PAPER www.rsc.org/analyst | Analyst

Evaluation of a gas sensor array and pattern recognition for the identification of bladder cancer from urine headspace†

Christina M. Weber,[a] Michael Cauchi,[a] Mitesh Patel,[a] Conrad Bessant,[a] Claire Turner,[b] Lezlie E. Britton[c] and Carolyn M. Willis*[c]

Received 5th June 2010, Accepted 30th September 2010
DOI: 10.1039/c0an00382d

Published on 22 October 2010 on http://pubs.rsc.org | doi:10.1039/C0AN00382D
Downloaded by Cranfield University on 08 February 2011

Previous studies have indicated that volatile compounds specific to bladder cancer may exist in urine headspace, raising the possibility that headspace analysis could be used for diagnosis of this particular cancer. In this paper, we evaluate the use of a commercially available gas sensor array coupled with a specifically designed pattern recognition algorithm for this purpose. The best diagnostic performance that we were able to obtain with independent test data provided by healthy volunteers and bladder cancer patients was 70% overall accuracy (70% sensitivity and 70% specificity). When the data of patients suffering from other non-cancerous urological diseases were added to those of the healthy controls, the classification accuracy fell to 65% with 60% sensitivity and 67% specificity. While this is not sufficient for a diagnostic test, it is significantly better than random chance, leading us to conclude that there is useful information in the urine headspace but that a more informative analytical technique, such as mass spectrometry, is required if this is to be exploited fully.

Introduction

Transitional cell carcinoma (TCC) of the bladder is the most common form of bladder cancer and simultaneously the second most common malignant tumour of the genito-urinary tract.[1] As with many cancers, untreated TCC can be fatal, so early detection is vitally important. Cystoscopy with biopsy is the "gold standard" for bladder cancer detection, but it is expensive, inconvenient and invasive. Urine cytology is a more acceptable alternative, and the most widely applied of the non-invasive procedures available. Although urine cytology has a high specificity for bladder cancer (90–98%), its sensitivity is low (20–50%), especially for low-grade tumours, since those shed proportionally few cells into the urine. Furthermore, the results take several days, requiring an expert to interpret the test.[1]

Utilisation of molecular biomarkers for bladder cancer is a tantalising prospect for making diagnosis more sensitive, rapid and convenient, provided such biomarkers can be discovered in urine and readily identified in a clinical setting. Two FDA approved protein markers, nuclear matrix protein 22 (NMP22) and bladder tumour antigen (BTAstat), have proven effective in this context.[2,3] They are more sensitive than urine cytology, having achieved reported sensitivities of 50–85% and 50–70%, respectively. However, their specificities, in the region of 60–70%, are inferior to that of urine cytology.

Recently, it has been suggested that volatile organic compounds (VOCs) present in the headspace of urine samples may be used as cancer biomarkers.[4,5] In particular, a previous study[6] showed that dogs could be trained to identify bladder cancer sufferers from the odour of their urine. The results of this proof of principle study suggested that one or more VOCs present within the urine headspace are indicative of bladder cancer and have the potential to be used as diagnostic biomarkers. As it is not realistic to use dogs in a clinical setting, we substituted the dogs with a more practical and objective gas sensor array, also called an electronic nose (eNose). Developed to mimic olfaction, these instruments have shown promise in lung cancer detection.[7]

Methods

Participant selection

A total of 30 patients, aged 50–88, presenting at Buckinghamshire Hospitals NHS Trust with new or recurrent transitional cell carcinoma (TCC) of the bladder supplied urine prior to surgical intervention. Grade and stage of the tumour were recorded. Fifty nine control subjects, categorised into one of three groups (Controls 1, 2 and 3, depending upon age and disease status), also provided urine samples. Twenty healthy individuals aged 18–31, with no urine abnormality on dipstick made up Control group 1 (C1). Control group 2 (C2) consisted of 20 subjects between 18 and 32, with any non-cancerous condition or disease, and/or one or more positive dipstick finding. Menstruating women with blood in their urine were included in this group, for example, as were individuals with suspected urinary tract infection, positive for leucocytes, blood and/or protein. Categorized into the Control 3 group (C3) were 19 patients, aged between 24 and 89, with confirmed non-cancerous urological disease, with or without urine dipstick abnormalities. Urological conditions included renal and ureteric stones, renal cysts and polypoid cystitis. A summary of the age and gender of the subjects in each cohort is included as ESI†.

