Non-invasive phenotyping using exhaled volatile organic compounds in asthma

Baharudin Ibrahim,1,2 Maria Basanta,1,2 Paul Cadden,1,2 Dave Singh,1,2 David Douce,3 Ashley Woodcock,1,2 Stephen J Fowler1,2,4

ABSTRACT

Background Breath volatile organic compounds (VOCs) may be useful for asthma diagnosis and phenotyping, identifying patients who could benefit from personalised therapeutic strategies. The authors aimed to identify specific patterns of breath VOCs in patients with asthma and in clinically relevant disease phenotypes.

Methods Breath samples were analysed by gas chromatography—mass spectrometry. The Asthma Control Questionnaire was completed, together with lung function and induced sputum cell counts. Breath data were reduced to principal components, and these principal components were used in multiple logistic regression to identify discriminatory models for diagnosis, sputum inflammatory cell profile and asthma control.

Results The authors recruited 35 patients with asthma and 23 matched controls. A model derived from 15 VOCs classified patients with asthma with an accuracy of 89%, and positive and negative predictive values of 0.85 and 0.89, respectively. Models also classified patients with asthma based on the following phenotypes: sputum (obtained in 18 patients with asthma) eosinophilia ≥2% area under the receiver operating characteristics (AUROC) curve 0.98, neutrophilia ≥40% AUROC 0.90 and uncontrolled asthma (Asthma Control Questionnaire ≥1) AUROC 0.96.

Conclusions Detection of characteristic breath VOC profiles could classify patients with asthma versus controls, and clinically relevant disease phenotypes based on sputum inflammatory profile and asthma control. Prospective validation of these models may lead to clinical application of non-invasive breath profiling in asthma.

INTRODUCTION

Current guideline-based treatment for asthma dictates stepwise increases in medication based on the concepts of disease control that are defined by symptoms and lung function.1 2 A personalised approach to treatment targeted at individual disease characteristics may well improve outcomes but is as yet impractical in the clinic. Sputum cell counts can predict response to inhaled steroids and importantly indicate when increasing the dose of these potentially harmful therapies may be of no benefit.3 They also have the potential to direct novel and expensive therapies to specific patient groups, where a benefit in unselected patients with asthma has not been shown.4 5 Unfortunately, sputum induction and processing is time consuming, unpleasant for the patient and requires significant technical expertise. Caution is warranted in those with unstable or severe asthma, and adequate samples cannot be obtained in at least 10% of attempts even in the most experienced centres.5

The detection of exhaled volatile organic compounds (VOCs) in the breath offers the opportunity for the discovery of novel disease-specific biomarkers. Patterns of VOCs, as detected by an electronic nose, can distinguish patients with chronic obstructive pulmonary disease from those with asthma.7 Profiling VOCs by gas chromatography—mass spectrometry has also shown potential as a diagnostic tool for asthma in children.8 Exhaled VOCs may also be able to discriminate breath samples by the dominant inflammatory cell type. Cultured white cells both release and use VOCs,9 10 and it is likely that the predominant airway inflammatory cell type will be associated with specific VOC profiles. Non-invasive assessment of metabolites in the breath therefore offers the potential for the discovery of novel biomarkers and for providing insights into inflammatory pathways.

The aim of the study was to investigate the discriminatory power of breath VOC profiles for differentiating the breadth of patients with asthma from healthy controls and for classifying patients with asthma based on their sputum inflammatory cell profile and degree of disease control. This was a hypothesis-generating study, a valid and useful approach in metabolomics studies where there are few available data to inform a priori assumptions regarding which VOC patterns may or may not be relevant to the comparisons of interest.11
Methods

Subjects
Volunteers were recruited from the research database of the Medicines Evaluation Unit, Wythenshawe Hospital. Inclusion and exclusion criteria are given in the online data supplement. Written informed consent was obtained, and the local ethics committee approved the study.

Study design
All subjects were asked to refrain from eating and drinking (except for unflavoured water) for 2 h prior to breath collection and from using their inhaled medication on the morning of the visit. A medical history and physical examination were performed before (in order): Asthma Control Questionnaire (ACQ),12 exhaled breath collection, exhaled nitric oxide (eNO) measurement, spirometry and sputum induction with hypertonic saline.

