

Published in final edited form as:

Meas Sci Technol. 2014 May 12; 25(6): . doi:10.1088/0957-0233/25/6/065108.

## Towards point of care testing for *C. difficile* infection by volatile profiling, using the combination of a short multi-capillary gas chromatography column with metal oxide sensor detection

N D McGuire<sup>1</sup>, R J Ewen<sup>1</sup>, B de Lacy Costello<sup>1</sup>, C E Garner<sup>1</sup>, C S J Probert<sup>2</sup>, K. Vaughan, and N M Ratcliffe<sup>1,\*</sup>

<sup>1</sup>Institute of Biosensing Technology, University of the West of England, Bristol, United Kingdom

<sup>2</sup>Institute of Translational Medicine, University of Liverpool, Liverpool, United Kingdom

### Abstract

Rapid volatile profiling of stool sample headspace was achieved using a combination of short multi-capillary chromatography column (SMCC), highly sensitive heated metal oxide semiconductor (MOS) sensor and artificial neural network (ANN) software. For direct analysis of biological samples this prototype offers alternatives to conventional GC detectors and electronic nose technology. The performance was compared to an identical instrument incorporating a long single capillary column (LSCC). The ability of the prototypes to separate complex mixtures was assessed using gas standards and homogenised in house 'standard' stool samples, with both capable of detecting more than 24 peaks per sample. The elution time was considerably faster with the SMCC resulting in a run time of 10 minutes compared to 30 minutes for the LSCC.

The diagnostic potential of the prototypes was assessed using 50 *C. difficile* positive and 50 negative samples. The prototypes demonstrated similar capability of discriminating between positive and negative samples with sensitivity and specificity of 85% and 80% respectively. *C. difficile* is an important cause of hospital acquired diarrhoea, with significant morbidity and mortality around the world. A device capable of rapidly diagnosing the disease at the point of care would reduce cases, deaths and financial burden.

### Keywords

artificial neural network; *Clostridium difficile*; gas chromatography; gastro-intestinal; metal oxide semiconductor; multicapillary column; sensor; volatile; stool

## 1 Introduction

Gas chromatography (GC) analysis, in particular when coupled with mass spectrometry, is one of the most common methods for analysing complex mixtures of volatile organic compounds (VOCs). The use of VOCs for identifying medical conditions is a rapidly expanding area in part due to the appeal of non-invasive testing and pilot data exists for a

---

\* norman.ratcliffe@uwe.ac.uk.

range of diseases (Probert *et al* 2009), with much of the focus on exhaled breath samples (Boots *et al* 2012, Dummer *et al* 2011, Kim *et al*, 2012, Thorn & Greenman 2012).

Traditional GC detectors have drawbacks in terms of cost, size, auxiliary equipment and gas required for operation. Electronic noses may provide solutions to some of these issues. Electronic noses generally consist of an array of sensors with varying but overlapping sensitivities to compounds; allowing the detection of a large number of unknown compounds and compositions. However, despite their potential as rapid diagnostic devices they often suffer from poor reproducibility, lack of selectivity and susceptibility to their background environment which currently limits their use (Oh *et al* 2011).

Recently, hybrids of these two technologies have begun to emerge, for example the zNose™, a miniaturised GC system with a surface acoustic wave detector, was recently used to assess volatiles from *Tuberculosis*-causing *Mycobacteria* (McNerney *et al* 2012). Metal oxide semiconductor (MOS) sensors, in addition to being standalone sensors, lend themselves to incorporation in electronic noses. They offer longevity and high sensitivity, can be miniaturised and are relatively inexpensive, consequently MOS sensors have been employed in a variety of applications (Fine *et al* 2010), including zinc oxide and zinc stannate sensors for sensing liquid petroleum gas (LPG), important for air monitoring (Sivapunniyan *et al* 2011), and doped tin oxide sensors for measuring trimethylamine, an indicator of fish freshness (Jung JY *et al* 2011).

Often it is not possible to identify a single volatile marker associated with a specific disease and instead statistical modelling is required to identify diagnostic “patterns” of VOCs. A quartz crystal microbalance combined with partial least squares discriminant analysis (PLSDA) diagnosed lung cancer from breath samples with 100% correct classification of lung cancer and 94% of controls (DiNatale *et al* 2003). As an alternative to traditional statistical analysis, chemical sensors can be combined with pattern recognition techniques. Artificial neural networks (ANN) are a computer based simulation of a collection of neurons, and can be particularly useful to discover patterns in complex data which may be difficult to discern by other means. Recent examples combining metal oxide array e-noses and ANNs were the classification of wines with 82-97% correct (Aguilera *et al* 2012); and aging of beer, with results 90-100% correct (Ghasemi-Varnamkhasti *et al* 2011).

Disadvantages of traditional single capillary GC columns include column overloading, high operating pressures and slower elution. Multi-capillary columns (MCC) consist of approximately 1000 short capillary columns in parallel, thereby potentially overcoming the traditional disadvantages of capillaries with faster peak elution; the ability to cope with larger sample volumes and lower operating pressures, thus lending themselves to incorporation in portable devices for point of care (POC) use. The uptake of multi-capillary columns has been limited by concerns regarding reproducibility and, to date, they have predominantly been used in combination with ion mobility spectroscopy (IMS), in applications such as analysis of volatiles from human urine (Rudinicka *et al* 2010) and the headspace of wine and cork samples (Marquez-Silero, Cardenas and Valcarcel 2012).

The combination of fast chromatography with sensor detection is potentially a powerful tool for rapid analysis of biological samples and also lends itself to the development of POC devices.

Our approach utilises VOCs emitted from faeces for the detection of the nosocomial infection *Clostridium difficile*, where there is a need for rapid detection. There is limited evidence that distinctive odours emanate from stool samples which can be noticed by clinical staff to support the diagnosis *C. difficile* associated diarrhoea (Burdette *et al* 2007).

Characteristic patterns of VOCs emanating from faecal samples from patients have been shown for a number of infectious and non-infectious diseases including; cholera (Garner *et al* 2009a), diarrhoea (Probert *et al* 2004), gastrointestinal diseases (Garner *et al* 2007), irritable bowel syndrome (Ahmed *et al* 2013 and Ratcliffe *et al* 2013), liver disease (Raman *et al* 2013 and Probert *et al* 2009), necrotising enterocolitis (Garner *et al* 2009b) and rotavirus infection (Al-Kateb *et al* 2012).

*C. difficile* infection (CDI) is the most important cause of hospital acquired diarrhoea and a significant cause of morbidity and mortality in hospitals, causing approximately 1% of deaths in UK hospitals, with 2053 deaths reported in 2011 (Office for National Statistics Statistical Bulletin, 2012). There were *circa* 18000 cases in England in 2011, 80% of which were in the over 65s. This causes a large burden not just clinically but also financially. The use of certain antibiotics can cause a change in the gut flora that allows *C. difficile* bacteria to proliferate, resulting in illness (Public Health England 2013). Risk factors include a weakened immune system, serious illness and long-term hospitalisation. CDI causes mild to severe of diarrhoea and can lead to more serious complications.

Diagnostic methods for *C. difficile* are problematic and tend to need to be undertaken in a specialist laboratory. This delays implementation of appropriate treatment for patients and infection control measures, such as isolation. Current tests include two reference methods: cytotoxicigenic culture which detects *C. difficile* that can produce toxins and the cell cytotoxicity assay (48 hours) which detects the presence of toxins in the samples, both are time consuming. Widely used alternatives are enzyme immunoassays (EIAs); these detect *C. difficile* toxins and are rapid, but have been deemed unreliable as stand-alone tests (Wilcox, 2012). The use of tests and combinations of tests varies across the UK and worldwide and there is much debate as to the optimum approach (Wilcox, 2011; 2012). Current guidelines within the National Health Service (NHS), UK recommend a two test screening protocol: a *C. difficile* test such as a glutamate dehydrogenase test followed by a sensitive toxin EIA and a complete refrain from a stand-alone EIA (Department of Health Report 2012).

