

Breath metabolomic profiling by nuclear magnetic resonance spectroscopy in asthma

B. Ibrahim^{1,2}, P. Marsden¹, J. A. Smith¹, A. Custovic¹, M. Nilsson³ & S. J. Fowler^{1,4}

¹The University of Manchester, Manchester Academic Health Science Centre, NIHR South Manchester Respiratory and Allergy Clinical Research Facility, University Hospital of South Manchester, Manchester, UK; ²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia; ³School of Chemistry, University of Manchester, Manchester; ⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

To cite this article: Ibrahim B, Marsden P, Smith JA, Custovic A, Nilsson M, Fowler SJ. Breath metabolomic profiling by nuclear magnetic resonance spectroscopy in asthma. *Allergy* 2013; DOI: 10.1111/all.12211.

Keywords

asthma; biomarkers; breath analysis; exhaled airway markers.

Correspondence

Stephen J. Fowler, Education and Research Centre, University Hospital of South Manchester, Southmoor Road, Manchester M23 9LT, UK.

Tel.: +44 161 291 5864 or

+44 1772 523237

Fax: +44 161 291 5730

E-mail: stephen.fowler@manchester.ac.uk

Accepted for publication 21 May 2013

DOI:10.1111/all.12211

Edited by: Michael Wechsler

Abstract

Background: Metabolomic profiling of exhaled breath condensate offers opportunities for the development of noninvasive diagnostics in asthma. We aimed to determine and validate discriminatory metabolomic profiles in adult asthma and to explore profiles in clinically relevant disease phenotypes.

Methods: Nuclear magnetic resonance spectroscopy was used to analyse breath condensate samples from 82 subjects with asthma and 35 healthy volunteers. Multivariate modelling was performed on a 'training set' (70% of the total sample) in order to produce a discriminatory model classifying asthmatics from healthy controls, and the model tested in the remaining subjects. Secondary analyses were performed to determine the models for the identification of asthmatic subgroups based on sputum eosinophilia, neutrophilia, asthma control and inhaled corticosteroid use.

Results: A classification model consisting of five discriminating spectral regions was derived using data from the training set with an area under the receiver operating curve (AUROC) of 0.84. In the test set (the remaining 30% of subjects), the AUROC was 0.91, thus providing external validation for the model. The success of the technique for classifying asthma phenotypes was variable, with AUROC for: sputum eosinophilia (3% cut-off) 0.69; neutrophilia (65% cut-off) 0.88; asthma control (cut-off Asthma Control Questionnaire score of 1) 0.63; and inhaled corticosteroid use 0.89.

Conclusion: Nuclear magnetic resonance spectroscopy of breath condensate successfully differentiates asthmatics from healthy subjects. With identification of the discriminatory compounds, this technique has the potential to provide novel diagnostics and identify novel pathophysiological mechanisms, biomarkers and therapeutic targets.

Exhaled breath condensate (EBC) contains many volatile and nonvolatile compounds, which may be altered in disease, and could provide a useful medium for biomarker discovery (1–3). In children, selected nuclear magnetic resonance (NMR) spectral regions have been able to classify EBC from asthmatics *versus* healthy controls with better accuracy than achievable using exhaled nitric oxide and forced expiratory volume in one-second (FEV₁) combined (4). These promising findings need validation in independent populations, as NMR spectra contain huge amounts of data, and model overfitting is a common pitfall (5).

Several techniques are available to analyse the chemical composition of EBC. In metabolomics, the most commonly used platforms are liquid chromatography–mass spectrometry (LC-MS) and ¹H-NMR spectroscopy. Nuclear magnetic resonance–based metabolomics provides a powerful method for investigating human disease pathophysiology, drug toxicity and disease phenotyping (6). It is nondestructive, provides important structural information and is suitable for high-throughput, all important for metabolomics studies. Combining metabolomic analytical platforms with advanced computer modelling and pattern recognition methods may

enable us to discover unknown or novel phenotypes of chronic disease such as asthma using an unbiased approach (7).

We aimed to determine whether ^1H -NMR spectroscopy could discriminate the EBC metabolic profile of adults with asthma from healthy controls, and to evaluate the technical reproducibility of the analytical methodology and the stability of EBC samples stored at room temperature. In addition, we explored the ability of this technique to classify clinically relevant asthma phenotypes, as defined by sputum cell profiles, asthma control and inhaled corticosteroid (ICS) use. We defined asthma phenotypes based on objective data and clinical history using the following criteria: sputum eosinophilia [cut-off 3% (8)], sputum neutrophilia [cut-off 65% (9)], asthma control, based on the Asthma Control Questionnaire-(ACQ) (10) ['uncontrolled' asthma defined as a score of at least one (11)], and inhaled corticosteroid (ICS) use.

Methods

Study design

ASessment of Manchester Asthmatics Longitudinally is a longitudinal asthma cohort study described in detail previously (12–14). Asthmatic patients fulfilled the following criteria: symptoms (wheeze, cough, chest tightness or breathlessness) in the previous 12 months, treatment with at least a short-acting bronchodilator, and no asthma exacerbation or respiratory infection within 4 weeks of the visit (12). Current smokers or ex-smokers with more than 10 pack-year history were excluded. Healthy controls met the following criteria: no history of respiratory or other diseases likely to interfere with the study, baseline $\text{FEV}_1 > 80\%$ predicted, FEV_1/FVC ratio > 0.7 and negative methacholine challenge. The study was approved by local ethics committee, and all participants gave written informed consent.

Measurements

Prior to study visits, subjects were asked to withhold short-acting bronchodilators for 6 h; long-acting bronchodilators, theophyllines and leukotriene receptor antagonists for 12 h; and antihistamines for 48 h. Spirometry and sputum induction were performed on a single visit, and then EBC collection occurred at a second visit 1 week later.

Spirometry was performed according to guidelines (15), using a dry wedge spirometer (Vitalograph, Buckinghamshire, UK). Sputum induction was performed as previously described (16). Further details are given in the Data S1. Exhaled breath condensate was collected using the EcoScreen (Jaeger, Hoechberg, Germany). The circuit included a saliva trap and subjects were also encouraged to swallow saliva during the procedure. Subjects wore a noseclip and were asked to breathe tidally into the mouthpiece until minimum 1.8 ml of condensate was collected. Samples were divided into aliquots and stored in a freezer at -80°C immediately. Metabolomic analysis was performed using 500 μl aliquots of

EBC and run in a 500-MHz Varian VNMR5 (Varian NMR Inc., Palo Alto, CA, USA). Full details of the NMR analysis, along with data preprocessing, are given in the Data S1.

Statistical analysis

Between-group comparisons were made using Pearson's chi-square, parametric (Student's *t*-test) and nonparametric tests (Mann–Whitney *U*-test) where appropriate using SPSS version 15 (SPSS Inc., Chicago, IL, USA). In proof-of-concept and validation analysis, the subjects were split into training and test sets using 70 : 30 ratios. Principal component analysis (PCA) was initially performed to explore the data and to detect any outliers and then used for random subject allocation to each set with SIMCA-P+ version 12.0.1 (Umetrics, Umea, Sweden) (Figure S1).

The process for data analysis was adapted from our previous methodology (17). Univariate logistic regression was used to remove redundant bins from the NMR spectra by selecting only the bins that could discriminate asthmatics from healthy controls with the *p* value set at 0.10 to limit the chance of excluding important data from the models. The NMR spectral peaks corresponding to the retained bins were then identified, and the normalized peak areas of the regions entered into multivariate logistic regression to generate a best-fit model for between-group discrimination in the training set. The model developed from these regions was validated in the test set and its performance evaluated by receiver operating characteristics (ROC) (18). This strategy was also applied to identify regions that could differentiate between different asthma phenotypes. In this hypothesis-generating part of the study, we did not split the data into training and test sets due to the small sample size, in particular those with viable induced sputum samples. Instead, discriminant function analysis (DFA) and leave-one-out cross-validation (LOOCV) were used in parallel to check the validity of the model.

Intraclass correlation coefficient (ICC) (19), which compares intra- and interindividual variations in the data set, was used to test analytical reproducibility.

Results

Participants

We recruited 89 asthmatics (76% atopic) and 35 healthy controls. Insufficient EBC volume was obtained from seven asthmatic subjects, who were thus excluded from the analysis. Data from one subject were excluded due to a distorted NMR spectrum, and those from three others which were clear outliers on PCA. Demographics of the study population are presented in Table 1, and those of patients by asthma phenotype in Table 2. Evaluable induced sputum samples were obtained in 53 asthmatics. Median eosinophil counts were 6.3% in the eosinophilic group and 1.0% in the noneosinophilic group; median neutrophil counts were 75.0% for the neutrophilic group and 52.9% for the non-neutrophilic group.

