Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples

Nicola M. Zetola[1,2,3], **Chawangwa Modongo**[3], **Ogopotse Matsiri**[3], **Tsaone Tamuhla**[3], **Bontle Mbongwe**[4], **Keikantse Matlhagela**[2], **Enoch Sepako**[2], **Alexandro Catini**[5], **Giorgio Sirugo**[6], **Eugenio Martinelli**[5], **Roberto Paolesse**[7], and **Corrado Di Natale**[5]

[1] Department of Radiation Oncology, University of Pennsylvania, Philadelphia, USA

[2] School of Medicine, University of Botswana, Gaborone, Botswana

[3] Botswana-UPenn Partnership, University of Pennsylvania, Gaborone, Botswana

[4] Department of Environmental Sciences, University of Botswana, Gaborone, Botswana

[5] Department of Electronic Engineering, University of Rome Tor Vergata, Rome, Italy

[6] Centro di Ricerca, Ospedale San Pietro Fatebenefratelli, Rome, Italy

[7] Department Chemical Science and Technology, University of Rome Tor Vergata, Rome, Italy

Abstract

Objectives—We determined the performance of a sensor array (an electronic nose) made of 8 metalloporphyrins coated quartz microbalances sensors for the diagnosis and prognosis of pulmonary tuberculosis (TB) using exhaled breath samples.

Correspondence: Nicola M. Zetola, PO Box AC 157 ACH, Gaborone, Botswana. Tel: (+267) 355 4855. Fax: (+267) 393 2267. nzetola@gmail.com; and, **Corrado Di Natale**, Department of Electronic Engineering, University of Rome Tor Vergata via del Politecnico 1, 00133 Roma, Italy. Tel: (+39) 0672597347. dinatale@uniroma2.it.

ntungwana@yahoo.co.uk
ogopotsem@bup.org.bw
tsaone.tamuhla@gmail.combontle
mbongwe@mopipi.ub.bw
matlhagelak@mopipi.ub.bw
enoch.sepako@mopipi.ub.bw
catini@ing.uniroma2.it
gsirugo@gmail.com
martinelli@ing.uniroma2.it
roberto.paolesse@uniroma2.it

Methods—TB cases and healthy controls were prospectively enrolled. Signals from volatile organic compounds (VOCs) in breath samples were measured at days 0, 2, 7, 14, and 30 of TB therapy and correlated with clinical and microbiological measurements.

Results—51 pulmonary TB cases and 20 healthy HIV-uninfected controls were enrolled in the study. 31 (61%) of the 51 pulmonary TB cases were coinfected with HIV. At day 0 (before TB treatment initiation) the sensitivity of our device was estimated at 94.1% (95% confidence interval [CI], 83.8-98.8%) and specificity was 90.0% (95% CI, 68.3-98.8%) for distinguishing TB cases from controls. Time-dependent changes in the breath signals were identified as time on TB treatment progressed. Time-dependent signal changes were more pronounced among HIV-uninfected patients.

Conclusion—The identification of VOCs signals in breath samples using a sensor array achieved high sensitivity and specificity for the diagnosis of TB and allowed following signal changes during TB treatment.

Keywords

Tuberculosis; HIV/AIDS; biomarkers; diagnosis; volatile organic compounds; sensors

INTRODUCTION

The natural course of tuberculosis (TB) disease is the result of highly dynamic changes at the human- and mycobacterial-level. Quantitative and qualitative variations of metabolites produced by host and pathogen could provide important information on the disease and its prognosis. Since many metabolic products are small molecules, volatile organic compounds (VOCs) analysis has the potential for providing a direct reflection of the physiological status of host-pathogen interactions (1).

Although VOCs could be detected in any clinical sample, analysis of VOCs in exhaled breath is particularly appealing for the diagnosis of pulmonary TB. The exhaled-breath test is painless and noninvasive, and therefore an ideal sample for children and critically ill patients (2). Breath samples are simple to collect and analyze, allowing for frequent measurements over time aimed at determining clinical progression or treatment response. Further, available sensor technology also allows exhaled-breath analysis at the point-of-care circumventing the need for high? infrastructure and expensive programs in low-income countries (3).

Broadly speaking, the analyses of VOCs for diagnostic or prognostic purposes could use two complementary, not mutually exclusive approaches. On one hand, one could target the identification and quantification of specific VOCs produced by the host or the mycobacteria. The identification of individual VOCs in a gas mixture is achieved through gas chromatography or mass spectroscopy (GC-MS) techniques. To date, most studies aimed at establishing the relationship between VOCs and infections, including *Mycobacterium tuberculosis* (Mtb) infection, have been carried out using this approach (4-14). GC–MS techniques are sensitive and reproducible and are currently recognized as the gold standard in breath VOC tests. However, such approaches are limited by technical complexity, high

costs, and the need for pre-processing of the samples before testing. In spite of recent advances in miniaturization, GC–MS technology is still not mobile and cannot perform real-time measurements, making it not suitable for point-of-care use (15). Alternatively, VOC analysis could be aimed at identifying the composite, unique signal resulting from the exposure of a complex mixture of VOCs to arrays of sensors (also called "electronic noses") (16, 17). After exposure to a complex VOC mixture, the physical or chemical properties of the sensors change. Those changes are then quantified into in a unique breath signal (or "breathprint"). Pattern recognition algorithms may then sort the data into classes discriminating those pertaining to specific microorganisms or diseases (18). In principle, identification of breath signals does not require identification of individual molecular constituents.

