



Contents lists available at ScienceDirect

Tuberculosis

journal homepage: <http://intl.elsevierhealth.com/journals/tube>

DIAGNOSTICS

Breath biomarkers of active pulmonary tuberculosis

Michael Phillips^{a,b,*}, Victoria Basa-Dalay^{c,h}, Graham Bothamley^{d,i}, Renee N. Cataneo^{a,j},
Phung Kim Lam^{e,k}, Maria Piedad R. Natividad^{f,l}, Peter Schmitt^{g,m}, James Wai^{a,j}

^a Menssana Research Inc., 1 Horizon Road, Suite 1415, Fort Lee, NJ 07024-6510, United States

^b Department of Medicine, New York Medical College, Valhalla, NY, United States

^c Tuberculosis Research Unit, Angelo King Medical Research Center, De La Salle Health Sciences Institute, Cavite, Philippines

^d Department of Respiratory Medicine, Homerton University Hospital NHS Foundation Trust, London E9 6SR, England, UK

^e Division of Pulmonary and Critical Care Medicine, University of California San Diego, 200 West Arbor Drive, San Diego, CA 92103-8374, USA

^f Center for Respiratory Medicine, The University of Santo Tomas Hospital (USTH), España Boulevard, Manila 1008, Philippines

^g Schmitt & Associates, 211 Warren St, Newark, NJ 07103, United States

ARTICLE INFO

Article history:

Received 27 October 2009

Received in revised form

18 January 2010

Accepted 26 January 2010

Keywords:

Tuberculosis

Breath

Volatile

Biomarker

Diagnosis

SUMMARY

Background: Volatile organic compounds (VOCs) in breath may contain biomarkers of active pulmonary tuberculosis derived from the infectious organism (metabolites of *Mycobacterium tuberculosis*) and from the infected host (products of oxidative stress).

Methods: We analyzed breath VOCs in 226 symptomatic high-risk patients in USA, Philippines, and UK, using gas chromatography/mass spectroscopy. Diagnosis of disease was based on sputum culture, smear microscopy, chest radiography and clinical suspicion of tuberculosis (CSTB). Chromatograms were converted to a series of 8 s overlapping time slices. Biomarkers of active pulmonary tuberculosis were identified with a Monte Carlo analysis of time-slice alveolar gradients (abundance in breath minus abundance in room air).

Results: Breath VOCs contained apparent biomarkers of active pulmonary tuberculosis comprising oxidative stress products (alkanes and alkane derivatives) and volatile metabolites of *M. tuberculosis* (cyclohexane and benzene derivatives). Breath biomarkers identified active pulmonary tuberculosis with C-statistic (area under curve of receiver operating characteristic) = 0.85 (i.e. 85% overall accuracy, sensitivity = 84.0%, specificity = 64.7%) when sputum culture, microscopy, and chest radiography were either all positive or all negative. Employing a single criterion of disease, C-statistic = 0.76 (smear microscopy), 0.68 (sputum culture), 0.66 (chest radiography) and 0.65 (CSTB).

Conclusion: A breath test identified apparent biomarkers of active pulmonary tuberculosis with 85% accuracy in symptomatic high-risk subjects.

© 2010 Elsevier Ltd. All rights reserved.

Active pulmonary tuberculosis (TB) is a leading cause of death from infectious disease throughout the world. Two billion people – one third of the world's population – are infected with *Mycobacterium tuberculosis* and 1.6 million died from the disease in 2005.¹

* Corresponding author. Menssana Research Inc., 1 Horizon Road, Suite 1415, Fort Lee, NJ 07024-6510, United States. Tel./fax: +1 201 886 7004.

E-mail addresses: mphillips@menssana-research.com (M. Phillips), victoriadalay@yahoo.com (V. Basa-Dalay), graham.bothamley@homerton.nhs.uk (G. Bothamley), rcataneo@menssana-research.com (R.N. Cataneo), plam@ucsd.edu (M.P.R. Lam), peachynatividad@yahoo.com (M.P.R. Natividad), peter.schmitt@schmittassociates.com (P. Schmitt), jwai2009@gmail.com (J. Wai).

^h Tel.: +63 46 416 0226x238.

ⁱ Tel.: +44 (0) 20 8510 7814.

^j Tel.: +1 973 643 5464.

^k Tel.: +1 619 543 5550; fax: +1 619 543 3276.

^l Tel.: +63 (02) 731 3001.

^m Tel.: +1 973 733 2225.

