Boundaries in volatile organic compounds in human breath

by

Matthew A Turner

A Doctoral Thesis

Submitted in partial fulfilment
of the requirements for the award of
Doctor of Philosophy of Loughborough University

Abstract

Exhaled breath is a rich and complex matrix containing many hundreds of compounds. Every breath offers the potential of a non-invasive measurement of the biochemical processes occurring in the human body and it is this notion that has led to the application of breath analysis for the detection of disease. With the majority of research in the field being focused on the detection of biomarkers, little has been presented on how the seemingly homeostatic matrix of breath varies during the course of normal life events. The research in this thesis describes how a subject's emotional state, physical state, and daily activities can alter the composition of exhaled breath.

The collection of exhaled breath samples was carried out using an adaptive breath sampler, which monitored the subject's breath profiles using a pressure transducer. The sampled breath was collected and stored using commercial thermal desorption tubes, containing multiple sorbents. The analysis of the exhaled breath samples was carried out using thermal desorption gas chromatography-mass spectrometry. The resultant chromatographic profiles were aligned through retention indexing which also provided the basis of a referencing system for the breath components. The data was processed through multivariate statistical analysis methods to establish potential markers of physical and emotional stress.

The pilot studies presented here examine the variability associated with exhaled breath. Using three different cohorts of subjects, three studies were conducted into how emotional state, physical stress and bowel evacuation affects the matrix of exhaled breath.

The homeostasis of breath was challenged by the application of a paced auditory serial addition task, which caused an emotional stress response in subjects and which was accompanied by elevated biological functions such as heart rate and blood pressure. Physical exercise was also investigated using an ergometer with breath samples being captured before and after physical exertion. These two

independent studies yielded a panel of potential markers associated with physical and emotional stress that had not been previously reported.

The third study involved the changes observed in the exhaled breath matrix in subjects after bowel evacuation. This study also highlighted the lower intestinal tract as a major contributor and potential origin of volatile organic compounds found in exhaled breath. The combination of how these normal, everyday functions may impact on the results of breath analysis has been evaluated here. As a result, these potential confounding factors, which have been largely unconsidered, highlight the need for rigorous participant pre-sampling protocols to be incorporated into biomarker discovery studies.

Acknowledgments

I would like to firstly acknowledge my supervisor, friend and colleague Prof Paul Thomas who has supported me throughout my research, helping, guiding and giving me the motivational boosts that I have needed from time to time. I would like to further acknowledge my sponsors from Astra Zeneca and Prof Ian Wilson without whom this research wouldn't have been possible. I would also like to thank those researchers that have been involved in specific aspects of the studies in this thesis.

The psychological stress study described in Chapter 4 was part of collaboration between the Centre for Analytical Science, Department of Chemistry, Loughborough University and the School of Sport, Exercise and Health Science, Loughborough University. The Author wishes to acknowledge the following collaborators, A. Malkar, P. Patel, H. J Martin, L. Edwards and S. Bandelow, who contributed to the collection of samples, use of equipment, participant recruitment and development of this collaborative project. Although multiple biological samples were collected in addition to exhaled breath, this thesis restricts the discussion to breath samples

The physiological stress study described in Chapter 5 was part of collaboration between the Centre for Analytical Science, Department of Chemistry, Loughborough University and the School of Sport, Exercise and Health Science, Loughborough University. The Author wishes to acknowledge the following collaborators, A. Malkar, P. Patel, H. J Martin, P. Watson and R. J. Maughan, who contributed to the collection of samples, use of equipment, participant recruitment and development of this collaborative project. Although multiple biological samples were collected in addition to exhaled breath, this thesis restricts the discussion to breath samples.

Finally I would like to thank Vikki Archibald for giving me the courage, strength and support to go back to university to undertake a PhD. I know I have been sometimes mad, sometimes sad, but I think you will agree that I am a different person because of it and that's why this thesis is dedicated to you.

HoBL.

Table of contents

Abstract i
Acknowledgments iii
Table of contents iv
List of tables ix
List of figures x
Glossary of terms xiii
Chapter 1 The problem of variability and the potential of breath 1
1.1 Proposition of the thesis 1
1.2 The potential of breath analysis 1
1.3 Gedanken experiment in breath complexity and biomarkers 2
1.4 Research objectives 7
Chapter 2 Homeostasis and VOCs in the human body 8
2.1 A brief introduction to homeostatic theory 8
2.2 A brief history of breath analysis 9
2.3 Expired breath and metabolites 11
2.3.1 Origins of VOCs in breath 11
2.3.2 Gas exchange in the lungs 12
2.3.3 Metabolic functions of the lungs 16
2.4 Variability of VOCs in human breath 16
2.4.1 Exogenous sources of VOCs 17
2.4.2 Endogenous sources of VOCs 18
2.5 Disease diagnostics 21
2.5.1 Lung cancer 21
2.5.2 Tuberculosis 22
2.6 Sampling exhaled breath 24
2.6.1 Single end tidal breath sampling 24

2.6.2 CO_2 and pressure controlled breath samplers 25
2.6.3 Whole breath collection 26
2.7 Methods of capture and analysis 26
2.7.1 Direct approach 27
2.7.2 Archival methods 28
2.7.3 Adsorbent based collection methods 29
2.8 Ambient air correction in breath analysis 30
2.9 Methods of analysis 32
2.9.1 Thermal desorption 33
2.9.2 Basics of gas chromatography 37
2.9.3 An introduction to quadrupole Ion trap mass spectrometers 40
2.10 Data processing 45
2.10.1 Multivariate analysis 45
2.10.2 Potential pitfalls of multivariate analysis methods when used in conjunction with metabolomic data sets. 50
Chapter 3 Method and systems 53
3.1 Overview 53
3.2 Ethics 53
3.3 Sampler preparation 55
3.3.1 Assembling and disassembling of full-face masks 55
3.3.2 Sterilization and cleaning of the full-face masks 57
3.3.3 Conditioning and evaluation of mixed bed thermal desorption tubes 58
3.4 Sampling exhaled breath 59
3.4.1 Exhaled breath collection 59
3.4.2 Adaptive breath sampler 60
3.5 Breath sample analysis by thermal desorption gas chromatography mass spectrometry (TD-GC-MS) 63

3.5.1 Thermal desorption sampler 63
3.5.2 Gas chromatography mass spectrometry 66
3.6 Method development. 68
3.6.1 Quality control procedures 70
3.7 Data processing 72
3.7.1 The need for data processing 72
3.7.2 Deconvolution of mass spectral data 72
3.7.3 Alignment of data by retention indexing 75
3.7.4 Initial component cataloguing 75
3.7.5 Component integration and breath-matrix formation 75
3.8 Statistical treatment of data 76
Chapter 4 It's all in the mind and sometimes in the breath 79
4.1 An introduction to psychological stress 79
4.2 Study objectives 80
4.3 Overview of the study 80
4.3.1 Volunteer participants 80
4.3.2 Heart-rate and blood-pressure measurements 81
4.3.3 Cognitive function tests 81
4.3.4 Paced Auditory Serial Addition Task (PASAT) 81
4.4 Results and discussion 82
4.4.1 Blood pressure and heart rate 82
4.4.2 PASAT response Score 83
4.4.3 Chromatographic data 84
4.4.4 Multi-variate analysis (MVA) 87
4.4.5 Preliminary identification of PASAT sensitive breath components 90
4.5 Study summary 97

