12-1-2021

Breath can discriminate tuberculosis from other lower respiratory illness in children

Carly A. Bobak
Thayer School of Engineering at Dartmouth

Lili Kang
Thayer School of Engineering at Dartmouth

Lesley Workman
Red Cross War Memorial Children's Hospital

Lindy Bateman
Red Cross War Memorial Children's Hospital

Mohammad S. Khan
Thayer School of Engineering at Dartmouth

See next page for additional authors

scientific reports

Check for updates

OPEN

Breath can discriminate tuberculosis from other lower respiratory illness in children

Carly A. Bobak[1,2], Lili Kang[1,7], Lesley Workman[3,7], Lindy Bateman[3], Mohammad S. Khan[1], Margaretha Prins[3], Lloyd May[1], Flavio A. Franchina[1,4], Cynthia Baard[3], Mark P. Nicol[5,6], Heather J. Zar[3,7] & Jane E. Hill[1,7]✉

Pediatric tuberculosis (TB) remains a global health crisis. Despite progress, pediatric patients remain difficult to diagnose, with approximately half of all childhood TB patients lacking bacterial confirmation. In this pilot study (n = 31), we identify a 4-compound breathprint and subsequent machine learning model that accurately classifies children with confirmed TB (n = 10) from children with another lower respiratory tract infection (LRTI) (n = 10) with a sensitivity of 80% and specificity of 100% observed across cross validation folds. Importantly, we demonstrate that the breathprint identified an additional nine of eleven patients who had unconfirmed clinical TB and whose symptoms improved while treated for TB. While more work is necessary to validate the utility of using patient breath to diagnose pediatric TB, it shows promise as a triage instrument or paired as part of an aggregate diagnostic scheme.

Tuberculosis (TB) is a leading cause of childhood mortality, with an estimated one million cases and 250,000 deaths reported each year[1–3]. While an accurate appraisal of underdiagnosis for childhood TB is unavailable, modelling estimates indicate that only 30% of childhood TB cases are diagnosed and notified[2]. Diagnosing pediatric TB is challenging. Children present with non-specific clinical symptoms, the available tests have poor sensitivity in this population, and there is often a lack of expertise and infrastructure available to obtain microbiologic confirmation in young children[1,4,5]. Even in well-resourced areas, the diagnostic yield from microbiological specimens is sub-optimal, with approximately 50% of pediatric patients diagnosed with TB not having bacterial confirmation[5]. Children co-infected with Human Immunodeficiency Virus (HIV) are particularly challenging to diagnose as the clinical presentation of TB is non-specific and immune deficiency often modifies the clinical presentation of TB disease[6]. The sensitivity of cartridge-based nucleic acid amplification assays for *Mycobacterium tuberculosis*, such as Xpert MTB/RIF and Xpert MTB/RIF Ultra (Xpert, Cepheid, Sunnyvale, California), in children is low: 62% and 75.2% respectively. Moreover, such specimens often rely on induced sputum, which, while safe, cheap and well tolerated, may be difficult to do in children in health care facilities especially in low and middle income country settings[7,8]. Moreover, induced sputum relies on trained personnel and laboratory infrastructure to test samples[8].

National Institutes of Health consensus guidelines for diagnostic studies of TB in children, classify children as having 'confirmed TB' (positive culture or Xpert MTB/RIF test for *M. tuberculosis*), 'unconfirmed TB' (negative microbiological results, but clinically diagnosed and treated for TB) or 'unlikely TB' (negative microbiologic investigations, not clinically diagnosed with PTB and improvement in the absence of TB treatment)[9]. Consistently, approximately half of pediatric patients diagnosed with TB fall into the unconfirmed TB category in studies[4,5]. There is a clear need for improved diagnostics for children, particularly for those in the unconfirmed TB category[3,10].

Previously, we and others have demonstrated that exhaled breath can be collected, analyzed, and mined for putative biomarkers for TB in adults[10–26]. The APOPO non-profit organization have demonstrated that African

[1]Thayer School of Engineering, Dartmouth College, Hanover, NH, USA. [2]Geisel School of Medicine, Dartmouth College, Hanover, NH, USA. [3]Department of Pediatrics and Child Health, MRC Unit on Child and Adolescent Health, University of Cape Town and Red Cross War Memorial Children's Hospital, Cape Town, South Africa. [4]Molecular Systems, Organic and Biological Analytical Chemistry Group, University of Liège, Liège, Belgium. [5]Division of Medical Microbiology and Institute for Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa. [6]School of Biomedical Sciences, University of Western Australia, Perth, Australia. [7]These authors contributed equally: Lili Kang and Lesley Workman; Heather J. Zar and Jane E. Hill. ✉email: jane.hill@ubc.ca

		Confirmed TB (n = 10)	Unconfirmed TB (n = 11)	Unlikely TB (n = 10)	Overall (n = 31)
Age (years)	Mean (SD)	6.3 (3.5)	5.64 (3.4)	6.2 (2.5)	6.0 (3.1)
Sex	Male	6 (60%)	5 (45%)	5 (50%)	16 (52%)
Weight-for-age	Underweight	1 (10%)	3 (27.3%)	4 (40%)	8 (25.8%)
Height-for-age	Stunted	1 (10%)	3 (27.3%)	3 (30%)	7 (22.6%)
HIV-infected	Positive	1 (10%)	0 (0%)	2 (20%)	3 (9.7%)
Tuberculin skin test	Positive	6 (60%)	8 (72.7%)	0 (0%)	14 (45.2%)
	Negative	2 (20%)	1 (9.1%)	10 (100%)	13 (41.9%)
	Missing	2 (20%)	2 (18.2%)	0 (0%)	4 (12.9%)

Table 1. Demographic and clinical characteristics across the TB study group.

giant pouched rats can be trained to 'sniff' TB, resulting in a substantial increase in case detections, even in children[27–41]. In this pilot study we investigate whether exhaled breath from children has diagnostic utility in a pilot South African study of pediatric patients diagnosed with confirmed TB, unconfirmed TB, or unlikely TB.