Our work focuses on determining whether there are changes in the VOC profiles that relate to *C. difficile* disease states and whether the combination of GC and sensor can detect those differences. Our aim was to develop an instrument and method that was able to accurately detect and measure volatile profiles characteristic of disease or absence of disease.

This paper describes the results of our comparison of a multi-capillary column and MOS sensor with a single capillary column and MOS sensor; followed by a study of the systems' ability to differentiate between *C. difficile* positive and negative diarrhoea samples.

The sensor in combination with GC could enable rapid diagnosis of a sample and has the potential to be undertaken in a ward or GP setting by non-specialists, as opposed to sending samples to a laboratory for initial diagnosis. The potential advantages of such a point-of-care device are faster diagnosis enabling patients to receive appropriate treatment more quickly, faster implementation of infection control measures, such as isolation and ward closure to reduce transmission, which would lead to a reduction in the number of cases, deaths and associated costs.

## 2 Methods

### 2.1 Prototypes - gas chromatography and sensor conditions

Two prototypes were tested, one using a long single capillary column and the other a short multicapillary column. The gas chromatograph (GC) was an SRI Instruments 8610 C (Torrance CA, USA) with a heated static headspace injector. This allowed sampling of a fixed volume of headspace. A bespoke insert (SRI Instruments) converted the sampler to accept a 10 mL version of the standard 40 mL vial with screw top and septa (Supelco, Sigma-Aldrich).

The carrier and purge gases were supplied by compressed, blended air (BOC, Manchester, UK) filtered through a 200cc hydrocarbon trap (Alltech Associates, Applied Science Ltd, Carnforth, UK). The split vent flow was set to 45 mL min<sup>-1</sup>. The purge gas pressure was 10 PSI (68.95 kPa) valve temperature 150 °C, vial temperature 50 °C, vial inlet and outlet piping temperature 80 °C.

The chromatographic detector used was a custom-designed metal oxide sensor operating at 450 °C. The electrode designs were made in-house and the electrodes printed onto 3 x 3 mm alumina substrates (ESL, Reading, UK). The gold interdigitated electrode design consisted of 8 interpenetrating bars, a track width of 150 µm and interdigitated gap of 100 µm. The gold electrode was hand-coated with a tin and zinc oxide paste using a previously described method (de Lacy Costello *et al* 2003). A platinum heater was printed onto the reverse of the substrates with the contact pads on the opposite side to the electrode pads for convenient attachment to the transistor outline header, TO-39 (Schott, Eltek semiconductors Ltd, Dartmouth, UK) used as a mount for the sensor. The four contacts were connected using gold wires to the 4 pin header.

The sensor was housed in a chamber, into which the end of the chromatography column was inserted and secured 1mm from the surface of the sensor, in order to detect compounds as they eluted from the column. The sensor chamber was a die cast aluminium enclosure of dimensions 89 x 35 x 30 mm with lid and gasket (RS Components UK). The chamber was modified to include inlets for the column, electronic connections, and inlet and outlet for purging air. A micro diaphragm pump (KNF Neuberger, Oxford, UK) was used to pass air through an activated charcoal trap (in-house) and through the chamber at a rate of 150 mL min<sup>-1</sup>.

### 2.1.1 The prototypes

The prototypes each consisted of the same aforementioned gas chromatograph and detector to facilitate standardisation and direct comparison of the two columns. The first column was a short coiled multi-capillary column (SMCC), 1 m in length, 2.2 mm diameter consisting of approximately 1200 capillaries per column. Each capillary had a diameter of approximately 40  $\mu\text{m}$  and was coated with stationary phase OV-5 (5% phenyl – dimethylpolysiloxane), thickness 0.2  $\mu\text{m}$  (Multichrom Ltd. Novosibirsk, Russia). It was operated at 30 °C with a carrier gas pressure of 10 PSI and column flow rate 12.2 mL min<sup>-1</sup>. The second was a long single capillary column (LSCC), length 30 m, stationary phase SPB<sup>TM</sup> – 1 Sulfur, poly(dimethylsiloxane), thickness 4  $\mu\text{m}$ , internal diameter 0.32 mm (Supelco, Sigma Aldrich, Gillingham, UK). The operating temperature was 40 °C, the carrier gas pressure was 30 PSI and the column flow rate was 7.8 mL min<sup>-1</sup>. The prototypes were powered continuously throughout, with the sensor heated, carrier gas flowing and sensor chamber purging.

### 2.2 Standards and samples

An in-house standard stool sample was used throughout the study to assess the intra- and inter-prototype stability. It comprised stool samples from 20 healthy volunteers. The samples were homogenised with an equal mass of water to give one consistent mixture, which was aliquotted into 1000 vials. These were stored frozen at -18°C.

All samples were from people presenting with diarrhoea, and the samples which had tested negative for other causes commonly tested: *Norovirus*, *Campylobacter*, *Salmonella* and *Shigella*. A sample was defined as *C. difficile* positive when diagnosed positive according to an enzyme immunoassay (EIA) *C. difficile* toxin test. A sample was defined as negative when diagnosed as negative according to the same EIA test.. These tests were undertaken by the Health Protection Agency (HPA) laboratory at Bristol Royal Infirmary. Samples were stored frozen at -18°C.

Ethical approval for this study was obtained from the Wiltshire Research Ethics Committee (REC ref. 06/Q2008/6).

### 2.3 Prototype testing

System sensitivity, stability and resolution were monitored each morning and evening, using a certified gas standard of 50 ppm ethanol and 50 ppm methanol balanced with dry blended air injected into the prototype, with the exception of the final study where a single certified calibration standard of 50 ppm ethanol was used. Both standards were certified +/-2.5% and from Cryoservice (AirProducts) Crewe, UK. The ability of the columns to separate complex mixtures was assessed using standard stool samples.

All samples were received frozen at -18°C and defrosted in a water bath at 37 °C for 30 minutes immediately prior to use. The sample was heated within the prototype for 10 minutes at 50°C prior to the headspace being injected. The chromatogram was recorded for 30 minutes on the LSCC and 15 minutes on the SMCC prototype following injection, after which a sample of the headspace of an empty vial was injected, to ensure no contamination

within the instrument or carryover from the previous sample. This gave a total run time for a sample of 40 and 55 minutes for the SMCC and LSCC prototypes, respectively.

A comparison of the LSCC and the SMCC was undertaken using 25 *C. difficile* positive and 28 *C. difficile* negative samples (hospital-based controls) from different patients. The diarrhoea samples were selected and run at random. All samples were run on the SMCC, after which the column was changed to the LSCC and the samples were re-run in the same order, to ensure as far as possible that the GC columns were exposed to the same conditions. A prototype was selected based on the outcome of this study. The selected prototype was fitted with a new column and evaluated using a larger, completely separate set of samples comprising: 50 positive and 50 negative *C. difficile* samples.

## 2.4 Artificial Neural Network analysis

Artificial neural network analysis was used to analyse the data. There are different models of ANN; here a model known as a Multi-Layer Perceptron (MLP) has been used to assess the diagnostic potential of the systems. The prototype output is a chromatogram, a trace of resistance change with time, with peaks corresponding to the sensor responding to compounds as they elute from the column. To generate a set of discrete inputs, the following process was applied: a three-point moving average was first applied to the sensor signal, and then the first derivative of the smoothed data with respect to time ( $dR/dT$ ) was taken to display the rate of change of the sensor resistance. Since exposure to VOCs results in a decrease in resistance, the negative portion of the  $dR/dt$  trace corresponds to the response of the sensor and the positive portion the subsequent recovery. As such the negative portion only was selected and used in the analysis, which gave a set of discrete peaks. To enable removal of baseline noise there was the option to set a minimum height threshold (HT) of the 1<sup>st</sup> differential,  $\Omega\text{ s}^{-1}$ , for which a variety of values were tested: 0, 50, 100 and 200. The first 10 minutes of the 1<sup>st</sup> differential chromatogram from the SMCC was divided into 120 x 5second bins and the first 30 minutes of the LSCC was split into 120 x 15 second bins. Sixty x 10 and 40 x 30 second bins were also investigated using the SMCC data.