Chapter 5 Changes in VOC profiles of exhaled breath before and after induced physiological stress 99
5.1 Introduction to VOCs and exercise 99
5.2 Study objectives 102
5.3 Study overview 102
5.4 Results 105
5.4.1 Heart rate results 105
5.4.2 Acetone and isoprene results 106
5.4.3 . Multivariate analysis of TD-GC-MS 110
5.4.4 Trends in VOC responses in breath post exercise. 115
5.5 Study summary 119
Chapter 6 The role of the lower intestinal tract in breath variability 120
6.1 Introduction 120
6.2 Study objectives 120
6.3 Overview of study 120
6.4 Results 122
6.4.1 Statistical evaluation of VOC breath profiles observed immediately before and after defecation 122
6.4.2 observations of VOC breath profiles before and after Defecation 126
6.5 Study summary 130
Chapter 7 Discussion, conclusions and future work 133
7.1 Overview of thesis 133
7.1.1 Chapters 1 and 2 133
7.1.2 Chapter 3 133
7.2 Overview of the studies 134
7.2.1 Breath VOCs and the paced auditory serial addition task 134
7.2.2 Changes to breath VOCs from induced physiological stress. 135
7.2.3 Influence of the lower intestinal tract on breath VOC profiles 136

7.3 Future work 140
7.3.1 Validation of potential candidate markers for psychological and physiological stress. 140
7.3.2 Developing future studies based on the current findings 141
7.4 Concluding remarks 141
Chapter 8 References 143
Appendix 1. Generic Breath Sampling Ethics 153
A1.1 Generic participant information sheet 153
Appendix 2. Retention Indexing 160
Appendix 3. Peer Reviewed Publications 173

List of tables

Table 1 Thermal desorption parameters for the Series 1 Unity Thermal Desorption. 64

Table 2 Gas chromatography conditions (Varian 3800) 67

Table 3 Deconvolution parameters 73

Table 4 PASAT response scores of participants during the neutral (PSR_N) and PASAT intervention sessions (PSR_S). 86

Table 5 Identification of PASAT sensitive breath components C1 to C6 with reverse and forward spectral match and de-convoluted spectrum lists. 93

Table 6 Summary of participant height, weight and calculated body mass index (BMI). 104

Table 7 Summary of recorded heart rate for the ten participants throughout each phase of the experiment. 105

Table 8. Candidate exercise response markers. 115

Table 10 VOCs observed to reduce post bowel evacuation 127

Table 11 List of VOCs highlighted in the studies presented in this thesis. 137

List of figures

Figure 1 Deconstructed examples of potential factors that contribute to the complexity of breath profiles. 3

Figure 2 Hypothetical experiment 2 5

Figure 3 Diagram of a normal single expired breath 12

Figure 4 Graphical representation of the diffusion of gas molecules 13

Figure 5 A plot of the relative number of compounds in each class that have been detected in various biological samples. 20

Figure 6 Example of the resultant Langmuir isotherms at different rates of adsorption(K) 35

Figure 7 Cross section of a typical quadrupole ion trap mass analyser 41

Figure 8 Mathieu stability diagram 44

Figure 9 Example of ion ejection at the stability limit. 45

Figure 10 Projection of a score plot from calculated principle components. 48

Figure 11 An overview of the breath analysis work-flow 54

Figure 12 Illustration of a participant wearing a full face mask 56

Figure 13 An image of a thermal desorption tube showing the sampling and conditioning flow directions flow 58

Figure 14 The adsorbent tube conditioning manifold 59

Figure 15 Schematic diagram of the adaptive breath sampler 61

Figure 16 Adsorbent sampler assembly 62

Figure 17 Example of a typical breath profile obtained from the adaptive breath sampler. 63

Figure 18 Overview of the thermal desorption process in the Unity thermal desorption platform. 65

Figure 19 Schematic diagram of TD-GC-MS 66

Figure 20. The deconvolution process. 74

Figure 21 The breath matrix summarising the observations recovered from a breath study. 76

Figure 22 Heart rate and blood pressure responses in PASAT and neutral sessions. 83

Figure 23 Examples of total ion current (TIC) chromatograms 87

Figure 24 S-Plot of variables used for discrimination of stressed and neutral states 89

Figure 25 Unsupervised principle component analysis (PCA) of six stress sensitive breath components identified from PLS-DA 92

Figure 26 Overlaid extracted ion chromatogram (XIC) responses for stressed experimental session (dashed line) and relaxed experimental session 96

Figure 27 Box-whisker summaries of the combined marker scores (CMS) for stressed and neutral participants 98

Figure 28 Bar Chart illustrating the approximate proportion blood of the total cardiac output 99

Figure 29 Experimental timeline of samples in each phase of the experiment 103

Figure 30 Plot of the mean % heart rate change above the initial recorded heart rate 106

Figure 31 Boxplot of observed isoprene responses in Phases A, C and D. 107

Figure 32 Boxplot of observed acetone responses in Phases A, C and D. 108

Figure 33 Boxplot of observed heart rate responses in Phases A, C and D 109

Figure 34 Correlation between average HR and average observed intensity of isoprene (top) and acetone (bottom) signals. 110

Figure 35. Heat maps of the breath data-matrices for air blanks, and sample collected in experimental phases A, C and D. 112

Figure 36 PCA score plot of Phase A and C 113

Figure 37 PCA Score plot of phases A, C and D 114

Figure 38 Box whisker plots of the peak areas observed for VOC breath components 117

Figure 39 Box whisker plots of the peak areas observed for VOC breath components 118

Figure 40 Box whisker plot of the peak areas observed for VOC breath component 118

Figure 41 Schematic diagram of the human cardiovascular system 121

Figure 42 Anatomical drawing showing the veins associated with lower intestine and rectum 122

Figure 43 OPLS-DA Score plot of 10 pre bowel evacuation (grey circles) breath samples and 10 post bowel evacuation (black square) breath samples.....124

Figure 44 PCA Score plot of 10 pre bowel evacuation (grey circles) breath samples and 10 post bowel evacuation (black square) breath samples using unsupervised model generated from the variables identified from the OPLSDA analysis.....125

Figure 45 Example of extracted ion response plotted against calculated retention index of detected analytes in pre (dashed line) and post (solid line) defecation breath samples 128

Figure 46 Example of extracted ion response plotted against calculated retention index of detected analytes in pre (dashed line) and post (solid line) defecation breath samples.....129