Very few studies have looked into the performance of electronic noses for TB diagnosis and prognosis. Fend *et al* showed that a conducting polymer array could be used for detecting Mtb in sputum samples and cultures (19). Other studies have successfully used sensors made of organically-capped gold nanoparticles for the diagnosis of *Mycobacterium bovis* in cattle and, more recently, for the diagnosis of pulmonary TB in humans (20). Those studies support the use of sensors for TB diagnostics and highlight the need for determining the most suitable sensor technology for the application.

In this study we exploited the sensing properties of metalloporphyrin-based sensors. Given their critical role in a wide variety of biological processes, metalloporphyrins are among the most suitable molecular platform for sensor arrays to capture the VOCs profiles produced by host and pathogens (21-23). Metalloporphyrins-based sensor arrays have been demonstrated to be sufficiently stable, sensitive and selective to diagnose different pathologies, such as lung cancer (23-25). We hypothesized that patients with active pulmonary TB will have different VOC signatures in breath samples compared with healthy controls. To test this hypothesis we first determined the diagnostic performance of our electronic nose comprised by an array of metalloporphyrins coated quartz microbalances for pulmonary TB. Then, we determined its performance monitoring TB treatment response.

METHODS

Study design

Patients with a new diagnosis of pulmonary TB (cases) and healthy controls were enrolled in the study. VOCs were measured on the patient's breath on days 0, 2, 7, 14, and 30 of therapy. The diagnostic performance of the device was assessed using data obtained at baseline (day 0, before TB treatment initiation). Follow up measurements were used to determine time-dependent changes in the breath signals after TB treatment initiation.

Participants

Participants were recruited prospectively at the Princess Marina Hospital in Gaborone, Botswana between May and October 2015. Following the World Health Organization recommendations for the development of diagnostic tests for TB, this study was conducted as an early proof-of-principle evaluation to demonstrate that our device can distinguish

patients with symptomatic, microbiologically confirmed pulmonary TB from healthy controls with reliable reproducibility (26). Accordingly, all pulmonary TB cases were adults (21 years of age or older) with: 1) symptoms and signs suggestive of active pulmonary disease; 2) an abnormal chest radiograph with findings consistent with pulmonary TB; and, 3) acid fast staining bacilli (AFB) seen on microscopy of sputa. To decrease the confounding effect of smoking over VOC signatures, self-reported smokers were not eligible for enrolment (27, 28). All cases were started on first-line anti-TB treatment (rifampin, isoniazid, ethambutol and pyrazinamide) on the day of enrolment, in accordance to Botswana national TB guidelines (29). HIV testing was performed on all cases without known HIV infection. Cases with new diagnosis of HIV infection and those with known HIV co-infection who were not on antiretroviral (ARV) therapy were started on ARV treatment within the first 2 weeks after TB treatment initiation according to Botswana national guidelines (30). Healthy controls were enrolled if they had no known history of respiratory disease, were asymptomatic at the time of assessment, were not taking medications and were non-smokers by self-report. All participants were tested for HIV test at the time of enrolment and only those who were HIV-uninfected were eligible as healthy controls for this study. Treatment success was defined as resolution of symptoms and negative sputum cultures at 2 months of TB treatment initiation.

Sputum samples and *M. tuberculosis* cultures

Induced sputum samples were obtained from all participants within 1 hour of obtaining consent (day 0, baseline) and at early hours of the morning during follow up appointments (days 2, 7, 14 and 30). Sputum samples were cultured in mycobacterial growth indicator tubes (MGIT). Routine antimicrobial susceptibility to the four primary TB drugs was carried out on all positive cultures. All clinical microbiology studies were performed at our TB research laboratory in Gaborone, Botswana.

Breath sampling and processing

Breath samples were obtained at the same time and under the same conditions for induced sputum samples. No food or drinks were allowed at least 2 h prior to collection of exhaled breath. Breath samples were collected by forced expiration of five breaths into a 2 L gas-sampling bulb (Supelco, Bellefonte, PA, USA). A three-way valve (QuinTron Instruments, Milwaukee, Wisc., USA), was used to separate the air expected to come from dead space from alveolar air (24). On each breath the first part (approx. 0.5 L) of the exhaled breath was discharged, while the remaining part of the breath was collected in the tedlar bag. Some rest was allowed between breaths but all breath samples were collected within 2 min. Collection bulbs were taken to our research laboratory within 1 hour after collection. Our procedures allowed all breath samples to be processed within 2 hours after collection. Each breath sample was analyzed 3 consecutive times and the average of those measurements was considered as the measure of the day for that patient.

The sensor array (electronic nose)

Given the pilot nature of this study, we designed a "general purpose" electronic nose for the diagnosis of TB using quartz microbalance (QMB) gas sensors. QMBs are mass sensors, which mean that slight changes of the mass of the sensitive film (?m) on the quartz surface