The magnitudes of all data points falling within the bin boundaries and above the HT were summed to give a value for each bin. The same threshold principle was applied to the bin values, with a threshold, called area threshold (AT),  $\Omega\text{ s}^{-1}$ , values tested were 0, 100, 200, 500, 1000 and 1500. The resulting set of bins values for a chromatogram were normalised such that the largest bin was assigned a value of one, empty bins 0 and all others proportionally between 0 and 1. The bins were then input into the ANN. The 'learning rate' and 'momentum' parameters used in calculating the correction factors to the weights during the learning process were fixed for all networks at a value of 0.5. All networks contained one hidden layer. To find the optimum number of neurons in the hidden layer, networks were generated using all numbers of neurons between the number of inputs in this case 120, and the number of outputs, two (*C. difficile* positive and *C. difficile* negative).



### 3 Results

#### 3.1 Prototype behaviour to standards

The prototypes sensitivity, stability and resolution were monitored through triplicate injections of 50 ppm ethanol and 50 ppm methanol twice daily. The SMCC prototype was able to separate methanol and ethanol but the resolution was inferior to the LSCC, as can be seen from Figure 1 a and b. The SMCC baseline has just begun to recover from methanol when the ethanol peak elutes, Figure 1c and d show the difference in separation of the first differential. The mean retention times (RTs) when using the SMCC prototype were 19.8 and 25.0 seconds for methanol and ethanol respectively, with an average relative standard deviation (RSD) each day of 1.5% and 1.3% respectively. The LSCC produced longer retention times, methanol 55.0 and ethanol 76.5 seconds, but with greater peak separation and average daily RSD 0.7% and 0.5% respectively, illustrating high stability of the columns (Figure 2 a).

The baseline resistance was the resistance of the sensor in the absence of standard or sample gases. For calculations the average was taken of 5 data points immediately prior to the injection of a sample. Sensor sensitivity was defined as  $(\text{baseline resistance} - \text{peak resistance}) / \text{peak resistance} \times 100$  and recovery,  $t_{95\%}$ , as the resistance returning to 95% of the baseline resistance value. The sensitivities of the SMCC prototype and LSCC prototype to ethanol are very similar, mean values 540% (3% RSD) and 539% (4% RSD) respectively. The LSCC was slightly more sensitive to methanol, 306% (4% RSD) compared to the SMCC, 274% (4% RSD). The LSCC system sensitivity was stable with no significant trends apparent. The SMCC system was slightly less stable with the sensitivity of the sensor increasing over the course of the study (Figure 2 b and c). The recovery times for each prototype was stable over the course of the study and, on average, the recovery time was approximately 20% faster when using the SMCC, mean of 190 (16% RSD), compared to 238 (12% RSD) seconds for the LSCC.

The baseline resistance prior to each methanol/ethanol injection are also shown in Figure 2 (b) and (c); the baselines of both prototypes were stable under rigorous testing conditions with daily drift of typically less than 10%. There was a trend producing a 'toothed' graph, where resistance was often greater at the start of a day. This suggests that a degree of sensor fouling occurs during the day, which the sensor recovers from during an extended period of time without sample exposure, such as overnight. This is to be expected as the test conditions are relatively intensive, with concentrations likely to be in the region 10-100 volume ppm of a range of VOCs.

#### 3.2 Standard stool samples

In addition to the ethanol and methanol standards, two standard stool samples were run on each prototype at the beginning and end of the study. The retention time of the first peak to elute from the standard stool sample was 18 seconds for the SMCC prototype and 40.5 seconds for the LSCC prototype (Figure 3). The number of peaks in each standard stool sample is shown in Table 1. A slightly greater number of peaks were detected with the LSCC prototype suggesting increased separation of the sample volatiles, with the majority

of peaks having eluted within 25 minutes of the injection. The SMCC, in contrast, had few peaks which eluted after 10 minutes. This illustrated the ability of the columns to separate VOCs and the sensors ability to respond to and recover from exposure to a range of VOCs.

### 3.3 Analysing diarrhoea samples & comparing prototypes

Twenty five positive and 28 negative *C. difficile* samples were run on each prototype; an example of a *C. difficile* positive sample chromatogram is shown in Figure 4. The numbers of peaks in a chromatogram were counted for each sample; greater numbers of peaks were obtained from the LSCC compared to the SMCC, and slightly greater numbers were found in the negative compared to the positive samples (Table 2).

In order to minimise sample analysis time the method has been developed so that only the first 10 (SMCC) and 30 minutes (LSCC) of the chromatogram is used for analysis as in many cases all if not the majority of the peaks had eluted by this point, thereby contained all the sample information that could be fed in to the artificial neural network.. However, the total run time is based on the presence of a later eluting peak. This occur in 56% of samples up to 50 minutes after injection on the SMCC and in 83% of samples up to 75 minutes after injection on the LSCC, indicating the minimum sample run time. This was a balance between collecting information and minimising runtime.

### 3.4 Artificial neural network

The raw resistance trace of each sample was converted to the 1<sup>st</sup> differential then split into bins as is illustrated in Figure 5. The 25 *C. difficile* positive and 28 negative samples were randomly split four times in to a training set (TS) and validation set (VS) producing four datasets. The ratio of positive: negative was 50:50 in the TS and VS. The TS:VS ratio was: 50:50 in dataset 1 and 75:25 TS: VS in datasets 2-4. The same datasets were used for the SMCC and LSCC to enable fair comparison of results. The highest percentage correct classification for each dataset from each prototype are summarised in Table 3. Overall the highest percentage correct classification was 78.6% (LSCC) and 92.3% (SMCC). The lowest overall percentage was 71.4% (LSCC) and 69.2% (SMCC), indicating differences between *C.difficile* and non *C.difficile* samples. In terms of sensitivity and specificity, the maximum sensitivity for the prototypes was 100% (both), the maximum specificity was 100% (SMCC) and 71% (LSCC).

### 3.5 Selection of a prototype for further testing

In deciding upon which prototype to develop for further testing, we first considered the prediction ability of each prototype followed by the device properties. The results of the ANN analysis indicate that, with one exception, the SMCC column returned better results than the LSCC, dataset 1 slightly favours the LSCC. Datasets 2-4 contain more samples in the training set which should result in a better trained network; conversely they are tested using fewer samples.

Other factors for consideration were the SMCC peaks elute faster and recover slightly more quickly, giving a shorter diagnosis time and fewer later eluting peaks. The SMCC is able to be operated at lower pressures, making the use of a compressor a possibility, important for a



future portable device. The LSCC prototype usually produced a few more peaks than the SMCC prototype, but this was not reflected as an advantage in the ANN results. Sensitivity and stability was similar for both prototypes.

After considering the ANN results and the ancillary factors, such as the actual and effective run times the decision was made to undertake the next study using the SMCC prototype.

### 3.6 SMCC prototype stability during testing of 100 patient samples

A new SMCC column was installed in the prototype. The mean retention time for ethanol throughout the study was 27.5 seconds (S.D. 0.40) and the RSD was 1.44% indicating little drift. This compared well to an overall mean of 25.0 seconds for the SMCC column used during the comparative testing study, both columns were from the same batch.

Sensitivity to ethanol exhibited a slight decrease during the course of the study (Figure 6) and a general trend of decreasing between the beginning and end of each day of testing. This trend was the same as that observed in the earlier sample testing.

