Non-invasive metabolomic analysis of breath using differential mobility spectrometry in patients with chronic obstructive pulmonary disease and healthy smokers

Maria Basanta,[a] Roger M. Jarvis,[b] Yun Xu,[b] Gavin Blackburn,[c] Ruth Tal-Singer,[d] Ashley Woodcock,[a] Dave Singh,[a] Royston Goodacre,[b] C. L. Paul Thomas[c] and Stephen J. Fowler*[ae]

Received 13th August 2009, Accepted 14th December 2009
First published as an Advance Article on the web 5th January 2010
DOI: 10.1039/b916374c

The rapid, accurate and non-invasive diagnosis of respiratory disease represents a challenge to clinicians, and the development of new treatments can be confounded by insufficient knowledge of lung disease phenotypes. Exhaled breath contains a complex mixture of volatile organic compounds (VOCs), some of which could potentially represent biomarkers for lung diseases. We have developed an adaptive sampling methodology for collecting concentrated samples of exhaled air from participants with impaired respiratory function, against which we employed two-stage thermal desorption gas chromatography-differential mobility spectrometry (GC-DMS) analysis, and showed that it was possible to discriminate between participants with and without chronic obstructive pulmonary disease (COPD). A 2.5 dm^3 volume of end tidal breath was collected onto adsorbent traps (Tenax TA/Carbotrap), from participants with severe COPD and healthy volunteers. Samples were thermally desorbed and analysed by GC-DMS, and the chromatograms analysed by univariate and multivariate analyses. Kruskal–Wallis ANOVA indicated several discriminatory ($p < 0.01$) signals, with good classification performance (receiver operator characteristic area up to 0.82). Partial least squares discriminant analysis using the full DMS chromatograms also gave excellent discrimination between groups ($\alpha = 19\%$ and $\beta = 12.4\%$).

Introduction

Chronic obstructive pulmonary disease (COPD) is a multi-phenotypic syndrome described by progressive airflow obstruction and inflammation, usually secondary to chronic inhalation of cigarette smoke. There is significant inter-individual variability in the clinical presentation and prognosis with many lifelong smokers never developing symptoms at all. The discovery of novel biomarkers that can identify disease phenotypes, assess response to treatment, and even identify "at-risk" smokers is a critical target for COPD research. These biomarkers may be single compounds or a metabolomic "fingerprint" unique to individual patients, but sharing common features within disease sub-phenotypes.

The analysis of volatile organic compounds (VOCs) in the exhaled breath in COPD may provide novel information about altered metabolic processes, and the extent and nature of tissue damage *via* a non-invasive method. A complete description of volatiles associated with COPD has yet to be developed. Nevertheless, studies have identified groups of volatiles and other biomarkers in COPD patients that have potentially useful relationships to cell degradation processes.[1,2]

Measuring specific volatiles requires a method that separates them from their immediate environment and then each other before they are detected. There are several analytical detection approaches available, including gas chromatography-mass spectrometry (GC-MS), proton transfer reaction-mass spectrometry (PTR-MS), selected ion flow tube-mass spectrometry (SIFT-MS) and ion mobility spectrometry (IMS)[3,4] and more recently electronic noses and sensors.[5,6] Differential mobility spectrometry (DMS) is a variant of IMS and has been used recently to complement mass spectrometry in peptide analysis.[7] However, it is also a compact (mass less than 5 kg), sensitive (ppt(v)) and robust (implemented in fieldable chemical agent detection systems) stand-alone analytical tool with a significant advantage of atmospheric pressure operation, and as such is ideal for future point-of-care applications. The principles and theory behind DMS technology have been reported previously,[8,9] and we have already demonstrated the feasibility of breath collection and VOC-analysis using DMS.[10] This new study aims to test the reproducibility of the sampling methodology when combined with GC-DMS and also exploits chemometrics in order to allow the unequivocal identification of participants with COPD, and compares their breath profiles to those of smokers without the disease.