results in changes of frequency (?f) of the electrical output signal (31, 32). The QMBs had a fundamental frequency of 20MHz which corresponds to a mass resolution of the order of a few nanograms. The free base of the 5,10,15,20-tetrakis-(4-butyloxyphenyl) porphyrin (TBPPH2) and corresponding metal complexes (TBPPCu, TBPPCo, TBPPZn, TBPPMg, TBPPMnCl, TBPPFeCl, TBPPSnCl2) were used to functionalize the QMBs composing the array. The coordinated metal ions were purposely selected to cover a wide spectrum of VOCs and those molecules were prepared following previously reported methods (33). Thin films of sensing materials were deposited by a spray-coating on both sides of the quartz disks. The sensors were housed in a stainless steel measurement chamber having a volume of 10 mL. Each sensor was connected to an individual oscillator circuit. Frequencies were measured by means of an integrated frequency counter and then stored on a computer. Sensors were calibrated measuring their response to a series of compounds representative of different chemical families including propionic acid, ethanol, triethylamine, hexane, toluene, and dimethysulfide (31). During each measurement session, the sensors were continually kept under a constant flow of reference air. The sensor response was evaluated as the frequency shift between the signal frequency measured immediately before the exposure and at the end of the exposure. The relative change of frequency for each of the 8 sensors was captured as raw data, which was subsequently used in the data analysis. We purposely performed VOCs analyses of individuals on the same day, at different points during follow up (e.g. participants on days 0 of follow up were assessed on the same days as participants on day 0, 2, 7, 14, or 30 of follow up) to decrease potential bias due to sensors drift.

Statistical analysis

Wilcoxon–Mann–Whitney U two-sample test and the Kruskal-Wallis test were used to compare the medians of 2 or more independent groups, as appropriate. Significance was defined as a value of $p < 0.05$. Principal component analyses (PCA) and pattern recognition analysis was employed to study the relationship between the sensors data and the pulmonary TB diagnosis (18). First, an exploratory evaluation using PCA was performed to investigate data structure and similarities between subjects and between. Then, a *k*-Nearest Neighbours (*k*-NN) classification model was calculated. k-NN is a simple algorithm that classifies new cases based on a similarity measure of k nearest neighbours selected among a pool of training examples. Confidence intervals (CI) were determined generating 10,000 random permutations of class labels and evaluating the statistical distribution of the classification rate obtained in each permutation. The classification model was internally validated using a permutation analysis and tested on the treatment data. A binary classifier was assessed, then the results of the model are either TB or control, the decision is taken according to the similarity of sensors signals to each class template. Results are coherent with the expected behavior of treatment. Namely, as the time progresses, the sensor signals are less similar to TB and more similar to controls.

Masking procedures

Clinicians and microbiologists who collected and cultured sputum samples had no knowledge of the breath test results. Breath samples were analysed in the laboratory without knowledge of the sputum smears or culture results.

Human subjects

This study was approved by the Institutional Review Boards of the University of Pennsylvania and the University of Botswana and by the Human Research and Development Committee of the Botswana Ministry of Health. All participants provided written informed consent for their participation.

RESULTS

Demographic and clinical characteristics of the participants

During the study period, 51 cases with microbiologically confirmed pulmonary TB and 20 healthy controls were enrolled (Table 1). 31 (60.8%) pulmonary TB cases were known to be co-infected with HIV. Of them, 17 (54.8%) were already on ARV treatment and were virologically suppressed at the time of enrolment. ARVs were initiated within the first 14 days after TB treatment on all remaining HIV-coinfected patients. CD4 cell counts ranged from 15 to 596 cells/mm^3 with a median of 301.2 (interquartile range [IQR], 121.5 – 457.5) cells/mm^3. All pulmonary TB cases were eventually confirmed to have culture-positive infections and all isolates were drug sensitive.

Assessment of the diagnostic performance of the sensor array

All sensors included in the array captured the baseline difference in breath VOCs between cases and controls. Figure 1 shows the distribution of the sensors signals from each of the 8 sensors included in the array at day 0. The signals produced by cases and controls were statistically different and with similar effect sizes in all sensors. However, there was some level of overlap between the distribution of such signals, suggesting that no individual sensor would be enough to fully differentiate cases and controls. The differentiation between the sensor responses produced by cases and controls on individual sensors remained significant even after aggregating all data points obtained during the course of TB treatment of cases (days 0, 2, 7, 14 and 30; Supplementary Figure 1). Within cases, the overall distribution of signals at baseline was not significantly different when stratified by HIV status.

Figure 2 shows the plot of baseline data projected in the plane of the first two principal components (PCs) where a near complete separation is achieved. The performance of the classification model based on k-NN is shown in Table 2. An optimal model limited to 8 neighbors was selected as a result of a leave-one-out cross validation procedure (15). Sensitivity was estimated at 94.1% (95% confidence interval [CI], 83.8-98.8%) and specificity was 90.0% (95% CI, 68.3-98.8%).

Assessment of the effect of TB treatment

All 51 pulmonary TB cases showed treatment success according to the study definitions. Same as what we observed at baseline, each sensor significantly contributed to determining time-dependent variations in the signals during the course of TB treatment. We observed a progressive decrease in the intensity of the signals produced by each sensor as the time from initiation of TB treatment increased (Figure 3A). In other words, as time on TB treatment progressed the signals produced by each individual sensor change following the treatment process. Although we found no statistical differences in the signals produced by HIV-

infected and uninfected cases at baseline, the negative trend in the signals produced by both individual sensors and the array were more pronounced among HIV-uninfected (Figures 3B and 3C).

Figure 4 shows the plot of the first two PCs of the complete set of data. The figure shows signals progress towards thoset of controls as the time in TB treatment increases. As expected, the largest separation between cases and controls occurred at baseline (day 0). Interestingly, while the signals obtained on day 30 of TB therapy have significant overlap with controls, the overlap is not complete indicating that differences in VOC patterns remain existent. The classification model was then tested on the TB cases measured from day 2 onwards. The percentage of TB prediction decreased as time on treatment increased indicating that measurements over time were sensitive to the effect of TB therapy (Supplementary Figure 2). The decay was different for HIV positive and HIV negative participants, suggesting a correlation between HIV infection and the evolution of TB signals under therapy.