### 3.7 Artificial Neural Network results for 100 samples

An independent set of fifty *C. difficile* positive and fifty *C. difficile* negative samples were run on the SMCC. The sample results were split 60:40 into TS:VS, with 30 positive and 30 negative samples in the TS, and 20 positive and 20 *C. difficile* negative samples in the VS. ANN networks were generated using various numbers of input bins and additionally with various height thresholds (HT) and area thresholds (AT) (Table 4), in order to optimise these parameters. Networks with 120 input bins gave consistently better results than networks generated with either 60 or 40 input bins, and there was minimal variation in results using different thresholds.

The highest overall percentage correct was 82.5%, obtained by several networks, but with different sensitivity and specificity. The highest sensitivity obtained across all networks was 90% (specificity 70%, overall 80%). The highest specificity was 95% (sensitivity 65% overall 80%). The best balanced networks had 85% sensitivity and 80% specificity (and vice versa) giving an overall classification of 82.5%; these all used 120 bins, and either zero thresholds or a height threshold of 100 with various area thresholds. Therefore one set of conditions, such as zero thresholds could be selected for future use.

## 4 Discussion

Gas chromatography combined with a MOS sensor detected numerous VOCs from the headspaces of 150 stool samples. Sensitivity to ethanol and methanol standards were consistent during the study. A small decrease in sensitivity was observed between each morning and evening, which then recovered by the following morning, indicating a degree of reversible fouling. Severe irreversible fouling of the sensor was not observed, which could have been expected from sulphur containing VOCs in the samples. A short multi capillary column in the GC produced resolution, stability and sensitivity comparable to the conventional single capillary column. Previous work (Sidelnikov *et al* 2010) found that poor resolution, as a result of peak broadening; can occur because of the difficulty in making each

capillary within the multicapillary column identical; resulting in slight variations in cross section and volume which alter the flow rate. Hence each individual capillary will have its own RT and the peak width will be a result of the difference in each RT. In this work however, it was not found to have a significant negative impact on the results, with resolution proving sufficient to detect features that could be used to classify samples. The LSCC prototype exhibits slightly higher RT stability but this might be expected because of the reasons cited above and possibly a higher stability of the stationary phase of the column, mainly due to comparative ease of manufacture and higher film thickness versus the SMCC.

The numbers of peaks in the chromatograms proved sufficient for giving diagnostic power, albeit the number of peaks was far fewer than the number of volatiles, that could be expected to be present in the headspace of a stool sample; in the region of 100 (Garner *et al* 2007). The diarrhoea samples averaged 24 and 15 peaks for the LSCC and SMCC respectively. This can be attributed in the main to the response and recovery time of the sensor being greater than the separation between some peaks, thereby resulting in multiple compounds being seen as one peak. An improvement of sensor response would likely lead to detection of more peaks, although the sensor may not be sensitive to all VOCs. The differences between prototypes are a result of several factors. The column length is important, with the LSCC allowing greater separation of VOCs, also to be considered is the stationary phase; both are poly dimethyl siloxane backbones, with the SMCC containing a 5% diphenyl substitution, for targeting aromatic compounds.

The ability to use ANN analysis of the results has been shown, the current approach enables future samples to be automatically assessed through a network after the run giving real-time diagnosis. Direct comparison of our results with the current tests in use is difficult due to widely varying reported values. The Clinical Practice Guidelines from the Society for Healthcare Epidemiology and the Infectious Diseases Society (Cohen 2010) reviewed many studies and summarised the reported sensitivities and specificities: Cytotoxin (cell cytotoxicity assay) detection sensitivity 66-100%; enzyme immunoassay for A, B or A/B sensitivity 63-94% and specificity 75-100%; GDH antigen tests sensitivity 85-95%, specificity 89-99, but also 76 and 32%; and a two-step GDH then EIA toxin sensitivity 77-87%. Considering these varying and sometime conflicting values, our results are comparable, achieving maximum sensitivity and specificity of 85% and 80% respectively (positive predictive value 84% and negative predictive value (81%) albeit in a relatively small study It should be noted that our study uses well-classified clinical samples and only hospital-based controls who were symptomatic but found to be *C. difficile* negative (not healthy individuals) thus despite being on a small scale is a reflection of the true diagnostic challenge. Eighty percent may be sufficient for a diagnostic test depending on its use, e.g. whether a presumptive test is appropriate; however the aim is to improve on this.

A larger study is to be conducted, including greater numbers of *C. difficile* positive and *C. difficile* negative samples which could also include other common diarrhoea causing infections such as Norovirus. A larger study will enable the device to be tested to a cohort of samples that is representative of samples submitted for testing, indicating the prevalence of *C. difficile* positive samples, currently approximately 10%, dependent on test used (Strachan *et al* 2013) A further improvement will be the use of triple testing for *C. difficile* using three

different tests: toxigenic culture, culture & EIA and EIA, to help compensate for false positives and false negatives with the various tests.

There is considerable scope for developing the device; improving chromatographic separation and sensor response, recovery, to increase the number of VOCs that can be detected. Areas being considered include purging and cleaning, temperature ramp, multiphase columns, multi-sensors and preconcentration. Additionally investigation of long term sensor stability and also reproducibility between sensors will become of greater importance and a reassessment of the fabrication process will be needed, to allow selection of sensors within set tolerances for use.

An important feature of the SMCC prototype is its use of air as the carrier gas, meaning the use of a compressor is feasible and preferential for portability and POC settings. This device has advantages over earlier GC-sensor devices such as the work by Khalid *et al* (2013), notably that the column is operated at low temperature, 40°C and without the need of a temperature ramp, and at a lower pressure, 10PSI as opposed to 35PSI. The use of a static headspace sampler simplifies sample introduction and ensures higher reproducibility.

The aim is to increase the detection of VOCs whilst limiting the overall analysis time, this is a balance improving separation and reducing analysis time, with the speed of response and recovery of the sensor being crucial in addition to the column.

## 5 Conclusions

In this paper it has been demonstrated that the combination of a GC column and MOS sensor enables detection of volatiles which can be used to discriminate between *C. difficile* and non-*C. difficile* samples. It has been shown that columns can be operated at low temperatures and air used as an alternative carrier gas whilst maintaining good stability and separation; both important for future smaller and more portable prototypes.

The subsequent pilot work has indicated the approach has the capability of detecting volatile compounds that can be used to discriminate between diseases, in this case *C. difficile*. Furthermore, comparison of a conventional capillary GC column with a short multi capillary column showed that although the output resolution was slightly inferior using the SMCC, the differentiation of samples based on the produced chromatograms was comparable.

ANN results of the device data showed comparable sensitivity and specificity to a range of commercial *C. difficile* tests, albeit the sample numbers were small. Future work will concentrate on running larger numbers of *C. difficile* positive samples to enhance the neural network models and further develop the prototypes.