[a]*Cranfield University, College Road, Cranfield, Bedfordshire, MK43 0AL, UK. E-mail: c.bessant@cranfield.ac.uk*
[b]*Department of Chemistry and Analytical Sciences, Open University, Milton Keynes, MK7 6AA, UK. E-mail: c.turner@open.ac.uk*
[c]*Department of Dermatology, Amersham Hospital, Amersham, Buckinghamshire, HP70JD, UK. E-mail: carolyn.willis@buckshosp.nhs.uk*
† Electronic supplementary information (ESI) available: Details on urine samples (provided in spreadsheet format). See DOI: 10.1039/c0an00382d

View Online

Published on 22 October 2010 on http://pubs.rsc.org | doi:10.1039/C0AN00382D
Downloaded by Cranfield University on 08 February 2011

As criteria for inclusion/exclusion, controls over 32 years of age were required to have had recent cystoscopy to exclude visible bladder malignancy. For both controls and the cancer positive group (TCC), men over 50 years were only included if recent cancer-negative prostate histology had been demonstrated. Individuals with pre-malignant urological disease or a history of urological carcinoma other than TCC were excluded. A history of malignancy in other organ systems (>5 years previously) was acceptable, providing the individual was now considered disease-free. All other past and/or present medical conditions were permissible. There were no exclusions on the basis of medication, menstrual cycle, diet, alcohol consumption, or chemical exposure. However, details of all of these factors were recorded for each participant, should their influence on the composition and odour of the urine need to be considered at any stage. Special attention was paid to smoking habits, with one third of those with bladder cancer being current cigarette smokers, as compared to 17 out of 59 control subjects. A spreadsheet containing all this information is provided as ESI†.

The study was approved by the Mid and South Buckinghamshire Local Research Ethics Committee, and all participants gave written informed consent.

Analysis and processing of urine samples

Following urinalysis (Multistix 10 SG, Bayer Corporation, NY, USA), fresh urine specimens were refrigerated immediately, and frozen 1–24 h later as 0.5 ml aliquots in glass vials. They were then stored at −80 °C until required.

Headspace analysis

An electronic nose was used to characterise the VOC (volatile organic compound) content of urine. Measurements were performed using an NST 3320 Lab Emission Analyser.[8] The instrument comprises 12 metal oxide semi-conductor (MOS) sensors and an array of 10 individual metal oxide semi-conductor field-effect transistor (MOSFET) sensors, together with a capacitance-based humidity sensor and an infrared-based CO_2 sensor.

The 89 urine samples were randomised, removed from storage at −80 °C and allowed to thaw at room temperature (approximately 21 °C). The thawed samples were then aliquoted into two headspace vials, each containing 2.5 ml of urine, providing duplicate samples. The samples were then incubated for 1 hour at 38 °C before they were analysed in batches of 12 samples using the eNose. Signals from each of the 24 sensors are generated by their response to the different chemical characteristics of the urine headspace. The analysis was performed in cycles, exposing the sensors to fresh reference air between the measurements. A cycle was composed of *Baseline*, *Sample* and *Recovery*. During runtime, all sensor signals were collected in periodic time intervals of one second. From these resulting raw data, eight characteristic signal parameters were estimated by the eNose software as *baseline*, *response*, *absolute response*, *on derivative*, *on integral*, *off derivative*, *off integral* and *difference*. This can be seen as data reduction, very likely resulting in a speed up of subsequent pattern recognition analysis. Fig. 1 shows the sensor parameters graphically. All samples were measured twice (89 samples and 89 duplicates) and the resulting extracted signal parameters—*features*—were finally exported in Microsoft Excel spreadsheet format.