Globally, the countries of sub-Saharan Africa suffer the highest incidence of TB as a consequence of the HIV/AIDS epidemic.² In these countries, the disproportionate amount of smear-negative disease has greatly complicated TB case detection and disease control.³ There has also been a resurgence of TB in developed countries. After approximately 30 years of decline, the number of TB cases reported in the United States increased 20% during 1985–1992, and this has been accompanied by the emerging threat of multidrug-resistant TB.¹

Microscopy and culture are still the mainstay of laboratory diagnosis of TB,⁴ and there is an urgent need for better diagnostic tools, especially in high-burden countries. An ideal diagnostic test would be sensitive and specific for active pulmonary TB, as well as rapid, cost-effective, non-invasive, and suitable for use in developing countries. A breath test might rationally fulfill these requirements because *Mycobacteria* manufacture unique volatile

organic compounds (VOCs) as metabolites with distinctive odors when cultured in vitro.^{5–7} The best-known application of breath testing is detection of alcohol consumption, but newer and more sensitive breath tests have detected products of normal physiological processes such as oxidative stress,⁸ and biomarkers of diseases including asthma,⁹ lung cancer,¹⁰ heart transplant rejection,¹¹ and infection with *Helicobacter pylori*.¹² However, detection of volatile disease biomarkers is technically difficult because most breath VOCs are excreted in picomolar concentrations (parts per trillion), and most analytical instruments in current use cannot detect VOCs in such low concentrations. Members of our group have approached this problem by developing a breath collection apparatus (BCA) that collects and concentrates breath VOCs on to sorbent traps for analysis by gas chromatography and mass spectroscopy (GC/MS).^{13,14} In a previous study employing this test, we observed apparent biomarker VOCs in the breath of patients with active pulmonary TB, and a number of these VOCs were similar to VOCs observed in vitro in cultures of *M. tuberculosis*.¹⁵

We performed a multicenter international study of breath VOCs in patients at high risk of active pulmonary TB in order to identify breath biomarkers of disease, and to determine the diagnostic accuracy of the test.

Materials and methods

Human subjects

Patients at high risk of active pulmonary TB

226 technically satisfactory breath samples and sputum cultures were obtained from high-risk patients at four sites: the University of California San Diego, California (38). The University of Santo Tomas, Manila, Philippines (100), De La Salle University Hospital, Cavite, Philippines (66), and The East London Tuberculosis Service, London, England (22). Patients aged 13 yr or older were defined as “high-risk” and included in the study if they had one or more symptoms or radiographic findings consistent with the diagnosis of active TB, and/or an epidemiologic risk factor for TB (e.g. immigrants or refugees from high-burden countries). Patients were excluded if they had more than one previous sputum smear that was positive for acid-fast bacilli during the past six months, if they were currently being treated for TB, if they had received more than seven days of treatment for TB during the past six months, or if there was evidence of extrathoracic TB.

Clinical evaluation

A medical history was obtained from all high-risk patients, followed by a physical examination and a chest radiograph. Sputum samples were collected on three different occasions employing sputum induction or bronchoalveolar lavage where indicated. Liquid cultures were performed (Bactec 460, Becton Dickinson, Sparks, MD) as well as microscopic examinations for acid-fast bacilli. Three smears and three cultures were performed for each patient. A chest radiograph was classified as abnormal if any of the following features were observed: cavity, infiltrate, fibrosis, bronchiectasis, pleural effusion, hilar or mediastinal adenopathy. The physician's clinical judgment of the final diagnosis was also determined, employing a standardized formal evaluation of clinical suspicion of tuberculosis (CSTB) based upon the diagnostic criteria listed above, as well as the medical history, patient demographics, physical examination, and response to treatment.³³

IRB approval and informed consent

The research was approved by an Institutional Review Board at all collaborating sites, and all subjects gave their signed informed consent to participate. Assent from adolescent subjects and consent

from a parent or legal guardian was obtained for subjects 13–16 yr in England or younger than 18 yr at sites in other countries.

Breath sample collection

The method has been described.^{12,18} A portable breath collection apparatus (BCA) was employed to capture the VOCs in 1.0 L breath and 1.0 L room air on to separate sorbent traps. The geometry of the breath reservoir of the BCA ensured that the sample comprised >99% alveolar breath. Subjects wore a nose-clip and respired normally for 2.0 min through a disposable and previously unused valved mouthpiece with a bacterial filter to prevent Mycobacterial contamination of the instrument. The mouthpiece and filter presented low resistance to respiration, ensuring that samples were collected without discomfort to patients.

Breath sample analysis

The method has been described.^{8,13} VOCs captured in the sorbent traps were analyzed in the laboratory by automated thermal desorption, gas chromatography and mass spectroscopy (ATD/GC/MS). In order to quantify peak areas and to control for drift in instrument performance, an internal standard was run with every chromatographic assay of breath and air (0.25 mL 2 ppm 1-bromo-4-fluorobenzene, Supelco, Bellefonte, PA).

Masking procedures

Neither the clinicians collecting breath samples nor the clinical staff performing sputum cultures were aware of the results of the breath test. Similarly, the laboratory staff performing the breath assays had no access to any clinical information about the subjects. Masking was not broken until data were analyzed.

Analysis of data and statistical methods

Diagnostic criteria

Patients were diagnosed with active pulmonary TB according to a single criterion i.e. positive sputum culture, positive sputum smear microscopy, or abnormal chest radiogram. A triple criterion was also employed (i.e. positive or negative for all three tests). Concordance between test results was evaluated with Kappa and McNemar's test.¹⁶ Clinical judgment (CSTB) was separately employed as a single criterion of active pulmonary TB.