Table 1 Demographics of asthmatics *versus* controls shown as mean (SD) except where indicated

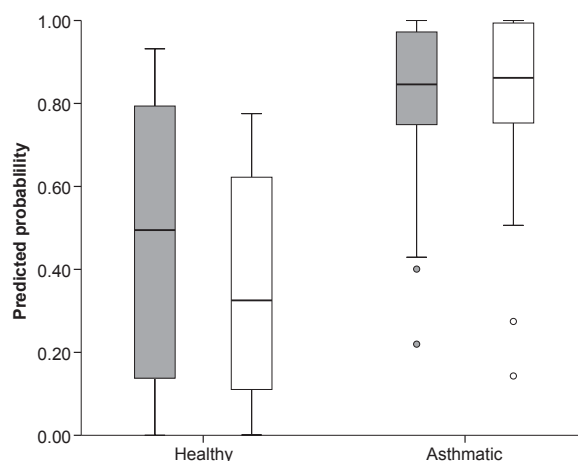
Demographics	Asthmatics (n = 79)	Controls (n = 34)
Age, years	57.7 (11.4)	43.6 (14.2)*
Gender (male/female)	35/44	12/22
Smoking status		
Ex	51%	24%*
Never	49%	76%
Pack years	7.2 (12.5)	1.1 (1.8)
FEV ₁ , % predicted	85.3 (22.2)	108.6 (10.8)*
FVC, % predicted	108.8 (20.3)	118.4 (11.9)
FEV ₁ /FVC	63.6 (12.1)	75.7 (4.9)*
Inhaled corticosteroid use (yes/no)	61/18	0/34*

* $P < 0.05$ for asthmatics vs healthy controls.

Exhaled breath condensate profile in asthmatics and control subjects

In the proof-of-concept and validation analyses, there were 57 asthmatics and 22 healthy controls in the training set (70% of total) and 22 asthmatics and 12 healthy controls in the test set (remaining 30%). Figure S1 shows the PCA plot representing components of the three different batches for training and test sets. There were no significant differences between the demographics of the training and test sets (data not shown).

Twenty-two bins were identified as discriminating asthmatics from healthy controls ($P \leq 0.10$), corresponding to 13 spectral regions. The normalized peak areas were used in multivariate logistic regression, and a model consisting of five regions predicted the asthmatics. The five discriminating spectral regions were 0.16–0.18 ppm, 0.78–0.84 ppm, 0.88–0.94 ppm, 7.36–7.42 ppm and 7.44–7.52 ppm. This model correctly classified 54 asthmatics and misclassified three (94.7% correct), while 11 (50.0%) of the healthy controls were classified correctly. The overall accuracy of the model was 82.3%, and area under receiver operating characteristic

**Figure 1** Boxplots of the predicted probabilities of asthmatics and healthy controls. Grey = training set, white = test set.

(AUROC) was 0.84. Positive predictive value (PPV) and negative predictive value (NPV) were 83.1% and 78.6%, respectively.

The regions identified in the training set were then used to differentiate between asthmatics and healthy controls in the independent test set. The PPV, NPV and AUROC were 83.3%, 80.0% and 0.91, respectively. Figure 1 shows the boxplots of the predicted probabilities of asthmatics and healthy controls for the training and test sets. Figure 2 shows an example of an NMR spectrum from an asthmatic, with the discriminatory regions enlarged.

Reproducibility/sample stability

We assessed the reproducibility of the five discriminating regions. The ICC varied by region with reproducibility classified as very good for regions 0.78–0.84 ppb and 0.88–0.94 ppb (ICC = 0.97 and 0.98), good for 7.44–7.52 ppb (ICC = 0.61), fair for 7.36–7.42 ppb (ICC = 0.37), but poor for region 0.16–0.18 ppb (ICC = −0.10).

Table 2 Demographic data for asthma phenotype comparisons shown as mean (SD) except where indicated

	Sputum Eosinophilia		Sputum Neutrophilia		Asthma control		Inhaled corticosteroid use	
	≥3% (n = 21)	<3% (n = 32)	>65% (n = 33)	≤65% (n = 20)	Well controlled (n = 36)	Not well controlled (n = 43)	Yes (n = 63)	No (n = 18)
Age, mean (SD)	60.8 (9.3)	58.7 (12.2)	62.8 (9.6)	54.1* (11.5)	56.8 (10.9)	58.9 (12.0)	59.5 (10.8)	52.6 (12.5)
Gender (male/female)	13/8	18/14	21/12	10/10	14/22	21/22	26/35	9/9
% Predicted FEV ₁ , mean (SD)	78.4 (22.7)	83.6 (24.6)	76.3 (24.5)	90.2 (20.2)	94.5 (2.6)	77.4 (23.6)*	82.8 (22.5)	92.7 (18.7)
% Predicted FVC, mean (SD)	109.4 (22.5)	106.1 (18.9)	102.3 (17.3)	115.8 (22.4)	117.2 (18.5)	101.3 (18.9)*	107.8 (22.0)	110.5 (12.7)
Inhaled corticosteroid use (yes/no)	19/2	23/9	26/7	16/4	24/12	37/6*	NA	NA

NA, not applicable.

* $P < 0.05$ for individual phenotype comparisons.

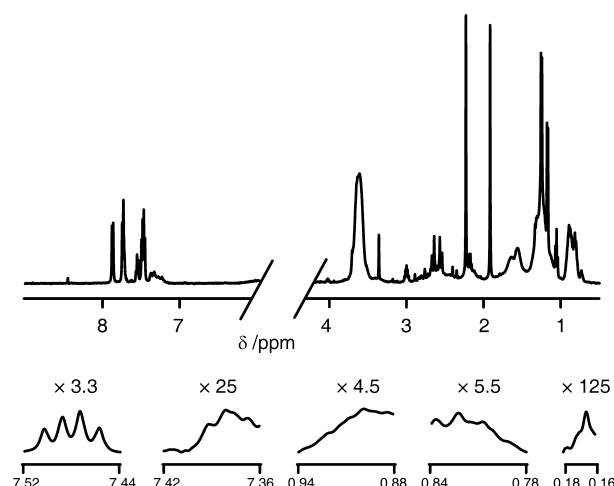


Figure 2 500-MHz ^1H -NMR spectrum (top) of exhaled breath condensate from an asthmatic patient. The expansions below show the five spectral regions used in the model discriminating asthmatics from healthy controls.

Exhaled breath condensate metabolomic profile in asthma phenotypes

Summary results of phenotype comparisons are presented in Table 3. Table 4 shows the related regions and correlation coefficient between the integrated peak areas and the variable of interest (e.g. percentage sputum eosinophils, inhaled corticosteroid dose).

Sputum eosinophilia

Multivariate analysis produced a model from two regions distinguishing the two groups with accuracy of 57% and AUROC of 0.69. There were no significant correlations between the two regions or the overall model and eosinophil count.

Sputum neutrophilia

The model discriminated groups with accuracy of 79% and AUROC of 0.88 and correlated with the neutrophil count, $r = 0.42$.

Table 4 Spectral regions included in the subgroup-comparison models

Phenotype	Spectral region, ppm	Correlation coefficient, r
Eosinophil (<3% vs $\geq$ 3%)	1.16–1.22	−0.21
	3.63–3.69	−0.16
	Overall model	0.20
Neutrophil ($\leq$ 65% vs >65%)	0.72–0.76	0.05
	0.78–0.84	0.04
	0.98–1.03	0.04
	1.12–1.16	0.14
	1.32–1.36	0.06
	2.64–2.68	0.11
	3.70–3.74	0.14
	Overall model	0.42*
Asthma control (yes vs no)	1.04–1.08	0.23*
Inhaled corticosteroid use (yes vs no)	0.84–0.88	−0.19
	0.88–0.94	−0.10
	0.98–1.03	−0.24*
	1.12–1.16	−0.14
	3.63–3.69	−0.15
	3.70–3.74	−0.14
	7.53–7.59	−0.36*
	7.86–7.90	−0.35*
	Overall model	0.55*

*indicates peaks or models with $P \leq 0.05$ for the correlation coefficient r (between peak areas of region vs the variable of interest).

Asthma control (ACQ)

The model consisted of one region, had accuracy of 61%, AUROC of 0.63 and correlated with the ACQ score, $r = -0.23$.

Inhaled corticosteroid use

The model of eight regions had accuracy of 85% and AUROC of 0.89. All the regions were negatively correlated with ICS dose and three significantly so. The overall model correlated with ICS dose, $r = 0.55$.

Table 3 Summary results of phenotype comparisons. The table shows for each phenotype comparison, numbers of discriminating bins between the two groups from the univariate analysis, numbers of regions that are associated with the bins and the outcomes from the multivariate analysis

Phenotypes	Significant bins	Significant regions	Multivariate logistic regression				AUROC	LOOCV accuracy (%)
			PPV (%)	NPV (%)	Accuracy (%)	Precision (%)		
Eosinophil (<3% vs $\geq$ 3%)	10	2	40	60	57	40	0.69	59
Neutrophil ($\leq$ 65% vs >65%)	24	7	82	74	79	82	0.88	57
Asthma control (yes vs no)	5	1	60	62	61	58	0.63	61
Inhaled corticosteroid use (yes vs no)	71	8	87	75	85	87	0.89	76

PPV, positive predictive value; NPV, negative predictive value; AUROC, area under receiver operating characteristic curve; LOOCV, leave-one-out cross-validation.

Discussion

This is the largest study investigating the potential application of exhaled breath metabolomics in asthma, and the first to provide validation for NMR-based modelling of EBC from asthmatics and healthy adults. We have identified spectral components that demonstrate good accuracy in discriminating asthmatics from controls but more importantly have validated the model in an independent test set of subjects. It is also the first study to use this technique to explore asthma phenotypes based on sputum inflammatory cells, asthma control and ICS use. While this method could discriminate asthmatics based on inhaled corticosteroid use and sputum neutrophilia, it was not able to differentiate the asthma phenotypes of sputum eosinophilia and asthma control with good accuracy. We have also shown that the reproducibility of the NMR analysis varies, and identified the most reproducible spectral regions.

