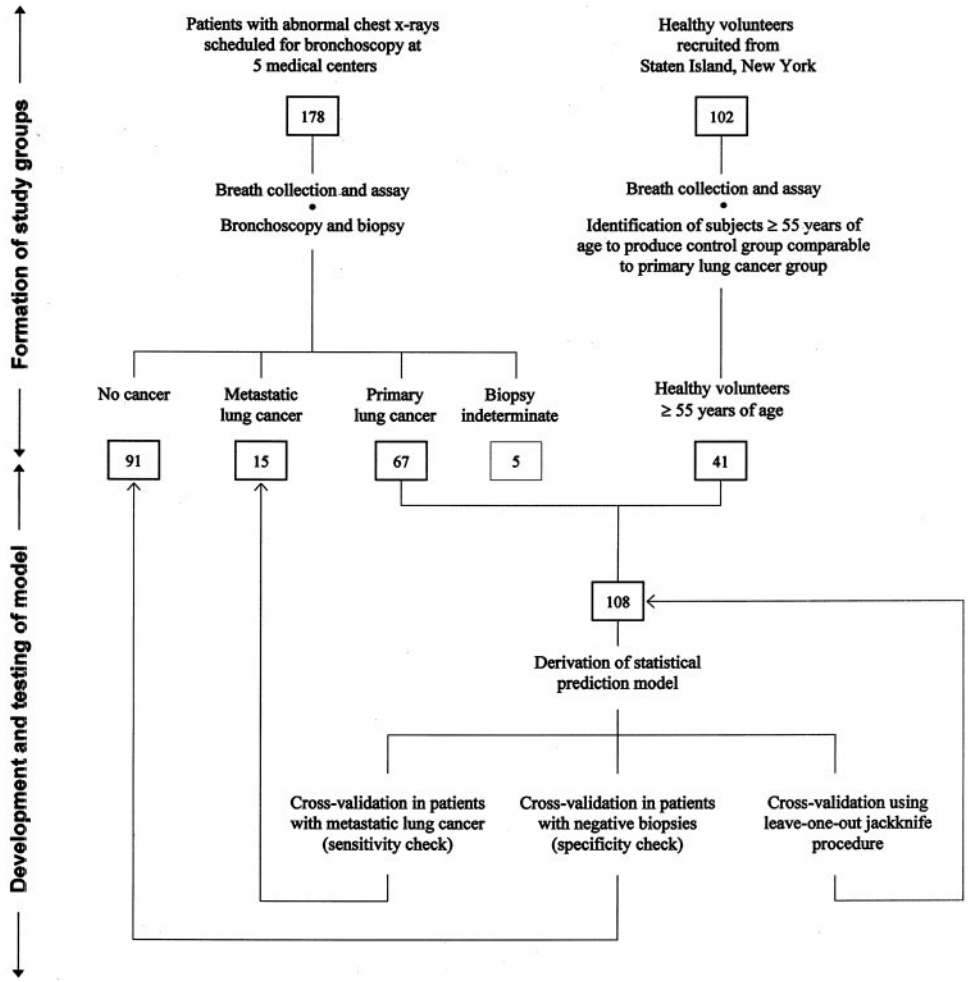


FIGURE 1. Study design flow diagram. Derivation of the four study groups used in the derivation and cross-validation of the statistical model are shown.

involved patients with PLC and healthy volunteers in the development and initial cross-validation of a statistical prediction model. The cross-sectional study involved two groups of bronchoscopy patients in additional, independent cross-validations (*ie*, patients with MLC and patients with negative biopsy findings).

Detection of Lung Cancer

Bronchoscopy was performed according to standard procedures.[27] Intraluminal lesions were lavaged or brushed for cytology and were biopsied directly using a standard alligator forceps. Parenchymal lesions were evaluated by the lavage of the appropriate airway segment and by transbronchial biopsy under direct fluoroscopic guidance. Lung biopsy specimens were preserved in formalin for microscopic examination by a pathologist. In patients with nondiagnostic findings at bronchoscopy, additional investigations (including serial CT imaging of the chest, needle biopsy, or surgical biopsy) were performed until the diagnosis of cancer was either established or excluded. The clinical stage of the disease was determined according to the International TNM staging system for lung cancer. MLC was defined as cancer from a nonpulmonary primary malignancy metastatic to the lung. Pathologists had no knowledge of the breath test results when the biopsy specimens were examined. The final diagnosis employed for data analysis, of cancer vs no cancer in the bronchoscopy group, was based on the reported histopathology of the biopsy specimens that had been obtained either during bronchoscopy or during subsequent biopsy procedures. Healthy control subjects were assumed to be free of cancer.