Results and discussion

Study population. Of the 31 children recruited, 10 (32.3%) were confirmed to have TB disease, 11 (35.4%) had unconfirmed TB, and 10 (32.3%) were unlikely to have TB disease. All confirmed TB patients had at least one positive culture or positive Xpert MTB/RIF test result. Two of these patients had TB confirmed on a cervical lymph node aspirate. Unlikely TB patients had both a negative culture and Xpert MTB/RIF test, and were diagnosed with a lower respiratory tract infection (LRTI) not due to TB. The mean (SD) age of the children was 6 (3.1) years and did not differ significantly by TB category (p value 0.874). Overall, 16 (52%) of patients were male, 8 (26%) were underweight for their age, and 7 were stunted (23%). These compositions did not vary significantly by TB category (p values 0.997, 0.305, and 0.507 respectively). There was 1 (10%) HIV-infected child in the confirmed TB category, 0 in the unconfirmed TB category and 2 (20%) in the unlikely TB category (p value 0.301). A positive tuberculin skin test occurred in 14 children (45.2%); this differed across TB category wherein 60% of confirmed TB patients, 73% of unconfirmed TB patients, and 0% of unlikely TB patients had positive tuberculin skin tests. Demographic and clinical characteristics of the study groups are shown in Table 1. More details about each patient can be found in Supplementary Table 1.

Four compounds in breath characterize children with a confirmed TB diagnosis from unlikely TB patients with an alternate lower respiratory tract infection. A Boruta feature selection algorithm was used to identify all relevant compounds for the task of classifying confirmed TB from unlikely TB patients. Four compounds were consistently ranked as more important than shadow features over 84 iterations (results shown in Supplementary Fig. 1). These analytes comprise: decane and 4-methyloctane (identities confirmed by comparing both retention indices and mass spectra with authentic standards) as well as two analytes (labelled Analyte A and B), whose retention times and mass spectra are consistent, but for which we could not find a suitable analytical standard for mass spectral confirmation. Chromatographic and mass spectral information on the four compounds is found in Supplementary Tables 2 and 3 and Supplementary Figs. 6–10.

The distribution of the each breathprint compound across confirmed TB and unlikely TB groups using normalized chromatographic peak area is shown in Fig. 1. Despite the small sample size, Analyte A and 4-methyloctane are statistically significant at a cut off of $\alpha = 0.1$ ($p = 0.052$ and $p = 0.023$, respectively). Decane and Analyte B did not reach statistical significance, however, different medians across both groups are observable. The unlikely TB group encapsulates a spectrum of non-TB lower respiratory tract infection (LRTI) cases, therefore it is reasonable to expect greater heterogeneity. The totality of these results supports the hypothesis that a multivariate signature of breath compounds is necessary and should be a focus of investigation in follow up studies.

Machine learning procedures allow us to build a predictive model to evaluate how accurately the four compounds categorized patients according to TB status. Here, we evaluated a random forest model and a support vector machine model with a polynomial kernel using the four features selected with the Boruta algorithm[12,42,43]. Random forest performed best and is discussed further. The SVM model, while complementary to the results from random forest, had slightly lower performance (see Supplementary Table 4 and Supplementary Fig. 3).

The observed area under the receivor operating curve across cross validation folds using the random forest model, shown in Fig. 2, was 0.99 with a 95% confidence interval of (0.961, 1). To better interpret the area under the receiver operating characteristic curve, we selected four compounds in the data at random and repeated the model building process. The observed area under the receiving operating curve across cross validation folds with four randomly selected compounds was 0.595 (0.329, 0.861), clearly demonstrating the utility of the Boruta-selected signature. The WHO's guidelines for a TB triage test recommend a specificity of 75% and a sensitivity equal to that of Xpert MTB/RIF (62% in children)[44,45]. Across cross validation folds, we observed an accuracy 90%, sensitivity of 80% and specificity of 100%. More performance statistics for the final model can be found in Supplementary Table 5. These data suggest that the four compound candidate biomarkers for pediatric subjects could be a promising route to investigate further.

Biologically, there is evidence to suggest the four compound breathprint characterizes TB. Decane in breath has been previously associated with TB in adults, and is also linked to isoniazid resistance in