DISCUSSION

Breath analyses has the potential to revolutionize TB diagnostics by providing cheap, sensitive and specific point-of-care technologies that provide real-time information on both metabolic activity of the mycobacterium and immune response produced by the host. In this study, we confirmed that VOC analysis in breath samples offers great potential as a diagnostic and prognostic approach for pulmonary TB. We first demonstrated that our sensor array has high sensitivity and specificity differentiating patients with pulmonary TB from healthy controls, indicating its potential as a diagnostic tool. Then, we demonstrated that our sensor array can identify longitudinal changes occurring during TB treatment, indicating its potential to monitor response to treatment.

The overall diagnostic performance of a sensor array depends on the sensing properties of the individual sensors receptors included in the array (13, 16). Individual sensors in the array can cross-react with more than one group of VOCs. Therefore, each composite breath signal captures spectra of partially overlapping VOCs. As a consequence, sensor arrays could be tailored to capture either large or narrow ranges of chemicals, allowing the optimization of the array for specific applications. In the case of pulmonary TB, the selection of sensors for any given array will determine the balance between several desirable characteristics of a diagnostic device and its diagnostic performance. "General purpose" devices for the initial screening and differentiation of pulmonary diseases into broad categories such as cancers, infections and inflammatory conditions will likely require a different array of sensors than devices designed for confirmation of a specific disease or even metabolically different stages of a given disease. Consistently, development of an electronic nose for the diagnosis of TB is particularly challenging given that different stages of disease (initial infection, latency and active disease) are determined by dynamic metabolic changes within pathogen and host, as well as interactions between them. Therefore, a sensor array for TB diagnosis should be able to react with VOCs produced by both the pathogen and the host and identify signal changes as the TB disease transitions from one stage to another (34).

Prior studies have identified specific VOCs associated with Mtb such as methyl-nicotinate, methyl-p-anisate, methyl phenylacetate, o-phenylanisole and nicotinic acid(12-14). Metal ions in metalloporphyrins strongly determine the response of porphyrin-based chemicals sensors (23). Theoretical chemistry have shown that the affinity between those Mtb-related VOCs and metal ions of different valency (particular Zn^{2+}, Cu^{2+}, and Co^{2+}) can be used for TB diagnosis (35).We used this knowledge to guide the selection of the sensors used in this study. While an in-depth characterization of specific Mtb- or human-related VOCs producing the sensor signals was out of the scopes of this study, the behavior of our sensors was fully consistent with the theoretical behavior of metal ions complexed in an organic framework. Besides, the metal ion is only one of the sensitive part of the porphyrin ring and then additional response might be expected.

In this early stage study, we also demonstrated that VOC analysis is a promising approach for following the response to TB treatment. In our study, sequential measurements during the course of TB treatment show a progressive departure from the signals produced at baseline, moving towards those produced by healthy controls. This observed trend suggests that the mixture of VOCs present in breath changes as TB treatment progresses, and could be used as a marker to follow treatment response. Interestingly, our longitudinal measurements suggests that the signals produced by HIV-infected and uninfected patients may proceed at a different pace, even in the setting of adequate TB and HIV treatments (36, 37). Even if not quantitatively expressed, our findings are consistent with prior observations that demonstrate not only different inflammatory responses to Mtb between HIV-infected and uninfected patients but also, different pace of resolution of immune responses even in the setting of appropriate treatment (38). Given that the VOC mixture in breath reflects host-pathogen interactions as they occur in real-time, we believe that VOC analyses also offers potential as a prognostic tool by allowing the identification of patients with VOC patters consistent with either decreased or exaggerated immune responses to Mtb infections and their changes while on treatment.

Our results need to be interpreted in the context of the limitations of our study design. First, our study should be considered an early-phase evaluation of the performance and reliability of our electronic nose for pulmonary TB diagnosis and prognosis as recommended by New Diagnostics Working Group of the Stop TB Partnership (26). The diagnostic performance of our device was determined using populations representing two extremes of the clinical spectrum of the natural history of TB disease: healthy, asymptomatic controls and symptomatic, smear positive pulmonary TB cases before treatment initiation and, thus, our study sample fails to represent the full spectrum of TB disease. Therefore, our results cannot be generalized to other populations or stages of disease. Failure to include cases representing the entire spectrum of disease may have also introduced "spectrum bias" into our study, which frequently leads to overly optimistic results (39). Late evaluation studies aimed at measuring our device's performance in real-life populations and settings are required.

It is important to remark that sensors capture a combination of quantitative and qualitative information about the composition of the sample. As a consequence, different breath compositions may result in similar signals. Similarly, our study design cannot determine causality between pulmonary TB and the VOC patterns identified by our device both for

diagnosis and prognostic purposes (31). However, we found a strong association between signals and TB disease, as well as resolution of breath signals associated with TB treatment response, increasing our confidence that such signals are related to metabolic events occurring within TB disease and TB treatment. In a prior study, we have also shown that the same sensor array can successfully differentiate a wide variety of bacteria and fungi in-vitro (34). Future studies should aim to determine the performance of our sensor array for differentiating patients with pulmonary TB from those with other pulmonary infections, as well as the identification of those patients co-infected with pulmonary TB and bacterial pneumonia (34).

