# High Sensitivity for Lung Cancer Detection by Analysis of Exhaled Carbonyl Compounds

Erin M. Schumer MPH, MD[1], Jaimin R. Trivedi, MPH, MD[1], Victor van Berkel MD[1], PhD, Matthew C. Black, MD[1], Mingxiao Li PhD[2], Xiao-An Fu PhD[2], Michael Bousamra II MD[1]

[1]Department of Cardiovascular and Thoracic Surgery

[2]Department of Chemical Engineering

University of Louisville School of Medicine, Abell Administration Center

323 E. Chestnut St., Louisville, Kentucky 40292

This work was supported by the Coulter Foundation, NSF, and Gates Foundation.

Drs. Bousamra, van Berkel, and Fu are founders of Breath Diagnostics Inc. All other authors declare no conflict of interest .

**Corresponding Author:** Michael Bousamra II, MD

Department of Cardiovascular and Thoracic Surgery, University of Louisville

201 Abraham Flexner Way, Suite 1200, Louisville, Kentucky 40202

Tel: 502-588-7601

Fax: 502-588-7701

E-mail: michael.bousamra@ulp.org

**Article Word Count: 2719**

**Abbreviations**

**2-HA** = 2-hydroxyhexanal

**4-HHE** = 4-hydroxyhexanal

**ATM** = 2-(aminooxy)-N, N, N trimethylethanammonium

**CT** = computed tomography

**COPD** = chronic obstructive pulmonary disease

**NLST** = National Lung Screening Trial

## Abstract

**Objective:** Several volatile carbonyl compounds in exhaled breath have been identified as cancer specific markers. The potential for these markers to serve as a screening test for lung cancer is reported.

**Methods:** Patients with CT-detected intra-thoracic lesions and healthy controls were enrolled from 2011 onwards. One liter of breath was collected from a single exhalation. The contents were evacuated over a silicon microchip, captured by oximation reaction, and analyzed by mass spectrometry. Concentrations of 2-butanone, 3-hydroxy-2-butanone, 2-hydroxyacetaldehyde, and 4-hydroxyhexanal were measured. The overall population was divided into lung cancer, benign disease, and control groups. An elevated cancer marker was defined as ≥1.5 standard deviations above the mean concentration of the control population. One or more elevated cancer markers constituted a positive breath test.

**Results:** There were 156 lung cancer, 65 benign disease, and 194 control subjects. A total of 103 (66.0%) lung cancer patients were early stage (Stage 0, I, and II). For ≥1 elevated cancer marker, breath analysis showed a sensitivity of 93.6% and a specificity of 85.6% for lung cancer patients. Additionally, 83.7% of stage I tumors ≤2 cm were detected while only 14% of the control population tested positive. When comparing cancer to benign disease, specificity was proportional to the number of elevated cancer markers present.

**Conclusion:** Screening using low-dose CT scan is associated with high cost, repeated radiation exposure, and low accrual. The high sensitivity, convenience,

and low cost of breath analysis for carbonyl cancer markers suggests it has the potential to become a primary screening modality for lung cancer.

**Word count of the abstract: 250**

## Perspective

Several volatile carbonyl compounds in exhaled breath have been identified as cancer specific markers. The high sensitivity of this clinical trial using carbonyl compounds to detect cancer suggests it has the potential to become a primary screening modality for lung cancer.

## Central Picture

The collection bag, microchip, and evacuation system are shown here.

## Central Message

Analysis of exhaled breath carbonyls has a high sensitivity for detecting lung cancer.

## Introduction

Lung cancer remains the leading cause of cancer death of both men and women in the United States, and the second most common cancer in each of these demographics, representing 13 % of all new cancer diagnoses[1]. It is an aggressive disease with a 5-year survival rate of 17.8%. Despite evolution in surgery, chemotherapy, and radiotherapy, long-term survival has remained poor compared to many other common cancers[2]. Improvements in survival of these cancer types can be partially attributed to effective screening modalities resulting in early detection[3,4]. However, screening is not uniformly effective and may lead to over diagnosis and increased morbidity, as has been observed with prostate cancer[5].