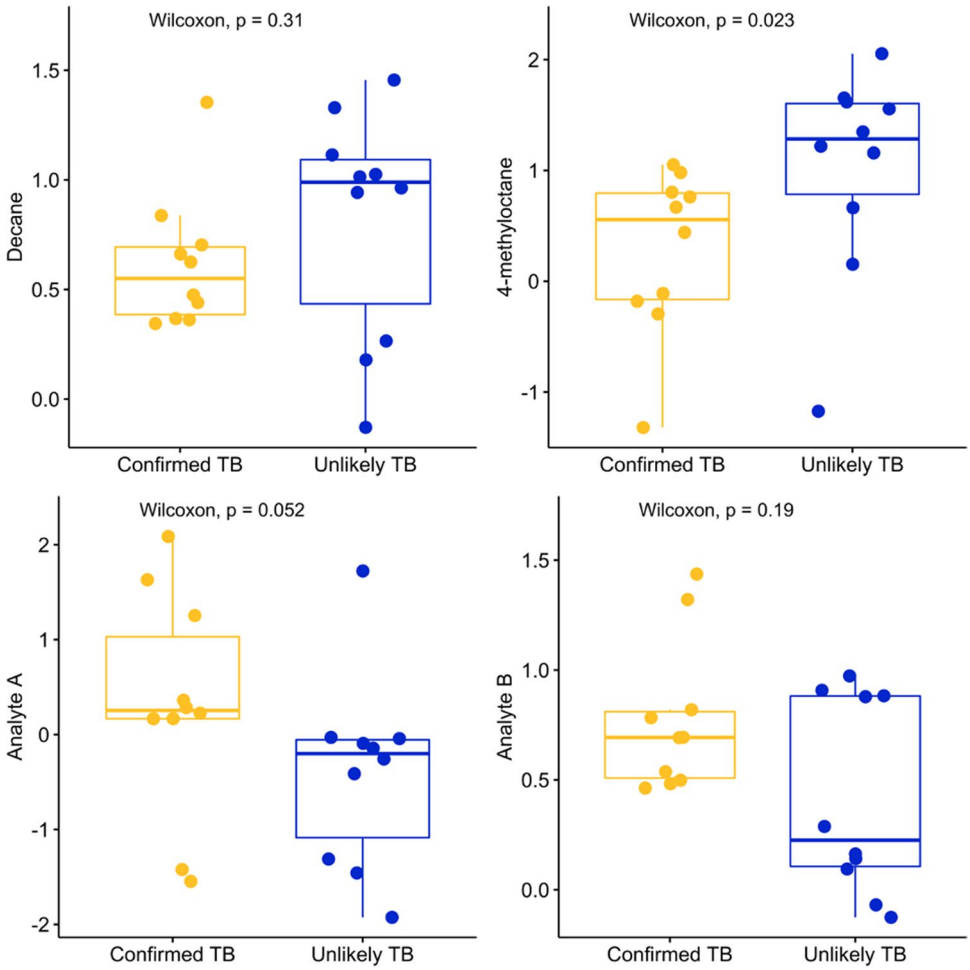

Figure 1. The distribution of the mean centered and normalized peak area of each of the four compounds selected in the breathprint across confirmed TB and unlikely TB patients. For each compound, the median observed peak area between the two groups is different, indicating univariate differences which may contribute to the discrimination of confirmed TB patients from unlikely TB patients. Boxplots show the quartiles of the data (first line is the first quartile, midline is the median, third line is the third quartile) where whiskers represent 1.5 × IQR (inter-quartile range). The distribution across all three TB groups is shown in Supplementary Fig. 2. Figure created in R[73] using 'ggplot2'[84] and 'ggpubr'[85].

Mtb[16,46]. Both decane and 4-methyloctane have also been identified as characterizing in the breath of adults with asthma, chronic obstructive pulmonary disease (COPD), or lung cancer compared to controls[47–52]. More research in a larger population will be necessary to quantify the differences in production of 4-methyloctane and decane across respiratory diseases.

Analyte A is a nitrogen-containing cyclic compound most likely to be a benzamide derivative. Benzamide derivatives have been detected in the breath of emphysema patients and smokers[53,54] and are often studied as possible inhibitors of *M. tuberculosis*[55,56]. Analyte B is an eleven-carbon alkene that is likely to be branched. Alkenes have been detected in the breath of patients with lung cancer[57]. To identify the precise molecular formula of analytes A and B, follow up studies will need to utilize a high-resolution mass-spectral instrument or equivalent. For more information about the chemical identity of analytes A and B, see Supplementary Table 3 and Supplementary Figs. 9,10.

The four compound breathprint classifies unconfirmed TB patients. Unconfirmed TB patients are suspected as having TB but do not have a positive culture or Xpert MTB/RIF test result. All unconfirmed TB patients in this study demonstrated improvement of symptoms and weight gain in response to TB treatment. Boxplots comparing the distribution of the mean centered and normalized peak area across the four analytes for each of the three TB categories is shown in Supplementary Fig. 2.

Using the four compound breathprint generated by the Boruta approach, the unconfirmed TB cases cluster closely to the confirmed TB cases and also share a similar pattern of relative compound presence in the breath samples (Fig. 3). Overall, 10 of the 11 unconfirmed TB patients cluster closely with the confirmed TB group and away from the unlikely TB patients. Patient 18 is the only unconfirmed patient that did not cluster closely with the

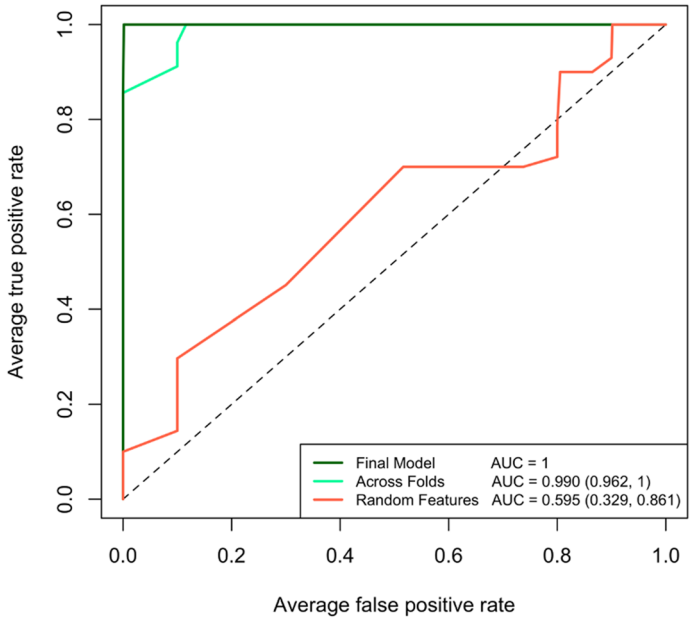

Figure 2. The receiver operating characteristic curves from the random forest model used to classify confirmed TB from unlikely TB patients. The final model demonstrates perfect classification but is almost certainly overfit to the data. The AUC observed across folds using the identified breathprint is 0.99, demonstrating very good sensitivity and specificity across all folds of the data. In comparison, a randomly selected 4-compound breathprint only demonstrated an AUC of 0.595 across cross validation folds of the data. Figure created in R[73].

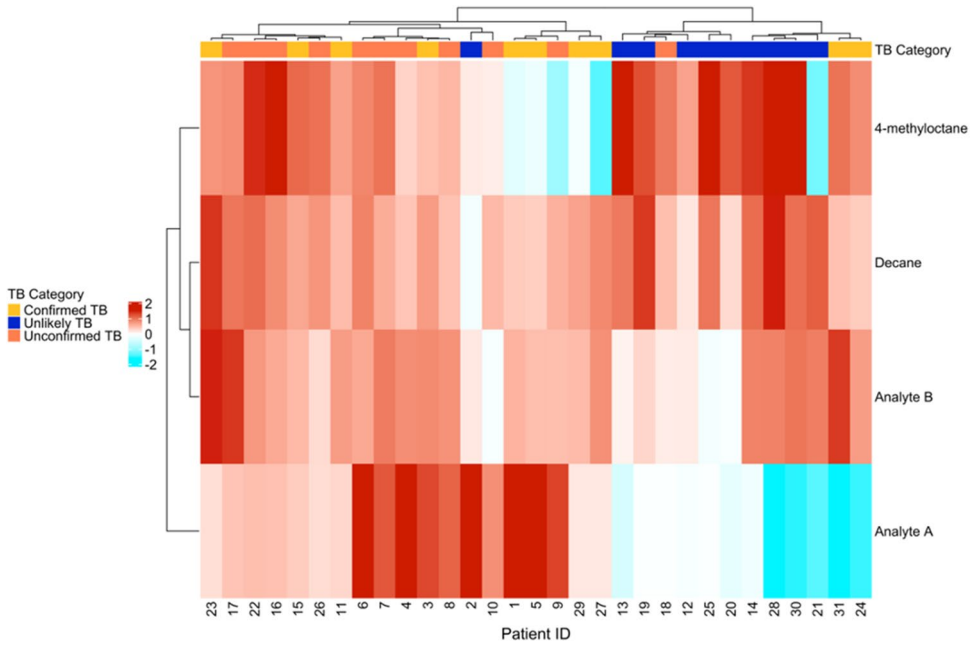

Figure 3. A dendrogram and heatmap demonstrating the unsupervised clustering of patients using the 4-compound breathprint. The annotation bar along the dendrogram indicates TB category. The heatmap shows the normalized peak area for each compound. Red indicates above average peak area, and blue indicates below average peak area. Figure created in R[73] using 'ComplexHeatmap'[86].