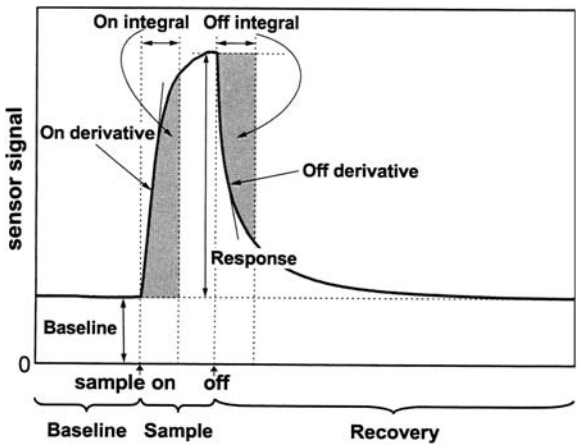

Fig. 1 Estimated sensor variables of an electronic nose. The analysis-cycle of an electronic nose is composed of Baseline, Sample phase and Recovery phase. During Sample phase and Recovery phase signal parameters are calculated as baseline, response, absolute response, on derivative, on integral, off derivative, off integral and difference. Those parameters describe the reaction of a sensor to a certain compound. Adapted from the manual of the eNose instrument.[6]

Data analysis

The provided spreadsheet files were analysed using Matlab R2009a (The Mathworks Inc.). Additionally the PLS Toolbox 3.5 (Eigenvector Research Inc.) was employed to classify the data *via* partial least squares-discriminant analysis (PLS-DA). Before the loaded data were analysed, the effect of data scaling was investigated, with the methods considered being auto-scaling (eqn (1)), mean-centring (eqn (2)), range-scaling between 0 and 1 (eqn (3)) and range-scaling between −1 and +1 (eqn (4)).[9]

$$y_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (1)$$

$$y_{ij} = x_{ij} - \bar{x}_j \quad (2)$$

$$y_{ij} = \frac{\left(x_{ij} - \min\left(x_j\right)\right)}{\left(\max\left(x_j\right) - \min\left(x_j\right)\right)} \quad (3)$$

$$y_{ij} = \frac{2 \times \left(x_{ij} - \min\left(x_j\right)\right)}{\left(\max\left(x_j\right) - \min\left(x_j\right)\right) - 1} \quad (4)$$

Exploratory Data Analysis was accomplished *via* principal component analysis (PCA), which is the most widely used multivariate statistical technique.[10–12] This part of the analysis was performed to reveal the characteristics that cause the greatest variance in the dataset.

Before the pre-processed data were sent to the classifier, a crucial step—*feature selection*—was performed to provide the classifier with the most significant variables, so as to achieve

 This journal is © The Royal Society of Chemistry 2011

View Online

Published on 22 October 2010 on http://pubs.rsc.org | doi:10.1039/C0AN00382D
Downloaded by Cranfield University on 08 February 2011

higher classification accuracy. This procedure resembles a filter method that automatically picks out variables within the data that are most distinctive of each class, or from another point of view deselects those variables that are not significant at all. For this purpose, univariate statistics—a *t*-test—was applied. The *t*-test is more or less a decision rule to determine if two samples belong to the same population, according to a specified significance level α.[13] Dependent on the significance level α, two variables, in this case, non-cancerous and cancerous samples, were either considered as similar or different. For data analysis, the significance level α was initially set to 0.05 and varied from 0.1 to 0.9 in steps of 0.1. Features whose probability value did not rise above the threshold α were taken as being different and were therefore regarded as relevant for disease-related discrimination. From a statistical point of view high α-values are not suitable to select significant features, but can be used to deselect insignificant variables as noise for example.

Next, partial least squares discriminant analysis (PLS-DA)[14] was used to build a classification model using the cancer status of the samples. PLS-DA is a supervised method. This calls for information about the parameter of interest (the cancer status) to be known in order to train the algorithm to identify which sensors capture the molecules that differentiate between the classes. PLS-DA is considered to be a dimensionality reduction method and can be seen as the regression extension of principal component analysis.[15] Unlike PCA, which attempts to describe the maximum variation in the measured data, PLS-DA tends to maximise the covariance between the input data and the output class. The information returned by PCA is that which was caused by the attribute with the biggest variance. In contrast, PLS-DA returns only data that were caused by the property under investigation.

It is known that PLS-DA is prone to overestimate the accuracy of classification if it is not accurately validated.[16] For this reason, a very thorough evaluation process, leave-one-out cross-validation (LOOCV),[17] was implemented to assess the performance of the PLS-DA-classifier. This method is based on the concept of classifying one sample against a model, built using the rest of the samples. This procedure is repeated with each sample until all samples have been classified. Information about the true positive rate and the true negative rate was established.