The National Lung Screening Trial (NLST) demonstrated a 20% reduction in lung cancer mortality and a 6.7% overall survival benefit for patients screened with CT scan[6]. However, CT scan resulted in a positive screening test in 24.2% of which 94% were falsely positive[7]. These false positives led to additional radiographic imaging or testing, sometimes invasive[8], with occasional major complications. Low dose CT scan detected a high proportion of early cancer (49% Stage IA); however 35% of detected cancers were advanced stage (III, IV). In turn, the number needed to screen to prevent one death from lung cancer was 340[6]. CT screening has recently been implemented in the United States and costs are anticipated to be in the billions of dollars annually[9]. Furthermore, the degree of

patient compliance is uncertain. These limitations imply that an alternative strategy for lung cancer detection would be welcome.

The analysis of exhaled breath is promising as a noninvasive diagnostic tool for detecting lung cancer. Previously, we have shown the ability of breath analysis to distinguish benign from malignant pulmonary disease[10,11]. In the current study, the cutoff level for an elevated cancer marker was set lower in order to capture more cases of lung cancer with the expectation that there would be more false positives. This study is an initial evaluation of the concept that breath analysis may eventually serve as an initial screening test for lung cancer to be followed by a CT scan in cases where the breath analysis is positive.

## Material and Methods

*Collection of Breath Samples*

The Institutional Review Board at the University of Louisville approved the research protocol for collection of exhaled breath samples. All study subjects signed informed consent before providing breath samples. One liter of breath was collected into a Tedlar ® bag (Sigma-Aldrich, St Louis, MO) from a single exhalation from each subject. Breath samples were collected from 415 patients. Lung cancer was confirmed by pathologic diagnosis in 149 (93.6%) patients. Benign pathology was confirmed by tissue diagnosis in 47 (72.3%) patients or by repeated CT scan with no change or decrease in size over at least two years.

*The silicon chip and mass spectrometry*

As previously described, the silicon microchips were fabricated from silicon wafers and consist of an array of micropillars[11]. A quaternary ammonium compound, 2-(aminooxy)-N, N, N trimethylethanammonium (ATM) iodide, was then used to coat the surfaces of the micropillars. ATM adsorbs to the silicon dioxide surfaces of the micropillars via electrostatic and hydrogen bond interactions. ATM selectively traps carbonyl compounds in exhaled breath by means of oximation reactions with capture efficiencies of ≥ 98%.

The procedure for capture of carbonyl compounds in air and exhaled breath has been previously described[10]. In brief, exhaled breath was collected in 1 Liter Tedlar bags and drawn through the silicon microreactor chip by applied vacuum (Figure 1). After this process, ATM iodide adducts in the microreactor chip were eluted with methanol from a slightly pressurized small vial. 99% of ATM adducts were recovered. The eluted solution was directly analyzed by Fourier transform-ion cyclotron resonance-mass spectrometry. A known amount of deuterated acetone completely reacted with ATM (ATM-acetone-d6) was added to the eluted solution as an internal reference. The concentrations of all carbonyl compounds in exhaled breath were determined by comparison of the relative abundance with that of added ATM-acetone-d6.

The analysis of the entire study population included the use of two different microchips with a different density of micropillars. Sub-analysis of these groups did not reveal a meaningful difference between the two microchips.

*Data Analysis*

We have previously described the initial identification of carbonyl cancer markers in breath samples of 88 healthy controls and confirmed lung cancer patients[10,11]. In this study, four carbonyl cancer markers, 2-butanone, 3-hydroxy-2-butanone, and 2-hydroxyacetadehyde, 4-hydroxyhexanal (4-HHE), out of a spectrum of 25 carbonyl compounds were again established as elevated in cancer patients and used for this analysis. Of note, heavier compounds, including 4-hydroxyhexanal, are fully evacuated within the last thirty minutes of complete evacuation, thus only cases using complete evacuation were included. Three cutoff values were used to determine a positive carbonyl marker: ≥1, ≥1.5, and ≥2 standard deviations about the mean of the control population. A positive breath test was defined as at least one positive carbonyl marker with a given patient having between zero and four elevated carbonyl cancer markers. Descriptive and univariate analyses were performed using SAS version 9 (SAS Institute, Cary, North Carolina). A p-value $< 0.05$ was considered significant.