confirmed TB group. Both confirmed TB patients 31 and 24 also cluster away from the main TB disease group, which may reflect the occurrence of extra-pulmonary TB (patient 31 had Spinal TB and patient 24 had lymph node TB; see Supplementary Table 1 for more clinical information about the patients). Notably, we observe no clustering by HIV status; patients 20, 19, and 27 are all HIV positive and cluster primarily with their likely TB

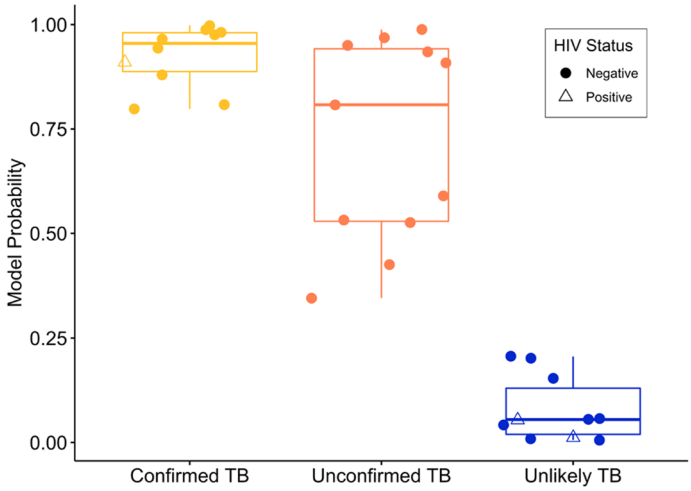

Figure 4. The output probabilities that each patient has TB disease from the random forest classifier across the TB categories. Patients with a probability of over 50% are assigned a label of having TB disease. Despite two unconfirmed TB patients having probabilities below 50%, there is clear differentiation in model probabilities between the unconfirmed and unlikely TB groups. Boxplots show the quartiles of the data (first line is the first quartile, midline is the median, third line is the third quartile) where whiskers represent 1.5 × IQR (inter-quartile range). Figure created in R[73] using 'ggplot2'[84] and 'ggpubr'[85].

category participants as opposed to each other. This suggests that our breathprint may be effective for patients with HIV co-infection. A larger cohort for further study will inform further interpretations with less speculation.

We evaluated how well the classification models described for confirmed TB and unlikely TB reference groups might predict the TB status of the unconfirmed TB group (Fig. 4). The random forest classifier correctly predicted 9 of the 11 unconfirmed TB patients (the results for the equivalent SVM analysis which also correctly classifies 9 of the unconfirmed TB patients are given in Supplementary Fig. 4). Six of those patients had probabilities indistinguishable from the confirmed TB cases. Despite two cases having model probabilities below 50%, there is obvious differentiation between the unconfirmed and unlikely TB categories. Specifically, the minimum probability among unconfirmed TB patients was 0.344, while the highest probability among unlikely patients is 0.242, indicating a TB cut-off between these two values exists that would perfectly classify every patient.

The clinical sensitivity of Xpert MTB/RIF is low in unconfirmed pediatric TB patients. If clinical diagnosis is considered the reference standard, the sensitivity of Xpert MTB/RIF in culture-negative samples from pediatric patients ranges from 4 to 15%[58]. Using the proposed 4-compound breathprint, the sensitivity among clinically diagnosed, but microbiologically-negative, pediatric patients is 82% (using a model probability cut-off value of 50%). Achieving 100% sensitivity and specificity is possible in this group if a model cut-off between 25 and 34% is used. Importantly, while confirmation status of patients in the unconfirmed TB patients in unavailable, all children in this study group demonstrated improvement of symptoms after completion of TB treatment. While this is preliminary data, the breathprint approach could be appealing as a clinically-relevant diagnostic tool for pediatric patients, especially to distinguish those with TB who have unconfirmed TB.

Previously, Zar and colleagues demonstrated an improvement in sensitivity of Ultra for culture confirmed TB disease in children by testing multiple samples for Ultra; a single induced sputum (sensitivity of 74.3%), two nasopharyngeal aspirates (sensitivity of an individual test is 46%) or combination of sputum and nasopharyngeal samples providing an overall sensitivity of 87.5%[4]. Given the 4-compound breathprint's sensitivity to both confirmed and unconfirmed pediatric TB cases, using it as a triage test prior to Ultra testing may further increase sensitivity in confirmed TB patients while adding further diagnostic evidence for unconfirmed TB patients.

These results, while positive, have limitations. The 4-compound breathprint may only applicable to mixed expiratory fixed-volume sampling method with patients breathing normally. Further evaluation will be needed if different breath sampling methods are used or different patient breathing patterns are employed, as some breath VOCs have been reported to be dependent on exhalation flow and the portion of the breath collected[59–64]. In addition, exhalation flow monitoring was not possible due to the design of our sampling kits. Sampling device with flow monitoring capabilities are currently under development in our laboratory. Further evaluation will be conducted when the flow monitoring sampling devices become available.

As a multi-center breath-analysis study, the effect of transportation and storage has always been a concern for breath samples using sorbent tubes, especially when no specific guideline has been established by European Respiratory Society (ERS)[65]. Other studies have indicated that the stability of breath compounds varies and may depend on sampling media (sorbent material), storage temperature and time, and the breath compositions[66–71]. Some molecules such as benzene, toluene and m-xylene are stable for 12 months on Tenax TA TD tubes[67], but in general, researchers suggested that analysis by day 14 in cold storage will minimize a potential 1–2 standard deviation gain or loss of VOC concentration[71]. For this and many other multi-center studies, sample analysis within 14 days of collection is usually not feasible. Integration of stability tests for novel breath molecules in

the current biomarker discovery study is even more challenging. Therefore, future independent studies on the transportation and storage stability of the 4-compound breathprint are required to ultimately validate this result.