We identified five regions differentiating asthmatics from controls. Carraro et al. (4) reported discriminating signals in 25 asthmatic children and 11 controls from two regions, namely 1.7–2.2 ppm and 3.2–3.4 ppm. We did not expect to exactly reproduce these results as specific features of sample collection, preparation and data acquisition methods differed. Different equipment, from the condensers to NMR spectrometers, was used. The samples in the Carraro study were dried before adding D₂O while we used the intact sample; drying might cause the loss of some of the metabolites, especially volatile and unstable compounds (20). We used a buffer to minimize pH-dependent chemical shifts; this is potentially very important, as subjects with active asthma are known to have a lower airway and EBC pH than nonasthmatics (21). Finally, we preselected discriminating bins using univariate logistic regression before determining the regions of interest and performing the multivariate analysis, while Carraro and colleagues utilized the bins from the whole NMR spectra. Further work should be undertaken with a view to developing standardized and optimized methodologies to enable findings to be compared between studies (22). One potential contributor to the classification may well be the use of ICS in the majority of asthmatics. Indeed, one region (0.88–0.94 ppm) was included in both discriminatory models.

In phenotype comparisons based on sputum eosinophilia of at least 3% and asthma control, the identified regions had very low accuracy discriminating the two groups of interest. The correlations between the peak areas of the identified regions and the eosinophil count were also low. This could be due to the relatively small sample size (false negative) or insufficient sensitivity of NMR spectroscopy as the concentration of the relevant metabolites in the EBC is relatively low (true negative). Future improvements in methodology could improve the sensitivity. Using a higher-field spectrometer with a cryogenic probe should increase the signal-to-noise ratio up to an order of magnitude. A larger sample volume could also be dried and reconstituted in a smaller sample, albeit at the risk of losing volatile compounds. Exhaled breath condensate contains a huge amount of water vapour that may dilute some of the important metabolites from the

airway lining fluid (23). Recent data have suggested that EBC analysis may be used to classify asthmatics by severity, but there are several important methodological differences between these and our data that make direct comparison difficult: theirs was a paediatric cohort with polarized groups (a group with low treatment requirements and good control was compared to one with high treatment requirements and poor control), and samples were analysed using LC-MS (24).

The models discriminating the phenotypes of sputum neutrophilia and ICS use were more accurate, with AUROC of 0.88 and 0.89, respectively, and also correlated with the variable of interest. The spectral regions making up the two models were not mutually exclusive however, with three of the discriminating regions for sputum neutrophilia also found in the ICS use comparison. Although inhaled corticosteroid use has been associated with tissue neutrophilia (25), the same is not necessarily true for sputum neutrophilia (9, 26), and indeed in the present study, ICS use was no more common in the neutrophilic group. The co-existence of these regions in the two models therefore may owe to the potentially confounding factors of older age and poor lung function, which tended to be more common in those on ICS or with sputum neutrophilia. Conversely, the spectral regions not represented in both groups may more directly link to the phenotype of interest and merit further investigation to investigate whether there is any pathophysiological link.

Reproducibility and repeatability are among the key issues in analytical chemistry and also in biomarker discovery studies involving human samples, due to the potential for machine and biological variability (20, 27). In our study, three regions had good to very good reproducibility, and two had fair to no reproducibility. Our results are consistent with the findings of two other studies looking at reproducibility of NMR spectroscopy analysis of EBC and urine specimens (4, 28). As the reproducibility varied by region, it seems likely that the regions with low ICCs were affected by prolonged sample storage at room temperature rather than the technical reproducibility of the analysis and that these regions represent particularly volatile or reactive compounds. We would recommend therefore that the current practice of analysing defrosted samples as soon as possible be adhered to.

Identification of the compounds in the discriminating spectral regions was not attempted at this stage as the concentrations were prohibitively low. Clearly this is of interest and will form part of future studies using a more sensitive NMR instrument and/or mass spectrometry (27). In childhood asthma, Carraro and colleagues speculated that the spectra of interest arose from acetylated and oxygen-containing compounds (4), while in exhaled breath gas we have previously shown methylated alkanes to be important in distinguishing between asthmatic and healthy breath (17). It may be desirable in future studies to combine analysis by NMR with LC-MS (29). Advantages of NMR include the potential for high-throughput, nondestructive analysis, but it is relatively low in sensitivity compared to LC-MS and identification of specific compounds is more difficult. While it is not always necessary to determine the identity of the compounds of interest, as a metabolomic fingerprint showing the pattern

of discriminating NMR spectra would be sufficient to diagnose and phenotype a disease (30), compound identification would nevertheless benefit methodology validation as well as raising interesting pathophysiological hypotheses.

There are justified concerns in breath analysis regarding the potential for contamination of samples by nonrespiratory volatile and nonvolatile compounds, arising, for example, from saliva, sampling equipment or ambient air (20, 31, 32). This is inevitable and will occur even when using a standardized air source and/or filtering. While salivary contamination is of particular concern in EBC samples, it does not appear to significantly alter the results of metabolomic analysis when performed by NMR (33). Our experimental design further minimizes this effect first because the case-control design means that the potentially confounding factors (at least those related to exogenous factors) should be randomly distributed between the groups of interest. More importantly, we have external validation of our findings for an independent cohort of asthmatics and healthy volunteers, thus minimizing the effect of these potential confounders. Nevertheless, should specific biomarkers be identified using these techniques, necessary validation would include the assessment of the potential effects of nonrespiratory contamination.

In summary, we have identified regions of NMR spectra that have the potential for discriminating asthmatics from healthy controls. The validation of the model developed from these regions in an independent test set is encouraging for further studies to include larger sample size and samples from other centres. Our exploratory findings in phenotyping asthma based on sputum neutrophilia and corticosteroid use are encouraging and, if validated, could have future

applications in clinical diagnostics as well as in identifying novel mechanisms and targets. Variable reproducibility of the regions suggests that careful sample storage and preparation is vital for metabolomic studies using EBC.

Acknowledgments

The authors gratefully thank Dr Julie Morris for statistical advice.

Author contributions

BI, PM, AC and SJF involved in study design. BI and PM collected the data. BI and MN involved in sample analysis. BI, JAS, MN and SJF analysed the data. BI, MN and SJF involved in manuscript writing. PM, JAS, AC and MN involved in reviewing the manuscript. SJF involved in study conception. All authors have approved the final manuscript.

Conflict of interest

None declared for all named authors.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Supplementary methods.

Figure S1. Principal components analysis plot representing the selection of samples from the three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles), batch 3 (squares).

References

- Nicolaou NC, Lowe LA, Murray CS, Woodcock A, Simpson A, Custovic A. Exhaled breath condensate pH and childhood asthma: unselected birth cohort study. *Am J Respir Crit Care Med* 2006;**174**:254–259.
- Montuschi P, Barnes PJ. Exhaled leukotrienes and prostaglandins in asthma. *J Allergy Clin Immunol* 2002;**109**:615–620.
- Corradi M, Rubinstein I, Andreoli R, Manini P, Caglieri A, Poli D et al. Aldehydes in exhaled breath condensate of patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2003;**167**:1380–1386.
- Carraro S, Rezzi S, Reniero F, Heberger K, Giordano G, Zancato S et al. Metabolomics applied to exhaled breath condensate in childhood asthma. *Am J Respir Crit Care Med* 2007;**175**:986–990.
- Seasholtz MB, Kowalski B. The parsimony principle applied to multivariate calibration. *Anal Chim Acta* 1993;**277**:165–177.
- Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica* 1999;**29**:1181–1189.
- Goodacre R. Metabolomics of a superorganism. *J Nutr* 2007;**137**:259S–266S.
- Schleich FN, Seidel L, Sele J, Manise M, Quaedvlieg V, Michils A et al. Exhaled nitric oxide thresholds associated with a sputum eosinophil count $\geq 3\%$ in a cohort of unselected patients with asthma. *Thorax* 2010;**65**:1039–1044.
- Green RH, Brightling CE, Woltmann G, Parker D, Wardlaw AJ, Pavord ID. Analysis of induced sputum in adults with asthma: identification of subgroup with isolated sputum neutrophilia and poor response to inhaled corticosteroids. *Thorax* 2002;**57**:875–879.
- Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *Eur Respir J* 1999;**14**:902–907.
- Juniper EF, Bousquet J, Abetz L, Bateman ED. Identifying 'well-controlled' and 'not well-controlled' asthma using the Asthma Control Questionnaire. *Respir Med* 2006;**100**:616–621.
- Langley SJ, Goldthorpe S, Custovic A, Woodcock A. Relationship among pulmonary function, bronchial reactivity, and exhaled nitric oxide in a large group of asthmatic patients. *Ann Allergy Asthma Immunol* 2003;**91**:398–404.
- Langley SJ, Goldthorpe S, Craven M, Woodcock A, Custovic A. Relationship between exposure to domestic allergens and bronchial hyperresponsiveness in non-sensitized, atopic asthmatic subjects. *Thorax* 2005;**60**:17–21.
- Langley SJ, Goldthorpe S, Craven M, Morris J, Woodcock A, Custovic A. Exposure and sensitization to indoor allergens: association with lung function, bronchial reactivity, and exhaled nitric oxide measures in asthma. *J Allergy Clin Immunol* 2003;**112**:362–368.
- Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A et al. Standardisation of spirometry. *Eur Respir J* 2005;**26**:319–338.

16. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary contamination. *Eur Respir J* 1996;**9**:1174–1180.
17. Ibrahim B, Basanta M, Cadden P, Singh D, Douce D, Woodcock A et al. Non-invasive phenotyping using exhaled volatile organic compounds in asthma. *Thorax* 2011;**66**:804–809.
18. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;**39**:561–577.
19. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;**86**:420–428.
20. Horvath I, Hunt J, Barnes PJ, Alving K, Antczak A, Baraldi E et al. Exhaled breath condensate: methodological recommendations and unresolved questions. *Eur Respir J* 2005;**26**:523–548.
21. Kostikas K, Papatheodorou G, Ganas K, Psathakis K, Panagou P, Loukides S. pH in expired breath condensate of patients with inflammatory airway diseases. *Am J Respir Crit Care Med* 2002;**165**:1364–1370.
22. Carraro S, Giordano G, Stocchero M, Baraldi E. Exhaled breath condensate: metabolomics. *Eur Respir Mon* 2010;**49**:231–236.
23. Barnes PJ. Future perspectives on exhaled biomarkers. *Eur Respir Mon* 2010;**49**:237–246.
24. Carraro S, Giordano G, Reniero F, Carpi D, Stocchero M, Sterk PJ et al. Asthma severity in childhood and metabolomic profiling of breath condensate. *Allergy* 2013;**68**:110–117.
25. Hauber HP, Gotfried M, Newman K, Danda R, Servi RJ, Christodoulopoulos P et al. Effect of HFA-flunisolide on peripheral lung inflammation in asthma. *J Allergy Clin Immunol* 2003;**112**:58–63.
26. Cowan DC, Cowan JO, Palmay R, Williamson A, Taylor DR. Effects of steroid therapy on inflammatory cell subtypes in asthma. *Thorax* 2010;**65**:384–390.
27. Lenz EM, Wilson ID. Analytical strategies in metabolomics. *J Proteome Res* 2007;**6**:443–458.
28. Dumas ME, Maibaum EC, Teague C, Ueshima H, Zhou B, Lindon JC et al. Assessment of analytical reproducibility of ¹H NMR spectroscopy based metabolomics for large-scale epidemiological research: the INTERMAP Study. *Anal Chem* 2006;**78**:2199–2208.
29. Pan Z, Raftery D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. *Anal Bioanal Chem* 2007;**387**:525–527.
30. Nicholson JK, Lindon JC. Systems biology: metabolomics. *Nature* 2008;**455**:1054–1056.
31. Izquierdo-Garcia JL, Peces-Barba G, Ruiz-Cabello J. Influence of ambient air on NMR-based metabolomics of exhaled breath condensates. *Eur Respir J* 2012;**40**:1294–1296.
32. Liu J, Conrad DH, Chow S, Tran VH, Yates DH, Thomas PS. Collection devices influence the constituents of exhaled breath condensate. *Eur Respir J* 2007;**30**:807–808.
33. de Laurentiis G, Paris D, Melck D, Maniscalco M, Marsico S, Corso G et al. Metabolomic analysis of exhaled breath condensate in adults by nuclear magnetic resonance spectroscopy. *Eur Respir J* 2008;**32**:1175–1183.

1 SUPPLEMENTARY MATERIAL

2

3 Breath Metabolomic Profiling by Nuclear Magnetic Resonance
4 Spectroscopy in Asthma

5

6 Baharudin Ibrahim ^{1,2}, Paul Marsden ¹, Jaclyn A Smith ¹, Adnan Custovic ¹, Mathias
7 Nilsson ^{1,3}, Stephen J Fowler ^{1,4}

8

9 ¹University of Manchester, Manchester Academic Health Science Centre, and NIHR
10 Respiratory and Allergy Clinical Research Facility, University Hospital of South
11 Manchester, Manchester, UK

12 ²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia

13 ³School of Chemistry, University of Manchester, UK

14 ⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Methods

Sputum induction and processing

Prior to sputum induction, subjects were given 200µg of salbutamol via a volumatic spacer and baseline FEV₁ was recorded. Sputum induction was performed using inhaled hypertonic saline from an ultrasonic nebuliser (Medix Sonix 2000, Clement Clarke, Essex, UK) in increasing strengths (3%, 4%, and 5%) as previously described (1).

The collected sputum was placed on ice and processed within two hours of the procedure. The salivary portion of the sample was discarded to reduce squamous cell contamination. The mucoid portion was selected and weighed to ensure at least 0.15g was available for further processing. The sample was mixed with freshly made 0.1% dithiothreitol (DTT) (Sigma, Poole, UK) with a ratio of 1g of sample to 4mL of DTT. The sputum was vortexed for 15 seconds and placed on ice on a rocker bench for 15 minutes. Phosphate buffered saline PBS (Sigma-Aldrich, UK) with the same volume as the DTT was then added and the sample again vortexed for 15 seconds and placed on rocker bench for 5 minutes. Cytospin and differential cell count of 400 cells were then performed by fully trained observers. Slides containing > 20% squamous cells were regarded as representing salivary contamination and excluded.

Exhaled breath condensate preparation and analysis by NMR spectroscopy

A volume of 25mL of buffer solution in deuterium oxide was prepared using Sorensen's phosphate buffer to give a buffer solution of 70mM with a pH of 7.3. A concentration of 0.5mg/ml of sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d₄ (TSP) was added to this solution as a chemical shift reference(2). Exhaled breath condensate samples were defrosted

immediately before processing and 100 μ L of the phosphate buffer was added to each sample to prevent any chemical shift variation due to pH.

The samples were run in a randomised order using an auto sampler and ^1H spectra were acquired non-spinning on a 500 MHz Varian VNMRS (Varian NMR Inc., USA) spectrometer at approximately 27°C, using a 5mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple probe. Prior to analysis, 30 minutes were allowed for temperature equilibration in each sample, before automatic gradient shimming and acquisition of 128 transients of 32768 complex data points using a spectral width of 12019.2 Hz, an acquisition time of 2.7 seconds, and a total recycling time of 10 seconds. Water suppression was achieved by the combination of pre-saturation (5 seconds) during the relaxation delay and a WET (water suppression enhanced through T_1 effects) element (3, 4). The auto sampler was able to accommodate a maximum of 50 samples at one time, so the samples were divided into three batches of up to 45 samples each. Each sample took approximately one hour to run. All the samples were run once again in the same order immediately after the end of the first run to test for technical reproducibility and investigate the stability of defrosted EBC (i.e. each sample was re-analysed after approximately 45 hours).

Each NMR spectrum was phase and baseline corrected and reference deconvoluted in the DOSY Toolbox (2, 5). The processed NMR spectra were binned to 0.02 ppm. The region containing the residual water peak was excluded as were empty spectral regions; a total of 367 bins were thus obtained. The peak areas were measured and normalised to TSP peak area before statistical analysis.

65

66 REFERENCES

- 67 1. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of
68 inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary
69 contamination. *Eur Respir J* 1996;**9**:1174-1180.
- 70 2. Nilsson M. The DOSY Toolbox: a new tool for processing PFG NMR diffusion data. *J Magn*
71 *Reson* 2009;**200**:296-302.
- 72 3. Ogg RJ, Kingsley PB, Taylor JS. WET, a T1- and B1-insensitive water-suppression method for
73 in vivo localized ¹H NMR spectroscopy. *J Magn Reson B* 1994;**104**:1-10.
- 74 4. Smallcombe SH, Patt SL, Keifer PA. WET Solvent Suppression and Its Applications to LC NMR
75 and High-Resolution NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* 1995;**117**:295-303.
- 76 5. Barjat H, Chilvers PB, Fetler BK, Horne TJ, Morris GA. A Practical Method for Automated
77 Shimming with Normal Spectrometer Hardware. *J Magn Reson* 1997;**125**:197-201.

78

79

80 **FIGURE LEGEND:**

81 Figure E1: Principal components analysis plot representing the selection of samples from the
82 three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles),
83 batch 3 (squares).