## Results

Patients were divided into three groups: healthy controls (n=194), benign pulmonary disease (single or multiple pulmonary nodules and/or mediastinal

adenopathy, n=65), and primary lung cancer (n=156). Age, smoking history, and tumor size of the three groups are summarized in Table 1. Lung cancer patients were older and had a higher proportion of current and former smokers when compared to the benign disease and control groups. There was a higher concentration of 3-hydroxy-2-butanone among smokers compared to non-smokers in the control and benign pulmonary disease groups; however these levels remained below those of cancer patients independent of smoking status. All other cancer markers were similar between smokers and non-smokers in each group (Table 2).

The lung cancer group was then divided into early (0, I, II) and advanced stage (III, IV) disease. The median (interquartile range) concentrations for each cancer marker are summarized for each group in Table 3. There is a progressive and significant increase in median concentration of each carbonyl cancer marker from the control group to the benign group to patients with lung cancer.

Sensitivity and specificity for the lung cancer and control populations are shown in Table 4. Sensitivity has a strong inverse correlation to the number of elevated cancer markers used to define a positive cancer result. It is less dependent on the carbonyl marker concentration cutoff point. At ≥1 standard deviations above the mean of the control population with at least one elevated carbonyl marker, sensitivity is 95.5%; at ≥1.5 standard deviations above the mean, sensitivity is 93.6%; and finally, at ≥2 standard deviations above the mean, sensitivity is

88.5%. When three cancer markers are used at all cutoff levels, sensitivity decreases but specificity increases to 100%. Therefore, specificity is highly and positively correlated with the number of elevated cancer markers while sensitivity is highest at the lowest cutoff point with at least one elevated cancer marker.

To compare to the NLST, a sub-analysis was performed using control subjects who approximate the criteria for the NLST (n=52), that is those at least 55 years of age and who are current (n=22, 42.3%) or former smokers (n=30, 57.7%) (Table 5). The results were similar to the comparison with the entire control group. At ≥1 standard deviation above the mean of the control population with at least 1 elevated cancer marker, sensitivity was 95.5% and specificity was 61.5%. At ≥1.5 standard deviations above the mean of the control population with at least 1 elevated cancer marker, sensitivity was 93.6% and specificity was 80.8%. At ≥2 standard deviations above the mean of the control population with at least 1 elevated cancer marker, sensitivity was 88.5% and specificity was 86.5%.

Overall stage distribution and histologies of the lung cancer group are summarized in Table 6. Ten patients did not have lung cancer confirmed pathologically and were presumed to have lung cancer based on radiologic characteristics. A total of 103 (66.0%) patients were early stage (Stage 0, I, and II) of whom 43 (41.7%) were stage 0 or I with tumor size <2 cm. A total of 36 (83.7%) of these patients tested positive at ≥1.5 standard deviations above the mean of the control population.

Within the benign disease group, there was a high prevalence of granulomatous disease (N=26, 40%). The specificity for lung cancer was strongly correlated with the number of elevated cancer markers defining a positive result. With one, two, and three elevated cancer markers, specificity was 44.6%, 78.5%, and 90.8%, respectively, when comparing the benign and cancer populations at 1.5 standard deviations. Notably, 74.8% (n=77) of early stage cancer patients have two or three elevated cancer markers.

**Conclusion**

Increased concentrations of volatile organic compounds (VOCs) in exhaled breath have been detected in patients with lung cancer[10-14]. Our diagnostic model utilizes four compounds, 2-butanone, 2-hydroxyacetaldehyde, 3-hydroxy-2-butanone, and 4-hydroxyhexanal, out of 25 carbonyl compounds separable by mass spectrometry. No other carbonyl compounds were specific for cancer or benign disease in this spectrum. These markers were not associated with smoking except for 3-hydroxy-2-butanone, which was higher in smokers in both the control and benign disease populations; however, control values were still significantly below the range of lung cancer patients.

The sensitivity for detecting lung cancer by breath analysis of exhaled carbonyl compounds is greater than 93% with at least one elevated carbonyl compound at ≥1.5 standard deviations above the mean of the control population. Additionally,

when the control population was limited to those ≥55 years with a history of smoking, sensitivity and specificity were 93.6% and 80.8%, respectively, under the same conditions. In both the entire control population and a subset of patients who meet NLST criteria, the specificity for lung cancer was greater than 80% when at least one carbonyl compound was elevated. For the purposes of potential screening it would follow that high sensitivity is needed. Thus choosing a lower carbonyl cancer marker concentration and accepting a single marker as positive would yield greater sensitivity with presumed reduction in specificity.