Experimental

Participants

Twenty six participants with a minimum 20 pack-year (*i.e.* one pack per day for 20 years) smoking history were enrolled in the

[a]*Respiratory Research Group, Manchester Academic Health Science Centre, University of Manchester, Wythenshawe Hospital, Manchester, UK. E-mail: stephen.fowler@manchester.ac.uk*
[b]*School of Chemistry & Manchester Interdisciplinary Biocentre, University of Manchester, UK*
[c]*School of Chemistry, Loughborough University, UK*
[d]*GlaxoSmithKline, King of Prussia, Pennsylvania, USA*
[e]*Lancashire Teaching Hospitals NHS Trust, Preston, UK*

study [20 COPD, 6 healthy smokers (HS)]. Participants with COPD fulfilled international consensus definitions for this disease[11] and had no other respiratory disease. HS had no known respiratory disease and normal lung function. Participants were recruited from the clinical trials database held by the Medicines Evaluation Unit, Wythenshawe Hospital, Manchester, UK. This study was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The local ethics committee approved the protocol before commencement of the study, and all subjects gave written informed consent.

Demographic data collected from all participants included gender, age, smoking status and pack-year history, height, weight, BMI, medical history, co-morbidities, medication, pulmonary function, alcohol consumption, hours of sleep before visit and details of diet and cosmetic-usage in 24 h prior to attending the laboratory (Table 1). Breath collections were made in the morning: participants were asked to fast for 8 h prior to sampling, not to wear make-up, perfume or any other facial products, not to smoke and to withhold their morning medication until after sampling. Spirometry was performed using a Vitalograph spirometer (Vitalograph Ltd, UK) according to American Thoracic Society guidelines.[12]

Breath collection

The adaptive breath sampler has been described previously.[10] In brief, a full face mask of the sort commonly used for non-invasive ventilation (Resmed Mirage™ NIV full face mask, UK) was fitted to the face of the subject and checks made to ensure comfort. Purified medical air (35 L min^{-1}) was supplied to the mask through a one-way non re-breathing valve and the subject acclimatised to the system for up to 10 min. The breathing mask had two luer fittings, one of which was used for the sampling point and connected to the sampling inlet of the sampling control unit through an adsorbent sampler assembly. The second luer fitting on the mask was connected, using silicone tubing, to an integrated-circuit pressure sensor. This enabled selected sampling during a pre-defined part of the respiratory cycle; in this study we programmed the sensor to trigger sampling only during late expiration, on the assumption that we would therefore be more likely to sample from the lower respiratory tract. Five replicates of 2.5 L breath samples were collected per subject.

Table 1 Demographic data shown as mean ± S.D. (range) except where indicated. Forced expiratory volume in 1 s (FEV1) and FEV1/forced vital capacity (FVC) ratio are both physiological measures of airflow obstruction typically used in the diagnosis of COPD

	COPD ($n = 20$)	Healthy smokers ($n = 6$)
Age (years)	64.3 ± 4.4 (58–70)	56.8 ± 13.1 (41–78)
Sex (male/female)	15/5	3/3
BMI (kg m^{-2})	27.6 ± 5.6 (21–42)	26.7 ± 2.6 (25–31)
Smokers (ex/current)	18/2	0/6
Pack years	52 ± 17 (23–78)	48 ± 37 (20–120)
FEV1 (L)	1.18 ± 0.40 (0.63–2.04)	2.69 ± 0.66 (2.23–3.64)
FEV1 (% predicted)	41.8 ± 11.8 (22–63)	100.8 ± 9.9 (88–112)
FEV1/FVC ratio	39.4 ± 10.2 (24–61)	78.5 ± 5.8 (74–87)
Co-morbidities	3 Gastro oesophageal reflux 8 Hypertension 2 Type II diabetes mellitus	2 Gastro oesophageal reflux 1 Hypertension