Breath Collection and Assay

A portable breath collection apparatus was employed to capture the VOCs in 1.0 L breath on a sorbent trap. The VOCs in 1.0 L room air then were captured on a separate sorbent trap (Fig 2). Subjects wore a nose clip while breathing in and out of the disposable mouthpiece of the apparatus for 2.0 min. Light flap valves in the mouthpiece presented low resistance to respiration, so that breath samples could be collected without discomfort to patients, including those who were elderly and/or had pulmonary disease. Breath samples from bronchoscopy patients were collected prior to the procedure, on the same day. All sorbent traps were sent to the laboratory for an analysis of the VOCs by automated thermal desorption, gas chromatography, and mass spectroscopy. This method has been described in detail elsewhere.[6,28] Laboratory personnel (RNC and JG) had no

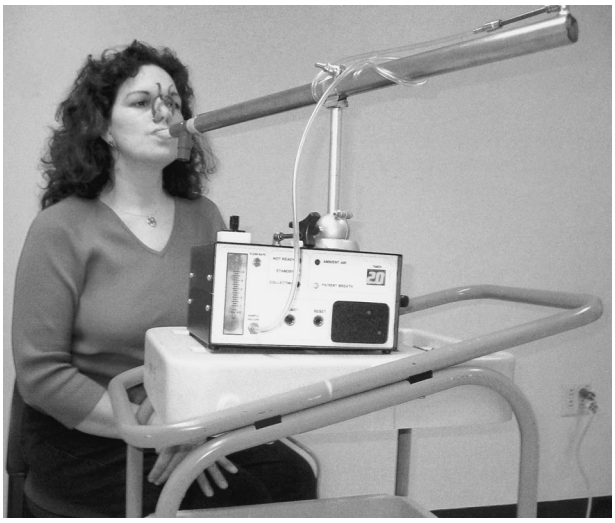


FIGURE 2. The breath collection apparatus is shown in use. The subject wears a nose clip and breathes quietly in and out through a disposable mouthpiece. The long tube is the breath reservoir, and the small tube affixed to its end is the sorbent trap, which captures the VOCs in the breath. The front panel of the breath collection apparatus shows the flowmeter on the left and a digital timer on the top right.

knowledge of the clinical or pathologic findings when the assay was performed. All subjects were able to contribute a breath sample without any adverse effects.

The physiologic interpretation of breath VOC chromatogram data requires the concept of the alveolar gradient.[6,29] For a given breath VOC, Vb denotes the area under the curve that is associated with the chromatogram peak for that VOC, and Ib denotes the analogous area associated with the internal standard used to calibrate the instrument (0.25 mL of 2 ppm 1-bromo-4-fluoro-benzene; Supelco; Bellefonte, PA). Va and Ia denote corresponding areas derived from the associated air sample. The alveolar gradient of the VOC is then determined as:

$$\text{alveolar gradient} = Vb/Ib - Va/Ia.$$

The two terms give the relative abundance of the VOC in breath and air, respectively. For each study subject, the above formula was used to compute the alveolar gradients of each of the C4 to C20 n-alkanes and their monomethylates. The mean alveolar gradients of these VOCs then were computed for the four study groups, and the results were displayed in a series of surface plots. These plots show the carbon chain length on the *x*-axis, the methylation site on the *z*-axis, and the mean alveolar gradient on the *y*-axis.

Statistical Analysis

Forward stepwise discriminant analysis[30] was used to identify the combination of VOCs that provided the best discrimination between patients with PLC and healthy volunteers. This multivariable technique produces a predictive model (or equation) that estimates the probability of disease for each study subject, predicting lung cancer in subjects having an estimated probability of disease of > 0.5.[31]

The accuracy of this model was first tested by cross-validation using a leave-one-out jackknife technique, in which each subject was classified using an equation derived from all other subjects.[32] This technique was used because the sample size was insufficient to divide the sample into a training set for model development and a validation set for testing that model. The model was further tested in the following two groups of patients that were not used in the derivation of the model: patients with MLC (*ie*, those with primary cancer of nonlung sites metastatic to lung); and bronchoscopy patients with negative biopsy findings.