The probability that a given patient has lung cancer increases with the increasing carbonyl marker concentration and the number of elevated carbonyl markers. If a CT scan is performed first and a pulmonary nodule is detected, then an algorithm yielding greater specificity is of interest. One could then apply the same data utilizing the higher carbonyl concentration and a greater number of carbonyl markers to serve as predictive of cancer versus benign disease. In our subset of patients comparing benign disease with malignant disease, evaluating two or more carbonyl markers at ≥2 standard deviations above the mean of the control population achieved a balance with 61.8% sensitivity and 84.6% specificity. At issue is whether or not this degree of specificity would prompt a change in clinical decision-making. A prospective, multi-institutional trial is proposed to evaluate the value added by breath analysis given an indeterminate pulmonary nodule detected by CT scan. A pretest probability based purely on CT scan and clinical findings could be rendered prior to evaluation of breath analysis data. Such a

comparison could ultimately determine whether breath analysis is more or less predictive of malignancy than established predictive models[15] based on clinical and radiographic data.

Our group has previously published results from analysis of exhaled breath distinguishing benign and malignant lung disease with a positive breath test defined as at least two of four cancer markers elevated[10]. The sensitivity and specificity to detect early stage lung cancer was 83% and 75%, respectively. This previous study treated breath analysis as a diagnostic test; however, to use breath analysis as a screening modality, the threshold to define an elevated carbonyl compound was lowered in order to capture a greater proportion of patients with lung cancer. The current study presents data at multiple cutoff values demonstrating the utility of breath analysis in different capacities both as a screening and diagnostic modality. With three elevated compounds at any threshold, the specificity is 100% indicating the ability of breath analysis to diagnose lung cancer. When at least one compound is elevated at ≥1 and 1.5 standard deviations, the sensitivity to detect lung cancer is high and similar to that of the NLST[7]. The stage distributions for breath analysis and the NLST are similar. A total of 66.0% of lung cancer patients were early stage in the breath analysis screening study, slightly higher than the proportion (57.1%) of early stage patients diagnosed during the screening trial. The NLST had a smaller proportion of patients with adenocarcinoma (36.3% vs. 41.0% in NLST and breath analysis, respectively) and squamous cell carcinoma (23.2% vs. 35.9% in

NLST and breath analysis, respectively) detected on initial CT scan screening; however, in our study population, all major cancer histologies were detected with high sensitivity. Because of the similar sensitivity and the high cost prediction of screening at-risk patients with low-dose CT scanning, breath analysis may be a possible substitute for the screening of this patient population. Additionally, patients who are false positives with breath analysis may have indolent cancers that may have contributed to over diagnosis with CT scanning[16]. The question can only be answered with a prospective trial comparing breath analysis to low-dose CT scanning. If it is born out that breath analysis has a similar sensitivity for early lung cancer detection, then one may presume an equivalent survival benefit with breath analysis to CT scan screening.

There are many advantages to the method of analysis of exhaled breath for volatile organic compounds described herein[11]. The patient need only supply one exhaled breath. Once the breath is collected, the bag could be immediately evacuated over the silicon microchip in an office or hospital setting and subsequent quantitative analysis could be performed at a central facility. The analytical process has multiple advantages. The amino-oxy compound coated on the microchip serves to isolate only carbonyl compounds and to concentrate them by 10,000 fold[11]. Analysis by mass spectrometry is then straightforward: it provides quantitative values of a limited number of easily identified carbonyl compounds. Another benefit includes the noninvasiveness of the test. It is inherently convenient for the patient and would avoid annual radiation exposure

associated with CT screening. Additionally, patient compliance may improve since breath analysis can be completed in the clinical setting during an office visit. Assuming nearly equal sensitivity, increased compliance would translate to a greater proportion of lung cancers detected at an early stage and, in turn, a greater improvement in overall lung cancer survival.