Breath analysis

Samples were analysed within one week of collection by gas chromatography-differential mobility spectrometry (GC-DMS) as previously described.[13] Adsorbent traps were packed with Tenax TA/Carbotrap (Markes International, UK). Prior to sampling, traps were conditioned with approximately 100 cm^3 min^{-1} pure N_2 at 300 °C overnight. After sampling adsorbent tubes were transported to the laboratory and stored at 4 °C. Adsorbent tubes were analysed in random order within one week of collection. Between each sample analysis, an instrument blank was run to ensure cross-contamination between samples was prevented. Thermal desorption (Unity™, Markes International, UK) was effected in two stages: from the adsorbent trap into a cold trap at 300 °C as a concentration step and then from the cold trap to the column at 360 °C in splitless mode. GC-DMS analysis was carried out using a Hewlett-Packard gas chromatograph coupled to a differential mobility spectrometer (Sionex, USA) with the dispersion field programmed from 14 to 30 kV cm^{-1} and the compensation field scanned from −860 to 300 V cm^{-1}. The scan duration was 2.45 s with a scan size of 100 steps. The exhaled breath volatiles were separated using a 5% biphenyl and 95% dimethylpolysiloxane capillary column [30 m long, 0.25 mm diameter with a 0.25 μm film thickness; Rtx-5MS low bleed (Restek)] held for 6 min at 40 °C, and then ramped to 220 °C at 4 °C min^{-1}. The final temperature was held for 3 min.

Statistical analysis

Outputs from the GC-DMS were pre-processed before proceeding to univariate and multivariate analyses. Fig. 1 represents examples of DMS exhaled breath outputs for a HS (top) and a subject with COPD (bottom). A large part of the visual output reflects the VOC-content of inspired air, which in the controlled environment was similar for all participants. In order to compensate for unavoidable baseline shifts the data were pre-processed by a three stage process involving interpolation, baseline correction and finally alignment of the chromatograms; see Fig. 2 for details. Interpolation is necessary because the sampling rate is relatively low and resulted in poor peak shape which may cause problems of alignment. Cubic spline interpolation was used. Baseline correction was achieved using asymmetric least squares (ALSs), an adaptive baseline correction algorithm as described by Eilers.[14] Alignment was applied to the whole dataset *via* the algorithm correlation optimisation warping (COW). The number of segments and the slacking size were optimised by using a simplex optimisation procedure.[15,16]

To test if the replicates from a single participant were reproducible a pair-wise distance matrix was produced using (1-correlation coefficient) measuring the dissimilarity between each pair of chromatograms. The dissimilarity matrix was then averaged according to the samples originating from the same individual and between individuals. Given m individuals in

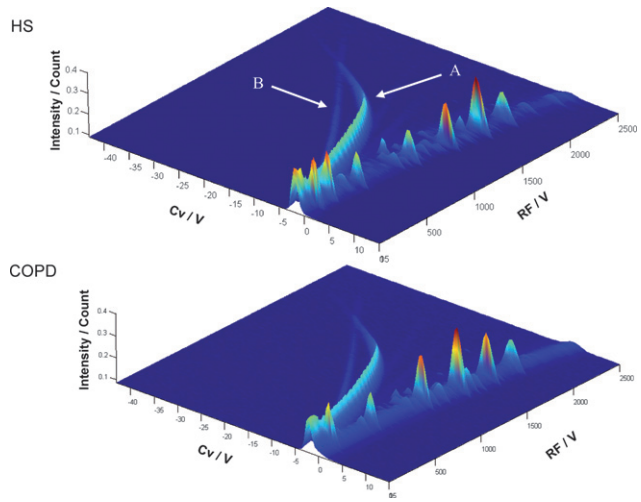

Fig. 1 Example contour plot of DMS output resultant from an exhaled breath sample of a healthy smoker (top) and a COPD subject (bottom). The trace that shifts from −5 to −40 V corresponds to the reactant ion response (A) together with the ammonia response (B). Note that these responses were removed during data pre-processing. Product compounds eluting from the column are detected and response output falls between −10 and 10 V.

a group, there are $m \times (m - 1)/2$ average dissimilarities between samples obtained from different individuals and m dissimilarities among the samples originated from the same individual. Thus, $m \times (m - 1)/2 + m$ dissimilarities were obtained for each group. Then dissimilarities were ranked from 1 (the least dissimilar) to $(m \times (m - 1)/2 + m)$ (the most dissimilar). The difference between the ranks originated from the same individuals and those from different individuals was evaluated by using Kolmogorov–Smirnov goodness-of-fit hypothesis test. The detailed description of this methodology can be found in ref. 17.