As final validation of the results, and to attain an indication of the significance of the results, a Monte Carlo Simulation was used to evaluate the obtained results.[18] This involves repeated random sampling. In this context a null model is generated from a set of data that is statistically similar to the data under study, but for which we do not expect to be able to build a meaningful classification model. A total of 250 datasets of this kind were generated, by simply assigning a random class to each sample (while maintaining the relative number of positive and negative samples). For a disease discriminating model trained on the real sample classes to be considered significant it needs to achieve a classification accuracy towards the extremities of those produced by the null models.

Results and discussion

The 24 sensors of the electronic nose instrument initially recorded 192 sensor variables. However, during a later stage of the eNose analysis, significant random noise was identified in five of the MOSFET sensors. As a consequence, the information provided by these sensors was omitted from the data analysis. Hence, 152 sensor variables (8 *sensor variables* × 19 *remaining sensors*) remained. To assess the role of the classification of each sensor technology, we used the full array for classification as well as only the MOS and only the MOSFET sensors to calculate different classification models.

As already mentioned in the Methods section, the 59 specimens belonging to the cancer-negative control group were subdivided into further subgroups (C1–C3). For this reason, we sectioned our work into four main experiments. Firstly, each of the three control subgroups was classified separately against the cancer group and, finally, all control subgroups were combined and classified together against the TCCs.

In each of the four experiments, we used the full sensor array as well as the MOS sensors only and the MOSFET sensors only. Furthermore we applied the same pre-processing methods to the data, including scaling and feature selection. Four different scaling methods and no scaling were evaluated, with range-scaling between 0 and 1 being found to be the best; therefore, only these results are shown in this paper. Given that we were looking for a pattern within the data and were therefore interested in relative quantities, not absolute quantities, it is not surprising that range-scaling was most appropriate.

Explorative data analysis, by way of PCA, was also applied in each experiment to reveal natural groupings based on the sensor responses of the electronic nose. The PCA scores plot derived from the experimental data is shown in Fig. 2. However, the visual output of this method did not disclose any relation to the cancer status of the samples, even though the principal components shown capture over 98% of the variance in the data. Other influences such as age, diet or gender may be responsible for the groupings obtained. However, this does not mean that the data do not contain any information concerning bladder cancer. The PCA only demonstrated that the cancer status is not responsible for the bigger part of the variance, captured by the first two or three principal components (PCs). Nevertheless, investigating

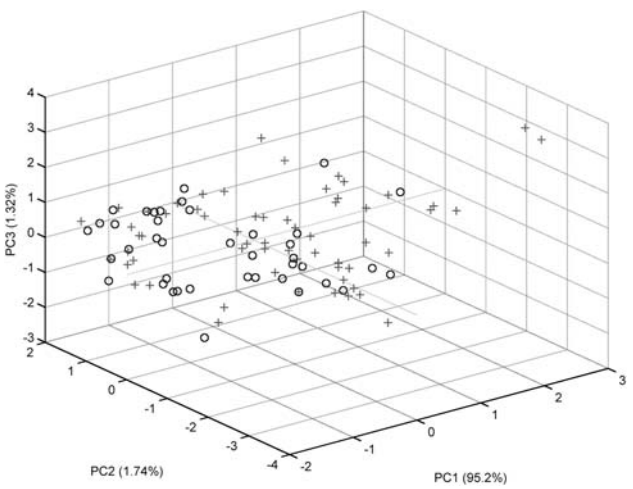

Fig. 2 The PCA scores plot does not appear to show any discrimination between cancerous (+) and control samples of group C1 (○). The data were pre-treated using range-scaling between 0 and 1.