Figure 47 Example of extracted ion response plotted against calculated retention index of detected analytes in pre (dashed line) and post (solid line) defecation breath samples.....130

Glossary of terms

APCI-MS	Atmospheric Pressure Chemical Ionization Mass Spectrometry
atm	Units of pressure 1 x Atmosphere is equivalent to 750 mm Hg
AUC	Area under the curve
Base ion	The ion fragment with the most intense signal in a mass spectrum.
block	A set of observations within a statistical matrix
Breath matrix	A table cataloguing all the VOC isolated in a study, organised by retention index and mass spectral fragments by participant.
BTEX	Collective acronym for the chemical compounds Benzene, Toluene, Ethyl benzene and the xylenes
distal	Away from the centre of the body
EESI	Extractive Electrospray ionization
GC-MS	Gas chromatography-mass spectrometry
Gedanken experiment	German term meaning “thought experiment”
MVA	Multi-variate analysis
Negative Feedback	Regulatory process in biology that causes a reaction to stimuli to maintain ideal conditions.
OPLSDA	Orthogonal partial least squares discriminant analysis
Pareto Variance (ParN)	A scaling method takes each variable, subtracts it from the mean of that variable set and divides it by the square root of the standard deviation for the variable set $\frac{(x_n - x_l)}{\sqrt{S_n}}$
PASAT	Paced audio serial addition task, set to induce stress in individuals who undertake this task.
PPE	Personal protective equipment
PSR	PASAT stress response score
PTR-MS	Proton Transfer Reaction Mass Spectrometry
RIU	Retention index unit, a standardised retention scale used to compare data obtained over many analytical runs, see Appendix 2
SIFT-MS	Selective Ion Flow Tube Mass Spectrometry
TD-GC-MS	Thermal desorption gas chromatography mass spectrometry.
TIC	Total ion current. The integrated signal obtained from the ion trap mass spectrometer across the whole m/z range
VOC	Volatile organic compounds.
Volatalome	All volatile organic compounds emanating from the human body
XIC	Extracted ion chromatograms

Chapter 1 The problem of variability and the potential of breath

1.1 Proposition of the thesis

The proposition of this thesis is that breath is more or less constant and it will only change if there is a perturbation, which causes metabolism to change and this subsequently causes the volatalome to change. It follows therefore that breath is susceptible to perturbation from many different sources such as disease, behavioural and environmental factors.

To evaluate these phenomena effectively we will need not only rigorous and harmonised sampling and analysis protocols but also participant protocols. To examine this proposition we need to establish the extent of the variability in breath from emotional, environmental and physical causes. This will then enable the evaluation of how these factors impact on breath protocols.

1.2 The potential of breath analysis

If breath analysis is subject to numerous factors that cause variability the question must arise 'why use it?'

Breath sampling and analysis has a number of advantages over other biological samples. Breath collection is non-invasive and carries a relative low risk of infection. Participants suffer no injury or pain, and are less reluctant to give breath samples than other invasive procedures such as biopsies for research purposes. In addition the collection of breath using the type of archival method described in Chapter 3 requires non-specialist medical training to take samples. The collected samples can be easily stored and transported to a central laboratory for subsequent analysis.

The analysis of VOC in breath has the potential to provide a window into the metabolic processes occurring in the body[1]. They can highlight processes from different compartments or organs and in addition the matrix of breath is less complicated than other biological matrices such as blood[2].

Although breath analysis offers many advantages there are still a number of

unresolved aspects that confound many applications of breath analysis. One of the most pressing issues with breath analysis is that there is currently no harmonised method for the capture and analysis of breath samples. With numerous methods presented in the literature[3–12] each with their own advantages and disadvantages a single harmonised method for collection or analysis may never be possible. A potential answer to this would be a proficiency assessment scheme for breath analysis where many research groups could take part. Further to this the issue that the physiological and biological origins of many VOC's remains elusive[13] and this can potentially add confounding factors into the interpretation of the VOCs in breath profiles. This is further compounded as VOC with exogenous origins such as sampling equipment or environmental sources can be metabolised in the body and excreted through the lungs in the breath[14]. This is in addition to endogenous confounders caused through lifestyle choices or medication. In spite of these issues breath analysis has found applications in the targeted analysis of breath alcohol levels for roadside testing[15], nitric oxide analysis as a marker of inflammation[16] and $^{13}CO_2$ for the determination of helicobacter pylori infections[17].

1.3 Gedanken experiment in breath complexity and biomarkers

Firstly consider what compounds are in breath and how these compounds would group together. If it was possible to deconstruct a breath profile obtained using TD-GC-MS, so that each factor contributing to the complexity of breath could be observed individually and it was then reconstructed one factor at a time. The resulting effect would be that with the addition of each factor the complexity observed would be compounded. This concept is attempted to be conveyed in Figure 1. Chromatogram (A) could be envisioned as how a breath profile might look if it was possible to look at a normal healthy human's baseline metabolism. The second image (B) is what might be observed if environmental factors were then introduced from exogenous sources such as air pollution, personal care products etc. The third example (C) then incorporates VOC compounds resulting from lifestyle factors, such as diet, medication and activity levels. The final chromatogram (D) could then potentially include components resulting from disease. What is apparent is that chromatograms C and D appear very similar, this then raises the question how much would a VOC or panel of VOC biomarkers need to change before they become significant and therefore useful as a diagnostic tool? Also, what factors can

contribute to variations in the profile that could confound this detection?