Chromatographic data

Chromatograms were converted into a series of data points by segmenting them into a series of 900 time slices, each with eight sec duration and four sec overlap. The alveolar gradient of each time slice was determined (i.e. abundance in alveolar breath minus abundance in ambient room air).^{13,17} In each time slice, alveolar gradient = $V_b/I_b - V_a/I_a$ where V_b denotes the integrated abundance of analytes detected by mass spectroscopy in breath, and I_b denotes the area under the curve (AUC) of the chromatographic peak associated with the internal standard. V_a and I_a denote corresponding values derived from the associated sample of room air.

Identification of biomarker time slices

Alveolar gradients were compared in patients who were positive or negative for active pulmonary TB, and ranked as candidate biomarkers according to the value of the C-statistic i.e. the AUC of the receiver operating characteristic (ROC) curve.¹⁸ We employed Monte Carlo simulations to select the chromatographic time slices that identified active pulmonary TB with better than random accuracy, in order to minimize the risk of including random identifiers of disease.^{19,20} The average random behavior of chromatographic time slices was determined by randomly assigning subjects to the “active pulmonary TB positive” or “active pulmonary TB

negative” groups, and performing 40 estimates of the C-statistic value of each time slice. For any given value of the C-statistic, it was then possible to compare the mean number of time slices exceeding that value by correct assignment or by random assignment. Conservative predictive algorithms were constructed with multivariate weighted digital analysis (WDA) models¹⁰ employing the time slices identified by multiple Monte Carlo simulations.

Tentative identification of biomarker VOCs

The chemical identity of the major VOC in each time slice was tentatively identified by the similarity of its mass spectrum to the mass spectrum of a known compound in a computer-based library (Turbo Mass Software, PerkinElmer Life And Analytical Sciences, Waltham, MA 02451).

Results

Human subjects

Table 1 displays demographics, prevalence of symptoms, results of chest radiography, sputum culture, and sputum microscopy, and subjects assessed as positive or negative for active pulmonary TB. No subject reported any adverse effects associated with breath sample donation.

Concordance between tests for active pulmonary TB

Table 3 displays concordance between results of smear microscopy, sputum culture and chest radiography. Kappa values and results of McNemar's test¹⁶ indicate low agreement between these three diagnostic methods, except for sputum culture versus smear microscopy for which agreement was moderate.

Construction of predictive algorithms

10 time slices with C-statistic ≥ 0.73 were identified of which only 5 were statistically significant. All combinations of 5/10 time slices were analyzed separately and the combination that yielded the lowest C-statistic was selected. This was the conservative lower limit of the analysis, and the true value may have been higher.

Identification of breath biomarkers of active pulmonary TB

Figure 1 displays the outcomes of single and multiple Monte Carlo simulations in subjects who were positive or negative for active pulmonary TB, employing triple criterion of disease. The

mean outcome of 40 Monte Carlo simulations demonstrated 10 time slices with an AUC of 0.73 or higher. The excess of correct over random time slices indicated that 5/10 of them comprised non-random biomarkers of active pulmonary TB. Table 2 displays tentative structural identification of the most abundant VOC in each of these time slices.

Sensitivity and specificity of breath biomarkers of active pulmonary TB

Figure 1 displays ROC curves for active pulmonary TB employing the single and the triple criterion of disease as well as CSTB. Multivariate WDA models employed the most conservative of all possible models (i.e. the model with the lowest C-statistic) derived from time slices identified by multiple Monte Carlo simulations. Employing triple criterion of disease, breath biomarkers identified active pulmonary tuberculosis with C-statistic = 0.85 (i.e. 85% overall accuracy, sensitivity = 84.0%, specificity = 64.7%). Employing single criterion of disease, C-statistic = 0.76 (smear microscopy), 0.68 (sputum culture), and 0.66 (chest radiography) and 0.65 (CSTB).

Discussion

Analysis of breath VOCs demonstrated apparent volatile biomarkers of active pulmonary TB, in an international multicenter study of high-risk patients. The VOC biomarkers may have been derived from the infective organism, the host, or from both.

These findings were consistent with the results of an earlier pilot study because the breath VOC biomarkers were similar to those previously observed in human breath, and also to the volatile metabolites of *M. tuberculosis* observed in vitro.¹⁵ The breath VOC biomarkers observed in both studies included derivatives of cyclohexane, benzene, decane and heptane. The structure of these VOCs was similar but not identical in the two reports – for example 1,3,5-trimethylbenzene was identified as a breath biomarker in this study and 1,2,3,4-tetramethylbenzene was observed in the pilot study. There are only minor differences between the mass spectra of these two compounds, and it may be difficult to differentiate them with current GC/MS instruments. The breath VOC biomarkers were also similar to volatile metabolites of *M. tuberculosis* observed in vitro. In the pilot study,¹⁵ reference samples of *M. tuberculosis* were cultured in sealed containers and VOC metabolites were sampled from the gaseous phase above the liquid incubation medium. Common VOCs observed in vitro and in breath included

Table 1

Human subjects.

| | |
|---|-------------|
| No. patients enrolled | 226 |
| Female/male | 101/125 |
| Age | |
| mean yr (SD) | 43.8 (16.8) |
| Range (yr) | 15–78 |
| History of tobacco or cannabis use (yes/no/unknown) | 107/118/1 |
| Symptoms | |
| cough (yes/no/unknown) | 178/47/1 |
| unintentional weight loss(yes/no/unknown) | 146/84/6 |
| fever (yes/no/unknown) | 84/138/4 |
| night sweats(yes/no/unknown) | 68/156/2 |
| Investigations | |
| sputum culture (positive/negative) | 65/161 |
| sputum microscopy (positive/negative/unknown) | 30/188/8 |
| chest radiogram (abnormal/normal/unknown) | 187/24/15 |