View Online

Published on 22 October 2010 on http://pubs.rsc.org | doi:10.1039/C0AN00382D
Downloaded by Cranfield University on 08 February 2011

Table 1 Best results of leave-one-out cross-validation runs with PLS-DA for each of the four experiments. The presented classification results were obtained after scaling the data between zero and one and performing feature selection

(a) C1 *vs.* TCC

Full sensor array	
Significance level α	0.8
No. of features	140
Total accuracy	70.0%
Cancer group TCC accuracy (sensitivity)	70.0%
Control group C1 accuracy (specificity)	70.0%
MOSFET sensors	
Significance level α	0.8
No. of features	35
Total accuracy	60.0%
Cancer group TCC accuracy (sensitivity)	61.7%
Control group C1 accuracy (specificity)	57.5%
MOS sensors	
Significance level α	0.8
No. of features	91
Total accuracy	69.0%
Cancer group TCC accuracy (sensitivity)	68.3%
Control group C1 accuracy (specificity)	70.0%

(b) C2 *vs.* TCC

Full sensor array	
Significance level α	0.1
No. of features	69
Total accuracy	67.0%
Cancer group TCC accuracy (sensitivity)	71.7%
Control group C2 accuracy (specificity)	60.0%
MOSFET sensors	
Significance level α	0.1
No. of features	3
Total accuracy	65.0%
Cancer group TCC accuracy (sensitivity)	65.0%
Control group C2 accuracy (specificity)	65.0%
MOS sensors	
Significance level α	0.1
No. of features	58
Total accuracy	59.0%
Cancer group TCC accuracy (sensitivity)	56.7%
Control group C2 accuracy (specificity)	62.5%

(c) C3 *vs.* TCC

Full sensor array	
Significance level α	0.8
No. of features	132
Total accuracy	62.2%
Cancer group TCC accuracy (sensitivity)	68.3%
Control group C3 accuracy (specificity)	52.6%
MOSFET sensors	
Significance level α	0.8
No. of features	29
Total accuracy	57.1%
Cancer group TCC accuracy (sensitivity)	61.7%
Control group C3 accuracy (specificity)	50.0%
MOS sensors	
Significance level α	0.8
No. of features	87
Total accuracy	60.2%
Cancer group TCC accuracy (sensitivity)	66.7%
Control group C3 accuracy (specificity)	50.0%

(d) C1, C2, C3 *vs.* TCC

Full sensor array	
Significance level α	0.4
No. of features	110
Total accuracy	64.6%
Sensitivity	60.0%
Specificity	66.9%
Control group C1 accuracy	75.0%
Control group C2 accuracy	72.5%
Control group C3 accuracy	52.6%
Cancer group TCC accuracy	60.0%
MOSFET sensors	
Significance level α	0.4
No. of features	13
Total accuracy	61.8%
Sensitivity	63.3%
Specificity	61.0%
Control group C1 accuracy	62.5%
Control group C2 accuracy	55.0%
Control group C3 accuracy	65.8%
Cancer group TCC accuracy	63.3%
MOS sensors	
Significance level α	0.4
No. of features	83
Total accuracy	60.7%
Sensitivity	43.3%
Specificity	69.5%
Control group C1 accuracy	70.0%
Control group C2 accuracy	67.5%
Control group C3 accuracy	71.1%
Cancer group TCC accuracy	43.3%

principal components of lower variance did not lead to an explicitly disease-related differentiation either. PLS-DA aims to maximise the co-variation between the measured data and the classification (non-cancerous *vs.* cancerous). This leads to the capability to discriminate between samples that could not be separated by PCA.

We started by training the PLS-DA-classifier with the two most disparate groups: Control 1 (C1), representing healthy males or females, and the TCC group incorporating people suffering from bladder cancer. Since group C1 possesses the most differences compared to the cancer group, we expected the classification outcome of this sample set to be the best, which proved to be correct. The best result obtained an overall total accuracy of 70% by range-scaling the data and configuring the significance level $\alpha = 0.8$ for feature selection. A summary of the best classification results of C1 *vs.* TCC is given in Table 1a).

Next, we trained the classifier with Control 2 (C2) and the cancer group (TCC) data. Urine samples within this control subgroup showed similar abnormalities on dipstick analysis to some cancer samples, such as blood, for example, and were therefore more difficult to distinguish from cancerous samples than Control 1 (C1) samples. Both specificity and the total accuracy, achieved with the full sensor array, reflect this. Only the sensitivity is marginally higher than that obtained for C1 *vs.* TCC in full sensor array mode. An overview of the results obtained is given in Table 1b).