Our study has limitations. The control and benign disease groups are both younger than the lung cancer group; however, the results using an approximation of the NLST criteria in the breath analysis control population revealed similar results to the original analysis with high sensitivity and specificity. However, the proportion of false positives using breath analysis is less than that of the NLST[7].

A multi-center trial across various geographic areas and environmental conditions is necessary to validate the sensitivity and specificity to detect lung cancer shown in this study. Other pulmonary diseases have the potential to produce false positive results; however, COPD, common in our control population, does not appear to be a confounding factor. As a screening modality, we would project that only patients in their baseline state of health should be screened. It would be reasonable to propose a prospective trial simultaneously comparing screening with CT scan vs. breath analysis in at-risk patients. Given the high rate of false positivity with both modalities, an accurate comparison of specificity could be made with a relatively low sample size. Conversely, the low incidence of lung cancer in the at-risk population of smokers, 1.5% per year,

could be augmented by including early stage cancer patients who present in an equivalent fashion with asymptomatic incidental findings.

In conclusion, the detection of elevated concentrations of 2-butanone, 3-hydroxy-2-butanone, 2-hydroxyacetaldehyde, and 4-HHE in exhaled breath can detect lung cancer with high sensitivity and acceptable specificity given the limitations of a lack of paired CT scans among control patients. A prospective trial comparing breath analysis and CT screening is being carried out currently. A reliable breath analysis method could spare annual CT imaging for a large proportion of subjects at risk for lung cancer, such that only those with a positive breath screen undergo radiographic imaging. The downstream savings from not evaluating indeterminate nodules may be profound.

## Acknowledgements

The authors acknowledge funding from Clinical and Translational Science Pilot Grant Program (CTSPGP) of the University of Louisville, the National Science Foundation (CBET:1159829), the University of Louisville Coulter Grant Program, and the Bill & Melinda Gates Foundation.

## References


1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA: a cancer journal for clinicians.* Jan-Feb 2014;64(1):9-29.
2. Oken MM, Hocking WG, Kvale PA, et al. Screening by chest radiograph and lung cancer mortality: the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized trial. *JAMA : the journal of the American Medical Association.* Nov 2 2011;306(17):1865-1873.
3. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. *CA: a cancer journal for clinicians.* Jan-Feb 2014;64(1):52-62.
4. DeSantis CE, Lin CC, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2014. *CA: a cancer journal for clinicians.* Jul-Aug 2014;64(4):252-271.
5. Castle PE. PSA testing for prostate cancer screening. *The Lancet. Oncology.* Jan 2015;16(1):e2-3.
6. National Lung Screening Trial Research T, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *The New England journal of medicine.* Aug 4 2011;365(5):395-409.
7. National Lung Screening Trial Research T, Church TR, Black WC, et al. Results of initial low-dose computed tomographic screening for lung cancer. *The New England journal of medicine.* May 23 2013;368(21):1980-1991.

8. de Koning HJ, Meza R, Plevritis SK, et al. Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the U.S. Preventive Services Task Force. *Annals of internal medicine.* Mar 4 2014;160(5):311-320.
9. Mauchley DC, Mitchell JD. Current estimate of costs of lung cancer screening in the United States. *Thoracic surgery clinics.* May 2015;25(2):205-215.
10. Bousamra M, 2nd, Schumer E, Li M, et al. Quantitative analysis of exhaled carbonyl compounds distinguishes benign from malignant pulmonary disease. *The Journal of thoracic and cardiovascular surgery.* Sep 2014;148(3):1074-1081.
11. Fu XA, Li M, Knipp RJ, Nantz MH, Bousamra M. Noninvasive detection of lung cancer using exhaled breath. *Cancer medicine.* Feb 2014;3(1):174-181.
12. van der Schee MP, Paff T, Brinkman P, van Aalderen WM, Haarman EG, Sterk PJ. Breathomics in lung disease. *Chest.* Jan 2015;147(1):224-231.
13. Wagner C, Munoz MA, Jareno J, et al. Volatile organic compounds, new biomarkers in exhaled breath samples of lung cancer patients with and without chronic obstructive pulmonary disease. *Chest.* Mar 1 2014;145(3 Suppl):332A.
14. Phillips M, Cataneo RN, Cummin AR, et al. Detection of lung cancer with volatile markers in the breath. *Chest.* Jun 2003;123(6):2115-2123.