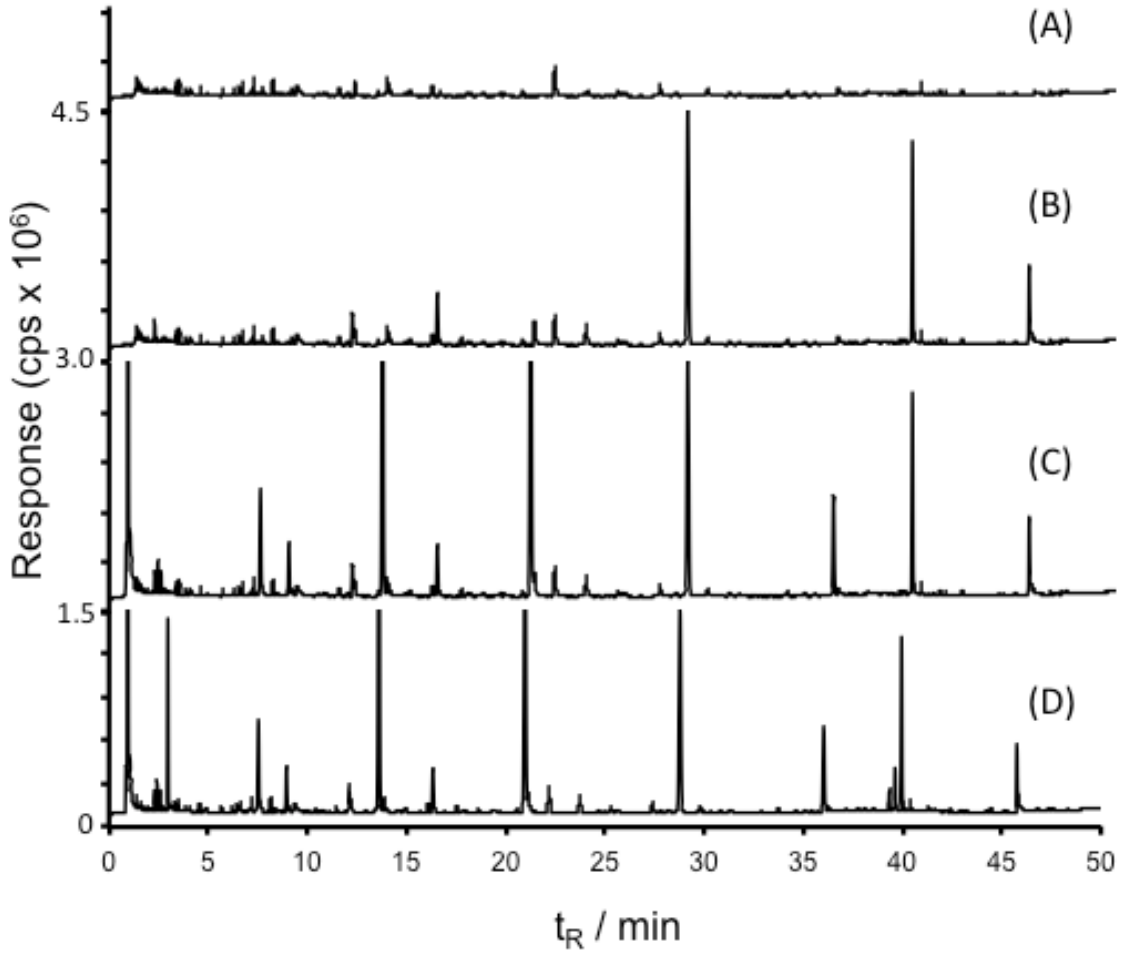


Figure 1 Deconstructed examples of potential factors that contribute to the complexity of breath profiles. (A) Hypothetical example of contributions from baseline metabolism in healthy human. (B) As for A, but with environmental factors superimposed. (C) As for B, but with lifestyle factor superimposed. (D) As for C but with potential biomarkers of disease superimposed.

If we further consider the variation observed in a single hypothetical VOC in breath it is feasible that we could quantify a target VOC in the breath of a single participant

every six hours for a day. This would possibly yield a type of result shown in Figure 2. A. This illustrates that this compound is stable over a short period of time. What would happen if the measurements were then extended not for 1 day but every six hours for one month? Figure 2B is what may be expected. In this example the emergence of diurnal cycles in the measurements starts to become apparent. If we also start to introduce endocrine control cycles on top of the diurnal cycles it is possible to establish statistical upper and lower limits on the normal concentration on this particular VOC in breath. Potentially responses outside these limits may correspond to early stage of a disease before physical symptoms have yet emerged, but more importantly the disease at this early stage may be curable (Figure 2C).

If this experiment is then extended to the timeframe of one year, again continuing to take a measurement every six hours the possible outcome is given in Figure 2D. This graph illustrates how the response potentially could drift due to the development of disease state over the course of a sampling campaign. As the course of the experiment approaches one year multiple measurements reveal that depending on when the sample is collected due to natural variation of the biological system the measurement could be outside of the upper limit 50 % of the time but still not be detected. However this example only considers an increase in a single participant with an illness that has been monitored many times over the course of a year. If these results are then amalgamated with that of a healthy participant the type of outcome is illustrated in Figure 2e. There is not a clear distinction between healthy and disease outside of the original upper and lower limits set during the month long experiment but with this intensive sampling plan the small overall differences observed would be enough to distinguish between the two participants, in the absence of this potential false negatives may be observed. The time and resources required to run such a sampling campaign is currently outside the scope of most research studies, therefore the need to understand what causes the variation in a normal breath profile needs to be addressed.