While assessing performance statistics across cross validation folds gives a more accurate indication of generalization than the final model, it has been suggested that estimates originating from cross validation may still be overly optimistic[72]. Due to the pilot nature of this study, validation of these results across a larger sample is necessary. Indeed, a larger population would allow assessment of additional co-morbidities (such as diabetes, childhood asthma, and more robust analyses in HIV + children), spectrum of TB disease, and other population characteristics that could influence the predictive ability of the TB pediatric breathprint. Moreover, this study cannot conclude if these results will generalize to populations outside of South Africa. Future work should consider a multi-site study aimed at evaluating breath as a diagnostic medium for pediatric TB across many endemic countries. Furthermore, while the unconfirmed TB group had clinical symptoms and chest radiographs suggestive of TB disease, microbiological confirmation was negative. Although unconfirmed patients improved while undergoing TB treatment, a gold standard diagnosis of TB is not possible in this group. Finally, the study is underpowered to confidently propose the 4-compound breathprint and subsequent random forest model as clinical instruments to diagnose TB in children. However, we confidently conclude that breath as a medium for diagnosis of pulmonary TB in pediatric patients in conjunction with machine learning models is feasible, demonstrates clinical utility, and warrants further investigation.

Methods

Study subjects and design. Study subjects were recruited, diagnosed, and treated in a prospective clinical study described previously[4]. In short, consecutive children hospitalized between April 4th 2017 and December 14, 2017 in Cape Town, South Africa with suspected TB were enrolled. Study eligibility criteria were age less than 15 years, cough of any duration, and at least one of the following: a household TB contact within the previous 6 months, weight loss or failure to gain weight within the previous 3 months, a positive tuberculin skin test or a chest radiograph suggesting pulmonary TB. All children had a chest radiograph, a tuberculin skin test if there was no known previous TB diagnosis, and HIV testing when HIV status was unknown. TB therapy was initiated at the discretion of the treating doctor. Response to treatment was assessed at follow up at 1, 3 and 6 months by recording signs and symptoms.

Children were classified according to diagnostic categories: 'confirmed TB' (culture or Xpert positive for *Mtb*), unconfirmed TB' (microbiologically negative, clinically diagnosed) or 'unlikely TB' (microbiologically negative, not clinically diagnosed, no tuberculosis treatment given, and documented improvement at follow up).

The Research Ethics Committee of the Faculty of Health Sciences, University of Cape Town (#045/2008) and the Committee for the Protection of Human Subjects at Dartmouth College approved the study (STUDY00030329). All methods were performed in accordance with relevant guidelines and regulations and identifying information is not presented in this report. Informed consent was obtained through parents or legal guardians.

Breath collection kits and procedure. A mixed expiratory fixed-volume sampling method was used, following the guidelines from European Respiratory Society (ERS) technical standard for exhaled biomarkers in lung disease[65]. Mixed expiratory breath and room air samples were collected using kits and protocols at the time of study enrollment as described previously[12]. In short, kits consist of a 1.5L Tedlar bags with a drinking straw mouthpiece for patients to breath into. Patients rinse mouth with water, and then are asked to breathe normally into the bag until it is full. Breath is then drawn through a 13 mm, 0.22 µm PTFE filter and into 3-bed thermal desorption tubes (TDT), using a vacuum pump. All samples were collected at time of enrollment, prior to commencement of treatment. Samples were shipped from Cape Town South Africa to Hanover, New Hampshire, United States of America and stored at 4 °C. Samples were processed within 6 months of collection.

Analytical instrumentation and initial processing. The breath compounds were collected on the TDT and desorbed at 330 °C into a cryogenically cooled (-120 °C) inlet liner of a GC × GC-TOFMS instrument (LECO Corporation, MI, USA). After desorption, the inlet is rapidly heated from − 120 to 270 °C and the trapped breath compounds are transferred onto an Rxi-624Sil-MS/Stabilwax chromatography columns. The TOFMS collected spectra over the range of m/z 30–500 at a rate of 200 Hz. For peak findings, a signal-to-noise (S/N) cutoff was set at 50:1 (with a minimum of three apexing masses) in at least one chromatogram and a minimum of 20:1 S/N in all others. The NIST 11 library was used for the initial identification of the analytes. A chemical formula was assigned if the analytes matched the following three criteria, (1) high mass spectral match, (2) group separation based on the structural formula and (3) the EIC ionization patterns among all observed samples. To verify the chemical formulas of discriminatory features, authentic standards were purchased, spiked into blank thermal desorption tubes, and run using the same analytical method as the breath samples. Retention indices were determined using C8–C20 n-alkane standard solution for both sample runs and standard runs. If both mass spectra and retention index of a feature is matched with the standard, the chemical structure the feature is confirmed. Alkane Standard Solution C8-C20 (~ 40 mg/L each in hexane) was purchased from Supelco (Darmstadt, Germany) and stored at 4 °C. 4-Mehyloctane was purchased from Toronto Research Chemicals (North York, ON, Canada,) and stored at 4 °C. The analytes that were not given a formula did not match on any of the previous criteria. Possible contaminants are manually removed before further data analysis (see Supplementary Table 6 for details).

Statistical analysis. A brief summary of our data cleaning and feature reduction process is shown in Supplementary Fig. 5. All statistical analyses were conducted in R 3.6.1 (R Core Team, Vienna, Austria)[73]. Data

cleaning was followed as described previously[12]. In short, a frequency of observation (FOO) cutoff of 80% in either the confirmed TB or unlikely TB categories was implemented. Remaining features were normalized using PQN, $\log_{10}$ transformed, and mean centered. Missing values were imputed using a random forest imputation[74]. Features were further reduced using a Mann–Whitney U-test to find features that were significantly different between patients and room air (Benjamini–Hochberg adjusted p value <0.05)[75,76]. A Boruta feature selection scheme was then used to find features which could discriminate between confirmed TB and unlikely TB groups[77,78]. It is recommended that pilot studies employ a more forgiving statistical threshold given that they are underpowered and designed for exploratory rather than confirmatory analysis. It is often recommended that pilot studies report findings as significant at a 75–85% confidence level and do not adjust for multiple comparisons[79,80]. Here, we consider a significance level of α=0.1 for statistical significance of the selected features to balance the pilot nature of this work while remaining appropriately conservative for follow-up studies.