The sample size of our study allowed the determination of significant differences between cases and controls with regard to VOC signals at baseline. However, given the limited size of our study it is possible that we were not able to identify differences between subgroups of TB patients included as cases. Particularly important is the relative overrepresentation of HIV-coinfected TB cases in our sample. Our analysis did not show any differences in VOC signals between HIV-infected and uninfected TB cases at baseline, suggesting that the diagnostic performance of our device for untreated HIV-infected and uninfected TB cases could be similar. However, we also showed that longitudinal changes and trends in VOC signals during TB treatment among HIV-infected and HIV-uninfected TB cases were different. While it is possible that our device performs differently as a diagnostic tool as opposed to as a tool destined to follow treatment response, it is equally possible that the lack of differences between HIV-infected and uninfected cases at baseline were due to insufficient power. Along the same lines, the effect of ARV treatment and time-to-ARV treatment initiation over our results is unknown and may have introduced further bias.

Since patients entered in the test at different time, the experimental design avoided the influence of any sensors drift. Finally, our study focuses on the recognition of "breathprints" or patterns of VOCs without caring for the specific compounds that produce those patterns. More precise determination of the specific VOCs leading to different signals at different points in the disease may allow for further optimization of the sensors to be included in the array. Consequently, specific arrays for specific purposes or for the diagnosis of specific stages of disease could be developed and, potentially, combined to achieve better diagnostic and prognostic performance.

In summary, our study demonstrates that breath VOC analyses using an array of partially selective sensors has good diagnostic accuracy for pulmonary TB in HIV-infected and HIV-uninfected patients. Data from studies aimed at better understanding of underlying pathophysiological pathways should be used for optimizing the selection of sensors to target the group of VOCs that provides the highest diagnostic and prognostic information for PTB, while allowing discrimination and diagnosis of co-existing infections and medical conditions. Late stage, independent validation of our classification models in different populations and different stages of the disease needs to be performed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

1. Tang Z, Liu Y, Duan Y. Breath analysis: technical developments and challenges in the monitoring of human exposure to volatile organic compounds. J Chromatogr B Analyt Technol Biomed Life Sci. 2015; 1002:285–99.

2. Nicol MP, Gnanashanmugam D, Browning R, Click ES, Cuevas LE, Detjen A, et al. 2015; 61(Suppl 3):S164–72.

3. Denkinger CM, Pai M. Point-of-care tuberculosis diagnosis: are we there yet? Lancet Infect Dis. 2012; 12(3):169–70. [PubMed: 22015304]

4. Mourao MP, Kuijper S, Dang NA, Walters E, Janssen HG, Kolk AH. Direct detection of Mycobacterium tuberculosis in sputum: A validation study using solid phase extraction-gas chromatography-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci. 2015; 1012-1013:50–4.

5. Dang NA, Mourao M, Kuijper S, Walters E, Janssen HG, Kolk AH. Direct detection of Mycobacterium tuberculosis in sputum using combined solid phase extraction-gas chromatography-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci. 2015; 986-987:115–22.

6. Dang NA, Kuijper S, Walters E, Claassens M, van Soolingen D, Vivo-Truyols G, et al. Validation of biomarkers for distinguishing Mycobacterium tuberculosis from non-tuberculous mycobacteria using gas chromatography-mass spectrometry and chemometrics. PLoS One. 2013; 8(10):e76263. [PubMed: 24146846]

7. Nicoara SC, Minnikin DE, Lee OC, O'Sullivan DM, McNerney R, Pillinger CT, et al. Development and optimization of a gas chromatography/mass spectrometry method for the analysis of thermochemolytic degradation products of phthiocerol dimycocerosate waxes found in Mycobacterium tuberculosis. Rapid Commun Mass Spectrom. 2013; 27(21):2374–82. [PubMed: 24097393]

8. O'Sullivan DM, Nicoara SC, Mutetwa R, Mungofa S, Lee OY, Minnikin DE, et al. Detection of Mycobacterium tuberculosis in sputum by gas chromatography-mass spectrometry of methyl mycocerosates released by thermochemolysis. PLoS One. 2012; 7(3):e32836. [PubMed: 22403716]

9. Cha D, Cheng D, Liu M, Zeng Z, Hu X, Guan W. Analysis of fatty acids in sputum from patients with pulmonary tuberculosis using gas chromatography-mass spectrometry preceded by solid-phase microextraction and post-derivatization on the fiber. J Chromatogr A. 2009; 1216(9):1450–7. [PubMed: 19171347]

10. Dorneanu O, Wittmer A, Diculencu D, Pelz K. Gas chromatographic detection of Mycobacterium tuberculosis complex in pure cultures from respiratory specimens. Rev Med Chir Soc Med Nat Iasi. 2000; 104(4):161–5.

11. Phillips M, Cataneo RN, Condos R, Ring Erickson GA, Greenberg J, La Bombardi V, et al. Volatile biomarkers of pulmonary tuberculosis in the breath. Tuberculosis (Edinb). 2007; 87(1):44–52. [PubMed: 16635588]

12. Syhre M, Chambers ST. The scent of Mycobacterium tuberculosis. Tuberculosis (Edinb). 2008; 88(4):317–23. [PubMed: 18296120]

13. Syhre M, Manning L, Phuanukoonnon S, Harino P, Chambers ST. The scent of Mycobacterium tuberculosis--part II breath. Tuberculosis (Edinb). 2009; 89(4):263–6. [PubMed: 19481976]

14. Nawrath T, Mgode GF, Weetjens B, Kaufmann SH, Schulz S. The volatiles of pathogenic and nonpathogenic mycobacteria and related bacteria. Beilstein J Org Chem. 2012; 8:290–9. [PubMed: 22423297]

15. Hahgighi F, Talebpour Z, Sanati-Nezhad A. Through the years with a on-a-chip gas chromatography: a review Lab Chip. 2015; 15:2559–57. [PubMed: 25994317]