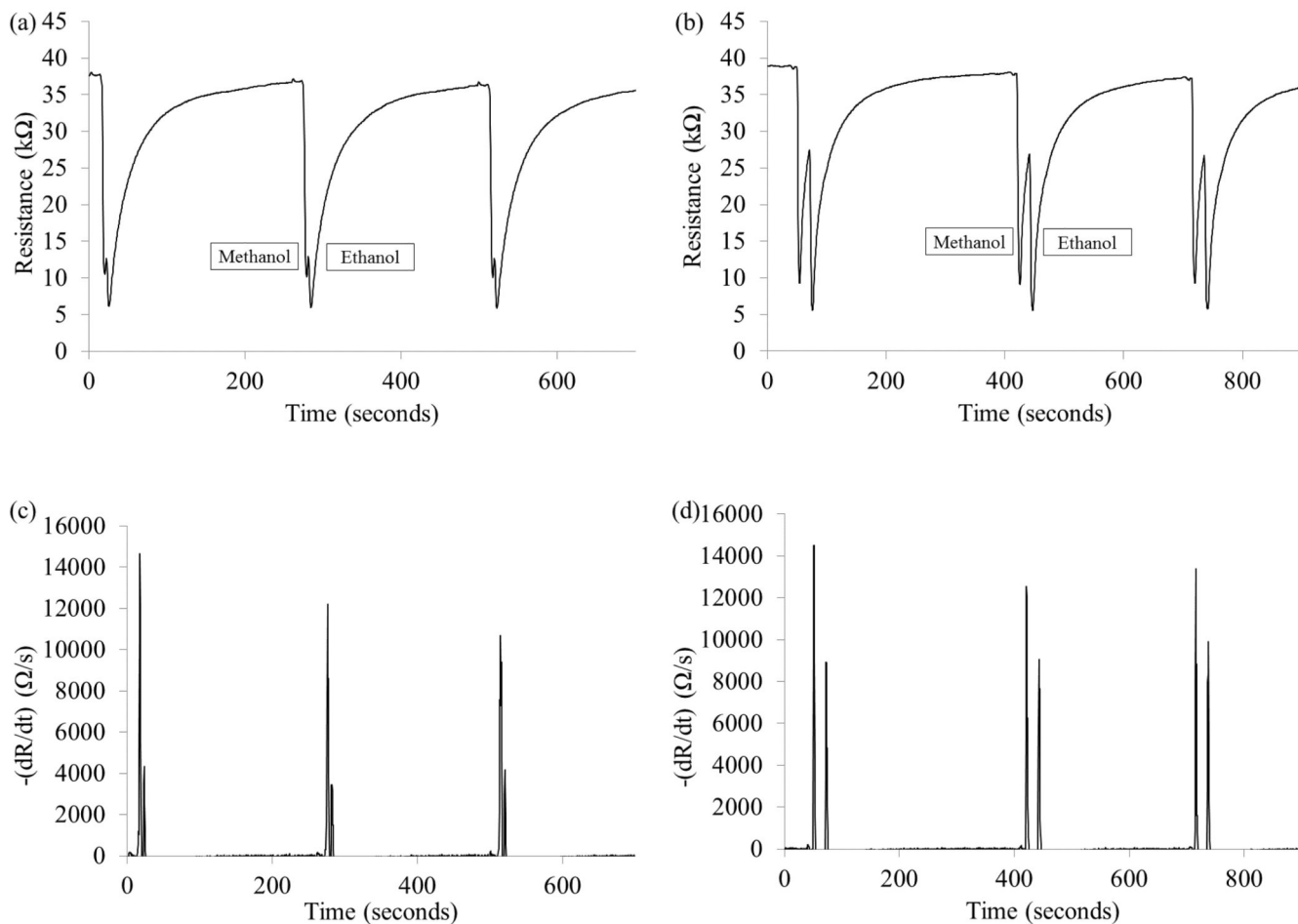
## Acknowledgments

The authors would like to thank the Wellcome Trust for funding the project, and the staff at the Bristol Royal Infirmary (Bristol, UK) and Health Protection Agency for all the effort in collecting and testing samples.

## References

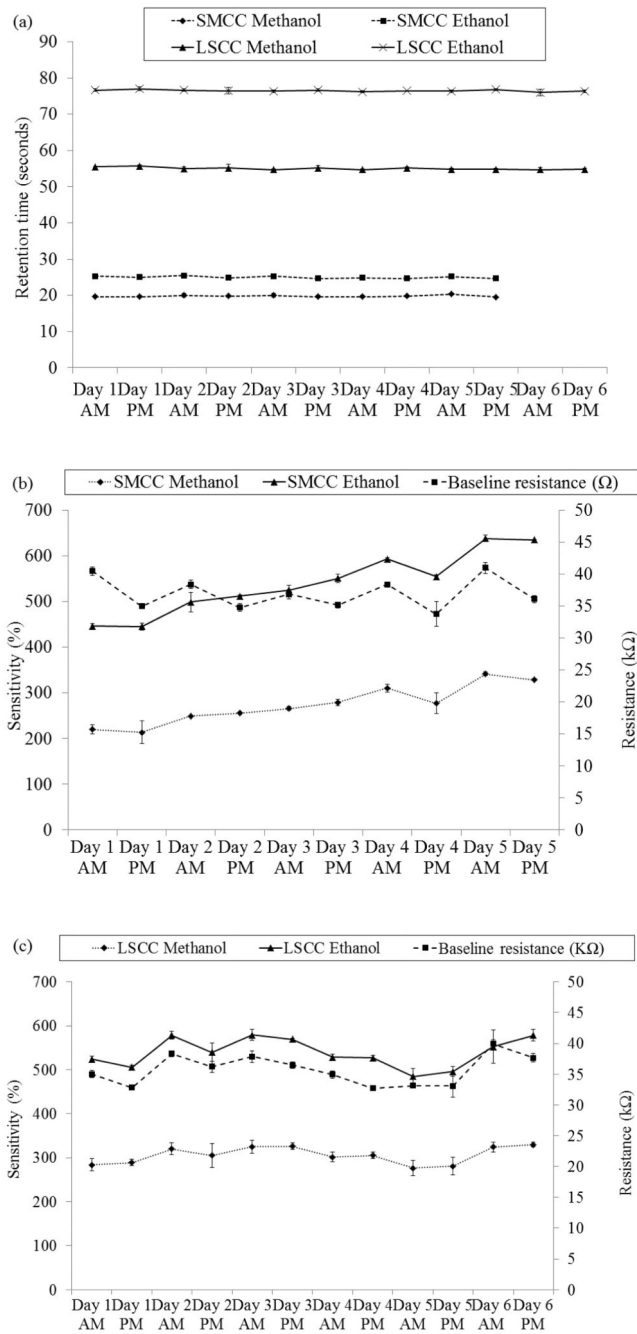
- Aguilera T, Lozano J, Paredes JA, Alvarez FJ, Suarez JI. Electronic Nose Based on Independent Component Analysis Combined with Partial Least Squares and Artificial Neural Networks for Wine Prediction. *SENSORS*. 2012; 12(6):805572.
- Ahmed I, Greenwood R, de Lacy Costello B, Ratcliffe NM, Probert CSJ. An Investigation of Fecal Volatile Organic Metabolites in Irritable Bowel Syndrome. *PLOS ONE*. 2013; 8(3):e58204. [PubMed: 23516449]
- Al-Kateb H, Cunliffe N, de Lacy Costello B, Probert C, Ratcliffe N. Analysis of faecal volatiles from young children infected with and without rotavirus. *GUT*. 2012; 61(s2):A364–365.
- Boots AW, van Berkel JJBN, Dallinga JW, Smolinska A, Wouters EF, van Schooten FJ. The versatile use of exhaled volatile organic compounds in human health and disease. *Journal of Breath Research*. 2012; 6:027108. [PubMed: 22621865]
- Burdette SD, Bernstein JM. Does the nose know? The odiferous diagnosis of *Clostridium difficile* associated diarrhoea. *Clinical infectious diseases*. 2007; 44:1142. (correspondence). [PubMed: 17366472]
- Cohen SH, Gerding DN, Johnson S, Kelly CP, Loo VG, McDonald LC, Pepin J, Wilcox MH. Clinical Practice Guidelines for Clostridium difficile Infection in Adults: 2010 Update by the Society for Healthcare Epidemiology of America (SHEA) and the Infectious Diseases Society of America (IDSA). *Infection Control and Hospital Epidemiology*. 2010; 31(5):431–55. [PubMed: 20307191]
- De Lacy Costello B, Ewen R, Ratcliffe NM, Sivanand PS. Thick film organic vapour sensors based on binary mixtures of metal oxides. *Sensors and Actuators B*. 2003; 92:159–66.
- Department of Health Report. Updated guidance on the diagnosis and reporting of Clostridium difficile. 2012. Published March 2012 [www.dh.gov.uk/publications](http://www.dh.gov.uk/publications) Gateway ref 17215
- Di Natale C, Macagnano A, Martinelli E, Paolesse R, D'Arcangelo G, Roscioni C, Finazzi-Agro A, D'Amico A. Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors. *Biosensors and Bioelectronics*. 2003; 18(10):1209–18. [PubMed: 12835038]
- Dummer J, Storer M, Swanney M, McEwan M, Scott-Thomas A, Bhandari S, Chambers S, Dweik R, Epton M. Analysis of biogenic volatile organic compounds in human health and disease. *Trends in Analytical Chemistry*. 2011; 30:960–967.
- Fine GF, Cavanagh LM, Afonja A, Binions R. Metal oxide semi-conductor gas sensors in environmental monitoring. *Sensors*. 2010; 10:5469–5502. [PubMed: 22219672]
- Garner CE, Smith S, Bardhan PK, Ratcliffe NM, Probert CSJ. A pilot study of faecal volatile organic compounds in faeces from cholera patients in Bangladesh to determine their utility in disease diagnosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2009; 103:1171–3. [PubMed: 19268999]
- a. Garner CE, Ewer AK, Elasoquad K, Power F, Greenwood R, Ratcliffe NM, de Lacy Costello B, Probert CS. Analysis of faecal volatile organic compounds in preterm infants who develop necrotising enterocolitis: a pilot study. *Journal of Pediatric Gastroenterology and Nutrition*. 2009; 49(5):559–65. [PubMed: 19668005]
- b. Garner CE, Smith S, de Lacy Costello B, White P, Spencer R, Probert CSJ, Ratcliffe NM. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *FASEB J*. 2007; 21:1675–1688. [PubMed: 17314143]
- Ghasemi-Varnamkhasti M, Mohtasebi SS, Siadat M, Lozano J, Ahmadi H, Razavi SH, Dicko A. Aging fingerprint characterization of beer using electronic nose. *Sensors and Actuators B- Chemical*. 2011; 159(1):51–9.
- Health Protection Agency. Health Protection Report. 2012; 6(45) URL: <http://www.hpa.org.uk/hpr/infections/enteric.htm#gofi> [03/12/2012].
- Jung JY, Lee CS. Characteristics of the TiO<sub>2</sub>/SnO<sub>2</sub> thick film semiconductor gas sensor to determine fish freshness. *Journal of industrial and engineering chemistry*. 2011; 17(2):237–42.
- Kim K-H, Jahan SA, Kabir E. A review of breath analysis for diagnosis of human health. *Trends in analytical chemistry*. 2012; 33:1–8.