In the third experiment, the classifier had to distinguish between samples with confirmed non-cancerous urological diseases (Control 3) and cancerous samples (TCC). This was

 This journal is © The Royal Society of Chemistry 2011

View Online

Published on 22 October 2010 on http://pubs.rsc.org | doi:10.1039/C0AN00382D
Downloaded by Cranfield University on 08 February 2011

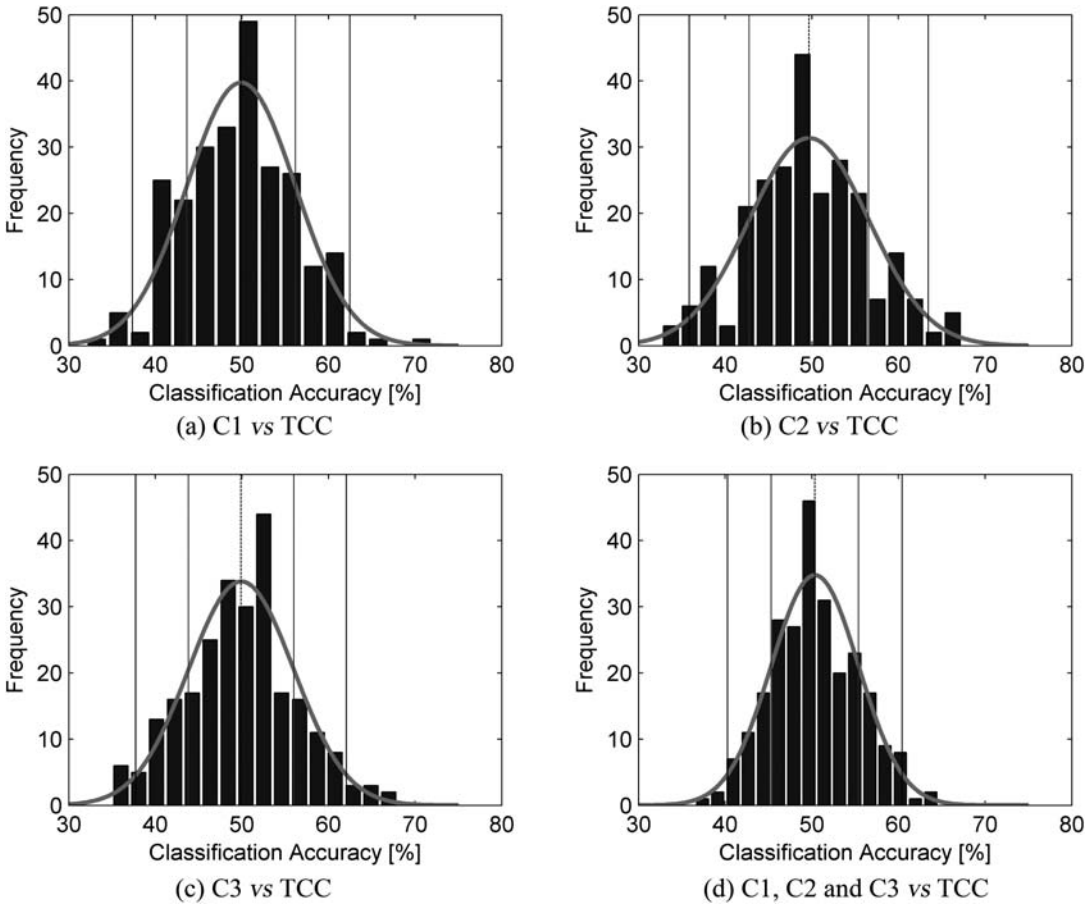

Fig. 3 Distribution of the overall percentage classified after randomised assignation of classes to the samples corresponding to each of the four experiments. Number of runs: 250. An achieved total accuracy that is beyond two standard deviations of the mean indicates that it is a significant result at the 95% confidence level. (a) C1 *vs.* TCC: Control 1 (C1), representing healthy males or females, and the TCC group incorporating people suffering from bladder cancer: total accuracy: 70%. (b) C2 *vs.* TCC: Control 2 (C2), representing urine samples showing similar abnormalities on dipstick analysis to some cancer samples, and the TCC: total accuracy: 67%. (c) C3 *vs.* TCC: Control 3 (C3), representing samples with confirmed non-cancerous urological diseases, and the TCC group: total accuracy: 67%. (d) C1, C2, C3 *vs.* TCC (using all available data): total accuracy: 64.61%.

expected to be the most difficult combination, as disease markers not specific to bladder cancer are likely to be present.