16. Rock F, Barsan N, Weimar U. Electronic nose: current status and future trends. Chem Rev. 2008; 108(2):705–25. [PubMed: 18205411]

17. Turner AP, Magan N. Electronic noses and disease diagnostics. Nat Rev Microbiol. 2004; 2(2): 161–6. [PubMed: 15040263]

18. Leopold JH, Bos LD, Sterk PJ, Schultz MJ, Fens N, Horvath I, et al. Comparison of classification methods in breath analysis by electronic nose. J Breath Res. 2015; 9(4):046002. [PubMed: 26669708]

19. Fend R, Kolk AH, Bessant C, Buijtels P, Klatser PR, Woodman AC. Prospects for clinical application of electronic-nose technology to early detection of Mycobacterium tuberculosis in culture and sputum. J Clin Microbiol. 2006; 44(6):2039–45. [PubMed: 16757595]

20. Nakhleh MK, Jeries R, Gharra A, Binder A, Broza YY, Pascoe M, et al. Detecting active pulmonary tuberculosis with a breath test using nanomaterial-based sensors. Eur Respir J. 2014; 43(5):1522–5. [PubMed: 24789953]

21. Di Natale C, Paolesse R, D'Amico A. Metalloporphyrins based artificially olfactory receptors. Sens Actuators B. 2007; 121:238–46.

22. Capuano R, Pomarico G, Paolesse R, Di Natale C. Corroles-porphyrins: a teamwork for gas sensor arrays. Sensors (Basel). 2015; 15(4):8121–30. [PubMed: 25856324]

23. Paolesse R, Nardis S, Monti D, Stefanelli M, Di Natale C. Porphyrinoids for chemical sensor applications. Chemical Reviews. 2016 doi 10.1021/acs.chemrev.6b00361.

24. Capuano R, Santonico M, Pennazza G, Ghezzi S, Martinelli E, Roscioni C, et al. The lung cancer breath signature: a comparative analysis of exhaled breath and air sampled from inside the lungs. Sci Rep. 2015; 5:16491. [PubMed: 26559776]

25. D'Amico A, Pennazza G, Santonico M, Martinelli E, Roscioni C, Galluccio G, et al. An investigation on electronic nose diagnosis of lung cancer. Lung cancer. 2010; 68(2):170–6. [PubMed: 19959252]

26. Organization. WH. Pathways to better diagnostics for tuberculosis: a blueprint for the development of TB diagnostics. 1st. World Health Organization; Geneva, Switzerland: 2009. 2009. http://wwwstoptborg/assets/documents/research/2009pdf

27. Gaida A, Holz O, Nell C, Schuchardt S, Lavae-Mokhtari B, Kruse L, et al. A dual center study to compare breath volatile organic compounds from smokers and non-smokers with and without COPD. J Breath Res. 2016; 10(2):026006. [PubMed: 27082437]

28. Allers M, Langejuergen J, Gaida A, Holz O, Schuchardt S, Hohlfeld JM, et al. Measurement of exhaled volatile organic compounds from patients with chronic obstructive pulmonary disease (COPD) using closed gas loop GC-IMS and GC-APCI-MS. J Breath Res. 2016; 10(2):026004. [PubMed: 27058460]

29. Botswana National Tuberculosis Programme Report. 2009

30. Government of Botswana. Botswana National HIV & AIDS Treatment Guidelines.

31. Sobel RM, Ballantine DS. 2D bitmapping approach for identification and quantitation of common base flavor adulterants using surface acoustic wave arrays and artificial neural network data analysis. Analytica chimica acta. 2008; 608(1):79–85. [PubMed: 18206997]

32. Rose-Pehrsson SL, Grate JW, Ballantine DS Jr. Jurs PC. Detection of hazardous vapors including mixtures using pattern recognition analysis of responses from surface acoustic wave devices. Anal Chem. 1988; 60(24):2801–11. [PubMed: 3245603]

33. Buchler, J. Synthesis and properties of metalloporphyrins. In: Dolphin, D., editor. The Porphyrins. Vol. 1. Academy Press; New York, NY, USA: 1978.

34. Zetola NM, Modongo C, Mathlagela K, Sepako E, Matsiri O, Tamuhla T, et al. Identification of a Large Pool of Microorganisms with an Array of Porphyrin Based Gas Sensors. Sensors (Basel). 2016; 16(4)

35. Ray R, Sarma B, Mohanty S, Prisbrey K, Misra M. Assessment of metals in detection of TB biomarkers: novel computational approach. Materials Chemistry and Physics. 2015; 161:1–8.

36. Lawn SD, Obeng J, Acheampong JW, Griffin GE. Resolution of the acute-phase response in West African patients receiving treatment for pulmonary tuberculosis. Int J Tuberc Lung Dis. 2000; 4(4):340–4. [PubMed: 10777083]

37. Djoba Siawaya JF, Bapela NB, Ronacher K, Veenstra H, Kidd M, Gie R, et al. Immune parameters as markers of tuberculosis extent of disease and early prediction of anti-tuberculosis chemotherapy response. J Infect. 2008; 56(5):340–7. [PubMed: 18359089]

38. Bell LC, Breen R, Miller RF, Noursadeghi M, Lipman M. Paradoxical reactions and immune reconstitution inflammatory syndrome in tuberculosis. Int J Infect Dis. 2015; 32:39–45. [PubMed: 25809754]

39. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med. 1978; 299(17):926–30. [PubMed: 692598]