After features were selected, models were built using a fivefold cross validation (CV) scheme in the 'caret' package[81,82]. CV splits the data into 5 equal size pieces, builds a model on 4 of the five pieces, and tests it on the remaining piece. It then leaves a different piece out and repeats this process[82]. This allows for parameter tuning across the models, as well as gives an estimate of model generalizability by examining accuracy statistics across the left-out pieces. All performance statistics are reported based on their performance across validation folds as these are more representative of performance and less influenced by overfitting[72]. Many models are sensitive to class imbalance, so an up-sampling scheme was used to split the data[81].

We fit two models on the data, random forest and a polynomial support vector machine[43,83]. Random forest models build a 'forest' of 'decision trees' where features are selected randomly in each tree according to how well they split the data[83]. Polynomial support vector machines fit a polynomial hyperplane between groups of interest in n-dimensional space[43]. Both models were built to classify between confirmed and unlikely TB patients and then used to predict the TB status of unconfirmed TB patients.

Data availability

The datasets generated during and/or analyzed in the current study are available from the corresponding author on reasonable request.

Received: 11 September 2020; Accepted: 28 December 2020
Published online: 01 February 2021

References

1. Martinez, L. & Zar, H. J. Tuberculin conversion and tuberculosis disease in infants and young children from the Drakenstein Child Health Study: A call to action. *S. Afr. Med. J.* **108**, 247 (2018).
2. Dodd, P. J., Gardiner, E., Coghlan, R. & Seddon, J. A. Burden of childhood tuberculosis in 22 high-burden countries: A mathematical modelling study. *Lancet Glob. Heal.* **2**, e453–e459 (2014).
3. *WHO Global tuberculosis report 2018. WHO* (World Health Organization, Geneva, 2019).
4. Zar, H. J. *et al.* Tuberculosis diagnosis in children using Xpert ultra on different respiratory specimens. *Am. J. Respir. Crit. Care Med.* https://doi.org/10.1164/rccm.201904-0772OC (2019).
5. Connell, T. G., Zar, H. J. & Nicol, M. P. Advances in the diagnosis of pulmonary tuberculosis in HIV-infected and HIV-uninfected children. *J. Infect. Dis.* **204**(Suppl 4), S1151-8 (2011).
6. Fry, S. H., Barnabas, S. & Cotton, M. F. Tuberculosis and HIV—An update on the 'cursed duet' in children. *Front. Pediatr.* **7**, 159 (2019).
7. Seong, G. M., Lee, J., Lee, J. H., Kim, J. H. & Kim, M. Usefulness of sputum induction with hypertonic saline in a real clinical practice for bacteriological yields of active pulmonary tuberculosis. *Tuberc. Respir. Dis.* **76**, 163 (2014).
8. Sakashita, K. *et al.* Efficiency of the Lung Flute for sputum induction in patients with presumed pulmonary tuberculosis. *Clin. Respir. J.* **12**, 1503–1509 (2018).
9. Graham, S. M. *et al.* Clinical case definitions for classification of intrathoracic tuberculosis in children: An update. *Clin. Infect. Dis.* **61**, S179–S187 (2015).
10. Nicol, M. P. *et al.* Microbiological diagnosis of pulmonary tuberculosis in children by oral swab polymerase chain reaction. *Sci. Rep.* **9**, 10789 (2019).
11. Beccaria, M. *et al.* Preliminary investigation of human exhaled breath for tuberculosis diagnosis by multidimensional gas chromatography—Time of flight mass spectrometry and machine learning. *J. Chromatogr. B* **1074–1075**, 46–50 (2018).
12. Beccaria, M. *et al.* Exhaled human breath analysis in active pulmonary tuberculosis diagnostics by comprehensive gas chromatography-mass spectrometry and chemometric techniques. *J. Breath Res.* **13**, 016005 (2018).
13. Maiga, M. *et al.* Stool microbiome reveals diverse bacterial ureases as confounders of oral urea breath testing for Helicobacter pylori and *Mycobacterium tuberculosis* in Bamako, Mali. *J. Breath Res.* **10**, 036012 (2016).
14. Morozov, V. N. *et al.* Non-invasive approach to diagnosis of pulmonary tuberculosis using microdroplets collected from exhaled air. *J. Breath Res.* **12**, 036010 (2018).
15. Phillips, M. *et al.* Point-of-care breath test for biomarkers of active pulmonary tuberculosis. *Tuberculosis* **92**, 314–320 (2012).
16. Phillips, M. *et al.* Breath biomarkers of active pulmonary tuberculosis. *Tuberculosis* **90**, 145–51 (2010).
17. Phillips, M. *et al.* Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis* **87**, 44–52 (2007).
18. Sahota, A. S. *et al.* A simple breath test for tuberculosis using ion mobility: A pilot study. *Tuberculosis* **99**, 143–146 (2016).
19. Saktiawati, A. M. I., Putera, D. D., Setyawan, A., Mahendradhata, Y. & van der Werf, T. S. Diagnosis of tuberculosis through breath test: A systematic review. *EBioMedicine* **46**, 202–214 (2019).
20. Bruins, M. *et al.* Diagnosis of active tuberculosis by e-nose analysis of exhaled air. *Tuberculosis* **93**, 232–238 (2013).
21. Kolk, A. H. J. *et al.* Breath analysis as a potential diagnostic tool for tuberculosis. *Int. J. Tuberc. Lung Dis.* **16**, 777–782 (2012).
22. Nakhleh, M. K. *et al.* Detecting active pulmonary tuberculosis with a breath test using nanomaterial-based sensors. *Eur. Respir. J.* **43**, 1522–1525 (2014).
23. Mohamed, E. I. *et al.* Qualitative analysis of biological tuberculosis samples by an electronic nose-based artificial neural network. *Int. J. Tuberc. Lung Dis.* **21**, 810–817 (2017).
24. Zetola, N. M. *et al.* Diagnosis of pulmonary tuberculosis and assessment of treatment response through analyses of volatile compound patterns in exhaled breath samples. *J. Infect.* **74**, 367 (2017).
25. Coronel Teixeira, R. *et al.* The potential of a portable, point-of-care electronic nose to diagnose tuberculosis. *J. Infect.* **75**, 441–447 (2017).