As far as the full sensor array is concerned sensitivity and specificity as well as total accuracy did not reach the results obtained in the two former experiments. A detailed overview of these results is presented in Table 1c).

Finally, in the fourth experiment, all three control subgroups (C1, C2, and C3) were merged to form one control group that was used as a counterpart to the TCC samples when training the binary classifier. The results of this experiment can be seen in Table 1d). As expected, the total accuracy, obtained with the entire sensor array, was not as high as in experiments 1 and 2, where only samples from young people with no urological disease (C1, C2) were included. The major contributor to this more moderate classification outcome was the inclusion of the samples from Control group 3. All subjects within this subgroup had confirmed non-cancerous urological disease, the pathological effects of which are likely to be similar to the secondary effects of bladder cancer. Within both these groups, varying amounts of metabolic products associated with inflammation, infection and/or necrosis will almost certainly be present.

Because of this, Control 3 samples from the most important control subset contain the most relevant information. Training the classifier with this kind of data is therefore fundamental in order to be able to subtract general disease compounds present in the urine from those specific for bladder cancer. Accurate diagnosis of the control subjects is, of course, paramount to this process, since the inclusion of false negative individuals would lead to incorrect classification rules.

Interestingly, within the TCC sample group, the majority of those incorrectly classified as negative were from patients with more advanced tumours. In these cases, it is possible that metabolic products generated secondarily to the tumour may overwhelm or mask the volatile cancer biomarkers within the urine, giving rise to a urine headspace more closely resembling that of Control 3 samples. Canine olfactory studies support this hypothesis; high grade TCCs with a significant level of invasion are missed more frequently by trained dogs than low-grade superficial tumours.[19]

To assess the significance of the presented results we carried out Monte Carlo Simulations. Fig. 3 shows the results attained for each of the four experiments each with 250 random runs. A model with an overall accuracy beyond two standard deviations

View Online

of the mean indicates that it is a significant result at the 95% confidence level. Thus, the first experiment that included totally healthy volunteers and cancer patients and produced a total accuracy of 70% is very significant (Fig. 3(a)). The result of the second experiment, comprising samples showing abnormalities at the dipstick analysis, can also be regarded as significant, with an overall accuracy of 67% (Fig. 3(b)). The third experiment included samples with confirmed non-cancerous urological diseases and was therefore the most difficult to build classifiers for. Not only did this experiment derive the lowest total accuracy (62%), the accuracy was found to be just below 62.5% that marks the 95% confidence limit in the null model results (Fig. 3(c)). The fourth experiment made use of all available data and attained a statistically significant classification result of 65%, which comfortably exceeds the 95% confidence limit (Fig. 3(d)).

Conclusions

By combining a commercially available gas sensor array with chemometrics techniques, we have made progress towards a new instrumental method of bladder cancer detection based on volatile biomarkers. As many as 70% of cancer patients and 70% of non-cancerous subjects were correctly classified when the classifier was trained with a combination of TCC positive urine samples and samples from healthy participants with no urine abnormality. PLS-DA-derived models gave an accuracy for patients presenting with other non-cancerous urological disease of 65%, with 60% sensitivity and 67% specificity. There is little difference between the performance of the two sensor types used in the array, but the results of the full array are consistently better than the individual sets of sensors.

Although the specificity achieved using this method was less than that of conventional urine cytology, sensitivity was appreciably higher.[1] This leads us to conclude that there is an anomalous composition to the VOC content of urine of patients suffering from bladder cancer, confirming previous canine olfactory data.[6] However, the gas sensor array may not be the ideal analytical technique to detect this abnormal composition of urine headspace. We therefore recommend a more revealing analytical technique such as mass spectrometry. More sophisticated pattern recognition techniques as support vector machines (SVMs)[20] or artificial neural networks (ANNs)[21] may also further advance the results obtained so far.

Published on 22 October 2010 on http://pubs.rsc.org | doi:10.1039/C0AN00382D
Downloaded by Cranfield University on 08 February 2011

Acknowledgements

The authors would like to thank the Amerderm Research Trust for generous financial support. We would like to thank Neus Planas Pont and Naresh Magan for access to, and assistance with, the electronic nose device used in this work.