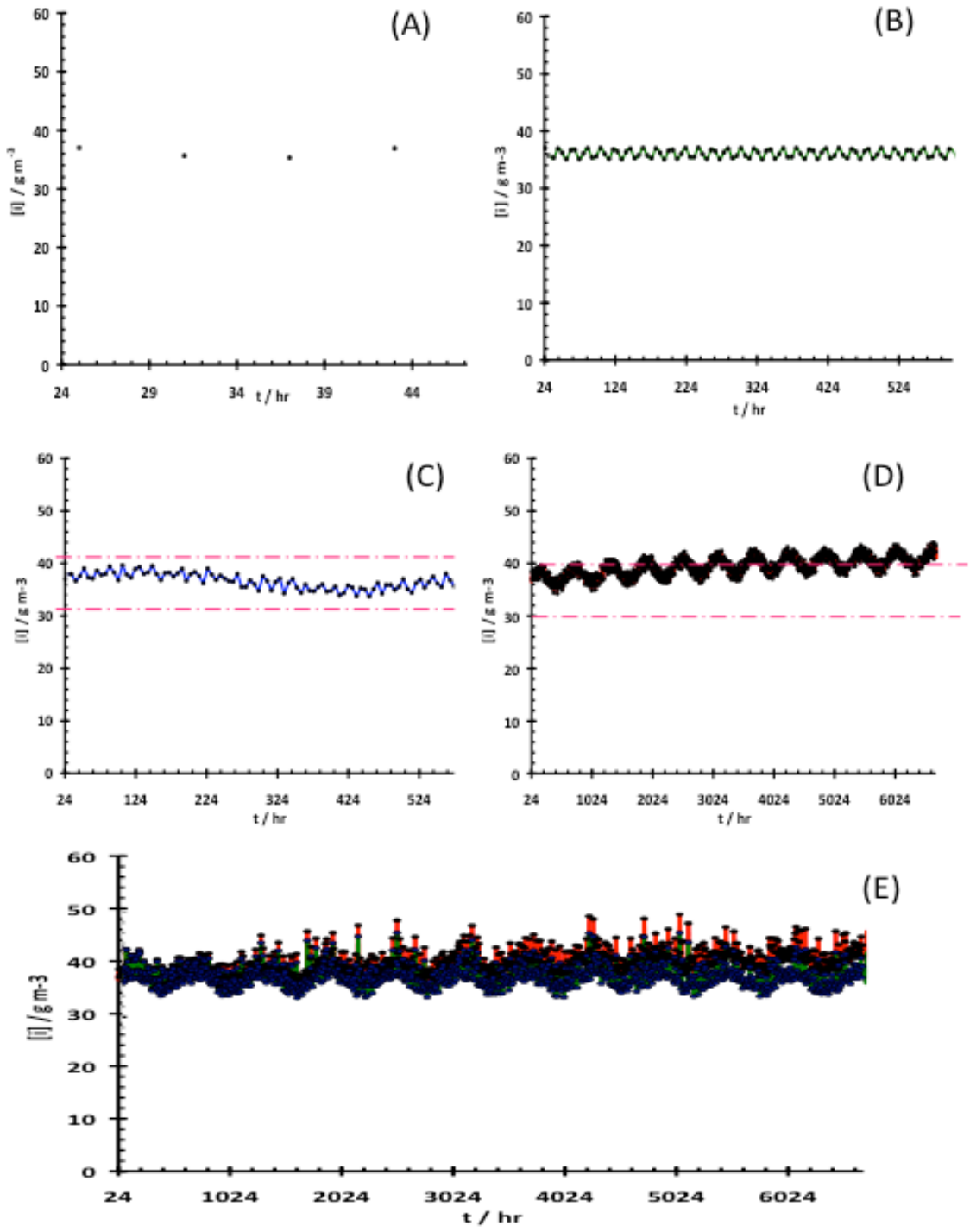


Figure 2 Hypothetical experiment 2 (A) measurement of a target VOC in human breath measured every 6 hours (B) continued measurements of the same target VOC starts to exhibit diurnal cycles. (C) Introduction of endocrine controls allows realistic statistical limits to be applied to the measurements. (D) Continued target VOC measurements for 1 year - a drift from the applied limits start to appear for patient as disease progresses. (E) Overlaid response of healthy and disease patients.

The concepts above illustrate how small changes can be significant and could be masked by the variability observed in the measurement of a particular breath VOC biomarker. Reducing and identifying potential sources of variability in breath is a key part of the analysis and the biomarker discovery process.

As it is not experimentally possible to interrogate the different factors discussed earlier in the absence of all other contributors, what we are left with is potentially in excess of 1000[18] components in a complex matrix. In addition many of the co-dependences and origins of the observed VOC in breath are not fully documented or completely understood[14].

How can this complexity and variation be combated when working with breath to ensure that the results and the changes observed are real and significant?

Moving away from global metabolomics profiling and biomarker discovery, if the wider field of breath analysis is now considered and in particular reflecting on the lessons learnt in the development of nitric oxide in clinical diagnostics for example, the sampling procedure used for nitric oxide has been standardized enabling it to be used as a clinical tool[19]. Consequently the results obtained are comparable and reproducible across different clinical centres.

To reduce variation, is it just a matter of harmonising sampling and analysis procedures?

Undoubtedly this would enable results obtained from different research and clinical centres to be amalgamated more easily. This could then lead to a repository or database of research, which could then be independently verified or utilised. While this would contribute greatly to biomarker discovery in breath, it is not the only factor that should be considered when trying to reduce variation in breath analysis. To fully understand the extent of variation in breath VOC and how to deal with it, the factors that affect breath must first be found.

1.4 Research objectives

The main aim of this research is to test the proposition that standardised participant protocols are needed for breath analysis. The specific experimental objectives were:

- Examine the effect of acute psychological stress on breath analysis.
- Investigate the impact of physiological exercise on a global metabolomics profiling approach of breath analysis.
- To examine how defecation could change breath profiles.

Before the individual experimental studies can be described, it would be helpful to discuss the background, fundamental principles and current literature on VOC in breath analysis. Chapter 2 will review these aspects and will also consider the methodologies current employed in the field of breath analysis.

Chapter 2 Homeostasis and VOCs in the human body

2.1 A brief introduction to homeostatic theory

The term homeostasis was first coined in the 1920's and is attributed to the American physiologist Walter Cannon[20]. Homeostasis refers to the body's natural tendency to maintain a near constant internal environment in the presence of external stimuli. It was Cannon's postulations on a consistent internal environment of the human body, which he envisaged as being in a steady state rather than an absolute constant that drove his research in the area. This steady state varies within limits while two opposing mechanisms are occurring within the internal environment. Cannon and his co-workers demonstrated an example of this though the opposing effects that insulin and adrenaline have on blood sugar[20].

The steady state of a biological system and the human body's ability to maintain this state within limits can be explained in terms of a rudimentary concept. The human body maintains an internal temperature of around 37^{o}C but it is not constant and can change if the body experiences an external stimulus. When exposed to heat, for example on a hot summer's day, receptors in the body will attempt will send an input to a control centre indicating a change has occurred. This will in turn send an output to an effector, which generates a response to address the change. In this example the response causes sweating and delivery of the blood closer to the skin in order to aid cooling. This control mechanism within the homeostatic environment is defined as negative feedback[21,22]. The resultant responses may be measurable (a decrease in body temperature) or leave behind an indicator or marker; (in this case sweat), that this response had occurred.

The ability to maintain the internal environment of a biological system as complex as a human, requires the coordination of all of the organ systems in the body[23]. The organs in turn require the cells that they are composed of to function effectively requiring many different biochemical reactions. The total of all of these reactions is termed metabolism and the chemical entities involved and produced in metabolism are termed metabolites[23].

2.2 A brief history of breath analysis

The concept of breath analysis as a diagnostic tool dates back to Ancient Greece. Hippocrates's treatise on breath aroma and disease, relate how the smell of breath and other bodily fluids can indicate what ailments a patient may be suffering with[24]. What these early physicians were smelling during their olfactory analysis of a patient's breath were scents such as rotten apples[25], stale urine[26], fishy-like or feculent scents[27],which we know now to be due to diabetic ketoacidosis, renal failure and liver failure respectively. These rudimentary aroma tests could be considered as the first detection of biomarkers of disease in breath.

The story of breath analysis is continued nearly two millennia later in 1784 when French chemist Lavoistier correctly identified that oxygen is a reactant consumed during respiration. He further demonstrated that carbon dioxide was a product of respiration and was exhaled in the breath of guinea pigs. Lavoistier then went on to confirm his findings in human volunteers[28]. Nearly a century after Lavoistier's discovery of carbon dioxide the first recorded volatile organic compounds (VOCs) were being discovered in the breath of humans. In 1874, British physician Anstie developed the first diagnostic test for the presence of ethanol in breath[29]. The test was colorimetric and involved bubbling breath through a solution of chromic acid, which changed colour from red/brown to green if an alcohol was present. The application of the detection of ethanol in breath remains the most common and wide spread diagnostic breath test in use to date. In 1897 the next VOC discovered in breath was one of the most abundant endogenous VOCs still often reported today. In Germany Nebelthau was working with diabetics and discovered that acetone was present in human breath[30]. Shortly after Nebelthaus' discovery, Haldane and Priestley described an apparatus designed for the collection of alveolar air. They were able to demonstrate the separation of the different phases of breathing through a rubber tube over a meter in length[31]. The concept of the collection of different fractions of exhaled breath developed by Haldane and Priestley in 1905, still play a role in modern approaches to breath collection.