84

85

86

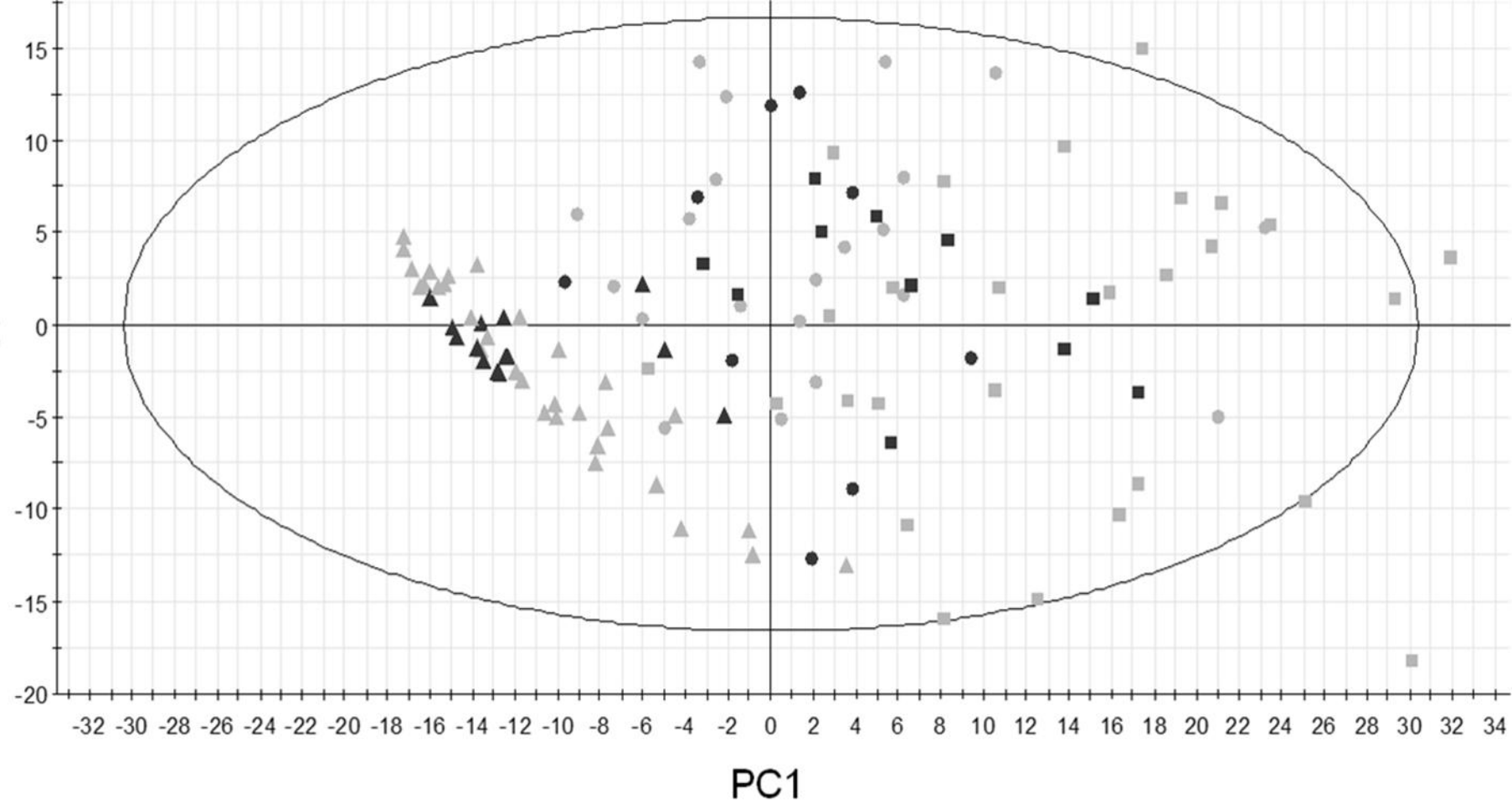
87

88

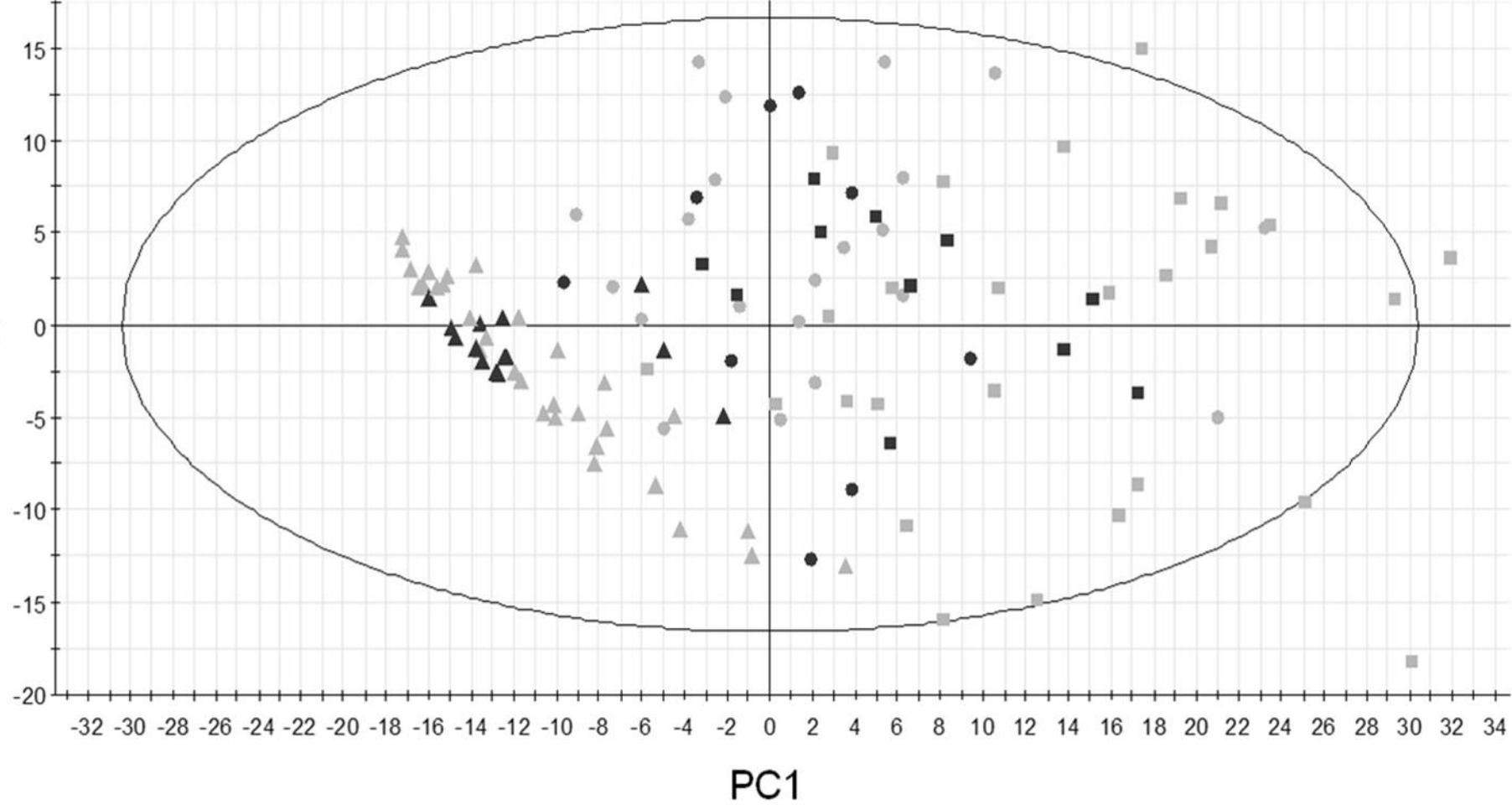
89

90

PC2



PC2



SUPPLEMENTARY MATERIAL

Breath Metabolomic Profiling by Nuclear Magnetic Resonance Spectroscopy in Asthma

Baharudin Ibrahim ^{1,2}, Paul Marsden ¹, Jaclyn A Smith ¹, Adnan Custovic ¹, Mathias Nilsson ^{1,3}, Stephen J Fowler ^{1,4}

¹University of Manchester, Manchester Academic Health Science Centre, and NIHR Respiratory and Allergy Clinical Research Facility, University Hospital of South Manchester, Manchester, UK

²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia

³School of Chemistry, University of Manchester, UK

⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Methods

Sputum induction and processing

Prior to sputum induction, subjects were given 200µg of salbutamol via a volumatic spacer and baseline FEV₁ was recorded. Sputum induction was performed using inhaled hypertonic saline from an ultrasonic nebuliser (Medix Sonix 2000, Clement Clarke, Essex, UK) in increasing strengths (3%, 4%, and 5%) as previously described (1).

The collected sputum was placed on ice and processed within two hours of the procedure. The salivary portion of the sample was discarded to reduce squamous cell contamination. The mucoid portion was selected and weighed to ensure at least 0.15g was available for further processing. The sample was mixed with freshly made 0.1% dithiothreitol (DTT) (Sigma, Poole, UK) with a ratio of 1g of sample to 4mL of DTT. The sputum was vortexed for 15 seconds and placed on ice on a rocker bench for 15 minutes. Phosphate buffered saline PBS (Sigma-Aldrich, UK) with the same volume as the DTT was then added and the sample again vortexed for 15 seconds and placed on rocker bench for 5 minutes. Cytospin and differential cell count of 400 cells were then performed by fully trained observers. Slides containing > 20% squamous cells were regarded as representing salivary contamination and excluded.

Exhaled breath condensate preparation and analysis by NMR spectroscopy

A volume of 25mL of buffer solution in deuterium oxide was prepared using Sorensen's phosphate buffer to give a buffer solution of 70mM with a pH of 7.3. A concentration of 0.5mg/ml of sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d₄ (TSP) was added to this solution as a chemical shift reference(2). Exhaled breath condensate samples were defrosted

immediately before processing and 100 μ L of the phosphate buffer was added to each sample to prevent any chemical shift variation due to pH.

The samples were run in a randomised order using an auto sampler and ^1H spectra were acquired non-spinning on a 500 MHz Varian VNMRS (Varian NMR Inc., USA) spectrometer at approximately 27°C, using a 5mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple probe. Prior to analysis, 30 minutes were allowed for temperature equilibration in each sample, before automatic gradient shimming and acquisition of 128 transients of 32768 complex data points using a spectral width of 12019.2 Hz, an acquisition time of 2.7 seconds, and a total recycling time of 10 seconds. Water suppression was achieved by the combination of pre-saturation (5 seconds) during the relaxation delay and a WET (water suppression enhanced through T_1 effects) element (3, 4). The auto sampler was able to accommodate a maximum of 50 samples at one time, so the samples were divided into three batches of up to 45 samples each. Each sample took approximately one hour to run. All the samples were run once again in the same order immediately after the end of the first run to test for technical reproducibility and investigate the stability of defrosted EBC (i.e. each sample was re-analysed after approximately 45 hours).

Each NMR spectrum was phase and baseline corrected and reference deconvoluted in the DOSY Toolbox (2, 5). The processed NMR spectra were binned to 0.02 ppm. The region containing the residual water peak was excluded as were empty spectral regions; a total of 367 bins were thus obtained. The peak areas were measured and normalised to TSP peak area before statistical analysis.