Notes and references

1 P. Bassi, V. D. Marco, A. D. Lisa, M. Mancini, F. Pinto, R. Bertoloni and F. Longo, *Urol. Int.*, 2005, **75**, 193–200.
2 V. Poulakis, U. Witzsch, R. D. Vries, H. M. Altmannsberger, M. J. Manyak and E. Becht, *BJU Int.*, 2001, **88**, 692–701.
3 A. S. Glas, D. Roos, M. Deutekom, A. H. Zwinderman, P. M. Bossuyt and K. H. Kurth, *J. Urol.*, 2003, **169**, 1975–1982.
4 P. Španěl, D. Smith, T. A. Holland, W. Al Singary and J. B. Elder, *Rapid Commun. Mass Spectrom.*, 1999, **13**, 1354–1359.
5 S. Smith, P. White, J. Redding, N. M. Ratcliffe and Chris S. J. Probert, *IEEE Sens. J.*, 2010, **10**, 92–96.
6 C. M. Willis, S. M. Church, C. M. Guest, W. A. Cook, N. McCarthy, A. J. Bransbury, M. R. Church and J. C. Church, *BMJ [Br. Med. J.]*, 2004, **329**, 712.
7 A. D'Amico, G. Pennazza, M. Santonico, E. Martinelli, C. Roscioni, G. Galluccio, R. Paolesse and C. Di Natale, *Lung Canc.*, 2010, **68**, 170–176.
8 Nordic Sensor Technologies AB Sweden, *NST 3220 Lab Emission AnalyserLab Emission Analyser (Manual)*, Nordic Sensor Technologies AB Sweden, Linköping, Sweden, 2009.
9 M. Otto, *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, Wiley-VCH, John Wiley distributor, Weinheim, Chichester, 2007, p. 328.
10 R. G. Brereton, *Applied Chemometrics for Scientists*, Wiley, John Wiley distributor, Hoboken, NJ, Chichester, 2007, p. 379.
11 J. W. Gardner and P. N. Bartlett, *Electronic Noses: Principles and Applications*, Oxford University Press, Oxford, 1999, p. 245.
12 R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, Wiley, New York, Chichester, 2001, p. 654.
13 S. A. Glantz, *Primer of Biostatistics*, McGraw-Hill, London, 2005, p. 520.
14 M. Barker and W. Rayens, *J. Chemom.*, 2003, **17**, 166–173.
15 K. Yuan, H. Kong, Y. Guan, J. Yang and G. Xu, *J. Chromatogr., B: Anal. Technol. Biomed. Life Sci.*, 2007, **850**, 236–240.
16 J. A. Westerhuis, H. C. J. Hoefsloot, S. Smit, D. J. Vis, A. K. Smilde, E. J. J. vanVelzen, P. M. van Duijnhoven John and F. A. van Dorsten, *Metabolomics*, 2008, **4**, 81–89.
17 T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York, 2001.
18 R. G. Brereton, *Chemometrics for Pattern Recognition*, Wiley-Blackwell, 2009.
19 C. M. Willis, R. Harris, L. E. Britton and C. M. Guest, *Canine olfactory detection of bladder cancer: sensitivity and specificity*, 2010, in preparation.
20 V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, London, 1995.
21 A. Fielding, *Cluster and Classification Techniques for the Biosciences*, Cambridge University Press, Cambridge, 2007, p. 246.

 This journal is © The Royal Society of Chemistry 2011

https://dspace.org/

Evaluation of a gas sensor array and pattern recognition for the identification of bladder cancer from urine headspace.

Weber, Christina M.; Cauchi, Michael; Patel, Mitesh; Bessant, Conrad M.; Turner, Claire ...

2011-01-21T00:00:00Z
Royal Society of Chemistry
http://dx.doi.org/10.1039/c0an00382d
Christina M. Weber, Michael Cauchi, Mitesh Patel, Conrad Bessant, Claire Turner, Lezlie E. Britton and Carolyn M. Willis, Evaluation of a gas sensor array and pattern recognition for the identification of bladder cancer from urine headspace, Analyst, Volume 136, Issue 2, 2011, Pages 359-364.

Downloaded from a DSpace repository. https://dspace.org/