Measurements
All breath samples were collected in the same room, used solely for this study, in an attempt to minimise the effect of variation in background air. Exhaled breath samples were collected using a breath sampler modified from that previously described.13 Subjects breathed VOC-filtered air (filter AX:456-06-20, 5M, Bracknell, UK) via a ResMed Mirage full-face mask of appropriate size (ResMed, Abingdon, UK). The air supply was generated by a continuous positive airway pressure device (Sullivan V, ResMed); a constant airflow through the system is necessary in order to minimise the effect of water vapour on the VOC profile. Respiratory pattern was tracked via a pressure transducer and visualised using bespoke software, enabling targeted sampling of breath from the lower respiratory tract, minimising contamination from the mouth, nose and upper airways as far as practicable. In short, the system tracked pressure changes at the mask, and the sampling software triggered sampling after a short prespecified delay following the pressure change from “baseline” (background mask pressure during inhalation) to “positive” (exhalation). Sampling stopped upon the signal returning to baseline, indicating the start of the next inhalation.

Subjects wore the mask for 5 min prior to sampling to acclimatise and to allow a degree of equilibration with the filtered air. For each breath sample, 3 litres of expiratory air was drawn across the absorbent bed via an Escort ELF laminar flow sampling pump (Kiesen Products, Chelmsford, UK), usually taking 5–7 min. Samples were collected in adsorbent tubes packed with Tenax TA/Carbotrap (Markes International, Llantrisant, UK), conditioned prior to sampling with pure N2 at 300°C and then at 320°C for 1 h, respectively. Samples were analysed within a maximum of 5 days (stored at 4°C) by gas chromatography–mass spectrometry (see online data supplement for details).

Exhaled nitric oxide was measured at 50 ml/s flow rate (Niox, Aerocrine, Sweden) with subjects exhaling slowly into the meter. The mean of three readings was taken as per ATS/ERS guidelines.14 Spirometry was performed using dry wedge spirometer (Vitalograph, Buckinghamshire, UK) according to standard guidelines.15 Finally, sputum induction was performed in accordance with European Respiratory Society recommendations.16 Further details are given in the online data supplement.

Statistical analysis
The aims of the study were to investigate whether the exhaled VOC profile could discriminate breath samples of (1) patients with asthma from healthy subjects, (2) patients with asthma having eosinophilic sputum from those having non-eosinophilic sputum and (3) patients with asthma having neutrophilic sputum from those having non-neutrophilic sputum. Cut-offs for sputum eosinophilia and neutrophilia of 2% and 40% were used.17 We also explored, as a secondary end point, whether the technique could classify patients with asthma on the basis of ‘asthma control’, defined by a score of <1 on the ACQ.12 18 Full details of the statistical methodology are given in the online data supplement. In brief, for each comparison of interest, logistic regression was performed for each VOC and only discriminatory VOCs with p<0.10 were retained for inclusion in the subsequent modelling. Principal component analysis (PCA) was then used to reduce the dimensionality of the data, and these principal components (PCs) were included in multivariate logistic regression, in order to develop a classification model. The validity of the model was checked using discriminant function analysis in parallel with leave-one-out cross-validation.

Results

Thirty-five patients with asthma and 23 healthy controls were recruited. Only one of the subjects had ever smoked: a patient with asthma having quit 5 years previously with a 2.5 pack year history. Ten of the patients with asthma had one or more co-morbidities requiring regular medication: gastro-oesophageal reflux (n=3), anxiety/depression (3), hypertension (4), hypothyroidism (1) and type II diabetes mellitus (1). Demographic details are shown in table 1.

Asthma versus healthy controls
Forty-seven exhaled VOCs classifying patients with asthma versus healthy subjects with p<0.10 were used in the PCA to generate 12 PCs explaining 82% of the variance. Stepwise logistic regression using these PCs as variables generated a predictive model of four PCs, which was able to classify patients with asthma with an accuracy of 86%. Discriminant analysis confirmed the classification power of the model, with leave-one-out cross-validation accuracy of 83%. The positive predictive and negative predictive values of the model were 0.85