- Khlaid T, White P, De Lacy Costello B, Persad R, Ewen R, Johnson E, Probert CSJ, Ratcliffe NM. A pilot study combining a GC-sensor device with a statistical model for the identification of bladder cancer from urine headspace. *PLOS One*. 2013; 8(7):1–8.
- Marquez-Sillero I, Cardenas S, Valcarcel M. Headspace-multicapillary column-ion mobility spectrometry for the direct analysis of 2,4,6-trichloroanisole in wine and cork samples. *Journal of chromatography A*. 2012; 1265:149–154. [PubMed: 23089513]
- McNerney R, Mallard K, Okolo PI, Turner C. Production of volatile organic compounds by mycobacteria. *FEMS microbiology letters*. 2012; 328:150–156. [PubMed: 22224870]
- Oh EH, Song HS, Park TH. Recent advances in electronic and bioelectronics noses and their biomedical applications. *Enzyme and microbial technology*. 2011; 48:427–37. [PubMed: 22113013]
- Office for National Statistics. [Accessed 30/11/2012] Statistical Bulletin. Deaths Involving *Clostridium difficile*, England and Wales, 2011. Report August 2012. <http://www.ons.gov.uk/ons/rel/subnational-health2/deaths-involving-clostridium-difficile/2011/stb-deaths-involving-clostridium-difficile-2011.html> 14:00
- Probert CSJ, Ahmed I, Khalid T, Johnson E, Smith S, Ratcliffe NM. Volatile organic compounds as diagnostic biomarkers in gastrointestinal and liver diseases. *Journal of gastrointestinal and liver diseases*. 2009; 18(3):337–43. [PubMed: 19795029]
- Probert CSJ, Jones PRH, Ratcliffe N. A novel method for rapidly diagnosing the causes of diarrhoea. *GUT*. 2004; 53:58–61. [PubMed: 14684577]
- Public Health England. *Clostridium difficile*. 2013. URL: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/ClostridiumDifficile/> [13/03/13]
- Raman M, Ahmed I, Gillevet I, Probert CSJ, Ratcliffe NM, Smith S, Greenwood R, Sikaroodi M, Lam V, Crotty P, Bailey J, et al. Fecal Microbiome and Volatile Organic Compound Metabolome in Obese Humans With Non alcoholic Fatty Liver Disease. *Clinical gastroenterology and hepatology*. 2013; 11(7):868–75. [PubMed: 23454028]
- Ratcliffe NM, de Lacy Costello B, Shepherd S. The use of MS for the investigation of irritable bowel syndrome and inflammatory bowel disease. *Clinical Laboratory International Magazine*. 2013; 37:18–20.
- Rudnicka J, Mochalski P, Agapiou A, Statheropoulos M, Amann A, Buszewski B. Application of ion mobility spectrometry for the detection of human urine. *Anal Bioanal Chem*. 2010; 398:2031–38. [PubMed: 20820995]
- Sidelnikov VN, Patrushev Yu V, Nikolaeva OA. Express gas chromatography on multicapillary columns and its potential. *Catalysis in Industry*. 2010; 2(3):206–16.
- Sivapunniam A, Wiromrat N, Myint MTZ, Dutta J. High-performance liquefied petroleum gas sensing based on nanostructures of zinc oxide and zinc stannate. *Sensors and Actuators B-Chemical*. 2011; 157(1):232–9.
- Strachan AJ, Evans NE, Williams OM, Spencer RC, Greenwood R, Probert CJ. Comparison of a frozen human foreskin fibroblast cell assay to an enzyme immunoassay and toxigenic culture for the detection of toxigenic *Clostridium difficile*. *Diagnostic microbiology and infectious disease*. 2013; 75(1):42–5. [PubMed: 23107315]
- Thorn RMS, Greenman J. Microbial volatile compounds in health and disease conditions. *Journal of Breath Research*. 2012; 6:024001. [PubMed: 22556190]
- Wiegand PN, Nathwani D, Wilcox MH, Stephens J. Clinical and economic burden of *Clostridium difficile* infection in Europe: a systematic review of healthcare-facility-acquired infection. *J Hospital Infection*. 2012; 81:1–14.
- Wilcox MH. Overcoming barriers to effective recognition and diagnosis of *Clostridium difficile* infection. *Clinical microbiology and infection*. 2012; 18:13–20. [PubMed: 23121550]
- Wilcox MH. Laboratory diagnosis of *Clostridium difficile* infection: in a state of transition or confusion or both? *J Hospital Infection*. 2011; 79:1–3.

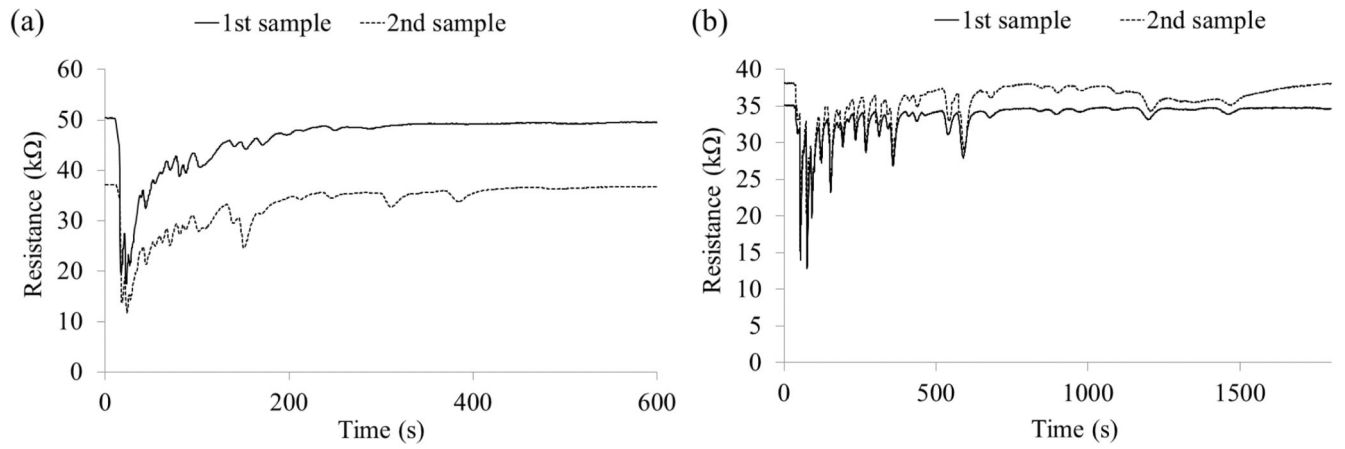


**Figure 1.** Three sequential ethanol and methanol injections (a) and (c) SMCC prototype, (b) and (d) LSCC prototype. (a) and (b) the chromatograms (c) and (d) the negative only portion of  $dR/dt$  corresponding to the sensor response.