Table 2

Breath VOC biomarkers of active pulmonary TB.

| |
|--------------------------------------|
| Oxetane, 3-(1-methylethyl)- |
| Dodecane, 4-methyl- |
| Cyclohexane, hexyl- |
| Bis-(3,5,5-trimethylhexyl) phthalate |
| Benzene, 1,3,5-trimethyl- |
| Decane, 3,7-dimethyl- |
| Tridecane |
| 1-Nonene, 4,6,8-trimethyl- |
| Heptane, 5-ethyl-2-methyl- |
| 1-Hexene, 4-methyl- |

Ten chromatographic time slices were identified as biomarkers of active pulmonary TB in patients who were positive for all three criteria of disease (sputum culture, smear microscopy, and chest radiography). The most abundant chemical species in each time slice was identified by mass spectroscopy. These VOCs may have been derived from either or both the host and the organism. Alkanes (e.g. tridecane) and methylated alkane derivatives (e.g. dodecane, 4-methyl-) are products of oxidative stress. Other VOCs on this list (including benzene, 1,3,5-trimethyl-, cyclohexane, hexyl-, and 1-hexene, 4-methyl-) are structurally similar to VOCs previously observed as metabolites of *M. tuberculosis* in vitro.

Table 3
Concordance between diagnostic tests for active pulmonary TB.

| <i>Sputum culture versus chest radiography</i> | | | | Kappa | McNemar's test |
|--|----------|------------------|-------------|-------|----------------|
| | | sputum culture | | | |
| | | positive | negative | | |
| chest radiography | positive | 59 (26.1%) | 128 (56.6%) | 0.05 | <0.0001 |
| | negative | 4 (1.8%) | 20 (8.8%) | | |
| <i>Sputum culture versus smear microscopy</i> | | | | Kappa | McNemar's test |
| | | sputum culture | | | |
| | | positive | negative | | |
| smear microscopy | positive | 26 (11.5%) | 4 (1.8%) | 0.47 | <0.0001 |
| | negative | 36 (15.0%) | 152 (67.3%) | | |
| <i>Chest radiography versus smear microscopy</i> | | | | Kappa | McNemar's test |
| | | smear microscopy | | | |
| | | positive | negative | | |
| chest radiography | positive | 29 (12.8%) | 154 (68.1%) | 0.04 | <0.0001 |
| | negative | 0 (0.0%) | 21 (9.3%) | | |

This table displays the concordance between results of sputum culture, sputum smear microscopy for acid-fast bacilli, and chest radiography in the patient population. Twice-measured data may be evaluated with two types of measures: Kappa indicates whether agreement was beyond chance, and McNemar's test assesses the symmetry of the misses about the diagonal.

For *sputum culture versus chest radiography*, Kappa = 0.05 indicated little to no agreement beyond chance, and a discrepancy was significantly more likely to be an abnormal chest radiogram with a negative sputum culture.

For *sputum culture versus smear microscopy*, Kappa = 0.47 indicated moderate agreement beyond chance, and a discrepancy was significantly more likely to be a negative smear microscopy with a positive sputum culture.

For *chest radiography versus smear microscopy*, Kappa = 0.04 indicated little to no agreement beyond chance, and a discrepancy was significantly more likely to be an abnormal chest radiogram with a negative sputum smear microscopy.

Overall, Kappa values indicated little to no agreement beyond chance between different criteria for diagnosis of active pulmonary TB, with the exception of moderate agreement between sputum culture and smear microscopy. The statistically significant outcomes of McNemar's tests indicated that the disagreement was represented by strong patterns of unequal discordance. Note that the total of the four boxes in each category does not add up the total number of patients studied (226) because results were unknown for 8 sputum microscopies and 15 chest radiographs.

derivatives of cyclohexane, benzene, heptane and hexane. Again, the chemical structures were similar but not identical. In addition, alkanes and alkane derivatives have been previously observed in human breath as VOC products of oxidative stress.^{8,23} These findings are consistent with previous reports that *M. tuberculosis* manufactures volatile metabolites,^{5–7} and that patients with active pulmonary TB exhibit increased oxidative stress.^{24,25}

The accuracy of a new diagnostic test is usually determined by comparing its performance to an accepted gold standard of disease. This study was limited by comparing the performance of the breath test to gold standards whose accuracy is known to be less than perfect e.g. a study of patients with a confirmed clinical diagnosis of active pulmonary TB demonstrated sensitivity of 70% for sputum culture and 61.5% for sputum smear microscopy, though both were 100% specific.³⁸ Chest radiographic criteria are nonspecific for active pulmonary TB and, except for cavitation, are common sequelae of nontuberculous causes. Kappa values and results of McNemar's test¹⁶ indicated low agreement between the three diagnostic criteria, except for moderate agreement between sputum culture and smear microscopy, consistent with previous reports of discrepant findings between these tests.^{26,27} These limitations may account, in part, for the results of this study, because no new diagnostic test could possibly exhibit better sensitivity and specificity for active pulmonary TB than the gold standard to which it is compared. In this study, when sputum culture was employed as the gold standard of disease, the ROC

curve in Figure 1 indicated that the breath test was approximately 70% sensitive and 70% specific. When a combined triple criterion of disease was employed, the breath test was approximately 70% sensitive and 92% specific, which approaches the reported accuracy of sputum culture.³⁸ Consequently, where the breath test appeared to misclassify patients with false-positive or false-negative findings, these errors may have arisen, at least in part, from deficiencies in the gold standards to which it was compared.