Table 1 Demographic details of subjects included in the study

<table>
<thead>
<tr>
<th></th>
<th>Asthma (n=25)</th>
<th>Healthy (n=23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>46.1 (14.4)</td>
<td>50.8 (14.0)</td>
</tr>
<tr>
<td>Gender (M/F)</td>
<td>17/18</td>
<td>10/13</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>28.6 (4.9)</td>
<td>26.2 (3.4)</td>
</tr>
<tr>
<td>% predicted FEV₁</td>
<td>91.5 (22.1)*</td>
<td>109.8 (17.7)</td>
</tr>
<tr>
<td>% predicted FVC</td>
<td>106.0 (14.4)</td>
<td>112.6 (16.2)</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>0.71 (0.11)*</td>
<td>0.79 (0.06)</td>
</tr>
<tr>
<td>Exhaled nitric oxide, ppb</td>
<td>29.2 (22.6)</td>
<td>NA</td>
</tr>
<tr>
<td>Treatment (n)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>As required, bronchodilators only</td>
<td>10</td>
<td>NA</td>
</tr>
<tr>
<td>ICS only</td>
<td>11</td>
<td>NA</td>
</tr>
<tr>
<td>Combination inhalers</td>
<td>14</td>
<td>NA</td>
</tr>
<tr>
<td>Daily inhaled steroid dose, mcg (beclomethasone dose equivalent)</td>
<td>568.6 (756.1).</td>
<td>NA</td>
</tr>
<tr>
<td>Sputum cell profile†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2% eos, ≥40% neuts</td>
<td>22.2%</td>
<td>NA</td>
</tr>
<tr>
<td>≥2% eos, &lt;40% neuts</td>
<td>22.2%</td>
<td>NA</td>
</tr>
<tr>
<td>&lt;2% eos, ≥40% neuts</td>
<td>33.3%</td>
<td>NA</td>
</tr>
<tr>
<td>&lt;2% eos, &lt;40% neuts</td>
<td>22.2%</td>
<td>NA</td>
</tr>
</tbody>
</table>

Data shown as mean (SD) except where indicated.

*p<0.05 versus healthy controls.
†n=18.
eos, eosinophils; FEV₁, forced expiratory volume in one second; FVC, forced vital capacity; ICS, inhaled corticosteroids; neuts, neutrophils.
and 0.89, respectively. Plots of the contributing compounds for each PC are shown in figure 1.

**Asthma phenotypes**

Eighteen subjects with asthma were able to provide induced sputum samples for cell counting. There were no significant demographic differences between those who could and could not produce an adequate sputum sample, for example, in terms of age, lung function, ACQ score, inhaled steroid use or dose. Those with sputum eosinophils ≥2% (n=8) had lower forced expiratory volume in one second (FEV1) and forced vital capacity compared with the non-eosinophilic group (table 2). Eleven discriminating compounds were used in the PCA, reduced to four PCs explaining 85% of the variance. Logistic regression produced a model using three of these PCs with a classification accuracy of 83%. Discriminant function analysis had a leave-one-out cross-validation accuracy of 85%, and the area under the receiver operating characteristic (AUROC) for the model was 0.98 (other ROC parameters are shown in table 3). For comparison, a high exhaled NO (cut-off based on19) was able to predict sputum eosinophilia in our cohort with an AUROC curve of 0.69 (data not shown). Model parameters and contributing compounds are shown in table 4 and the ROC curve in figure 2.

The eight subjects with a sputum neutrophil count ≥40% were older than those without (table 2). Univariate analysis identified 14 discriminating compounds, reduced to five PCs by PCA explaining 81% of the variance. Logistic regression generated a model comprising one PC with an accuracy of 79%, discriminant function leave-one-out cross-validation accuracy 72% and AUROC curve of 0.90 (tables 3 and 4, figure 2).

Eighteen subjects had uncontrolled asthma (ACQ≥1); they had poorer lung function and were on higher levels of treatment compared with the 17 controlled subjects (table 2). Thirteen discriminating compounds produced five PCs in PCA explaining 65% of the variance. The logistic regression model included four of these PCs and had an accuracy of 89%. Discriminant function analysis had a leave-one-out cross-validation accuracy of 80% and the AUROC curve was 0.97 (tables 3 and 4, figure 2).

**Figure 1** Box-and-whisker plots of relative intensity for the individual discriminating volatile organic compounds used for each principal component in the model predicting healthy controls (grey boxes) versus patients with asthma (white boxes).
also likely to be alkanes or methylated alkanes (both C13H28) identical list of discriminating VOCs to previous asthma methodologies, it is unsurprising that we have not identified differing selection criteria, breath collection and analytical a potential biomarker for asthma.

their methylated derivatives) merits further exploration as 2,6,11-trimethyl-dodecane24) are longer chain alkanes, which may be more specific to asthma.25 Another potential source of oxidative stress may result from conditions coexistent with asthma, such as obesity and hypertriglyceridaemia,20 although in our cohort the mean BMI was not different between sputum producers and non-producers (data not shown). It is unclear why we were only able to obtain viable sputum samples in just over half of these subjects, although one could speculate that the airway-drying effect of unhumidified continuous positive airway pressure may have had an impact on subsequent successful sputum induction. We would propose that this group of compounds (C7 to C12 alkanes and their methylated derivatives) merits further exploration as a potential biomarker for asthma.