RESULTS

Table 1 summarizes the demographics of the study subjects. The breath test identified 80 different C4 to C20 alkanes and monomethylated alkanes that had been either synthesized or catabolized by at least one subject (Fig. 3). Forward stepwise discriminant analysis identified nine of these VOCs as the best set of markers of disease (Table 2 and Fig 4, 5). This combination yielded a sensitivity of 89.6% (60 of 67 patients) and a specificity of 82.9% (34 of 41 patients) when a 0.5 probability of disease was used as the dividing point between a positive and a negative breath test.

Cross-validation using the leave-one-out jackknife procedure (and keeping the 0.5 probability of disease as the threshold) yielded a sensitivity of 85.1% (57 of 67 patients) and a specificity of 80.5% (33 of 41 patients). In patients with MLC, the sensitivity of the model was 66.7% (10 of 15 patients). Bronchoscopy patients with negative biopsy findings were classified as being *cancer-negative* with a specificity of 37.4%

Table 1—*Demographics of Study Subjects

Variables	No.	Age, yr Mean (SD)	Male No. (%)	Female No. (%)
Bronchoscopy negative for cancer	91	58.4 (14.2)	41 (45.1)	50 (54.9)
Primary lung cancer	67†	68.2 (9.9)	48 (71.6)	19 (28.4)
Metastatic lung cancer	15	66.6 (9.2)	4 (26.7)	11 (73.3)
Lung cancer, undetermined	5	63.0 (28.3)	3 (60.0)	2 (40.0)
Healthy volunteers	41	69.6 (12.6)	16 (39.0)	25 (61.0)

*Lung cancer was classified as *undetermined* when it was not possible to determine with certainty whether it was metastatic to the lung or it had arisen as a lung primary (*eg*, an adenocarcinoma of indeterminate primary origin).
†Ten small cell and 57 non-small cell cancers.