The accuracy of the breath test was assessed with the C-statistic, or AUC of the ROC curve, which is typically used to assess how well a test or model separates individuals into two classes, such as diseased and non-diseased.^{18,28} The value of the C-statistic may vary between 0.5 (discriminatory value no better than chance) and 1.0 (a perfect test, with no false-positive or false-negative results). The value of the breath test C-statistic varied according to the diagnostic criterion employed: positive sputum smear microscopy (0.76), positive sputum culture (0.68), abnormal chest radiography (0.66) or CSTB (0.65). However, when a more rigorous diagnostic criterion was employed, in which sputum culture, smear and chest radiography were all positive or all negative, the breath test C-statistic value was increased to 0.85 i.e. the breath test identified patients with active pulmonary TB with an overall accuracy of 85%, with sensitivity = 84% and specificity = 64.7% (i.e. true positive rate = 84%, false-positive rate = 35.3%). In clinical practice, similar results might be reasonably anticipated in primary screening of other high-risk populations. However, clinicians have the option of selecting a different test cutoff value that would yield a higher true positive rate at the expense of a higher false-positive rate. The expected positive and negative predictive values of the breath test were not determined in this study, since they will vary with the prevalence of active pulmonary TB in the populations where the test is employed in the future.

In 1971, Pauling reported the first evidence of a highly complex set of low-molecular weight VOCs in human breath. He cryogenically concentrated samples of normal breath, analyzed the samples employing the then-new technology of gas chromatography, and demonstrated the presence of large numbers of VOCs in low concentrations.²⁹ Subsequent studies employing mass spectroscopy with gas chromatography have tentatively identified the chemical structure of more than 3000 different VOCs in human breath.³⁰ However, mass spectroscopy may identify a breath VOC incorrectly if the mass spectrum is contaminated by noise in the chromatogram. This can result from several causes including the simultaneous coelution of two different VOCs. For this reason, analysis of the VOCs in a gas chromatogram by retention times alone may provide an intrinsically more robust technique that is less susceptible to error. 30 years ago, Stern et al employed time-slice chromatography to differentiate between different strains of Enterobacteriaceae and *Yersinia enterocolitica*.^{31,32} The principle is straightforward: the chromatogram is partitioned into a series of time slices, each of equal duration, and the integrated abundance of chromatographic peaks in each time slice is determined. The chromatogram is thereby converted to a time series of data points. Chromatograms derived from two populations may then be compared employing multivariate analysis of data, in order to identify the specific time slices that distinguish between the two groups. The advantage of this approach is that it minimizes the risk of experimental error in primary identification of breath biomarkers, while it also permits tentative identification of their chemical structure by mass spectroscopy.

The clinical value of diagnostic breath testing lies in its ability to sample blood VOCs in a non-invasive manner, after they have passively diffused from blood into breath across the alveolar