Our models were able to predict both neutrophilic and eosinophilic sputum cell phenotypes with very high diagnostic accuracy. Furthermore, the intensity of many of the compounds contributing to the models (and the overall models themselves) correlated with the percentage of cells of interest, strengthening the association. There were no clear patterns in terms of classes of compounds seen in this part of the analysis, but should these be confirmed in an independent cohort, then investigation into the relevance and origin of the discriminating compounds would be indicated. Despite the evidence that sputum analysis can be used to personalise asthma treatment,6 6 28 it has not been adopted into routine management outside specialist asthma units for several reasons: sputum induction, processing and interpretation are time consuming and require experienced and skilled personnel; samples cannot be obtained in a significant minority of subjects; samples must be analysed within a few hours of collection and care must be taken in those with uncontrolled or severe disease. By contrast, breath analysis offers the opportunity for surrogates to be detected in a non-invasive, straightforward way. Future developments may also lead to online detection as has been achieved with eNO. In fact, we found that the exhaled breath VOC profile predicted sputum eosinophilia more accurately than eNO in our cohort. Although ours was a small sample, similar results have recently been reported for predicting sputum eosinophilia ≥3% using eNO with an AUROC of 0.78.28 eNO may in fact provide complementary information to exhaled VOCs and a combined model may be even more powerful. It should be noted that just over half of our participants were able to provide sputum for analysis and thus may not be entirely representative of the group as a whole, although demographic data were not different between sputum producers and non-producers (data not shown). It is unclear why we were only able to obtain viable sputum samples in just over half of these subjects, although one could speculate that the airway-drying effect of unhumidified continuous positive airway pressure may have had an impact on subsequent successful sputum induction. We will investigate this possibility in future studies.

**DISCUSSION**

We have demonstrated that breath VOC analysis can discriminate patients with asthma from healthy controls and also identify clinically important disease phenotypes related to sputum inflammation and a measure of asthma control. While these findings need to be validated in an independent cohort of patients, we have provided a degree of internal validation using discriminant function analysis. Furthermore, our analysis has identified a number of potential new biomarkers for asthma and its phenotypes.

Several of the compounds that were raised in patients with asthma compared with controls were alkanes. Alkanes arise as products of lipid peroxidation,20 and raised levels of exhaled pentane have previously been demonstrated in acute asthma.21 However, pentane appears to be relatively non-specific as levels are also raised in other inflammatory conditions such as rheumatoid arthritis22 and inflammatory bowel disease.23 The alkanes found to be discriminatory in our study and others (such as 2,6,11-trimethyl-dodecane24) are longer chain alkanes, which may be more specific to asthma. Two unidentified compounds also likely to be alkanes or methylated alkanes (both C13H28) were also found to be raised in children with asthma.8 With differing selection criteria, breath collection and analytical methodologies, it is unsurprising that we have not identified an identical list of discriminating VOCs to previous asthma studies,8 24 although the repeated discovery of similar compounds (alkanes and methylated alkanes) under differing experimental conditions adds strength to the evidence that these may well be important. Methylated alkanes have been shown to increase with age and probably further relate to oxidative stress.25 Another potential source of oxidative stress may result from conditions coexistent with asthma, such as obesity and hypertriglyceridaemia,26 although in our cohort the mean BMI of the healthy and asthmatic groups was similar. We would propose that this group of compounds (C7 to C12 alkanes and their methylated derivatives) merits further exploration as a potential biomarker for asthma.
Excellent diagnostic accuracy was also demonstrated identifying patients with asthma with poor disease control. Asthma control is a measure of symptoms related to current disease activity and, to a degree, is independent of disease severity. As such, recent guidelines have emphasised the importance of achieving disease control. The seven-item ACQ includes six such, recent guidelines have emphasised the importance of achieving disease control.1 The seven-item ACQ includes six activity and, to a degree, is independent of disease severity. As such, recent guidelines have emphasised the importance of achieving disease control.1 The seven-item ACQ includes six ACQ score

<table>
<thead>
<tr>
<th>Model parameters and compounds for subgroup comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Logistic regression model</strong></td>
</tr>
<tr>
<td><strong>PCs</strong></td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td><strong>Eosinophils ≥2%</strong></td>
</tr>
<tr>
<td>PC2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PC3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PC4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Overall model</strong></td>
</tr>
<tr>
<td><strong>Neutrophils ≥40%</strong></td>
</tr>
<tr>
<td>PC1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Overall model</strong></td>
</tr>
<tr>
<td><strong>ACQ score ≥1</strong></td>
</tr>
<tr>
<td>PC2</td>
</tr>
<tr>
<td>PC3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PC4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>PC5</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Overall model</strong></td>
</tr>
</tbody>
</table>

Asterisk indicates a significant correlation (p<0.05) between the compound or overall model and the variable of interest (eg, camphene is inversely correlated with sputum eosinophil count, r=−0.47, p<0.05).