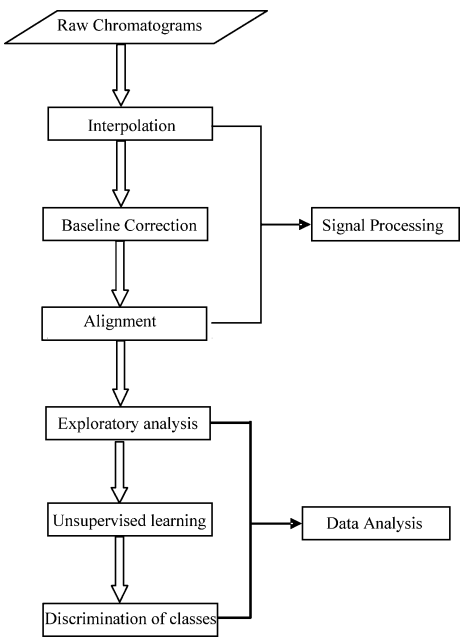

Fig. 2 Data processing/analysis flow chart.

Clinical metadata relating to individual participants were evaluated by Kruskal–Wallis (KW) analysis of variance (ANOVA) to determine which demographic attributes discriminated significantly between groups. KW-ANOVA tests the null hypothesis that there is no difference between the group medians, for which a χ^2 probability or p-value can be calculated, a $p < 0.05$ would indicate a significant difference between-group medians within a 95% confidence level. In addition to the calculation of p-values by KW-ANOVA, the area under the receiver operating characteristic (ROC) curve for all chromatographic DMS time points was calculated and a scatter plot of these two metrics was used to determine the chromatographic bins which showed greatest differentiation between individuals with COPD and HS. The ROC is a calculation of the rate of false positive (1-sensitivity) to true positive (specificity) classifications which makes no assumption of class boundaries.[18] An area under ROC of 0.5 would indicate a model with a classification accuracy no better than chance; the closer this value is to unity, the better the classification power of a particular variable. These p-values and ROC areas were used to determine the chromatographic peaks that gave the best discrimination and classification accuracy.

Finally, a multivariate predictive model designed to discriminate between COPD and HS was derived from DMS chromatograms using partial least squares discriminant analysis (PLS-DA).[19] PLS-DA is a supervised fixed linear regression algorithm which can be used to formulate empirical models from multivariate datasets. The algorithm was implemented using a random re-sampling methodology[20] over 200 iterations. First, the two groups were balanced by random selection from the larger set of COPD samples, and then the training and test samples for PLS-DA were selected by a second random split to give 4 training samples and 1 test sample for each model. For each "training" and "test" dataset, replicates from individuals were kept together. Using these sub-sampled data the PLS-DA outputs were encoded as '1' for COPD and '0' for HS.

Results and discussion

Metadata

Demographic details are shown in Table 1. The breath sampling procedure was well tolerated in all participants, and notably including those with severe impairment of respiratory function.

Breath analysis can be influenced by a multitude of factors both endogenous (gender, age, weight, *etc.*) and exogenous (exposure to VOCs in the place of work, diet, medication, *etc.*).[4,5,21] At this early stage of metabolomic investigation of the breath in respiratory disease, we feel it is vital to control for as many of the many potential confounders as possible. In this regard our methodology controls the volatiles in the background so all the participants breathe similar inspiratory air and thus we are able to control and/or monitor other potential exogenous sources of VOCs.

The clinical study presented herein was evaluated to look at the effect and the influence of these factors. Principal component analysis of preliminary GC-DMS data (data not shown), where a group of replicate samples from the same individual were very different (outliers) from all other participants, suggested that one such factor, the wearing of make-up, could have a significant

influence on the results. Therefore the samples from this individual are not representative of the breath from a human lung and its airways and so must be discarded. Other factors were also considered and when inspecting the clinical metadata for the completed study we used KW-ANOVA to test the null hypothesis that there is no difference between the group medians of the metadata with respect to COPD *versus* HS and found that using a p-value <0.05 spirometry and whether the participant had consumed caffeinated drinks that day to be the only measured significant confounders (Table 2). For future studies it is clear that we need to define which exogenous factors need to be controlled for, and that again detailed recording of all factors that may influence the VOC profile should be recorded.