26. McNerney, R. *et al.* Field test of a novel detection device for *Mycobacterium tuberculosis* antigen in cough. *BMC Infect. Dis.* **10**, 161 (2010).
27. Van Beek, S. C. *et al.* Measurement of exhaled nitric oxide as a potential screening tool for pulmonary tuberculosis. *Int. J. Tuberc. Lung Dis.* **15**, 66 (2011).
28. Mgode, G. F., Cox, C. L., Mwimanzi, S. & Mulder, C. Pediatric tuberculosis detection using trained African giant pouched rats. *Pediatr. Res.* **84**, 99–103 (2018).
29. Mgode, G. F. *et al. Mycobacterium tuberculosis* volatiles for diagnosis of tuberculosis by Cricetomys rats. *Tuberculosis* **92**, 535–542 (2012).
30. Weetjens, B. J. *et al.* African pouched rats for the detection of pulmonary tuberculosis in sputum samples. *Int. J. Tuberc. Lung Dis.* **13**, 66 (2009).
31. Mahoney, A. *et al.* Using giant African pouched rats to detect tuberculosis in human sputum samples: 2010 findings. *Pan Afr. Med. J.* **9**, 66 (2011).
32. Poling, A. *et al.* Tuberculosis detection by giant African pouched rats. *Behav. Anal.* **34**, 47–54 (2011).
33. Poling, A. *et al.* Active tuberculosis detection by pouched rats in 2014: More than 2,000 new patients found in two countries. *J. Appl. Behav. Anal.* **50**, 165–169 (2017).
34. Mulder, C. *et al.* Accuracy of giant African pouched rats for diagnosing tuberculosis: Comparison with culture and Xpert MTB/RIF. *Int. J. Tuberc. Lung Dis.* **21**, 1127–1133 (2017).
35. Mgode, G. F. *et al.* Mycobacterium genotypes in pulmonary tuberculosis infections and their detection by trained African giant pouched rats. *Curr. Microbiol.* **70**, 212–218 (2015).
36. Ellis, H., Mulder, C., Valverde, E., Poling, A. & Edwards, T. Reproducibility of African giant pouched rats detecting *Mycobacterium tuberculosis*. *BMC Infect. Dis.* **17**, 298 (2017).
37. Poling, A. *et al.* Using giant african pouched rats to detect human tuberculosis: A review. *Pan Afr. Med. J.* **21**, 66 (2015).
38. Mgode, G. F. *et al.* Diagnosis of tuberculosis by trained African giant pouched rats and confounding impact of pathogens and microflora of the respiratory tract. *J. Clin. Microbiol.* **50**, 274–280 (2012).
39. Reither, K. *et al.* Evaluation of giant African pouched rats for detection of pulmonary tuberculosis in patients from a high-endemic setting. *PLoS ONE* **10**, e0135877 (2015).
40. Mahoney, A. *et al.* Giant African pouched rats (*Cricetomys gambianus*) as detectors of tuberculosis in human sputum: Two operational improvements. *Psychol. Rec.* **63**, 583–594 (2013).
41. Poling, A. *et al.* Using giant African pouched rats to detect tuberculosis in human sputum samples: 2009 Findings. *Am. J. Trop. Med. Hyg.* **83**, 1308–1310 (2010).
42. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
43. Suykens, J. A. K. & Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **9**, 293–300 (1999).
44. Van Der Linden, D. *et al.* Xpert MTB/RIF ultra for tuberculosis testing in children: A mini-review and commentary. *Front. Pediatr.* **7**, 34 (2016).
45. van't Hoog, A. H. *et al.* Optimal triage test characteristics to improve the cost-effectiveness of the Xpert MTB/RIF assay for TB diagnosis: A decision analysis. *PLoS ONE* **8**, e82786 (2013).
46. Loots, D. T. An altered *Mycobacterium tuberculosis* metabolome induced by katG mutations resulting in isoniazid resistance. *Antimicrob. Agents Chemother.* **58**, 2144–9 (2014).
47. Chen, X. *et al.* Association of smoking with metabolic volatile organic compounds in exhaled breath. *Int. J. Mol. Sci.* **18**, 66 (2017).
48. Caldeira, M. *et al.* Profiling allergic asthma volatile metabolic patterns using a headspace-solid phase microextraction/gas chromatography based methodology. *J. Chromatogr. A* **1218**, 3771–80 (2011).
49. Van Berkel, J. J. B. N. *et al.* A profile of volatile organic compounds in breath discriminates COPD patients from controls. *Respir. Med.* **104**, 557–563 (2010).
50. Dent, A. G., Sutedja, T. G. & Zimmerman, P. V. Exhaled breath analysis for lung cancer. *J. Thorac. Dis.* **5**(Suppl 5), S540-50 (2013).
51. Cazzola, M. *et al.* Analysis of exhaled breath fingerprints and volatile organic compounds in COPD. *COPD Res. Pract.* **1**, 7 (2015).
52. Caldeira, M. *et al.* Allergic asthma exhaled breath metabolome: A challenge for comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* **1254**, 87–97 (2012).
53. Cristescu, S. M. *et al.* Screening for emphysema via exhaled volatile organic compounds. *J. Breath Res.* **5**, 046009 (2011).
54. Kushch, I. *et al.* Compounds enhanced in a mass spectrometric profile of smokers' exhaled breath versus non-smokers as determined in a pilot study using PTR-MS. *J. Breath Res.* **2**, 026002 (2008).
55. Naz, S. *et al.* Identification of new benzamide inhibitor against α-subunit of tryptophan synthase from *Mycobacterium tuberculosis* through structure-based virtual screening, anti-tuberculosis activity and molecular dynamics simulations. *J. Biomol. Struct. Dyn.* **37**, 1043–1053 (2019).
56. Joshi, S. D. *et al.* Pharmacophore mapping, molecular docking, chemical synthesis of some novel pyrrolyl benzamide derivatives and evaluation of their inhibitory activity against enoyl-ACP reductase (InhA) and *Mycobacterium tuberculosis*. *Bioorg. Chem.* **81**, 440–453 (2018).
57. Queralto, N. *et al.* Detecting cancer by breath volatile organic compound analysis: A review of array-based sensors. *J. Breath Res.* **8**, 027112 (2014).
58. Dunn, J. J., Starke, J. R. & Revell, P. A. Laboratory diagnosis of *Mycobacterium tuberculosis* infection and disease in children. *J. Clin. Microbiol.* **54**, 1434–1441 (2016).
59. Thekedar, B., Oeh, U., Szymczak, W., Hoeschen, C. & Paretzke, H. G. Influences of mixed expiratory sampling parameters on exhaled volatile organic compound concentrations. *J. Breath Res.* **5**, 66 (2011).
60. Boshier, P. R., Priest, O. H., Hanna, G. B. & Marczin, N. Influence of respiratory variables on the on-line detection of exhaled trace gases by PTR-MS. *Thorax* **66**, 919–920 (2011).
61. Bikov, A. *et al.* Standardised exhaled breath collection for the measurement of exhaled volatile organic compounds by proton transfer reaction mass spectrometry. *BMC Pulm. Med.* **13**, 43 (2013).
62. Lärstad, M. A. E., Torén, K., Bake, B. & Olin, A. C. Determination of ethane, pentane and isoprene in exhaled air—Effects of breath-holding, flow rate and purified air. *Acta Physiol.* **189**, 87–98 (2007).
63. Dragonieri, S. *et al.* An electronic nose in the discrimination of patients with asthma and controls. *J. Allergy Clin. Immunol.* **120**, 856–862 (2007).
64. Montuschi, P. *et al.* Diagnostic performance of an electronic nose, fractional exhaled nitric oxide, and lung function testing in asthma. *Chest* **137**, 790–796 (2010).
65. Horváth, I. *et al.* A European respiratory society technical standard: Exhaled biomarkers in lung disease. *Eur. Respir. J.* **49**, 66 (2017).
66. Van Der Schee, M. P. *et al.* Effect of transportation and storage using sorbent tubes of exhaled breath samples on diagnostic accuracy of electronic nose analysis. *J. Breath Res.* **7**, 016002 (2013).
67. Peters, R. J. B. & Bakkeren, H. A. Sorbents in sampling Stability and breakthrough measurements. *Analyst* **119**, 71–74 (1994).
68. Brown, V. M., Crump, D. R., Plant, N. T. & Pengelly, I. Evaluation of the stability of a mixture of volatile organic compounds on sorbents for the determination of emissions from indoor materials and products using thermal desorption/gas chromatography/mass spectrometry. *J. Chromatogr. A* **1350**, 1–9 (2014).