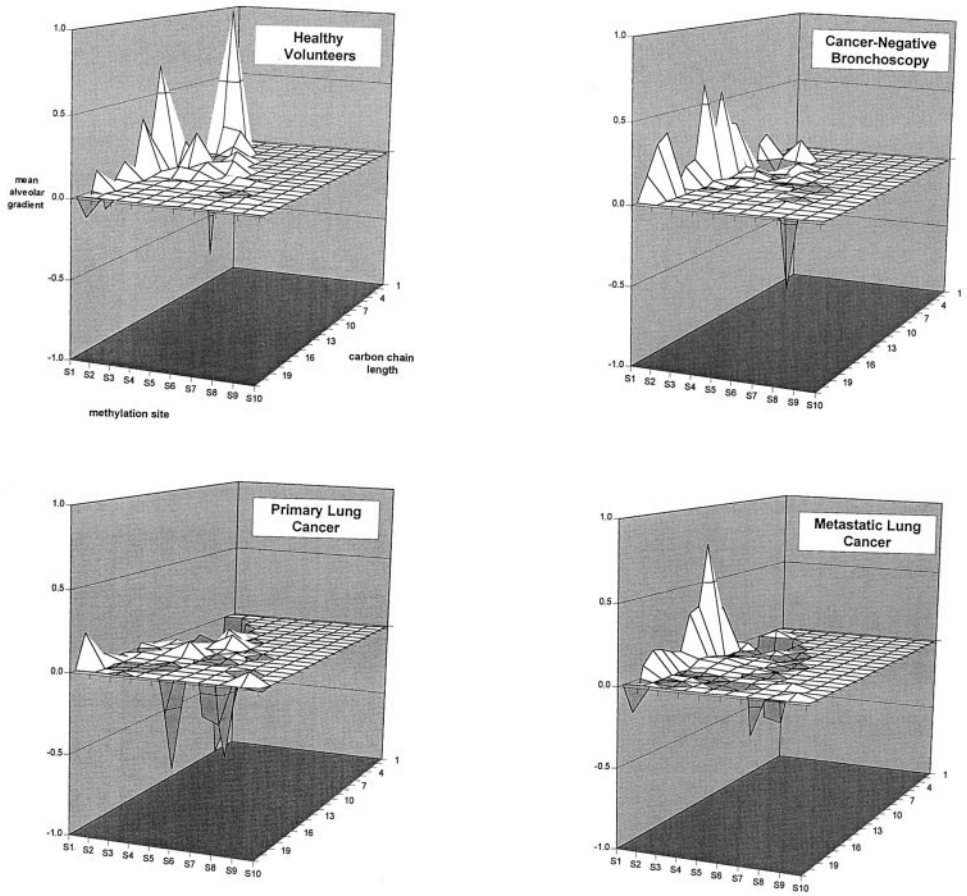


FIGURE 3. Surface plots of breath test results. The mean alveolar gradient is shown on the vertical axis for C4 to C20 alkanes and their monomethylated derivatives. The horizontal axes identify the specific VOC (*eg*, the combination of a carbon chain length of 4 and a methylation site of S2 corresponding to 2-methylbutane). Several of the mean alveolar gradients appeared to be reduced in patients with lung cancer compared to age-matched healthy volunteers and to patients with negative bronchoscopy findings. The hypothetical mechanism accounting for these differences is the accelerated clearance of alkanes and methylalkanes by induced CYP polymorphs in patients with lung cancer.

(34 of 91 patients). The results of cross-validation in healthy volunteers and patients with PLC were stratified by smoking status, histology, and TNM staging (Table 3). In smokers and ex-smokers, the sensitivity was 85.9% (55 of 64 patients) and the specificity was 82.6% (19 of 23 patients), indicating that tobacco smoking does not adversely affect the accuracy of the breath test. Stratification by histologic type and TNM staging of the cancer revealed no marked effects, suggesting that the model's performance is consistent across these strata. The sample size is insufficient, however, to reach definitive conclusions for individual subgroups.

Table 2—*VOCs Used to Identify Patients with Lung Cancer*

Order of Entry into Model
Butane*
Tridecane, 3-methyl
Tridecane, 7-methyl
Octane, 4-methyl
Hexane, 3-methyl
Heptane
Hexane, 2-methyl
Pentane
Decane, 5-methyl

*Best single discriminator.

DISCUSSION

This study demonstrated two main findings. First, a predictive model employing nine breath VOCs was sensitive and specific for lung cancer. Second, the statistical characteristics of the breath test showed only minor changes when subjects were stratified according to a history of tobacco smoking, the histologic type of the cancer, or the TNM stage of the cancer.

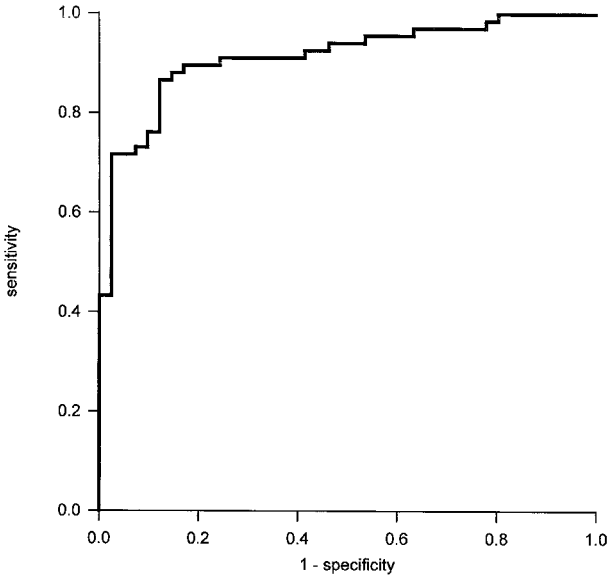


FIGURE 4. Detection of PLC with the breath test: receiver operating characteristic curve. The BMACs in patients with PLC were compared to the BMACs of age-matched healthy control subjects. Using discriminant analysis, the probability of lung cancer based on each subject's breath test result was determined as a value between 0 and 1. This ROC curve shows the results obtained with a model using nine VOCs to discriminate between patients with PLC (n = 67) and healthy volunteers (n = 41). Using a cutoff point of p = 0.5, the sensitivity was 89.6% (60 of 67 patients) and the specificity was 82.9% (34 of 41 patients). The contour of a receiver operating characteristic curve indicates the overall accuracy of a diagnostic test.