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.* Nov 2005;237(2):395-400.
16. Patz EF, Jr., Pinsky P, Gatsonis C, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA internal medicine.* Feb 1 2014;174(2):269-274.

## Tables

**Table 1. Demographic characteristics of study population.**

| | Lung Cancer (N=156) | Benign (N=65) | Control (N=194) | P-value |
|---|---|---|---|---|
| **Patients** | | | | |
| Age (years) | 65.1±10.4 | 54.2±13.1 | 49.4±16.4 | <0.001 |
| Gender (Male) | 81 (51.9) | 32 (49.2) | 108 (55.7) | 0.610 |
| **Smoking History** | | | | |
| Current | 69 (44.2) | 25 (38.5) | 73 (37.6) | <0.001 |
| Former | 80 (51.3) | 20 (30.8) | 41 (21.1) | <0.001 |
| Never | 7 (4.5) | 20 (20.8) | 80 (41.2) | <0.001 |
| **Size of nodule** | | | | |
| <4 mm | 0 | 2 | | |
| 4-6 mm | 4 | 5 | | |
| 7-10 mm | 8 | 7 | | |
| 11-20 mm | 49 | 28 | | |
| 21-30 mm | 27 | 8 | | |
| >30 mm | 62 | 11 | | |
| Unknown | 6 | 4 | | |

Summary of demographic information for the entire study population. Former smokers include patients who quit > 1 week prior to sample collection. Results are reported in n (%). A p-value <0.05 is considered significant.

**Table 2. Concentrations of carbonyl compounds by disease state. and smoking status**

| | 2-Butanone | 2-hydroxy-3-butanone | 2-HA | 4-HHE |
|---|---|---|---|---|
| **Control** | | | | |
| Non-smoker (N=80) | 1.26 (0.93,1.93) | 0.07 (0.03, 0.10) | 0.19 (0.11, 0.28) | 0.001 (0.000, 0.003) |
| Smoker (N=114) | 1.40 (0.86, 1.92) | 0.09 (0.04, 0.13) | 0.17 (0.07, 0.25) | 0.002 (0.001, 0.004) |
| P-Value | 0.693 | 0.037 | 0.060 | 0.308 |
| **Benign** | | | | |
| Non-smoker (N=20) | 1.45 (1.23, 2.43) | 0.13 (0.09, 0.17) | 0.17 (0.06, 0.27) | 0.001 (0.000, 0.004) |
| Smoker (N=45) | 1.77 (1.31, 2.48) | 0.16 (0.10, 0.31) | 0.21 (0.12, 0.33) | 0.002 (0.001, 0.005) |
| P-Value | 0.143 | 0.043 | 0.231 | 0.491 |
| **Cancer** | | | | |
| Non- smoker (N=7) | 2.47 (1.51, 3.13) | 0.15 (0.13, 0.36) | 0.29 (0.08, 0.38) | 0.007 (0.003, 0.027) |
| Smoker (N=149) | 3.04 (2.39, 3.97) | 0.31 (0.20, 0.54) | 0.33 (0.21, 0.46) | 0.007 (0.002, 0.015) |
| P-Value | 0.565 | 0.08 | 0.177 | 0.517 |

Values are reported in median (interquartile range). 2-HA = 2-hydroxyacetaldehyde, 4-HHE = 4-hydroxyhexanal.

**Table 3. Concentrations of carbonyl compounds by disease state.**

| | 2-Butanone | 2-hydroxy-3-butanone | 2-HA | 4-HHE |
|---|---|---|---|---|
| **Control (N=194)** | 1.36 (0.88, 1.92) | 0.07 (0.04, 0.12) | 0.18 (0.09, 0.26) | 0.001 (0.001, 0.004) |
| **Benign (N=65)** | 1.72 (1.29, 2.48) | 0.14 (0.10, 0.26) | 0.20 (0.09, 0.33) | 0.001 (0.001, 0.005) |
| **Cancer (N=156)** | | | | |
| Early (N=103) | 2.89 (2.17, 3.60) | 0.27 (0.17, 0.51) | 0.33 (0.20, 0.43) | 0.007 (0.002, 0.014) |
| Late (N=53) | 3.44 (2.55, 4.63) | 0.33 (0.20, 0.67) | 0.31 (0.20, 0.55) | 0.009 (0.003, 0.016) |
| **P-value** | <0.001 | <0.001 | <0.001 | <0.001 |

Median concentrations (nmol/L) of 2-butanone, 2-hydroxy-3-butanone, 2-hydroxyacetaldehyde (2-HA), 4-hydroxyhexanal (4-HHE) reported in median (interquartile range).