In addition to the clinical metadata, KW-ANOVA was performed on the sampling metadata associated with each participant (Table 2). Parameters such as the individual components used in the sampling system, such as the connecting tubes and masks used, had p-values >0.7 indicating no evidence of any differences between COPD and HS in terms of the sampling apparatus used and the health status of the individuals analysed. This demonstrates that minimal bias from the sampling system was introduced through the experimental intervention. Both dynamic change and endogenous–exogenous equilibrium in exhaled breath analysis present difficulties in achieving reproducibility from replicate breath samples; an essential prerequisite for a metabolomic approach. We propose our careful methodology which controls for as many of these factors as is reasonable, is vital for achieving such essential reproducibility

Reproducibility

Data from 130 chromatograms were analysed, five from each of the 26 participants. Reproducibility of the sampling methodology was studied by ranking the overall similarities (pair-wise distance matrices) within and between individuals as detailed above. As shown in Fig. 3 within-subject variability (AA) was significantly lower than between-subject variability (AB) ($p = 6.23 \times 10^{-23}$). Hence, it can be concluded that there are significant amount of very reproducible signals in the exhaled breath samples collected from the same individuals, and that different individuals displayed different (but reproducible) VOC profiles.

Table 2 Between-group p-values from Kruskal–Wallis ANOVA for each of the clinical metadata

Attribute	Kruskal–Wallis p-value
Spirometry	0.0003
Coffee/tea	0.0199
Age (years)	0.067
Pack years	0.2408
Sex (M/F)	0.2538
Morning cigarette	0.2598
Cosmetics	0.3564
Mask no.	0.7324
Height (m)	0.7374
Capillary no.	0.8529
BMI (kg m^{-2})	0.86
Weight (kg)	0.9031

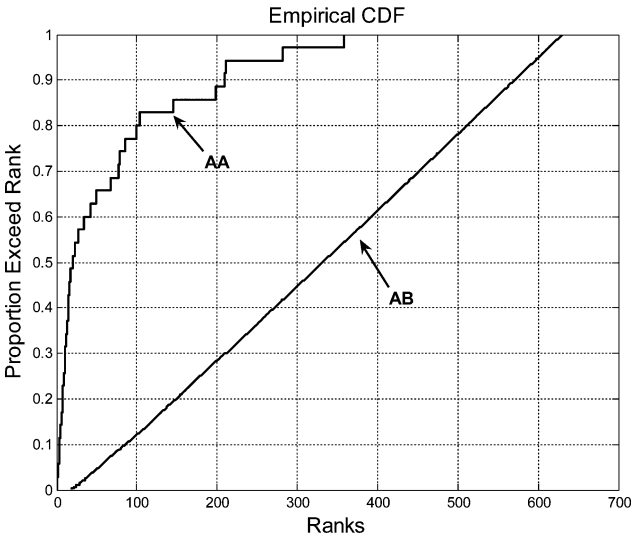

Fig. 3 Rank test: AA represents the cumulative distribution function (CDF) of the ranks of the averaged dissimilarity of the samples coming from the same individual and AB represents that of the samples coming from different individuals.

Univariate analysis

The performance as a classifier of intensity measured at each time point in the DMS chromatogram was assessed by KW-ANOVA and ROC (Fig. 4). As can be seen in this figure the best chromatographic region for discrimination was selected based upon a low p-value/large ROC area trade-off. This chromatographic region corresponded to a retention time of 179.33 s, and discriminated between COPD and HS with a $p < 5 \times 10^{-5}$, and the ROC area under the curve of 0.82 (Table 3). By further analysis of subgroups based on the metadata (such as age, smoking history, *etc.*) we demonstrated that the main (recorded) factor influencing the result was indeed the diagnosis of COPD,