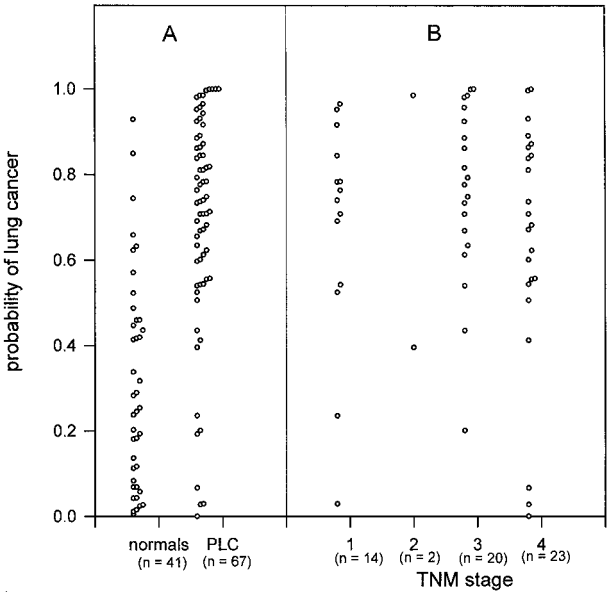


FIGURE 5. Predictions of the discriminant model in patients with PLC and healthy volunteers. The scatter diagram indicates the probability of lung cancer according to the cross-validation of the discriminant model employing the nine breath VOCs. *Left*, *A*: the probability for each of the patients with PLC and the healthy volunteers is shown. *Right*, *B*: the PLC data were stratified according to TNM staging in the 59 patients for whom staging data were available.

The mean alveolar gradients of the alkanes and monomethylated alkanes observed in the breath were predominantly negative in patients with PLC and predominantly positive in the age-matched healthy volunteers (Fig 3, *left panels*), which is consistent with the findings of previous reports of increased activity of polymorphic CYP in patients with lung cancer. The alveolar gradient of a VOC varies with its rate of synthesis minus its rate of clearance,[6] so that the observed differences in PLC were consistent either with a decreased rate of synthesis (caused by an unknown mechanism) or with an increased rate of clearance (caused by the induction of mixed oxidase enzymes). CYP activation is the basis of an emerging hypothesis for the etiology of lung cancer, which is based on an interaction between inborn risk (a genotype containing several polymorphous CYP enzymes) and environmental toxins (including components of tobacco smoke).

The predictive model for PLC yielded superior specificity when the control group comprised healthy volunteers instead of cancer-negative bronchoscopy patients with abnormal chest radiograph findings. There are at least two possible reasons for this finding. First, not all members of the cancer-free group may have been truly free of cancer. Some may have harbored small foci of cancer that were too small for detection. Bronchoscopy and biopsy are seldom diagnostic in lesions that are < 1.1 cm in diameter, and false-negative findings are not uncommon, particularly in lesions that are < 2.0 cm in diameter.[33–35] Second, some patients with abnormal chest radiograph findings may have harbored a precancerous lesion that induced changes in CYP me-

Table 3—*Effect of Smoking Status, Histology, and TNM Staging on Performance of the Breath Test

Variables	Sensitivity	Specificity
Smokers and ex-smokers	85.9 (55/64)	82.6 (19/23)
Nonsmokers	66.7 (2/3)	77.8 (14/18)
Total	85.1 (57/67)	80.5 (33/41)
Non-small cell cancer	87.8 (50/57)	
Small cell cancer	70.0 (7/10)	
Total	85.1 (57/67)	
Stage 1	85.7 (12/14)	
Stage 2	50.0 (1/2)	
Stage 3	90.0 (18/20)	
Stage 4	82.6 (19/23)	
Total†	84.7 (50/59)	

*Values given as % (No. of patients in whom cancer was detected/total No. of patients).
†Staging data were available for 59 of the 67 cancer patients.

tabolism. These possibilities could be evaluated in future studies by employing more advanced tumor imaging with spiral chest CT scanning, and by longitudinal observation of these patients for the subsequent development of lung cancer.

The predictions of the breath test in patients with PLC were not affected by the stage of the tumor (Fig 5). This finding was similar to our previous observations and is consistent with the postulated mechanism (*ie*, the results of the breath test are altered by the induction of CYP polymorphs, not by the mass of the tumor itself). Since enzyme induction is associated with the earliest stages of carcinogenesis, the breath test may provide a rational method for the detection of lung cancer in its earliest stages.