**Table 4. Sensitivity and specificity of breath analysis for lung cancer.**

| ≥1 standard deviation | | |
|---|---|---|
| **All Lung Cancer** | | |
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 95.5% | 64.4% |
| ≥2 | 83.3% | 92.3% |
| ≥3 | 55.8% | 100.0% |
| **Early Stage Lung Cancer** | | |
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 93.1% | 64.4% |
| ≥2 | 80.2% | 92.3% |
| ≥3 | 51.5% | 100.0% |

| ≥1.5 standard deviations | | |
|---|---|---|
| **All Lung Cancer** | | |
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 93.6% | 85.6% |
| ≥2 | 76.9% | 99.5% |
| ≥3 | 40.4% | 100.0% |
| **Early Stage Lung Cancer** | | |
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 91.1% | 85.6% |
| ≥2 | 74.3% | 99.5% |
| ≥3 | 36.6% | 100.0% |

| ≥2 standard deviations | | |
|---|---|---|
| **All Lung Cancer** | | |
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 88.5% | 91.8% |
| ≥2 | 61.5% | 100.0% |
| ≥3 | 26.9% | 100.0% |
| **Early Stage Lung Cancer** | | |
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 85.1% | 91.8% |
| ≥2 | 58.4% | 100.0% |
| ≥3 | 18.8% | 100.0% |

Early stage lung cancer includes patients with stage 0, I, and II lung cancer.

**Table 5. Sensitivity and specificity for controls who meet National Lung Screening Trial Criteria**

| ≥1 standard deviation<br>All Lung Cancer | | |
|---|---|---|
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 95.5% | 61.5% |
| ≥2 | 83.3% | 86.5% |
| ≥3 | 55.8% | 100.0% |

| ≥1.5 standard deviations<br>All Lung Cancer | | |
|---|---|---|
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 93.6% | 80.8% |
| ≥2 | 76.9% | 100.0% |
| ≥3 | 40.4% | 100.0% |

| ≥2 standard deviations<br>All Lung Cancer | | |
|---|---|---|
| Elevated cancer markers, n | Sensitivity | Specificity |
| ≥1 | 88.5% | 86.5% |
| ≥2 | 61.5% | 100.0% |
| ≥3 | 26.9% | 100.0% |

Sensitivity and specificity for selected control subjects and lung cancer patients.

The control group includes patients ≥55 years who are current or former smokers.