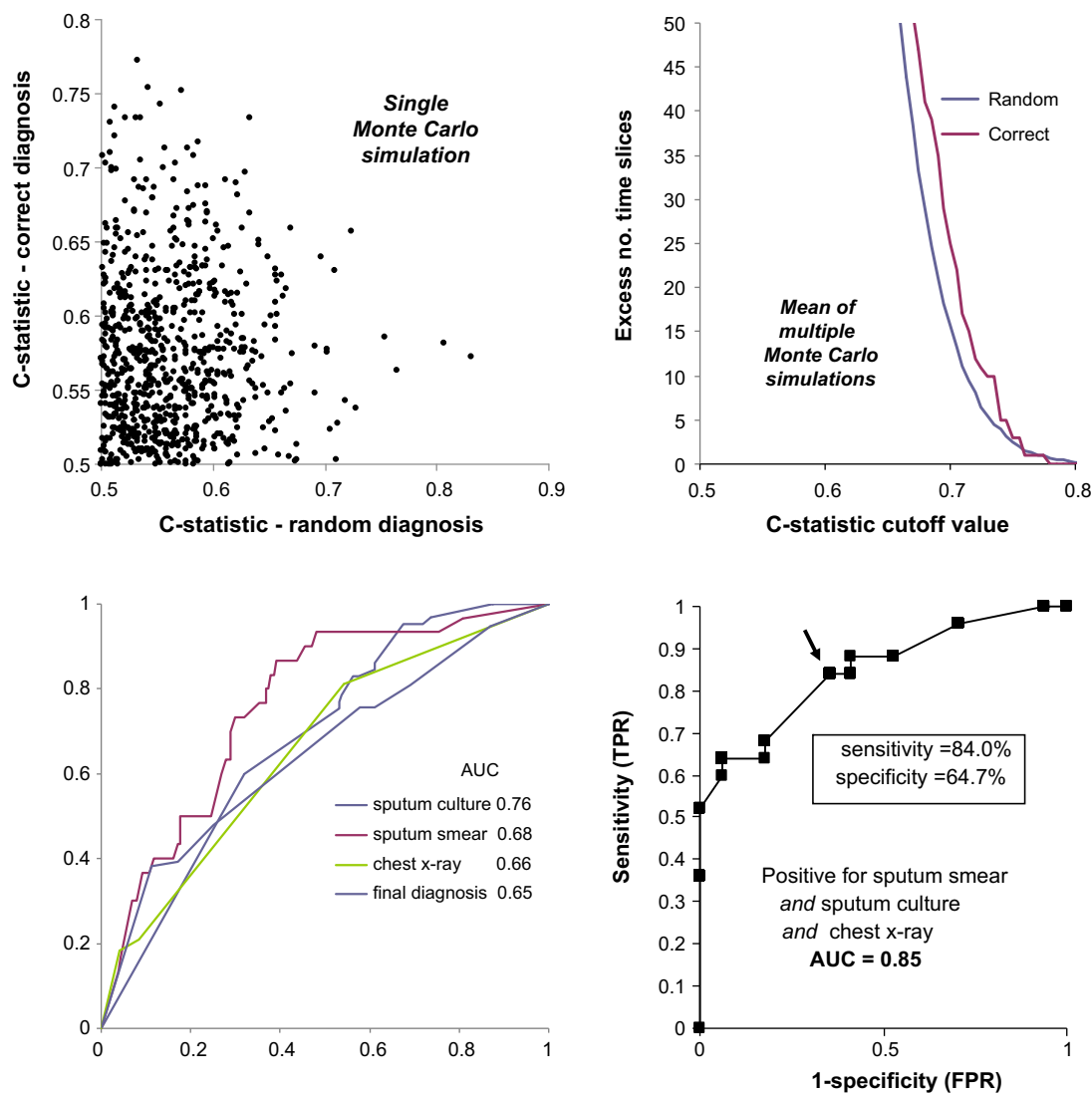


Figure 1. Breath biomarkers of active pulmonary TB. *Single Monte Carlo simulation (top left):* Patients were classified as positive for active pulmonary TB if all of three criteria of disease were positive (sputum culture, sputum smear, and abnormal chest radiogram), and negative for disease if all three criteria were negative. Each chromatographic time slice was evaluated as a candidate biomarker of active pulmonary TB by determination of its C-statistic value, derived from differences between alveolar gradients in subjects who were positive or negative for disease. This figure displays the C-statistic of each chromatographic time slice employing either randomized or correct assignment of diagnosis of active pulmonary TB on the x- and y-axes respectively. In this particular simulation, correct diagnosis was associated with 10 time slices with C-statistic ≥ 0.73 , compared to 5 with randomized diagnosis i.e. when the diagnosis was randomly assigned. There was an excess of 5 time slices over the random outcome. *Multiple Monte Carlo simulations (top right):* This figure displays the mean outcome of 40 Monte Carlo simulations performed in the same fashion as the single simulation shown in the upper panel. Curves display the mean number of chromatographic time slices exceeding a given AUC cutoff with random or correct assignment to the diagnostic group (positive or negative for active pulmonary TB). If the breath contains a signal of active pulmonary TB, the correct assignment curve should be significantly higher than the randomized curve, depending upon the number of true biomarkers present in the data. At C-statistic cutoff value = 0.73, a mean excess of 5 (i.e. 10 – 5) time slices was observed with correct assignment of diagnostic group. This excess of correct over random time slices indicated that 5/10 VOCs comprised non-random biomarkers of active pulmonary TB. *Accuracy of the breath test using a single criterion of disease (bottom left):* The ROC curves displays the accuracy of the breath test for active pulmonary TB when a single criterion of disease was either positive or negative (sputum culture, sputum smear, abnormal chest radiogram, or CSTB respectively). *Accuracy of the breath test using three criteria of disease (bottom right):* The ROC curve displays the accuracy of the breath test for active pulmonary TB when all three criteria of disease were either positive or negative (sputum culture, sputum smear, and abnormal chest radiogram). The predictive model was a multivariate WDA algorithm employing 5/10 VOCs in the breath chromatogram that were identified by multiple Monte Carlo simulations. The ROC curve employed the most conservative model possible i.e. the multivariate model with the lowest C-statistic derived from all possible combinations of 5 out of 10 of them. The VOCs in these time slices are shown in Table 2. The lowest C-statistic was 0.85, and the sensitivity and specificity of the breath test is shown at the shoulder of the ROC curve where their sum was maximal.