The sensitivity and specificity of the breath test were not significantly affected by tobacco smoking (Table 3). The reason for this is unclear since a history of smoking should predispose the patient to the induction of polymorphic CYP activity and should accelerate the clearance of breath markers of oxidative stress. However, the absence of an observed effect of tobacco smoking possibly may have resulted from wide interindividual variations in susceptibility to the induction of these enzymes. Also, we may anticipate the existence of a subset of smokers who have developed the phenotype of induced CYP activity but who have not yet progressed to detectable lung cancer. This study identified a number of subjects whose breath test results were identified as being false-positives because the results were consistent with induced CYP activity, but imaging studies and bronchoscopy did not reveal a tumor. This group may be at increased risk of the future development of lung cancer, but it would require a long-term prospective study to test this hypothesis.

Cross-validation of the statistical model based on nine breath VOCs revealed that the model would be expected to predict PLC with a sensitivity of 85.1% and a specificity of 80.5%. Based on these findings, it is possible to estimate the potential value of the breath test as a primary screen for lung cancer in apparently healthy subjects. A good screening test must produce very few false-negative results without producing too many false-positive results. That is, a screening test should exhibit a very high negative predictive value (NPV) and a reasonable positive predictive value (PPV), where NPV is the percentage of subjects testing negative who do not have lung cancer, and PPV is the percentage of subjects testing positive who have lung cancer.

One way to define *high* and *reasonable* is to examine the test characteristics of other commonly used screening tests, such as mammography, for detecting breast cancer. A 1993 mammography study[36] involving 31,000 women reported a PPV of 9% when the test was used to screen women aged 50 to 59 years, and a 1997 mammography study[37] involving 1,007 women aged 14 to 82 years reported a NPV of 98%. These figures suggest that a cancer screening test with comparable statistical properties (*ie*, PPV, 9%; NPV, 98%) would be considered clinically useful. Table 4 shows the expected outcome of screening 1,000 asymptomatic smokers who were ≥ 60 years of age with the breath test. Of those smokers, 27 will have previously undetected lung cancer, based on the findings of Henschke et al[38] who used chest CT scanning to screen a similar group of subjects. The PPV and NPV of the breath test were 10.8% and 99.5%, respectively, demonstrating that a breath test employed as a primary screen for lung cancer could potentially exhibit greater accuracy than a mammogram employed as a primary screen for breast cancer.

It may be asked, what is the potential role of the breath test in a clinical setting in which highly sensitive and specific techniques for imaging lung cancer, such as spiral CT scanning, are already available? The answer lies in seeing these different technologies as being complementary rather than competitive, each occupying its own niche of safety, cost, and efficacy. High-risk patients could potentially be screened for lung cancer in a rational and cost-effective program, as follows: primary screening with the breath test (low cost); if the results are positive, go to secondary screening with spiral CT scan of the chest (intermediate cost); if that has a positive result, go to final testing with bronchoscopy and biopsy (high cost). In summary, what the breath test potentially may add to existing technology is a safe, noninvasive, sensitive, and specific tool for the primary screening for lung cancer in high-risk populations at a comparatively low cost. We have found in practice that hospital and office staff members can be trained rapidly to collect technically satisfactory breath VOC samples. It would be straightforward to

Table 4—*Lung Cancer Screening With the Breath Test

Test Result	Lung Cancer	No Cancer	Total
Positive test result	23 TP	190 FP	213
Negative test result	4 FN	783 TN	787
Total	27	973	1,000

*TP = true positive; FP = false positive; FN = false negative; TN = true negative. Sensitivity = TP/(TP + FN) = 85.1% (23/27); specificity = TN/(TN + FP) = 80.5% (783/973); PPV = TP/(TP + FP) = 10.8% (23/213); NPV = TN/(TN + FN) = 99.5% (783/787).

implement a breath test screening program for at-risk populations because each site would require only a portable breath collection apparatus. Samples are sent to a laboratory for assay with standard instruments that are widely available.