ACQ, Asthma Control Questionnaire; PCs, principal components.

Figure 2 Receiver operating characteristic curves for the models predicting eosinophilic versus non-eosinophilic asthma (solid line); neutrophilic versus non-neutrophilic asthma (dashed line) and controlled versus uncontrolled asthma (dotted line).
breath VOCs include those from alveolar air, reflecting systemic exposure. The kinetics of inhaled VOCs are therefore not dependent simply on equilibration between lung air and atmospheric air, but will behave according to concepts familiar from basic pharmacokinetics, that is, adsorption, distribution, metabolism and excretion.32 These same principles also inform our practice in dealing with the potential effect of environmental VOCs within the sampling room. One approach is to sample the environment at the time of breath sampling and subtract this background from the breath profile.53 This assumes equilibrium exists between the breath and the environment at all times, which is not likely to be the case where subjects will have moved through several environments in the hours and days prior to sampling. We have attempted to control this potential problem as far as reasonable by using a VOC filter in our sampling circuit and sampling in the same room (used exclusively for this purpose) throughout the study. Furthermore, we would expect any influence of the background signal to be balanced between the cases and the controls with this study design, as long as subjects were not sampled in group-related sequence (e.g., cases first, controls second) over the course of the study.

It is important to recognise the characteristics of the subjects with asthma who were included in this study. These were patients with a doctor’s diagnosis of asthma, with current symptoms of the disease and requiring at least one asthma medication. We felt it important to address ‘real world’ subjects with asthma rather than those strictly defined by objective tests such as bronchodilator reversibility, which are often negative in patients treated with regular inhaled steroids. Nevertheless, many of our subjects did have some objective evidence of asthma as recommended in the most recent British Thoracic Society guidelines:71% had at least one of bronchodilator reversibility (≥400 ml), bronchial hyper-responsiveness (methacholine PC20 <8 mg/ml), sputum eosinophilia (≥2%) or raised exhaled NO (≥25 ppb at 50 ml/sec). Of the remainder, all but two had some suggestive but less definite evidence (atopy, requirement for prednisolone treatment in the previous year, borderline eosinophilia (≥1%) or reversibility (≥200 ml)).

In conclusion, we have shown that breath VOCs can be used to discriminate patients with asthma from healthy controls as well as important disease phenotypes related to airway inflammation and disease control. Further prospective studies are now required to validate these models and should include patients where the diagnosis is as yet unknown. Such validation will allow future work to be directed towards developing point-of-care devices for breath phenotyping that can be used in clinical trials as well as in the outpatient clinic.

Acknowledgements Our thanks go to GlaxoSmithKline for contributing to the funding of this study and Waters Corporation for loan of mass spectrometry equipment. We also acknowledge Richard Knowles for his review of the manuscript and Julie Morris for statistical advice. BI gratefully acknowledges Universiti Sains Malaysia for funding.

Funding GlaxoSmithKline kindly contributed towards the funding of this study.

Competing interests None.

Ethics approval This study was conducted with the approval of the Tameside and Glossop REC (now known as North West B REC—Greater Manchester East).

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


Non-invasive phenotyping using exhaled volatile organic compounds in asthma

Baharudin Ibrahim, Maria Basanta, Paul Cadden, Dave Singh, David Douce, Ashley Woodcock and Stephen J Fowler

Thorax 2011 66: 804-809 originally published online July 11, 2011
doi: 10.1136/thx.2010.156695

Updated information and services can be found at:
http://thorax.bmj.com/content/66/9/804

These include:

Supplementary Material
Supplementary material can be found at:
http://thorax.bmj.com/content/suppl/2011/06/22/thx.2010.156695.DC1

References
This article cites 32 articles, 7 of which you can access for free at:
http://thorax.bmj.com/content/66/9/804#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections
Asthma (1782)
Child health (843)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/