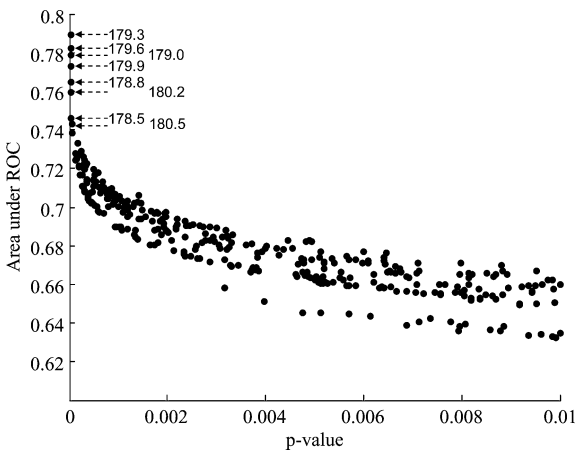

Fig. 4 Kruskal–Wallis ANOVA p-values *versus* area under the ROC (AUROC) for all chromatographic time points with $p < 0.01$. The best-discriminating points at the left of the plot (with low p-values and high AUROC) are labeled with their retention times (s) from the chromatogram, and it can be seen that they arise from one specific region of the chromatogram between 178.5 and 180.5 s.

Table 3 Model performance parameters and area under ROC (AUROC) for the chromatographic region which gave the best discrimination between groups. Also shown for the PLS-DA are the test predictions across all 200 randomly sub-sampled models

	Response for region 179.33/s	Response for PLS-DA test predictions
Sensitivity	0.88	0.88
Specificity	0.58	0.81
Precision	0.88	0.73
Accuracy	0.76	0.84
AUROC	0.79	0.91
AUROC convex hull	0.82	0.92

rather than any demographic differences between the groups (Fig. 5). One limitation of our study demonstrated here is the possibility of a type I error occurring given our small sample size. Whilst we have tried to minimise this by combining ANOVA with ROC,[22] it is well recognised that such an error is still possible. One potential confounder which we have not been able to fully control for is current smoking status, as all our healthy controls were current smokers, and the majority of our COPD cases ex-smokers. We attempted to minimise this by asking patients not to smoke on the morning of the test, and by sampling early in the day.

Multivariate analysis

Finally, the application of PLS-DA allowed for discrimination between COPD and HS with $\alpha = 19\%$ and $\beta = 12.4\%$ (the type I and type II errors, respectively). The model predictions for the randomly sub-sampled training and test samples (200 iterations; randomly split 4 : 1, training–test) are summarised by a box-plot in Fig. 6. Most of the outputs of the test samples were within the 95% confidence interval of those of the training set which suggest that the risk of overfitting was low and the predictive model has good generalisation performance. The performance parameters for this model were excellent, demonstrating an area under the ROC curve of 0.91 and sensitivity 0.88 for discriminating COPD *versus* HS (Table 3).

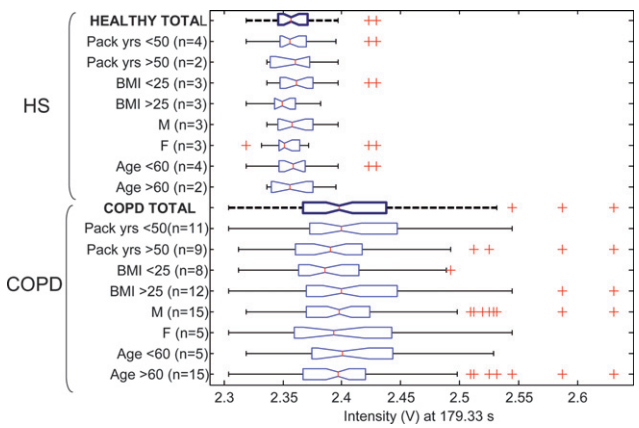

Fig. 5 Box-plot showing signal intensity for chromatographic region 179.33 s, for healthy smokers (HS) (top) and COPD (bottom), with further sub-groupings based on metadata. *n* represents the number of participants in each category. Boxes represent the lower, median and upper quartiles. Whiskers represent the range or 1.5× the length of the box, whichever is shorter. Outliers (+) are the values out of the whisker range.