Breath testing has a number of advantages over other proposed methods for the early detection of lung cancer. It is noninvasive, intrinsically safe, comparatively inexpensive, and highly acceptable to patients. Another advantage derives from the hypothetical mechanism, as follows: changes in the breath test should accompany CYP activation that has progressed sufficiently to convert procarcinogens to carcinogens. Hence, changes in the breath test could be observed while the cancer was still in its earliest and most treatable stages. Potentially, this could translate into a reduction in mortality from lung cancer, but confirmation would require another clinical study to determine prospectively the effects of breath test screening on intervention and survival.

We conclude that a breath test for C4 to C20 alkanes and monomethylated alkanes provided a rational new set of markers that identified lung cancer in a group of patients with histologically proven disease. However, this study was limited by the following three main factors: the limited range of presenting disorders among the patients; the comparatively small number of patients with lung cancer; and the comparatively large number of variables in the breath-methylated alkane contour. The latter two factors necessitated cross-validation with a leave-one-out jackknife technique. A preferable cross-validation procedure is to randomly allocate patients to the following two groups: a training set, to derive the statistical model; and a test set, to validate the model. However, this test of cross-validation requires a larger number of patients than were available for this study. Further studies are required to validate the breath test more stringently in a greater number of patients with lung cancer who present with a more diverse range of disorders. Connelly and Inui[39] have identified the criteria for a screening test, which depend both on the disease and the method of screening. The disease must be sufficiently burdensome to the population that a screening program is warranted, the disease must have a long preclinical latent period, and efficacious treatment must be available. The screening method must have acceptable technical performance parameters and must detect the disease at an earlier stage than would be possible without screening, while minimizing false-positive and false-negative results. In addition, early detection must improve disease outcome, and the cost, feasibility, and acceptability of screening and early treatment should be established. The breath test is comparatively low in cost, technically feasible, and acceptable to patients. Based on this study, it appears likely that it could provide earlier detection of lung cancer with an acceptably low rate of false-positive and false-negative results. However, it is not yet known whether it can improve disease outcome. Further studies are needed to confirm these findings and to evaluate the potential value of breath testing in screening for early lung cancer.

ACKNOWLEDGMENTS: We thank Dr. Eugen Sersen for statistical consultation and Dr. Anton Sinniah for assisting with the collection of breath samples.

References

1 Minna JD. Neoplasms of the lung. In: Fauci AS, Braunwald E, Isselbacher KJ, et al, eds. Harrison's principles of internal medicine. New York, NY: McGraw Hill, 1998; 552–562
2 Lam S, Lam B, Petty TL. Early detection for lung cancer: new tools for casefinding. Can Fam Physician 2001; 47:537–544
3 Mulshine JL, Scott F. Molecular markers in early cancer detection: new screening tools. Chest 1995; 107(suppl):280S–286S
4 Henschke CI, Yankelevitz DF. CT screening for lung cancer. Radiol Clin North Am 2000; 38:487–495
5 Pauling L, Robinson AB, Teranishi R, et al. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. Proc Natl Acad Sci U S A 1971; 68:2374–2376
6 Phillips M, Herrera J, Krishnan S, et al. Variation in volatile organic compounds in the breath of normal humans. J Chromatogr B Biomed Sci Appl 1999; 729:75–88
7 Phillips M, Gleeson K, Hughes JM, et al. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. Lancet 1999; 353:1930–1933
8 Gordon SM, Szidon JP, Krotoszynski BK, et al. Volatile organic compounds in exhaled air from patients with lung cancer. Clin Chem 1985; 31:1278–1282
9 Preti G, Labows JN, Kostelc JG, et al. Analysis of lung air from patients with bronchogenic carcinoma and controls using gas chromatography-mass spectrometry. J Chromatogr 1988; 432:1–11
10 O'Neill HJ, Gordon SM, O'Neill MH, et al. A computerized classification technique for screening for the presence of breath biomarkers in lung cancer. Clin Chem 1988; 34:1613–1618
11 Kneepkens CM, Ferreira C, Lepage G, et al. The hydrocarbon breath test in the study of lipid peroxidation: principles and practice. Clin Invest Med 1992; 15:163–186
12 Kneepkens CM, Lepage G, Roy CC. The potential of the hydrocarbon breath test as a measure of lipid peroxidation. Free Radic Biol Med 1994; 17:127–160
13 Loft S, Poulsen HE. Cancer risk and oxidative DNA damage in man. J Mol Med 1996; 74:297–312
14 Ray G, Batra S, Shukla NK, et al. Lipid peroxidation, free radical production and antioxidant status in breast cancer. Breast Cancer Res Treat 2000; 59:163–170
15 Hietanen E, Bartsch H, Bereziat JC, et al. Diet and oxidative stress in breast, colon and prostate cancer patients: a case-control study. Eur J Clin Nutr 1994; 48:575–586
16 Risby TH, Sehnert SS. Clinical application of breath biomar-