51

52 REFERENCES

- 53 1. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of
54 inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary
55 contamination. *Eur Respir J* 1996;**9**:1174-1180.
- 56 2. Nilsson M. The DOSY Toolbox: a new tool for processing PFG NMR diffusion data. *J Magn*
57 *Reson* 2009;**200**:296-302.
- 58 3. Ogg RJ, Kingsley PB, Taylor JS. WET, a T1- and B1-insensitive water-suppression method for
59 in vivo localized ¹H NMR spectroscopy. *J Magn Reson B* 1994;**104**:1-10.
- 60 4. Smallcombe SH, Patt SL, Keifer PA. WET Solvent Suppression and Its Applications to LC NMR
61 and High-Resolution NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* 1995;**117**:295-303.
- 62 5. Barjat H, Chilvers PB, Fetler BK, Horne TJ, Morris GA. A Practical Method for Automated
63 Shimming with Normal Spectrometer Hardware. *J Magn Reson* 1997;**125**:197-201.

64

65

66 **FIGURE LEGEND:**

67 Figure E1: Principal components analysis plot representing the selection of samples from the
68 three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles),
69 batch 3 (squares).

70

71

72

73

74

75

76

SUPPLEMENTARY MATERIAL

Breath Metabolomic Profiling by Nuclear Magnetic Resonance Spectroscopy in Asthma

Baharudin Ibrahim ^{1,2}, Paul Marsden ¹, Jaclyn A Smith ¹, Adnan Custovic ¹, Mathias Nilsson ^{1,3}, Stephen J Fowler ^{1,4}

¹University of Manchester, Manchester Academic Health Science Centre, and NIHR Respiratory and Allergy Clinical Research Facility, University Hospital of South Manchester, Manchester, UK

²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia

³School of Chemistry, University of Manchester, UK

⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Methods

Sputum induction and processing

Prior to sputum induction, subjects were given 200µg of salbutamol via a volumatic spacer and baseline FEV₁ was recorded. Sputum induction was performed using inhaled hypertonic saline from an ultrasonic nebuliser (Medix Sonix 2000, Clement Clarke, Essex, UK) in increasing strengths (3%, 4%, and 5%) as previously described (1).

The collected sputum was placed on ice and processed within two hours of the procedure. The salivary portion of the sample was discarded to reduce squamous cell contamination. The mucoid portion was selected and weighed to ensure at least 0.15g was available for further processing. The sample was mixed with freshly made 0.1% dithiothreitol (DTT) (Sigma, Poole, UK) with a ratio of 1g of sample to 4mL of DTT. The sputum was vortexed for 15 seconds and placed on ice on a rocker bench for 15 minutes. Phosphate buffered saline PBS (Sigma-Aldrich, UK) with the same volume as the DTT was then added and the sample again vortexed for 15 seconds and placed on rocker bench for 5 minutes. Cytospin and differential cell count of 400 cells were then performed by fully trained observers. Slides containing > 20% squamous cells were regarded as representing salivary contamination and excluded.

Exhaled breath condensate preparation and analysis by NMR spectroscopy

A volume of 25mL of buffer solution in deuterium oxide was prepared using Sorensen's phosphate buffer to give a buffer solution of 70mM with a pH of 7.3. A concentration of 0.5mg/ml of sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d₄ (TSP) was added to this solution as a chemical shift reference(2). Exhaled breath condensate samples were defrosted

immediately before processing and 100 μ L of the phosphate buffer was added to each sample to prevent any chemical shift variation due to pH.

The samples were run in a randomised order using an auto sampler and ^1H spectra were acquired non-spinning on a 500 MHz Varian VNMRS (Varian NMR Inc., USA) spectrometer at approximately 27°C, using a 5mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple probe. Prior to analysis, 30 minutes were allowed for temperature equilibration in each sample, before automatic gradient shimming and acquisition of 128 transients of 32768 complex data points using a spectral width of 12019.2 Hz, an acquisition time of 2.7 seconds, and a total recycling time of 10 seconds. Water suppression was achieved by the combination of pre-saturation (5 seconds) during the relaxation delay and a WET (water suppression enhanced through T_1 effects) element (3, 4). The auto sampler was able to accommodate a maximum of 50 samples at one time, so the samples were divided into three batches of up to 45 samples each. Each sample took approximately one hour to run. All the samples were run once again in the same order immediately after the end of the first run to test for technical reproducibility and investigate the stability of defrosted EBC (i.e. each sample was re-analysed after approximately 45 hours).

Each NMR spectrum was phase and baseline corrected and reference deconvoluted in the DOSY Toolbox (2, 5). The processed NMR spectra were binned to 0.02 ppm. The region containing the residual water peak was excluded as were empty spectral regions; a total of 367 bins were thus obtained. The peak areas were measured and normalised to TSP peak area before statistical analysis.

51

52 REFERENCES

- 53 1. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of
54 inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary
55 contamination. *Eur Respir J* 1996;**9**:1174-1180.
- 56 2. Nilsson M. The DOSY Toolbox: a new tool for processing PFG NMR diffusion data. *J Magn*
57 *Reson* 2009;**200**:296-302.
- 58 3. Ogg RJ, Kingsley PB, Taylor JS. WET, a T1- and B1-insensitive water-suppression method for
59 in vivo localized ¹H NMR spectroscopy. *J Magn Reson B* 1994;**104**:1-10.
- 60 4. Smallcombe SH, Patt SL, Keifer PA. WET Solvent Suppression and Its Applications to LC NMR
61 and High-Resolution NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* 1995;**117**:295-303.
- 62 5. Barjat H, Chilvers PB, Fetler BK, Horne TJ, Morris GA. A Practical Method for Automated
63 Shimming with Normal Spectrometer Hardware. *J Magn Reson* 1997;**125**:197-201.

64

65

66 **FIGURE LEGEND:**

67 Figure S1: Principal components analysis plot representing the selection of samples from the
68 three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles),
69 batch 3 (squares).

70

71

72

73

74

75

76

1 SUPPLEMENTARY MATERIAL

2

3 Breath Metabolomic Profiling by Nuclear Magnetic Resonance
4 Spectroscopy in Asthma

5

6 Baharudin Ibrahim ^{1,2}, Paul Marsden ¹, Jaclyn A Smith ¹, Adnan Custovic ¹, Mathias
7 Nilsson ^{1,3}, Stephen J Fowler ^{1,4}

8

9 ¹University of Manchester, Manchester Academic Health Science Centre, and NIHR
10 Respiratory and Allergy Clinical Research Facility, University Hospital of South
11 Manchester, Manchester, UK

12 ²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia

13 ³School of Chemistry, University of Manchester, UK

14 ⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Methods

Sputum induction and processing

Prior to sputum induction, subjects were given 200µg of salbutamol via a volumatic spacer and baseline FEV₁ was recorded. Sputum induction was performed using inhaled hypertonic saline from an ultrasonic nebuliser (Medix Sonix 2000, Clement Clarke, Essex, UK) in increasing strengths (3%, 4%, and 5%) as previously described (1).

The collected sputum was placed on ice and processed within two hours of the procedure. The salivary portion of the sample was discarded to reduce squamous cell contamination. The mucoid portion was selected and weighed to ensure at least 0.15g was available for further processing. The sample was mixed with freshly made 0.1% dithiothreitol (DTT) (Sigma, Poole, UK) with a ratio of 1g of sample to 4mL of DTT. The sputum was vortexed for 15 seconds and placed on ice on a rocker bench for 15 minutes. Phosphate buffered saline PBS (Sigma-Aldrich, UK) with the same volume as the DTT was then added and the sample again vortexed for 15 seconds and placed on rocker bench for 5 minutes. Cytospin and differential cell count of 400 cells were then performed by fully trained observers. Slides containing > 20% squamous cells were regarded as representing salivary contamination and excluded.

Exhaled breath condensate preparation and analysis by NMR spectroscopy

A volume of 25mL of buffer solution in deuterium oxide was prepared using Sorensen's phosphate buffer to give a buffer solution of 70mM with a pH of 7.3. A concentration of 0.5mg/ml of sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d₄ (TSP) was added to this solution as a chemical shift reference(2). Exhaled breath condensate samples were defrosted

immediately before processing and 100µL of the phosphate buffer was added to each sample to prevent any chemical shift variation due to pH.

The samples were run in a randomised order using an auto sampler and ^1H spectra were acquired non-spinning on a 500 MHz Varian VNMRS (Varian NMR Inc., USA) spectrometer at approximately 27°C, using a 5mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple probe. Prior to analysis, 30 minutes were allowed for temperature equilibration in each sample, before automatic gradient shimming and acquisition of 128 transients of 32768 complex data points using a spectral width of 12019.2 Hz, an acquisition time of 2.7 seconds, and a total recycling time of 10 seconds. Water suppression was achieved by the combination of pre-saturation (5 seconds) during the relaxation delay and a WET (water suppression enhanced through T_1 effects) element (3, 4). The auto sampler was able to accommodate a maximum of 50 samples at one time, so the samples were divided into three batches of up to 45 samples each. Each sample took approximately one hour to run. All the samples were run once again in the same order immediately after the end of the first run to test for technical reproducibility and investigate the stability of defrosted EBC (i.e. each sample was re-analysed after approximately 45 hours).

Each NMR spectrum was phase and baseline corrected and reference deconvoluted in the DOSY Toolbox (2, 5). The processed NMR spectra were binned to 0.02 ppm. The region containing the residual water peak was excluded as were empty spectral regions; a total of 367 bins were thus obtained. The peak areas were measured and normalised to TSP peak area before statistical analysis.