As the mid 20th century is approached, breath analysis starts to develop the basis of all breath tests to follow. In the early 1950's a breathalyzer developed by Borkenstein for the quantification of ethanol in breath sparks the oldest and most wide spread application of breath analysis, this application also demonstrated the quantitative

relationship between blood and exhaled breath[32]. In the 1960's and 1970's clinical breath tests for gastroenterology start to appear which measure breath hydrogen and radiolabeled CO_2 levels for the diagnosis of carbohydrate malabsorption[33].

Then Linus Pauling's seminal paper entitled; Quantitative analysis of urine vapour and breath by gas-liquid partition chromatography published in 1971[34] reports over 200 components to be present in breath. This ground breaking work used a gas chromatograph equipped with a stainless steel open tubular column to separate the breath matrix. The detection of breath components was performed using a flame ionization detector. Introduction of the breath sample to the chromatograph was accomplished by first trapping 10-15 exhaled breaths per sample onto a 5 ft. coiled stainless steel column, cooled using an isopropyl alcohol ice bath. The sample was then flash heated and transferred to the chromatography column. Pauling's work with breath was to begin the era of modern breath analysis.

Following on from Pauling's work, throughout the 1970's and 1980's breath analysis was being explored for both clinical and environmental applications. Kaji *et al.* examined sulfur-containing VOCs in the breath of patients with liver disease[35] and Wallace and co-workers reported a number of environmental contaminants in the exhaled breath of their volunteers[36]. Wallace's research highlighted that our environment has an effect on the VOC profiles obtained from the breath matrix. In 1985 the first report of breath analysis being used to profile VOCs in lung cancer patients was published[37].

The early 1990's Phillips published his first method for the collection and analysis of breath using TD-GC-MS[38]. This sampling system used adsorbent tubes to trap VOCs from exhaled breath. The system incorporated a charcoal filter that supplied the participant with clean air. The exhaled breath was diverted through a one-way valve and the end phase of breath containing the alveolar fraction was drawn through a water trap to the adsorbent sampling tube. These tubes were then transferred to a GC system where they were subsequently analyzed. In the following 20 years the field of breath analysis has expanded to include the detection of potential suites of biomarkers for breast cancer[39], colorectal cancer[40], asthma[41], COPD[42] and tuberculosis[43]. As the field has expanded so too has the instrumentation used in the analysis of breath. The analytical systems to date have

included but have not been limited to: GC-IMS[44], GC-MS[45] GCxGC-TOF-MS[46], PTR-MS[47], SIFT-MS[48] and sensor systems[10]. As the detection capabilities of these systems has evolved over time, so the amount of detectable breath components has also increased 10 fold from the initial observations made by Pauling almost 50 years ago[34].

2.3 Expired breath and metabolites

Human breath has a distinctive chemical composition that is different from inspired air[49]. The bulk matrix of breath is composed of nitrogen, water vapour, carbon dioxide and oxygen[1]. All these components are present in percentage concentrations. In the small remaining fraction several hundred different volatile organic compounds have been observed[49]. Volatile organic compounds can be defined as any compound that has the element carbon and one or more of the following species; hydrogen, oxygen, nitrogen, halogens, sulfur, phosphorous or silicone. Moreover these compounds also have sufficient vapour pressure ($\geq$0.01 kPa) at a temperature of 293.15 K to volatilize and enter the atmosphere[50]. The VOCs that are associated with breath are able to transverse the alveolar membrane of the lungs, and it is those produced through a range of different metabolic states that are of particular interest in this thesis. However VOCs found in breath do not all originate from in vivo mechanisms, other sources of VOCs in breath profiles can be exogenous and due to the inspiratory air, food and drink, and the uptake of chemicals through inhalation or dermal absorption[38].

2.3.1 ORIGINS OF VOCS IN BREATH

Exhaled breath is a mixture of both “dead space” and alveolar air. Alveolar air is that fraction of breath that has been in contact with blood via the alveoli. “Dead space”[1] breath is made up of the air that has been in contact with the trachea, bronchi and the upper respiratory tract[51]. The mixture of gases that we exhale is complicated and its value as a sample can be largely dependent on where, or rather what phase of, the expired breath is taken[1]. The majority of VOCs reported to have been found in breath are hydrocarbons but also incorporate acids, esters, aldehydes, ketones, furans, ethers, and sulfur-, nitrogen-, and halogen- containing compounds[52]

A single expired breath maybe broken into four parts and is best illustrated by the schematic drawing of a single expired breath shown in Figure 3.

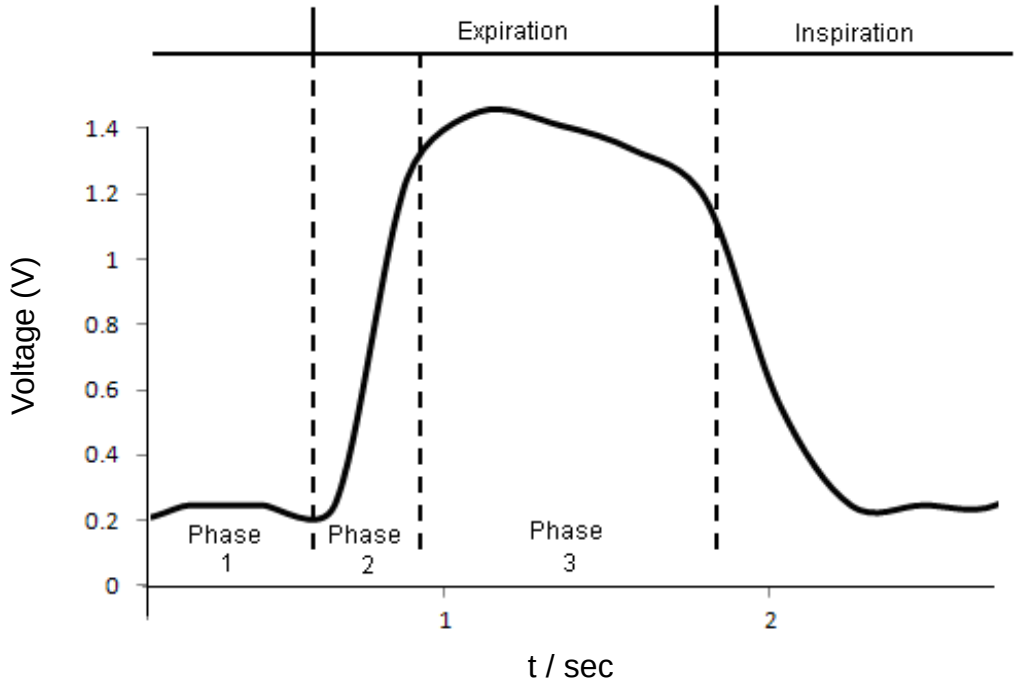


Figure 3 Diagram of a normal single expired breath illustrating the four phases of a single breath. **The response in volts is representative of the pressure response from a pressure transducer from a sampler using pressure to control breath collection.**

From Figure 3, phase 1 represents the start of the expired breath and gas sampled at this point would not contain alveoli breath but would comprise dead space air. Phase 2 shows the steep increase in pressure compared to the baseline, samples taken from this region show a mixture of both dead space air and alveolar air. Phase 3 on the diagram shows a plateau in the pressure and is attributed to alveolar evacuation. Inspiration is the terminal phase of the exhaled breath[1]

2.3.2 GAS EXCHANGE IN THE LUNGS

The exchange of oxygen and carbon dioxide is required for life. The way in which the lungs achieve this can be described by the following concepts.