HIGHLIGHTS

- Analysis of volatile organic compounds (VOCs) in breath samples using sensor technology (electronic nose) is a promising non-invasive approach for the diagnosis and prognosis of pulmonary tuberculosis (TB)

- In this proof-of-concept study, our electronic nose showed high performance differentiating cases with pulmonary TB from healthy controls

- Serial measurements of VOCs also allowed for determining VOC signals change during TB treatment among patients with pulmonary TB

- The sensors used in our study proved to be stable enough to allow for longer, larger-scale studies of different designs looking into the diagnostic and prognostic value of VOC analysis for the diagnosis and prognosis of pulmonary TB during different stages of disease and within different populations
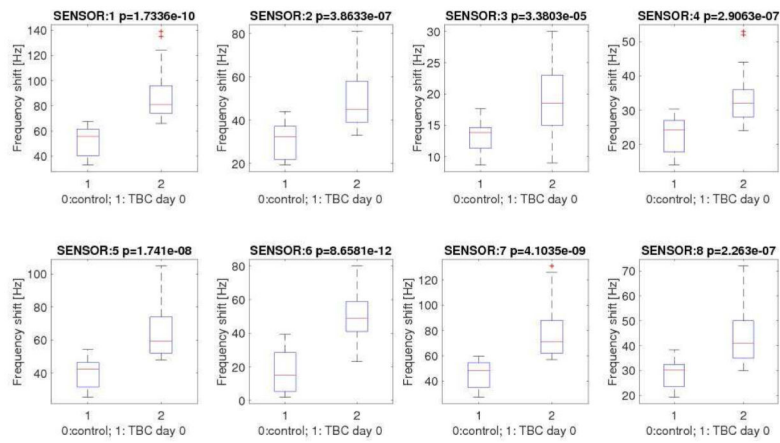
Figure 1.
Distribution of sensors signals among controls and cases with pulmonary tuberculosis at baseline (day 0, before treatment initiation).

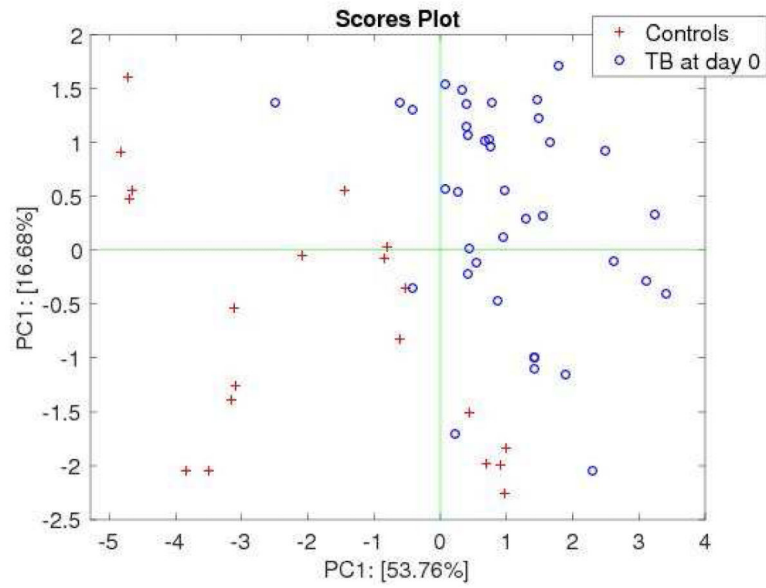**Legend:** * p values are reported in the header of each plot

Figure 2.
Plot of the first two principal components of cases with pulmonary tuberculosis and controls at baseline (before treatment initiation)

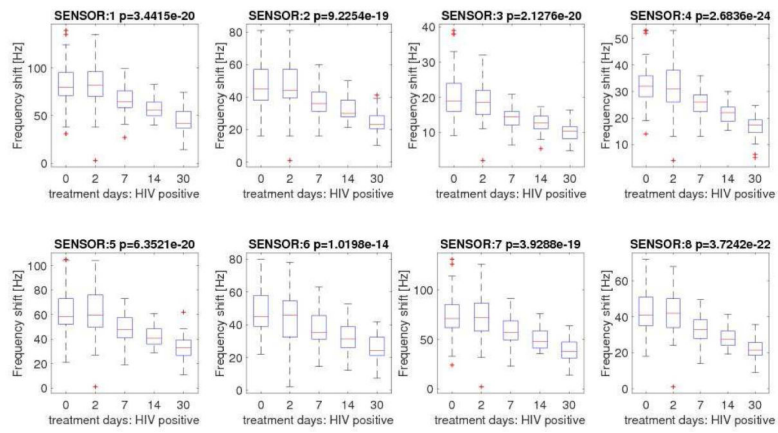**Legend:** * p values are reported in the header of each plot