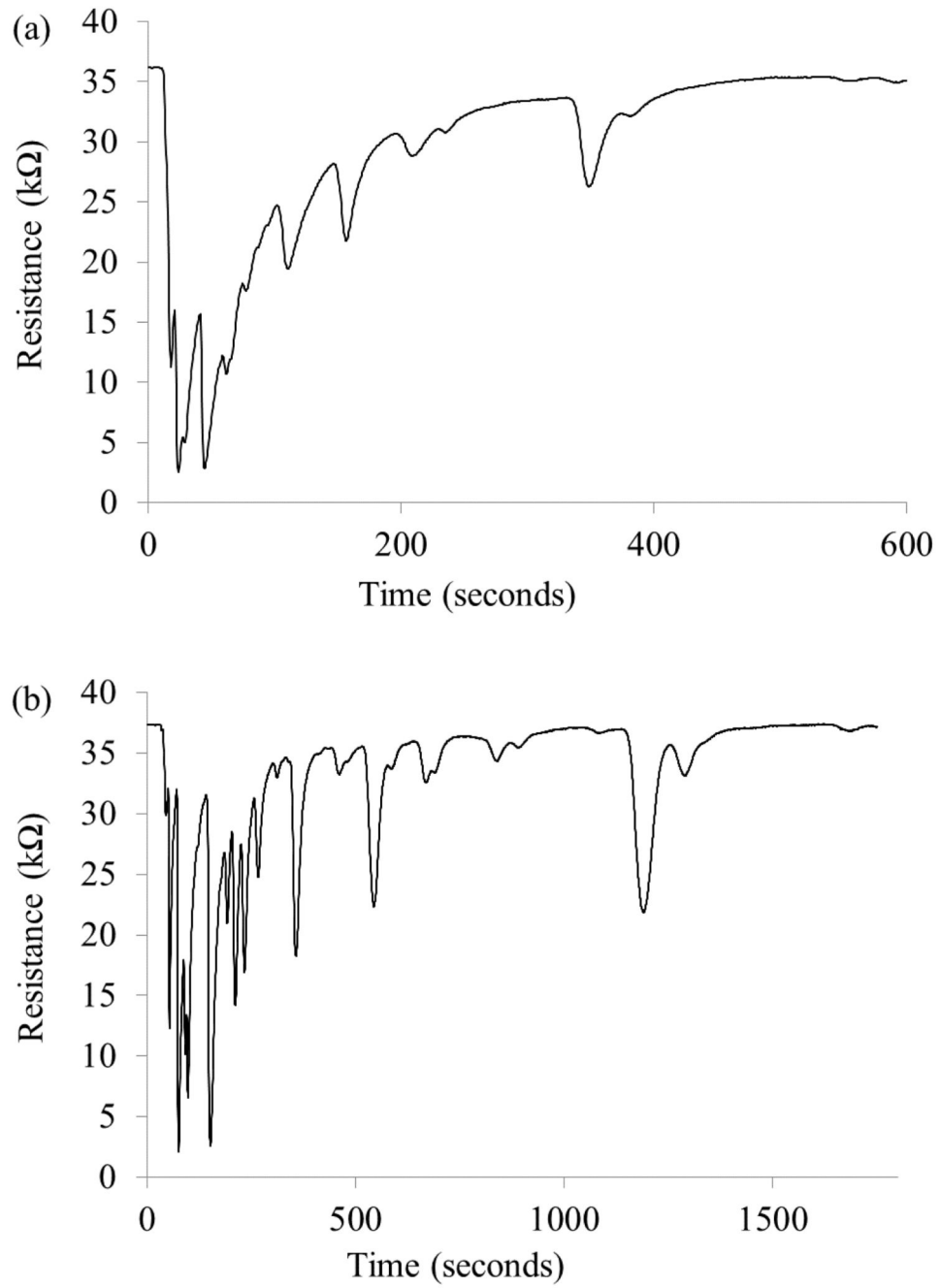




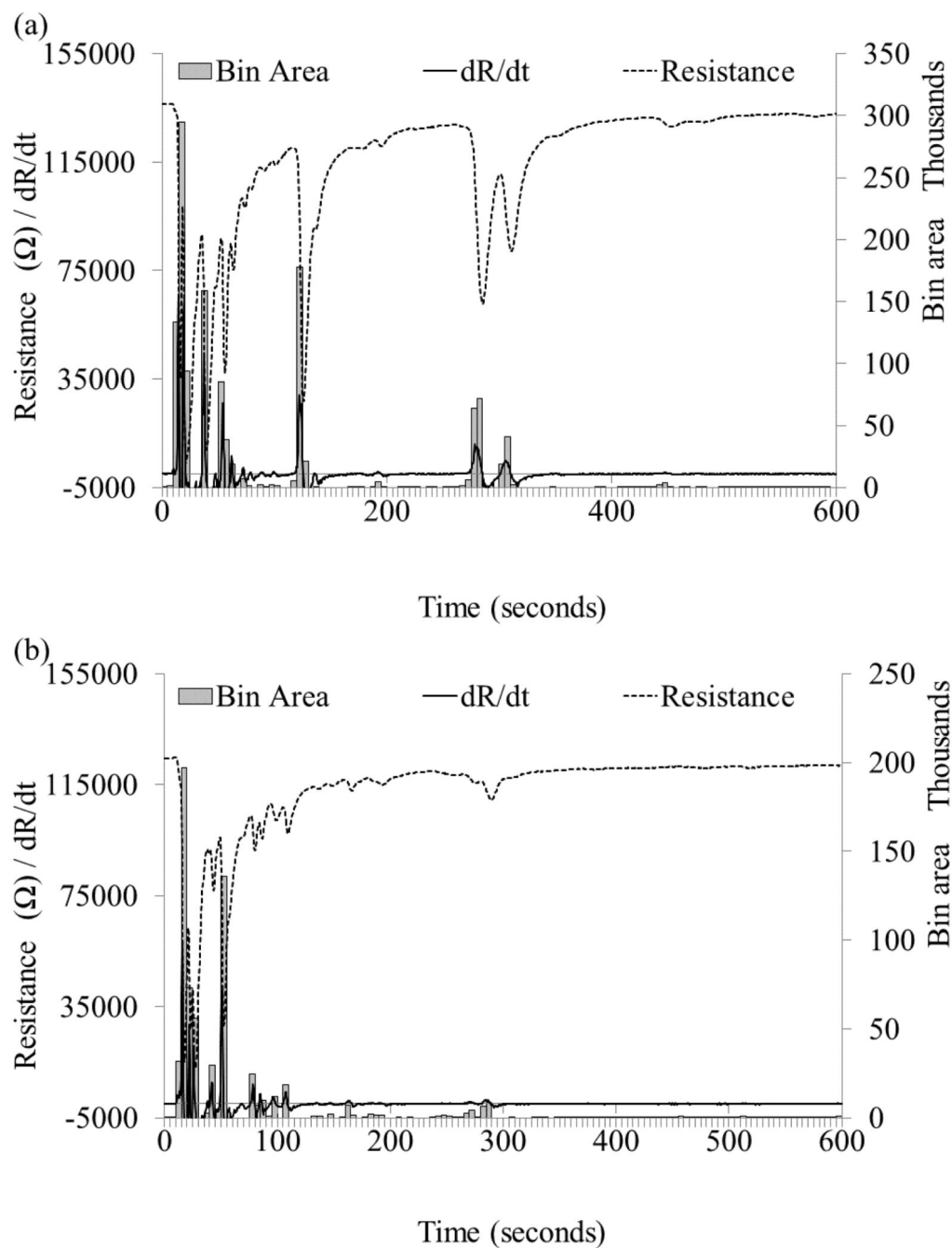
**Figure 2.** The 50 ppm ethanol and 50 ppm methanol injections using the SMCC and LSCC prototypes. (a) The retention times for both. (b) and (c) the sensitivity and baseline resistance of the SMCC and LSCC prototypes respectively. The mean of three injections is shown. The error bars are  $\pm 1$  S.D.