The concept of diffusion is illustrated in Figure 4 and can be thought of as the movement of molecules from a region of high concentration to a region of low concentration, and maybe described as a concentration of molecules ([M]) that move through an area (A) with a parallel z-axis (z) per unit time (t). This concept was described by Adolf Fick[53] is known as flux (J) and is driven by a decrease in the Gibb's free energy[54]

$$\frac{[\mathrm{M}]}{\mathrm{A.\,t}} = \mathrm{J} = -\mathrm{D}.\frac{\delta \mathrm{N}}{\delta \mathrm{z}}$$

Equation 1

Where: J = Flux (mol/m^2/sec)

D = Diffusion coefficient (m^2/sec)

δN = Change in number density of particles or concentration (mol/m^3)

δz = distanced travelled (m)

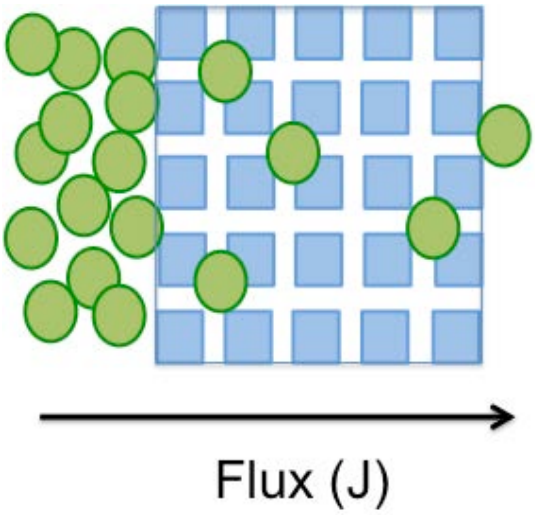


Figure 4 Graphical representation of the diffusion of gas molecules (green circles) from a region of high concentration to a region of low concentration through a membrane (blue squares).

This is Fick's first law and can be used to describe the diffusion of gases across the alveoli membranes[53]. However when dealing with gases and biological fluids the use of concentration is insufficient. Concentration of a particular gas would incorporate all of that gas that is available for diffusion (free) and that which is bound to other chemical entities such as haemoglobin and is not freely available to take part in the

diffusion process. In this case the use of partial pressures is employed. The partial pressure of a gas can be thought of as the pressure exerted by that gas alone on the external walls of a container. As the partial pressure of gas is proportional to its concentration it can therefore be substituted for the concentration term in Equation 2.

$$J \propto \frac{A}{T} D \Delta P$$

Equation 2

The above equation states that the number of molecules per unit time per unit membrane surface area or flux (J) is inversely proportional to the thickness (T) of the alveoli membrane and proportional to the membrane surface area(A), the diffusion constant (D) for a particular compound and the change in partial pressure(ΔP).

If we further consider the diffusion constant (D) we find that it is proportional to the solubility of the compound (SOL) and is inversely proportional to the square root of the compound's molecular weight (MW).

$$D \propto \frac{SOL}{\sqrt{MW}}$$

Equation 3

The difference in the rates of diffusion of carbon dioxide and oxygen through the alveoli membrane can be explained using this equation[53]. Carbon dioxide diffuses much more rapidly (approximately 20 times) than oxygen, which is due to its greater solubility even though the molecular weights are not vastly different.

The solubility of gases in liquids is described by Henry's law, which states that the solubility of a gas in a liquid is dependent on the temperature, the chemical properties of the gas and the liquid, and the partial pressure over the liquid. The relationship of these factors may be summarised in the following equation[23]

$$C = kP$$

Equation 4

Where:

C = The molar concentration of the gas

P = The partial pressure of the gas

k = Henry's law constant at a given temperature

Therefore at a constant temperature the concentration of gas in a liquid is proportional to its partial pressure[23].

In the alveoli the gas pressure differs from that of atmospheric pressure because of the exchange of gases occurring continuously between alveolar air and capillary blood. In addition the air in the alveoli is saturated with water vapour and as the lung is inflated drawing atmospheric air in, mixing occurs between the fresh atmospheric air and the carbon dioxide rich, oxygen deficient alveolar air. As blood enters the pulmonary capillaries, gases in the blood and lungs then diffuse along their respective partial pressure gradients[53,23].

Using oxygen and carbon dioxide as examples, P_{O_2} alveolar = 100 mm Hg and P_{O_2} blood = 40 mm Hg allowing the oxygen to diffuse into the blood whilst P_{CO_2} alveolar = 40 mm Hg and P_{CO_2} blood = 46 mm Hg allowing for the diffusion of carbon dioxide out of the blood[23].

The examples of oxygen and carbon dioxide are the most common, but the principles illustrate how both endogenous and exogenous VOCs can find routes of excretion through the lungs.

2.3.3 METABOLIC FUNCTIONS OF THE LUNGS

The lungs contain a large number of the total vascular endothelial cells found in the human body and are the only other organs apart from the heart that receives all of the blood from the body[23]. It is this attribute that makes the lungs an appropriate site for the modification of blood borne substances[53]. In addition to gas exchange the lungs perform metabolic functions. In particular a number of vasoactive substances are metabolised in the lung. Peptides, amines and arachidonic acid metabolites can undergo deactivation and activation mechanisms in the lung, and peptides such as Angiotensin I and bradykinin are metabolised. Angiotensin I is the only known example of biological activation by passage through the pulmonary system. It is converted by Angiotensin Converting Enzyme (ACE) to the potent vasoconstrictor Angiotensin II which is up to 50 times more active than its precursor and is subsequently unaffected by passage through the lungs. In the case of serotonin (amine) the lungs are the main site of inactivation by means of uptake and storage[53].

2.4 Variability of VOCs in human breath

From the established definition of homoeostasis, it could be assumed that the matrix of exhaled breath is in a constant steady state with little variation occurring in the observed concentration of VOCs. As there is a constant exchange of gases occurring between the blood in the pulmonary capillaries and the alveoli of the lungs, the detection of VOCs in breath must therefore reflect the levels of VOCs in the blood and by extension reflect the metabolic processes occurring in the body. It therefore stands to reason that if these metabolic processes were disrupted, that this would also impact on the concentrations of VOCs in the expired breath. In fact we know this to be true: the most well-known example of exhaled breath analysis is that of the breathalyzer used by police forces in the UK on a daily basis. This breath test indicates the level of ethanol in expired breath and as a result reflects the concentration of ethanol in the blood. Here we have a demonstration of a deviation in homoeostasis occurring through the consumption of alcoholic drinks and as a result the body's normal metabolic pathways becoming saturated and the concentration of ethanol observed in expired breath increasing. The concept described above illustrates how external stimulus disrupts the normal levels of VOCs in expired

breath. As an extension to this concept it is therefore not unreasonable to assume that disease may also cause variations in VOCs in human breath. To date the bulk of the research involving VOC breath analysis has been concentrated in the field of disease diagnostics.