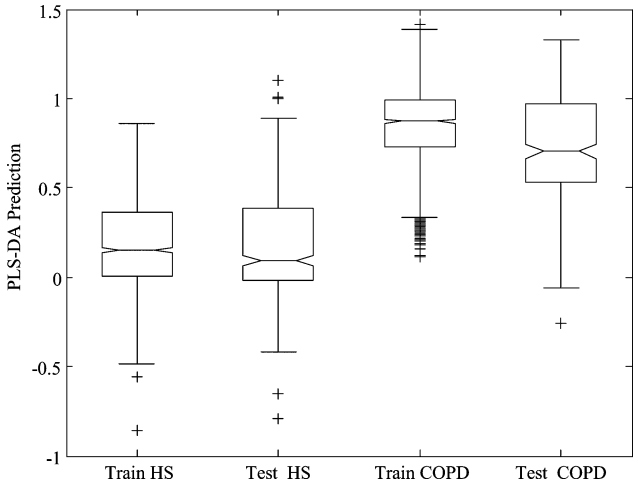

Fig. 6 PLS-DA model predictions for randomly selected test samples based upon 200 sub-sampling iterations, with COPD participants encoded as "1" and healthy smokers (HS) "0". Definition of box-plot parameters as for Fig. 5.

As described above, PLS-DA is a supervised modeling algorithm that can be used to generate predictive models from multivariate data. The benefit of using such a methodology over a simple univariate analysis is that combinations of spectral features that define the class separation are taken into account, and therefore such tools can act as more powerful classifiers for spectral data. In addition, where cohorts have a small number of participants, random re-sampling of training and test data avoids model bias due to sampling effects, allowing for the calculation of confidence limits for the model predictions. Further, mapping of the distributions of model predictions enables visualisation of the model skewness as well as calculation of other summary statistics such as: accuracy, precision, sensitivity and specificity as detailed in Table 3.

Conclusion

In summary, we have demonstrated a reproducible method for collecting exhaled breath and measuring the profile of VOCs in healthy smokers and participants with COPD, including those with severe impairment of respiratory function. Differentiation between COPD and asymptomatic smokers has been achieved by applying a selective and controlled sampling methodology for the collection of breath and subsequent analysis by GC-DMS with univariate and multivariate chemometric analyses (with appropriate re-sampling to assess overfitting). Further investigations will require the use of mass spectrometry to confirm and quantify the identity of any discriminative biomarkers. With the identification of such biomarkers, we may gain novel insights into mechanistic pathways of disease, and further they may merit investigation in the clinical setting as potential aids in the diagnosis and monitoring of COPD.

Acknowledgements

The authors gratefully acknowledge Theo Koimtzis for his work in developing the breath sampler, Rachel Dockry for her

invaluable help recruiting volunteers, the support provided by the Sionex Corporation through the loan of a differential mobility spectrometer, and GlaxoSmithKline for funding. RG thanks BBSRC for financial support, RMJ is grateful to HOSDB for support and YX thanks EU Framework 6 programme Biotracer for funding.