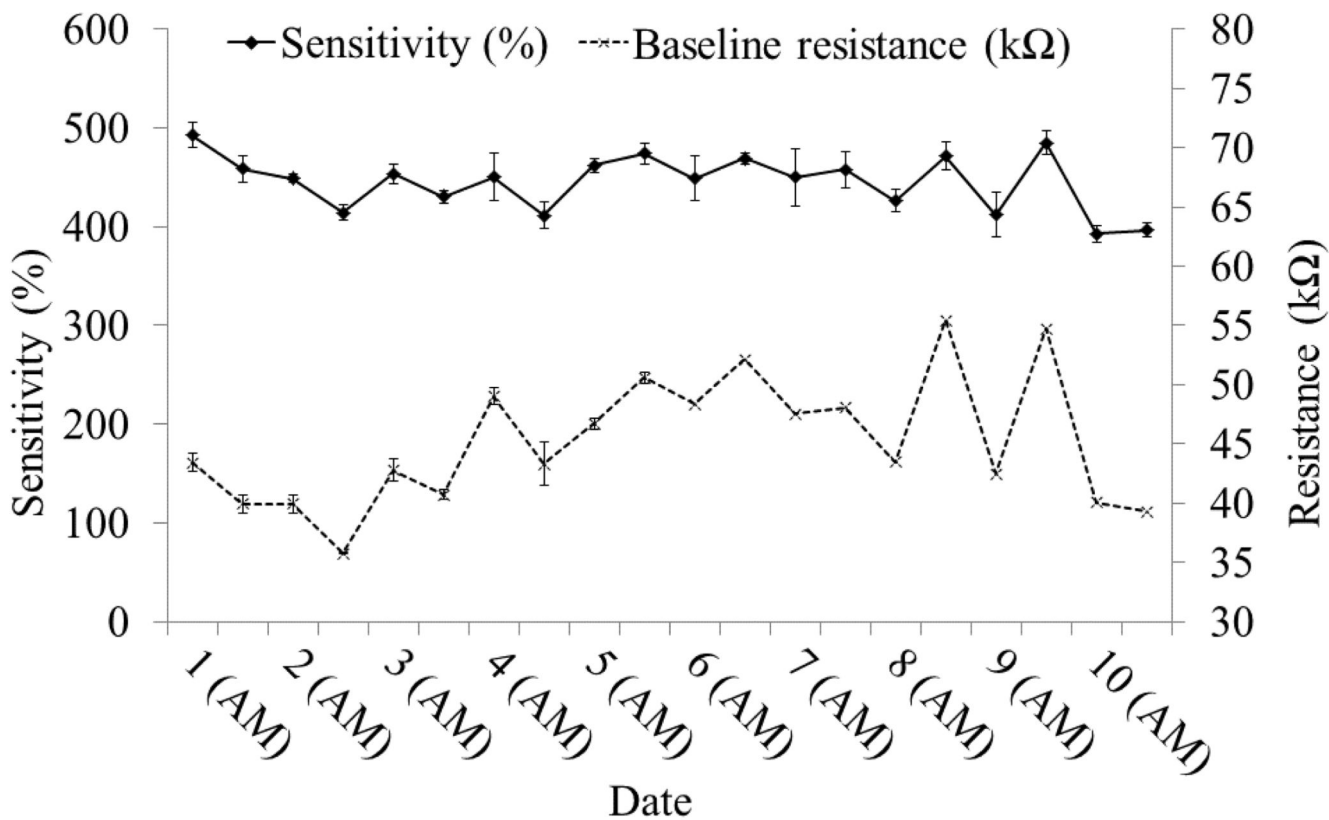
**Figure 3.** Two chromatograms of the standard stool samples run on each prototype (a) SMCC (b) LSCC.



**Figure 4.** A *C. difficile* positive sample split into two aliquots and run on (a) the SMCC and (b) the LSCC prototype.



**Figure 5.** Illustration of the conversion of the raw resistance trace to the 1<sup>st</sup> differential and then to bins for input into the ANNs. Samples shown were run on the SMCC prototype (a) *C. difficile* positive (b) *C. difficile* negative.



**Figure 6.** Retention times and sensitivity for ethanol during the course of the study. The mean of three injections is shown. The error bars are  $\pm 1$  S.D.

**Table 1**  
The table shows the number of peaks observed in the chromatograms from standard stool samples.

	LSCC		SMCC	
	1st sample	2nd sample	1st sample	2nd sample
Total number of peaks	32	32	25	27
Peaks between 0-600 seconds	23	23	21	23
Peaks between 601-1800 seconds	9	9	4	4



**Table 2**  
**The number of peaks from the chromatograms of each *C. difficile* positive and negative sample run using both prototypes.**

	<i>C.difficile</i> positive		<i>C.difficile</i> negative	
	SMCC chromatograms	LSCC chromatograms	SMCC chromatograms	LSCC chromatograms
Mean	14	23	16	25
Minimum	8	16	8	15
Maximum	23	30	27	33
Range	15	14	19	18
S.D.	3.57	4.07	4.86	5.36
R.S.D.	25.8	18.0	31.2	21.3

**Table 3**  
**Results summary of Artificial Neural Network processing.**

Dataset	Percentage correct (%)					
	LSCC prototype			SMCC prototype		
	<i>C. difficile</i> positive (Sensitivity)	<i>C. difficile</i> negative (Specificity)	Overall	<i>C. difficile</i> positive (Sensitivity)	<i>C. difficile</i> negative (Specificity)	Overall
1	91.7	57.1	73.1	83.3	57.1	69.2
2	100	57.1	76.9	83.3	100	92.3
3	85.7	71.4	78.6	100	71.4	85.7
4	71.4	71.4	71.4	85.7	85.7	85.7

**Table 4**  
ANN results for 100 samples using the SMCC prototype; 60:40 TS:VS split.

Number of Bins	Bin thresholds		Percentage correct		Overall percentage correct
	Height	Area	<i>C. difficile</i> positive (Sensitivity)	<i>C. difficile</i> negative (Specificity)	
40	0	0	65	80	72.5
60	0	0	70	80	75
			65	85	75
60	100	0	65	85	75
120	0	0	80	85	82.5
120	0	200	75	85	80
			80	80	80
120	0	500	80	80	80
			75	85	80
			85	75	80
120	0	1500	70	90	80
120	50	0	75	85	80
			80	80	80
120	50	200	75	85	80
			70	90	80
120	100	0	85	80	82.5
120	100	100	85	80	82.5
120	100	200	80	85	82.5
120	100	500	80	80	80
120	100	1000	65	95	80
			90	70	80
120	200	0	70	90	80
			75	85	80
			90	70	80