**Table 6. Stage and histology of lung cancer patients.**

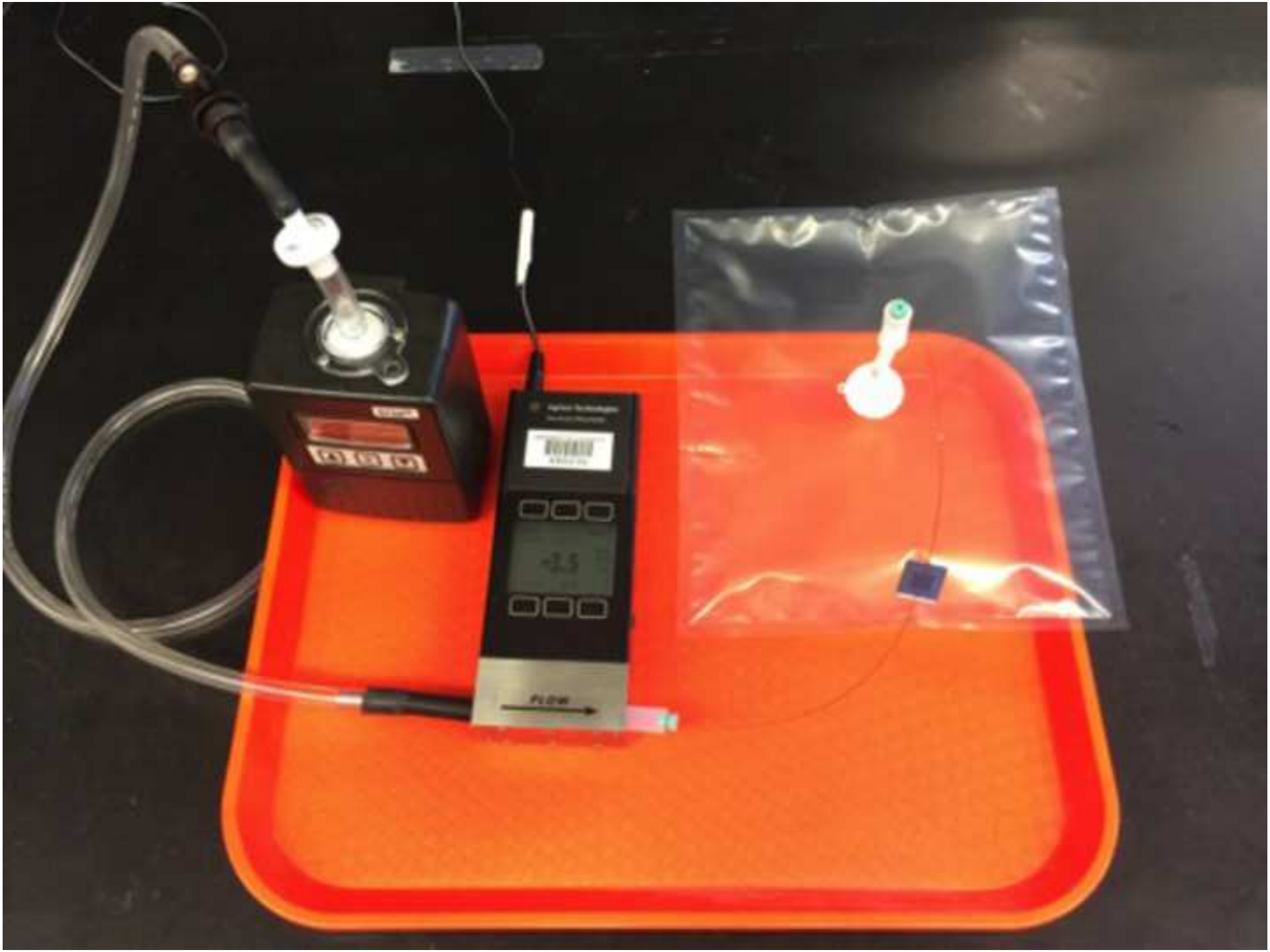
| | Breath Analysis (N=156) | NLST[6] (N=1060) |
|---|---|---|
| **Stage** | | |
| 0 | 1 (0.6) | 0 (0) |
| IA | 52 (33.3) | 416 (40.0) |
| IB | 23 (14.7) | 104 (10.0) |
| IIA | 8 (5.1) | 35 (3.4) |
| IIB | 19 (12.2) | 38 (3.7) |
| IIIA | 21 (13.5) | 99 (9.5) |
| IIIB | 9 (5.8) | 122 (11.7) |
| IV | 23 (14.7) | 226 (21.7) |
| **Histology** | | |
| Adenocarcinoma in situ | 1 (0.6) | 110 (10.5) |
| Adenocarcinoma | 64 (41.0) | 380 (36.3) |
| Squamous cell | 56 (35.9) | 243 (23.2) |
| Small cell | 13 (8.3) | 137 (13.1) |
| Non-small cell, other | 9 (5.8) | 131 (12.5) |
| Carcinoid | 2 (1.3) | 6 (0.6) |
| Large cell | 1 (0.6) | 41 (3.9) |
| Unknown | 10 (6.4) | 0 (0) |

Results reported in n(%). NLST = National Lung Screening Trial

## Figure Legends

**Figure 1.** The collection bag, microchip, and evacuation system are shown here. The contents of the bag are evacuated over the silicon microchip using the vacuum pump and is regulated by the flow meter.

Figure(s) (see Info for Authors for details)
Click here to download Figure(s) (see Info for Authors for details): Figure.tif



Central Picture

Click here to download Central Picture: Figure 1.tif