membrane. The most familiar application is breath alcohol testing, which has been widely employed for many years for the estimation of blood alcohol levels. However, progress in medical applications of breath testing has been comparatively slow until recent years because of the technical challenges. Unlike breath ethanol which is excreted in parts per thousand or higher, most breath VOC biomarkers of disease are excreted in very low concentrations:

parts per billion or parts per trillion i.e. six to nine orders of magnitude more dilute than ethanol. Consequently, microanalysis of breath VOCs requires specialized instrumentation for the collection and analysis of samples. Recent progress in this technology has enabled breath testing to move into the mainstream of medical diagnostic applications, The United States Food & Drug Administration has approved clinical use of the Heartbreath test

for heart transplant rejection,^{11,37} and clinical studies have reported breath VOC biomarkers in a number of diseases including asthma,⁹ lung cancer,¹⁰ and infection with *H. pylori*.¹²

Breath testing is rapid, non-invasive, and completely safe. It could potentially be employed as a primary screening tool in populations at increased risk of TB. Since the combination of sensitivity and specificity of a test is dependent upon the choice of the diagnostic cutpoint value in the ROC curve, it may be tailored to the circumstances of clinical practice.^{21,22} For example, in a setting with limited resources, the cutpoint might be adjusted for higher specificity and lower sensitivity. This would identify patients at high risk of active pulmonary TB without overburdening the health care system with an excessive number of false-positive results. The breath test might also be employed with two cutpoints (e.g. at 90% sensitivity and 90% specificity respectively) so that patients could be triaged into groups at high, intermediate, or low risk of disease.

The cost of breath testing for active pulmonary TB in clinical practice may not be inferred from this study because we employed a laboratory-based ATD/GC/MS instrument for VOC biomarker discovery. MS is expensive because it requires sophisticated pumping and sealing technology to generate and maintain a high vacuum, and an instrument modified for breath VOC assays may cost approximately \$200,000. Currently, only MS technology can provide definitive structural identification of breath VOCs, but that refinement will not be necessary in clinical practice because a point-of-care instrument employing ATD/GC with a less expensive detector may be optimized to detect the VOC biomarkers identified in this study. Members of our group have developed point-of-care instruments that collect, concentrate, and analyze breath VOCs for approximately one tenth the cost of laboratory ATD/GC/MS. They are currently initiating a follow-up study of breath testing for active pulmonary TB employing a point-of-care instrument with an ATD/GC/surface acoustic wave detector that costs approximately \$20,000. This point-of-care breath test might eventually provide a highly cost-effective primary screening tool for the detection of active pulmonary TB, with a low per-test cost in a busy clinical setting.

For this reason, an important objective of future research will be to adapt the breath test to a less expensive analytical platform that delivers results within minutes at the point-of-care. This objective has yet to be fulfilled, though its feasibility is supported by recent advances in breath testing technology.^{34–36}

We conclude that a breath test detected apparent volatile biomarkers of active pulmonary TB, and that this test has potential applications in clinical practice.

Acknowledgements

This research was supported by SBIR award R44 AI52504-03 from NIH/NIAID. Michael Phillips is President and CEO of Menssana Research, Inc.; he had access to and takes responsibility for the integrity of the data and the accuracy of the data analysis. Other coauthors have no financial interest in the subject matter or any conflict of interest. Michael Phillips designed and supervised the study and drafted the manuscript. Victoria Basa-Dalay, Graham Bothamley, Phung Kim Lam, and Maria Piedad R. Natividad recruited patients and collated data at the clinical research sites, Renee N Cataneo analyzed the samples in the laboratory, and Peter Schmitt, and James Wai analyzed the data statistically. James J. Grady, DrPH (Office of Biostatistics, The University of Texas Medical Branch at Galveston) provided independent statistical review of the report. All authors made

significant contributions to the study and approve the manuscript.

Funding: None.

Competing interests: None declared.

Ethical approval: Not required.