kers of oxidative stress status. Free Radic Biol Med 1999; 27:1182–1192
17 Phillips M, Cataneo RN, Greenberg J, et al. Effect of age on the breath methylated alkane contour, a display of apparent new markers of oxidative stress. J Lab Clin Med 2000; 136:243–249
18 Remmer H, Hintze T, Frank H, et al. Cytochrome P-450 oxidation of alkanes originating as scission products during lipid peroxidation. Xenobiotica 1984; 14:207–219
19 Watanabe M. Polymorphic CYP genes and disease predisposition–what have the studies shown so far? Toxicol Lett 1998; 102–103:167–171
20 Hecht SS. Approaches to cancer prevention based on an understanding of N-nitrosamine carcinogenesis. Proc Soc Exp Biol Med 1997; 216:181–191
21 Benhamou S, Bonaiti-Pellie C. Susceptibility to bronchial cancer: an example of genetic-environmental interaction. Ann Biol Clin (Paris) 1995; 53:507–513
22 Anderson LM, Beebe LE, Fox SD, et al. Promotion of mouse lung tumors by bioaccumulated polychlorinated aromatic hydrocarbons. Exp Lung Res 1991; 17:455–471
23 Zevin S, Benowitz NL. Drug interactions with tobacco smoking: an update Clin Pharmacokinet 1999; 36:425–438
24 Phillips M, Greenberg J, Cataneo RN. Effect of age on the profile of alkanes in normal human breath. Free Radic Res 2000; 33:57–63
25 Zarling EJ, Mobarhan S, Bowen P, et al. Pulmonary pentane excretion increases with age in healthy subjects. Mech Ageing Dev 1993; 67:141–147
26 Jones M, Shiel N, Summan M, et al. Application of breath pentane analysis to monitor age-related change in free radical activity [abstract]. Biochem Soc Trans 1993; 21:485S
27 Prakash UB. Advances in bronchoscopic procedures. Chest 1999; 116:1403–1408
28 Phillips M. Method for the collection and assay of volatile organic compounds in breath. Anal Biochem 1997; 247:272–278
29 Phillips M, Greenberg J, Sabas M. Alveolar gradient of pentane in normal human breath. Free Radic Res 1994; 20:333–337
30 SPSS, Inc. SPSS for Windows, standard version. Cary, NC: SPSS Inc, 1998
31 Jennrich R. Stepwise discriminant analysis. In: Enslein K, Ralston A, Wilf H, eds. Statistical methods for digital computers. New York, NY: Wiley, 1977
32 Efron B. The jackknife, the bootstrap, and other resampling plans. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1982
33 Mori K, Yanase N, Kaneko M, et al. Diagnosis of peripheral lung cancer in cases of tumors 2 cm or less in size. Chest 1989; 95:304–308
34 Torrington KG, Poropatich RK. Utility of repeated fiberoptic bronchoscopy for suspected malignancy. Chest 1992; 102: 1080–1084
35 Collins CD, Breatnach E, Nath PH. Percutaneous needle biopsy of lung nodules following failed bronchoscopic biopsy. Eur J Radiol 1992; 15:49–53
36 Kerlikowske K, Grady D, Barclay J, et al. Positive predictive value of screening mammography by age and family history of breast cancer. JAMA 1993; 270:2444–2450
37 Hindle WH, Chen EC. Accuracy of mammographic appearances after breast fine-needle aspiration. Am J Obstet Gynecol 1997; 176:1286–1290
38 Henschke CI, McCauley DI, Yankelevitz DF, et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. Lancet 1999; 354:99–105
39 Connelly MT, Inui TS. Principles of disease prevention. In: Fauci AS, Braunwald E, Isselbacher KJ, et al, eds. Harrison's principles of internal medicine. New York, NY: McGraw-Hill, 1998; 46–48