Figure 3A.
Distribution of sensors signals among cases with pulmonary tuberculosis at baseline and days 2, 7, 14 and 30 after tuberculosis treatment initiation
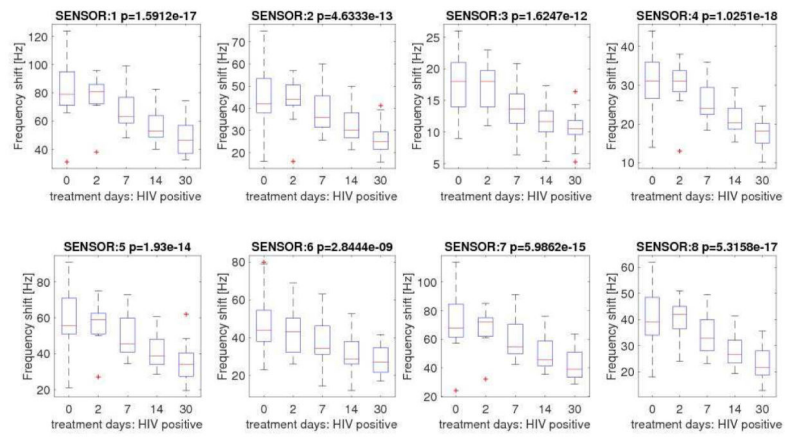
Figure 3B.
Distribution of sensors signals among HIV-infected cases with pulmonary tuberculosis at baseline and days 2, 7, 14 and 30 after tuberculosis treatment initiation.
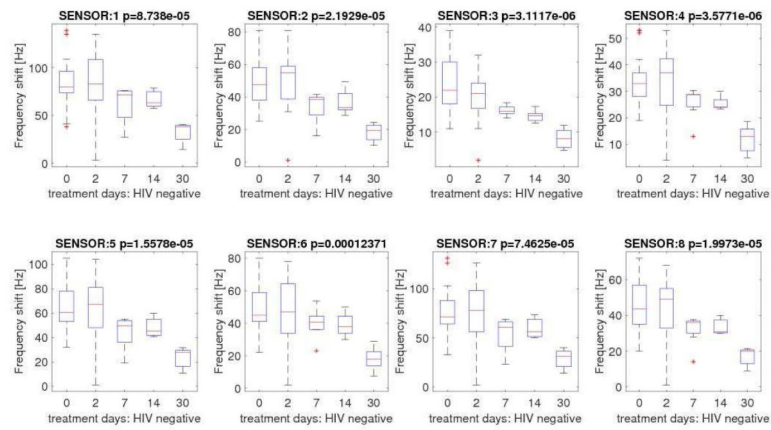
Figure 3C.
Distribution of sensors signals among HIV-uninfected cases with pulmonary tuberculosis at baseline and days 2, 7, 14 and 30 after tuberculosis treatment initiation.

Legend: * p-value of null hypothesis is reported in the header of each plot.
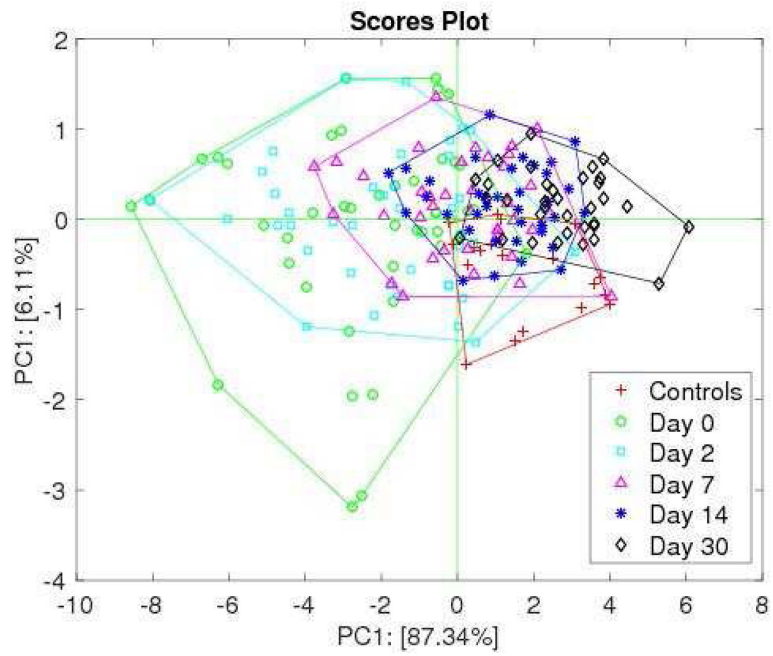
Figure 4.
Plot of the first two principal components of controls and pulmonary tuberculosis cases.
With respect to Fig. 2 here samples collected from day 2 to day 30 from the start of therapy
have been included in the data set.

Table 1

Demographic, clinical and microbiological characteristics of the participants

	Pulmonary tuberculosis cases			Healthy controls (n=20)
	All (n=51)	HIV infected (n=31)	HIV-uninfected (n=20)	
Age (mean ± SD), years	36.1 (±9.8)	28.7 (±7.2)	39 (±9.3)	33 (±11)
Sex (%)				
Male	30 (57.7%)	13 (41.9%)	16 (81.0%)	7 (40%)
Semi-quantitative AFB in sputum				
Scanty	9 (17.7%)	7 (22.6%)	2 (10.0%)	
1+	6 (11.7%)	3 (9.7%)	3 (15.0%)	
2+	12 (23.5%)	6 (19.4%)	6 (30.0%)	
3+	24 (47.1%)	15 (48.4%)	9 (45.0%)	
CD4 cells/mm3 (median, IQR)		301.2 (121.5 – 457.5)		
On HIV antiretroviral treatment at the time of TB diagnosis (n, %)		17 (55%)		

Table 2

Confusion matrix of the k-Nearest Neighborhood classifier in training.

		Classification according to the volatile compound analysis (at day 0)		Total
		Negative	**Positive**	
Disease status	**Healthy controls***	18	2	20
	Pulmonary tuberculosis	3	48	51
	Total	21	50	

*
20 healthy controls were enrolled and 3 measurements per control are included in the analysis. Sensitivity was estimated at 94.1% (95% confidence interval [CI], 83.8-98.8%) and specificity was 90.0% (95% CI, 68.3-98.8%).