65

66 REFERENCES

- 67 1. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of
68 inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary
69 contamination. *Eur Respir J* 1996;**9**:1174-1180.
- 70 2. Nilsson M. The DOSY Toolbox: a new tool for processing PFG NMR diffusion data. *J Magn*
71 *Reson* 2009;**200**:296-302.
- 72 3. Ogg RJ, Kingsley PB, Taylor JS. WET, a T1- and B1-insensitive water-suppression method for
73 in vivo localized ¹H NMR spectroscopy. *J Magn Reson B* 1994;**104**:1-10.
- 74 4. Smallcombe SH, Patt SL, Keifer PA. WET Solvent Suppression and Its Applications to LC NMR
75 and High-Resolution NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* 1995;**117**:295-303.
- 76 5. Barjat H, Chilvers PB, Fetler BK, Horne TJ, Morris GA. A Practical Method for Automated
77 Shimming with Normal Spectrometer Hardware. *J Magn Reson* 1997;**125**:197-201.

78

79

80 **FIGURE LEGEND:**

81 Figure E1: Principal components analysis plot representing the selection of samples from the
82 three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles),
83 batch 3 (squares).

84

85

86

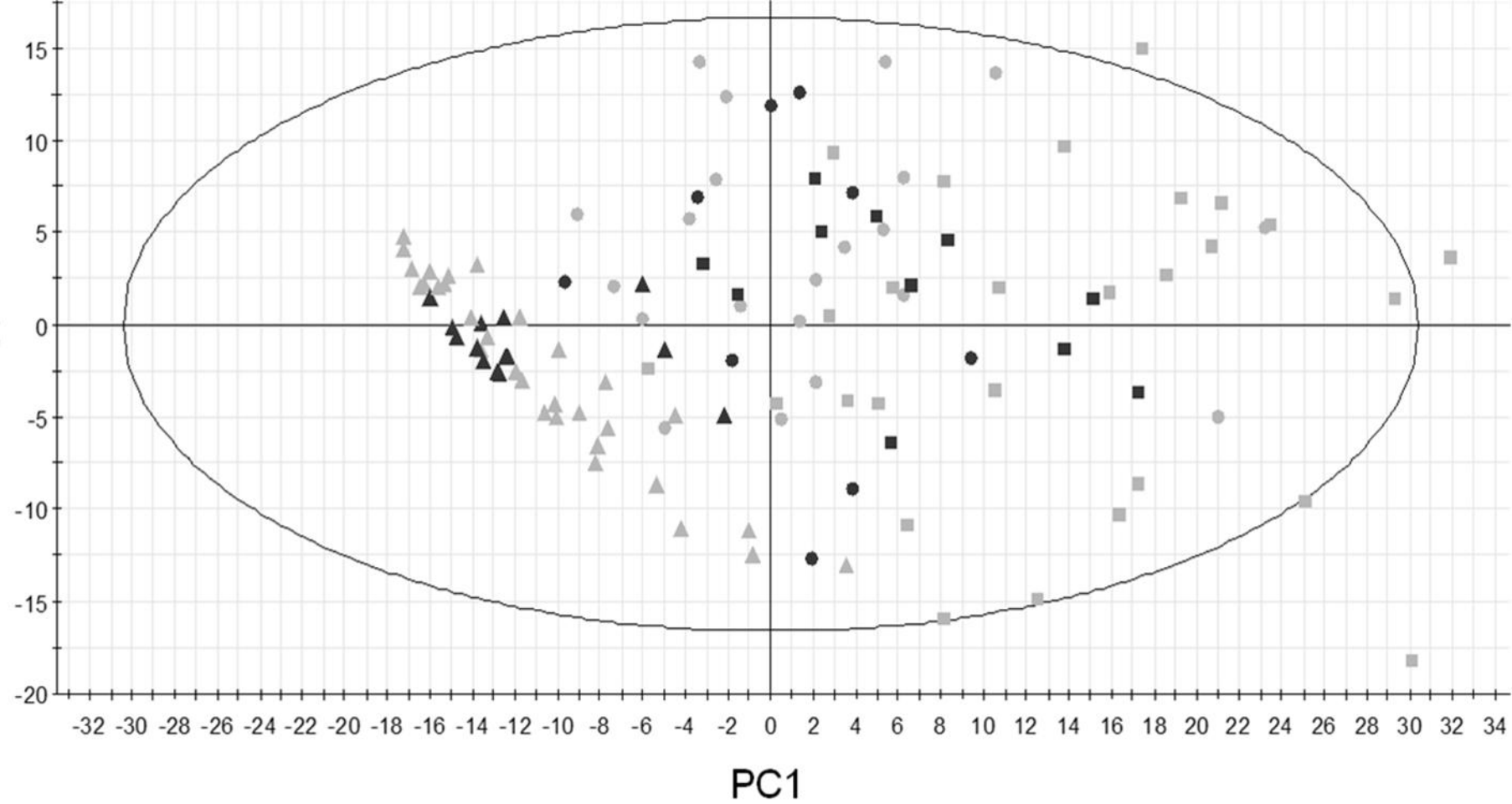
87

88

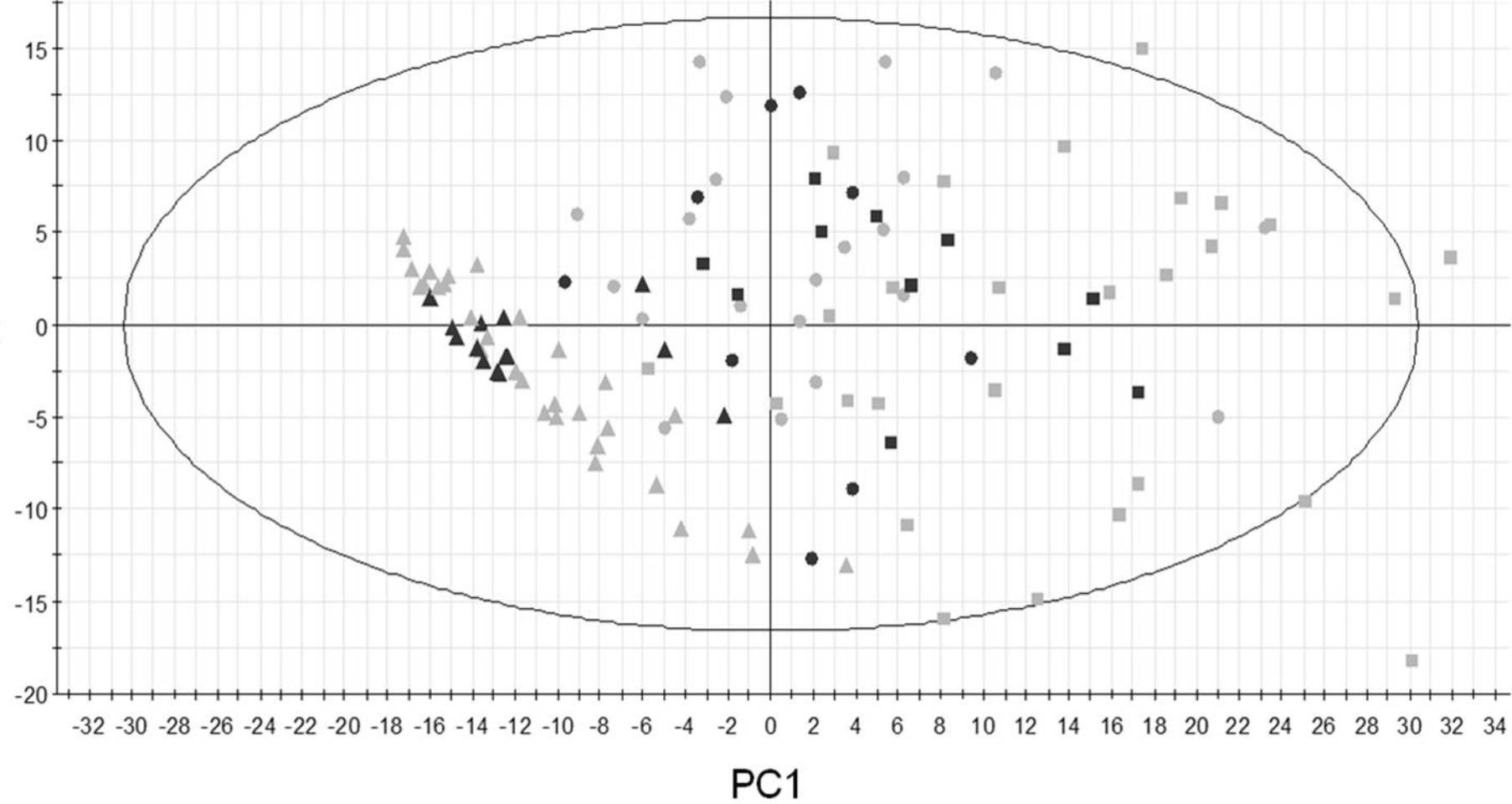
89

90

PC2



PC2



SUPPLEMENTARY MATERIAL

Breath Metabolomic Profiling by Nuclear Magnetic Resonance Spectroscopy in Asthma

Baharudin Ibrahim ^{1,2}, Paul Marsden ¹, Jaclyn A Smith ¹, Adnan Custovic ¹, Mathias Nilsson ^{1,3}, Stephen J Fowler ^{1,4}

¹University of Manchester, Manchester Academic Health Science Centre, and NIHR Respiratory and Allergy Clinical Research Facility, University Hospital of South Manchester, Manchester, UK

²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia

³School of Chemistry, University of Manchester, UK

⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Methods

Sputum induction and processing

Prior to sputum induction, subjects were given 200µg of salbutamol via a volumatic spacer and baseline FEV₁ was recorded. Sputum induction was performed using inhaled hypertonic saline from an ultrasonic nebuliser (Medix Sonix 2000, Clement Clarke, Essex, UK) in increasing strengths (3%, 4%, and 5%) as previously described (1).