References

1 M. Corradi, I. Rubinstein, R. Andreoli, P. Manini, A. Caglieri, D. Poli, R. Alinovi and A. Mutti, Aldehydes in exhaled breath condensate of patients with chronic obstructive pulmonary disease, *Am. J. Respir. Crit. Care Med.*, 2003, **167**, 1380–1386.
2 M. Phillips, R. N. Cataneo, J. Greenberg, R. Gunawardena, A. Naidu and F. Rahbari-Oskoui, Effect of age on the breath methylated alkane contour, a display of apparent new markers of oxidative stress, *J. Lab. Clin. Med.*, 2000, **136**, 243–249.
3 R. Goodacre, Metabolomics of a superorganism, *J. Nutr.*, 2007, **137**, 259S–266S.
4 D. Smith and P. Spanel, The challenge of breath analysis for clinical diagnosis and therapeutic monitoring, *Analyst*, 2007, **132**, 390–396.
5 A. Wehinger, A. Schmid, S. Mechtcheriakov, M. Ledochowski, C. Grabmer, G. A. Gastl and A. Amann, Lung cancer detection by proton transfer reaction mass-spectrometric analysis of human breath gas, *Int. J. Mass Spectrom.*, 2007, **256**, 49–59.
6 S. Dragonieri, R. Schot, B. J. Mertens, S. Le Cessie, S. A. Gauw, A. Spanevello, O. Resta, N. P. Willard, T. J. Vink, K. F. Rabe, E. H. Bel and P. J. Sterk, An electronic nose in the discrimination of patients with asthma and controls, *J. Allergy Clin. Immunol.*, 2007, **120**, 856–862.
7 P. J. Mazzone, J. Hammel, R. Dweik, J. Na, C. Czich, D. Laskowski and T. Mekhail, Diagnosis of lung cancer by the analysis of exhaled breath with a colorimetric sensor array, *Thorax*, 2007, **62**, 565–568.
8 D. S. Levin, R. A. Miller, E. G. Nazarov and P. Vouros, Rapid separation and quantitative analysis of peptides using a new nanoelectrospray-differential mobility spectrometer-mass spectrometer system, *Anal. Chem.*, 2006, **78**, 5443–5452.
9 H. Borsdof and G. A. Eiceman, Ion mobility spectrometry: principles and applications, *Appl. Spectrosc. Rev.*, 2006, **41**, 323–375.
10 R. A. Miller, E. G. Nazarov, G. A. Eiceman and A. T. King, A MEMS radio-frequency ion mobility spectrometer for chemical vapour detection, *Sens. Actuators, A*, 2001, **91**, 301–312.
11 M. Basanta, T. Koimtzis, D. Singh, I. Wilson and C. L. P. Thomas, An adaptive breath sampler for use with human subjects with an impaired respiratory function, *Analyst*, 2007, **132**, 153–163.
12 American Thoracic Society, Standardization of spirometry, 1994 update, *Am. J. Respir. Crit. Care Med.*, 1995, **152**, 1107–1136.
13 M. Basanta, D. Singh, S. Fowler, I. Wilson, R. Dennis and C. L. P. Thomas, Increasing analytical space in gas chromatography-differential mobility spectrometry with dispersion field amplitude programming, *J. Chromatogr., A*, 2007, **1173**, 129–138.
14 P. H. C. Eilers, Parametric time warping, *Anal. Chem.*, 2004, **76**, 404–411.
15 N. P. V. Nielsen, J. M. Carstensen and J. Smedsgaard, Aligning of single and multiple wavelength chromatographic profiles for chemometric data analysis using correlation optimised warping, *J. Chromatogr., A*, 1998, **805**, 17–35.
16 T. Skov, F. van den Berg, G. Tomasi and R. Bro, Automated alignment of chromatographic data, *J. Chemom.*, 2006, **20**, 484–497.
17 Y. Xu, F. Gong, S. J. Dixon, R. G. Brereton, H. A. Soini, M. V. Novotny, E. Oberzaucher, K. Grammer and D. J. Penn, Application of dissimilarity indices, principal co-ordinates analysis and rank tests to peak tables in metabolomics of the gas chromatography mass spectrometry of human sweat, *Anal. Chem.*, 2007, **79**, 5633–5641.
18 C. E. Metz, Basic principles of ROC analysis, *Semin. Nucl. Med.*, 1978, **8**, 283–298.
19 H. Martens and T. Næs, *Multivariate Calibration*, John Wiley, Chichester, 1989.
20 R. K. H. Galvao, M. C. U. Araujo, M. D. N. Martins, G. E. Jose, M. J. C. Pontes, E. C. Silva and T. C. B. Saldanha, An application of subagging for the improvement of prediction accuracy of multivariate calibration models, *Chemom. Intell. Lab. Syst.*, 2006, **81**, 60–67.
21 D. Poli, P. Carbognani, M. Corradi, M. Goldoni, O. Acampa, B. Balbi, L. Bianchi, M. Rusca and A. Mutti, Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross-sectional and nested short-term follow-up study, *Respir. Res.*, 2005, **6**, 71–80.
22 D. I. Broadhurst and D. B. Kell, Statistical strategies for avoiding false discoveries in metabolomics and related experiments, *Metabolomics*, 2006, **2**, 171–196.