References

1. Anon. World TB Day – March 24, 2008. *MMWR* 2008;**57**(11):281.
2. Zar HJ. Global paediatric pulmonology: out of Africa. *Paediatr Respir Rev* 2006;**7**(Suppl. 1):S226–S228.
3. Perkins MD, Cunningham J. Facing the crisis: improving the diagnosis of tuberculosis in the HIV era. *J Infect Dis* 2007;**196**(Suppl. 1):S15–S27.
4. Cho SN, Brennan PJ. Tuberculosis: diagnostics. *Tuberculosis (Edinb)* 2007;**87**(Suppl. 1):S14–S17.
5. Pavlou AK, Turner AP. Sniffing out the truth: clinical diagnosis using the electronic nose. *Clin Chem Lab Med* 2000;**38**(2):99–112.
6. Zhang Y, Zhuang Y, Liu Z, Ruan J. Identification of twenty-eight species mycobacteria with their cellular fatty acids by capillary gas chromatography. *Wei Sheng Wu Xue Bao* 1991;**31**(3):187–97.
7. Syhre M, Chambers ST. The scent of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 2008;**88**(4):317–23.
8. Phillips M, Cataneo RN, Greenberg J, Gunawardena R, Naidu A, Rahbari-Oskoui F. Effect of age on the breath methylated alkane contour, a display of apparent new markers of oxidative stress. *J Lab Clin Med* 2000;**136**(3):243–9.
9. Kharitonov SA, Barnes PJ. Biomarkers of some pulmonary diseases in exhaled breath. *Biomarkers* 2002;**7**(1):1–32.
10. Phillips M, Altorki N, Austin JH, et al. Detection of lung cancer using weighted digital analysis of breath biomarkers. *Clin Chim Acta; Int J Clin Chem* 2008;**393**(2):76–84.
11. Phillips M, Boehmer J, Cataneo R, et al. Heart allograft rejection: detection with breath alkanes in low levels (the HARDBALL study). *J Heart Lung Transpl* 2004;**23**:701–8.
12. Braden B, Lembcke B, Kuker W, Caspary WF. 13C-breath tests: current state of the art and future directions. *Dig Liver Dis* 2007;**39**:795–805.
13. Phillips M. Method for the collection and assay of volatile organic compounds in breath. *Anal Biochem* 1997;**247**(2):272–8.
14. Phillips M, Altorki N, Austin JH, et al. Prediction of lung cancer using volatile biomarkers in breath. *Cancer Biomark* 2007;**3**(2):95–109.
15. Phillips M, Cataneo RN, Condos R, et al. Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis (Edinb)* 2007;**87**(1):44–52.
16. Suzuki S. Conditional relative odds ratio and comparison of accuracy of diagnostic tests based on 2 × 2 tables. *J Epidemiol/Jpn Epidemiol Assoc* 2006;**16**(4):145–53.
17. Phillips M, Greenberg J, Awad J. Metabolic and environmental origins of volatile organic compounds in breath. *J Clin Pathol* 1994;**47**(11):1052–3.
18. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem* 2008;**54**(1):17–23.
19. Zou YY, Yang J, Zhu J. A robust statistical procedure to discover expression biomarkers using microarray genomic expression data. *J Zhejiang Univ Sci* 2006;**7**(8):603–7.
20. Ma S, Song X, Huang J. Regularized binormal ROC method in disease classification using microarray data. *BMC Bioinform* 2006;**7**:253.
21. Phillips M, Altorki N, Austin J, et al. Detection of lung cancer using weighted digital analysis of breath biomarkers. *Clin Chim Acta* 2008;**393**:76–84.
22. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005;**365**(9458):488–92.
23. 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**(2):127–60.
24. Lamsal M, Gautam N, Bhatta N, Toora BD, Bhattacharya SK, Baral N. Evaluation of lipid peroxidation product, nitrite and antioxidant levels in newly diagnosed and two months follow-up patients with pulmonary tuberculosis. *Southeast Asian J Trop Med Public Health* 2007;**38**(4):695–703.
25. Kwiatkowska S, Piasecka G, Zieba M, Piotrowski W, Nowak D. Increased serum concentrations of conjugated dienes and malondialdehyde in patients with pulmonary tuberculosis. *Respir Med* 1999;**93**(4):272–6.
26. Gatner EM, Burkhardt KR. Correlation of the results of X-ray and sputum culture in tuberculosis prevalence surveys. *Tubercle* 1980;**61**(1):27–31.
27. Wilcke JT, Kok-Jensen A. Diagnostic strategy for pulmonary tuberculosis in a low-incidence country: results of chest X-ray and sputum cultured for *Mycobacterium tuberculosis*. *Respir Med* 1997;**91**:281–5.
28. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Inform* 2005;**38**:404–15.
29. Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc Natl Acad Sci U S A* 1971;**68**:2374–6.
30. Phillips M, Herrera J, Krishnan S, Zain M, Greenberg J, Cataneo RN. Variation in volatile organic compounds in the breath of normal humans. *J Chromatogr B Biomed Sci Appl* 1999;**729**(1–2):75–88.

31. Stern NJ, Kotula AW, Pierson MD. Differentiation of selected Enterobacteriaceae by pyrolysis–gas–liquid chromatography. *Appl Environ Microbiol* 1979;**38**(6):1098–102.
32. Stern NJ, Kotula AW, Pierson MD. Virulence prediction of *Yersinia enterocolitica* by pyrolysis gas–liquid chromatography. *Appl Environ Microbiol* 1980;**40**(3):646–51.
33. Catanzaro A, Lam P, Bothamley G, et al. Using clinical suspicion of tuberculosis (CSTB) to evaluate a rapid lateral flow (RLF) serologic immunoassay. *Am J Respir Crit Care Med* 2007;**175**:A576.
34. Laakso O, Haapala M, Jaakkola P, et al. The use of low-resolution FT-IR spectrometry for the analysis of alcohols in breath. *J Anal Toxicol* 2000;**24**(4):250–6.
35. Ueno M, Shinada K, Yanagisawa T, et al. Clinical oral malodor measurement with a portable sulfide monitor. *Oral Dis* 2008;**14**(3):264–9.
36. Hockstein NG, Thaler ER, Lin Y, Lee DD, Hanson CW. Correlation of pneumonia score with electronic nose signature: a prospective study. *Ann Otol Rhinol Laryngol* 2005;**114**(7):504–8.
37. FDA website for Heartsbreath: <http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/DeviceApprovalsandClearances/Recently-ApprovedDevices/ucm081213.htm>.
38. Prasad R, Lath SK, Mukerji PK, Agrawal SK, Srivastava R. Clinical utility of polymerase chain reaction in patients of pulmonary tuberculosis. *Ind J Tub* 2001;**48**:135–8.