The collected sputum was placed on ice and processed within two hours of the procedure. The salivary portion of the sample was discarded to reduce squamous cell contamination. The mucoid portion was selected and weighed to ensure at least 0.15g was available for further processing. The sample was mixed with freshly made 0.1% dithiothreitol (DTT) (Sigma, Poole, UK) with a ratio of 1g of sample to 4mL of DTT. The sputum was vortexed for 15 seconds and placed on ice on a rocker bench for 15 minutes. Phosphate buffered saline PBS (Sigma-Aldrich, UK) with the same volume as the DTT was then added and the sample again vortexed for 15 seconds and placed on rocker bench for 5 minutes. Cytospin and differential cell count of 400 cells were then performed by fully trained observers. Slides containing > 20% squamous cells were regarded as representing salivary contamination and excluded.

Exhaled breath condensate preparation and analysis by NMR spectroscopy

A volume of 25mL of buffer solution in deuterium oxide was prepared using Sorensen's phosphate buffer to give a buffer solution of 70mM with a pH of 7.3. A concentration of 0.5mg/ml of sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d₄ (TSP) was added to this solution as a chemical shift reference(2). Exhaled breath condensate samples were defrosted

immediately before processing and 100 μ L of the phosphate buffer was added to each sample to prevent any chemical shift variation due to pH.

The samples were run in a randomised order using an auto sampler and ^1H spectra were acquired non-spinning on a 500 MHz Varian VNMRS (Varian NMR Inc., USA) spectrometer at approximately 27°C, using a 5mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple probe. Prior to analysis, 30 minutes were allowed for temperature equilibration in each sample, before automatic gradient shimming and acquisition of 128 transients of 32768 complex data points using a spectral width of 12019.2 Hz, an acquisition time of 2.7 seconds, and a total recycling time of 10 seconds. Water suppression was achieved by the combination of pre-saturation (5 seconds) during the relaxation delay and a WET (water suppression enhanced through T_1 effects) element (3, 4). The auto sampler was able to accommodate a maximum of 50 samples at one time, so the samples were divided into three batches of up to 45 samples each. Each sample took approximately one hour to run. All the samples were run once again in the same order immediately after the end of the first run to test for technical reproducibility and investigate the stability of defrosted EBC (i.e. each sample was re-analysed after approximately 45 hours).

Each NMR spectrum was phase and baseline corrected and reference deconvoluted in the DOSY Toolbox (2, 5). The processed NMR spectra were binned to 0.02 ppm. The region containing the residual water peak was excluded as were empty spectral regions; a total of 367 bins were thus obtained. The peak areas were measured and normalised to TSP peak area before statistical analysis.

51

52 REFERENCES

- 53 1. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of
54 inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary
55 contamination. *Eur Respir J* 1996;**9**:1174-1180.
- 56 2. Nilsson M. The DOSY Toolbox: a new tool for processing PFG NMR diffusion data. *J Magn*
57 *Reson* 2009;**200**:296-302.
- 58 3. Ogg RJ, Kingsley PB, Taylor JS. WET, a T1- and B1-insensitive water-suppression method for
59 in vivo localized ¹H NMR spectroscopy. *J Magn Reson B* 1994;**104**:1-10.
- 60 4. Smallcombe SH, Patt SL, Keifer PA. WET Solvent Suppression and Its Applications to LC NMR
61 and High-Resolution NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* 1995;**117**:295-303.
- 62 5. Barjat H, Chilvers PB, Fetler BK, Horne TJ, Morris GA. A Practical Method for Automated
63 Shimming with Normal Spectrometer Hardware. *J Magn Reson* 1997;**125**:197-201.

64

65

66 **FIGURE LEGEND:**

67 Figure E1: Principal components analysis plot representing the selection of samples from the
68 three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles),
69 batch 3 (squares).

70

71

72

73

74

75

76

SUPPLEMENTARY MATERIAL

Breath Metabolomic Profiling by Nuclear Magnetic Resonance Spectroscopy in Asthma

Baharudin Ibrahim ^{1,2}, Paul Marsden ¹, Jaclyn A Smith ¹, Adnan Custovic ¹, Mathias Nilsson ^{1,3}, Stephen J Fowler ^{1,4}

¹University of Manchester, Manchester Academic Health Science Centre, and NIHR Respiratory and Allergy Clinical Research Facility, University Hospital of South Manchester, Manchester, UK

²School of Pharmaceutical Sciences, Universiti Sains Malaysia, Penang, Malaysia

³School of Chemistry, University of Manchester, UK

⁴Lancashire Teaching Hospitals NHS Foundation Trust, Preston, UK

Methods

Sputum induction and processing

Prior to sputum induction, subjects were given 200µg of salbutamol via a volumatic spacer and baseline FEV₁ was recorded. Sputum induction was performed using inhaled hypertonic saline from an ultrasonic nebuliser (Medix Sonix 2000, Clement Clarke, Essex, UK) in increasing strengths (3%, 4%, and 5%) as previously described (1).

The collected sputum was placed on ice and processed within two hours of the procedure. The salivary portion of the sample was discarded to reduce squamous cell contamination. The mucoid portion was selected and weighed to ensure at least 0.15g was available for further processing. The sample was mixed with freshly made 0.1% dithiothreitol (DTT) (Sigma, Poole, UK) with a ratio of 1g of sample to 4mL of DTT. The sputum was vortexed for 15 seconds and placed on ice on a rocker bench for 15 minutes. Phosphate buffered saline PBS (Sigma-Aldrich, UK) with the same volume as the DTT was then added and the sample again vortexed for 15 seconds and placed on rocker bench for 5 minutes. Cytospin and differential cell count of 400 cells were then performed by fully trained observers. Slides containing > 20% squamous cells were regarded as representing salivary contamination and excluded.

Exhaled breath condensate preparation and analysis by NMR spectroscopy

A volume of 25mL of buffer solution in deuterium oxide was prepared using Sorensen's phosphate buffer to give a buffer solution of 70mM with a pH of 7.3. A concentration of 0.5mg/ml of sodium 3-(trimethylsilyl)-propionate-2,2,3,3-d₄ (TSP) was added to this solution as a chemical shift reference(2). Exhaled breath condensate samples were defrosted

immediately before processing and 100 μ L of the phosphate buffer was added to each sample to prevent any chemical shift variation due to pH.

The samples were run in a randomised order using an auto sampler and ^1H spectra were acquired non-spinning on a 500 MHz Varian VNMRS (Varian NMR Inc., USA) spectrometer at approximately 27°C, using a 5mm $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ triple probe. Prior to analysis, 30 minutes were allowed for temperature equilibration in each sample, before automatic gradient shimming and acquisition of 128 transients of 32768 complex data points using a spectral width of 12019.2 Hz, an acquisition time of 2.7 seconds, and a total recycling time of 10 seconds. Water suppression was achieved by the combination of pre-saturation (5 seconds) during the relaxation delay and a WET (water suppression enhanced through T_1 effects) element (3, 4). The auto sampler was able to accommodate a maximum of 50 samples at one time, so the samples were divided into three batches of up to 45 samples each. Each sample took approximately one hour to run. All the samples were run once again in the same order immediately after the end of the first run to test for technical reproducibility and investigate the stability of defrosted EBC (i.e. each sample was re-analysed after approximately 45 hours).

Each NMR spectrum was phase and baseline corrected and reference deconvoluted in the DOSY Toolbox (2, 5). The processed NMR spectra were binned to 0.02 ppm. The region containing the residual water peak was excluded as were empty spectral regions; a total of 367 bins were thus obtained. The peak areas were measured and normalised to TSP peak area before statistical analysis.

51

52 REFERENCES

- 53 1. Pizzichini E, Pizzichini MM, Efthimiadis A, Hargreave FE, Dolovich J. Measurement of
54 inflammatory indices in induced sputum: effects of selection of sputum to minimize salivary
55 contamination. *Eur Respir J* 1996;**9**:1174-1180.
- 56 2. Nilsson M. The DOSY Toolbox: a new tool for processing PFG NMR diffusion data. *J Magn*
57 *Reson* 2009;**200**:296-302.
- 58 3. Ogg RJ, Kingsley PB, Taylor JS. WET, a T1- and B1-insensitive water-suppression method for
59 in vivo localized ¹H NMR spectroscopy. *J Magn Reson B* 1994;**104**:1-10.
- 60 4. Smallcombe SH, Patt SL, Keifer PA. WET Solvent Suppression and Its Applications to LC NMR
61 and High-Resolution NMR Spectroscopy. *Journal of Magnetic Resonance, Series A* 1995;**117**:295-303.
- 62 5. Barjat H, Chilvers PB, Fetler BK, Horne TJ, Morris GA. A Practical Method for Automated
63 Shimming with Normal Spectrometer Hardware. *J Magn Reson* 1997;**125**:197-201.

64

65

66 **FIGURE LEGEND:**

67 Figure S1: Principal components analysis plot representing the selection of samples from the
68 three batches for training (grey) and test (black) sets: batch 1 (triangles), batch 2 (circles),
69 batch 3 (squares).

70

71

72

73

74

75

76