69. Patil, S. F. & Lonkar, S. T. Evaluation of Tenax TA for the determination of chlorobenzene and chloronitrobenzenes in air using capillary gas chromatography and thermal desorption. *J. Chromatogr. A* **684**, 133–142 (1994).
70. Gjølstad, M., Bergemalm-Rynell, K., Ljungkvist, G., Thorud, S. & Molander, P. Comparison of sampling efficiency and storage stability on different sorbents for determination of solvents in occupational air. *J. Sep. Sci.* **27**, 1531–1539 (2004).
71. Harshman, S. W. *et al.* Storage stability of exhaled breath on Tenax TA. *J. Breath Res.* **10**, 046008 (2016).
72. Tabe-Bordbar, S., Emad, A., Zhao, S. D. & Sinha, S. A closer look at cross-validation for assessing the accuracy of gene regulatory networks and models. *Sci. Rep.* **8**, 6620 (2018).
73. R Core Team. *R: A Language and Environment for Statistical Computing* (2019).
74. Pantanowitz, A. & Marwala, T. *Missing Data Imputation Through the Use of the Random Forest Algorithm*, in 53–62 (Springer, Berlin, 2009). https://doi.org/10.1007/978-3-642-03156-4_6
75. McKnight, P. E. & Najab, *J. Mann–Whitney U Test*, in *The Corsini Encyclopedia of Psychology* 1–1 (Wiley, New York, 2010). https://doi.org/10.1002/9780470479216.corpsy0524
76. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).
77. Kursa, M. B., Jankowski, A. & Rudnicki, W. R. Boruta—A system for feature selection. *Fundam. Inform.* **101**, 271–285 (2010).
78. Kursa, M. B. & Rudnicki, W. R. Feature selection with the Boruta package. *J. Stat. Softw.* **36**, 1–13 (2010).
79. Lee, E. C., Whitehead, A. L., Jacques, R. M. & Julious, S. A. The statistical interpretation of pilot trials: Should significance thresholds be reconsidered?. *BMC Med. Res. Methodol.* **14**, 41 (2014).
80. Althouse, A. D. Adjust for multiple comparisons? It's not that simple. *Ann. Thorac. Surg.* **101**, 1644–1645 (2016).
81. Kuhn, M. *et al.* caret: Classification and Regression Training (2018).
82. Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection (1995).
83. Hastie, T., Tibshirani, R. & Friedman, J. *Random Forests*, in 587–604 (2009). https://doi.org/10.1007/978-0-387-84858-7_15
84. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2016).
85. Kassambara, H. *ggpubr: 'ggplot2' Based Publication Ready Plots* (2020).
86. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

Acknowledgements

The authors thank the children who participated in the study, the children's carers, and the staff at Red Cross War Memorial Children's Hospital for their support. C.A.B. was supported by the Burroughs Wellcome Fund institutional program grant unifying population and laboratory based sciences to Dartmouth College (Grant#1014106). This study was also supported by the Bill and Melinda Gates Foundation (to J.E.H.). HZ is supported by the SA MRC. Funding for the TB diagnostic study is from the NIH. J.E.H is supported by the US NIH, the US Cystic Fibrosis Foundation, and the Australian MRC.

Author contributions

H.J.Z. and J.E.H. outlined the study design. H.Z. and M.N. wrote and obtained funding for the core TB study. J.E.H. obtained funding for the breath study. J.E.H., C.B., and H.J.Z. completed ethical approval processes. L.M. generated the breath collection kits. L.M. and L.B., M Prins, collected the breath samples. C.B. oversaw clinical aspects. L.J.W. was responsible for clinical data management. L.M. and F.A.F. conducted the experiments and acquired the data on the analytical instrumentation. M.S.K preprocessed the chromatographic data. C.A.B. conducted the statistical analyses and prepared all figures and tables. L.K. and M.S.K. performed biomarker identification activities on the 4-compound breathprint. C.A.B. completed first draft of the manuscript. C.A.B., L.K. and J.E.H. completed the bulk of the manuscript writing. C.A.B, L.K., M.S.K., F.A.F, H.Z. and J.E.H edited and revised the manuscript. All authors approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-80970-w.

Correspondence and requests for materials should be addressed to J.E.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.