2.4.1 EXOGENOUS SOURCES OF VOCS

There are two main sources of VOC variability that contribute to differences within the profiles observed in humans: exogenous and endogenous. Exogenous sources of VOC exposure can be further classified into personal and social contributions to VOCs, and housing and demographic exposure.

Personal and social contributions to VOCs include factors such as occupational exposure through working with chemicals and proximity to vehicular exhaust systems. Cleaning products used in the home and the use of scented body sprays and perfumes can also add variability[55,56] to an individual's profile. Social activities, for example smoking and consumption of alcohol, also contribute to VOCs found in breath.

Demographic factors can affect the variability within breath profiles. The location of residential housing relative to petrol stations, airports, industrial sites, landfill and incinerators can contribute to the uptake of numerous VOCs and environmental contaminants, which can present themselves in breath profiles as either the parent compound from the initial exposure or its corresponding metabolites[55,56,57]

The types of compounds associated with these two categories of exogenous factors include but are not limited to; Benzene, toluene, ethylbenzene and xylene (BTEX) which are ubiquitous exogenous VOCs associated with gasoline, paints, adhesives and vehicular exhaust emissions. Methyl *tert*-butylether (MTBE) is another VOC that is connected with exposure to gasoline products as it is used as an additive in fuels. Chloroform is a by-product created when water is treated with chlorine: swimming pools and showers are both sources of chloroform found in breath profiles. The use of products such as deodorizers, disinfectants and air fresheners in confined areas contribute to the uptake of 1,4-dichlorobenzene. The six compounds mentioned above are just a few examples of the exogenous VOCs that people are commonly exposed to when carrying out routine daily tasks, for example housekeeping and commuting by car[55,56].

2.4.2 ENDOGENOUS SOURCES OF VOCS

The second source of variability associated with inter-individual VOC breath profiles are endogenous origins. Numerous factors such as diet, age, hormonal state, sex and disease effect endogenous sources of VOCs. The presence of acetone in breath profiles is common and is formed via the decarboxylation of acetoacetate, in addition in vivo studies have indicated possible routes by the dehydrogenation of isopropanol[58]. The concentration of acetone in breath has been seen to be elevated in individuals who have been fasting. In contrast elevated concentrations are also associated with uncontrolled diabetes mellitus[1,59].

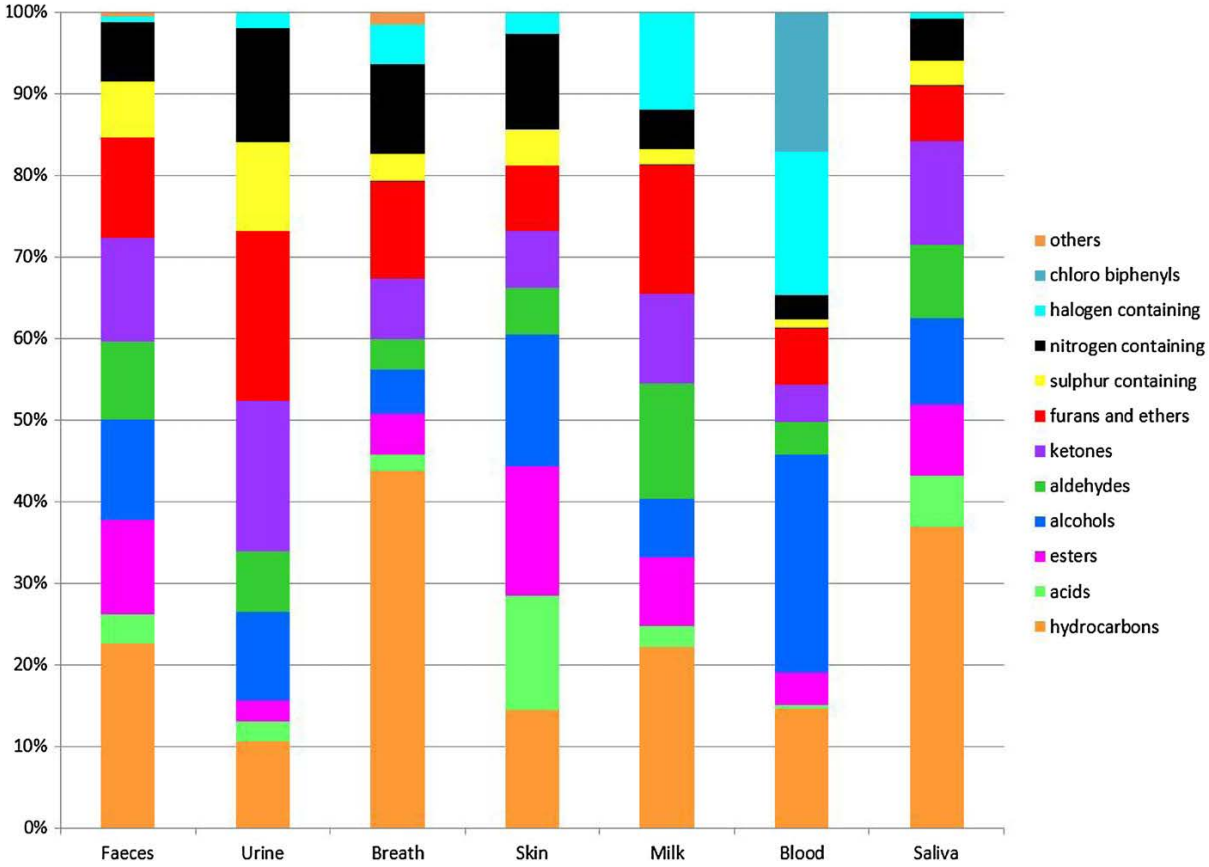
The unsaturated hydrocarbon isoprene is a common constituent of human breath and is formed along the mevalonic pathway during cholesterol synthesis[1]. The parallel decrease of isoprene in exhaled breath when sterol synthesis is impeded through the administration of lovastatin has been documented by Stone *et al*[60] and provides evidence for isoprene production through this biological mechanism. In addition further evidence suggests that oxidative damage to the body and fluid lining of the lungs can contribute to isoprene concentrations in breath[61]. Although persistent in human breath the concentration of isoprene does vary with a circadian rhythm which shows a maximum at 6 am and a minimum at 6 pm[1]. Therefore there may be significant variation in isoprene levels between an individual's breath profile taken on a particular day at a particular time, and a further breath sample taken a few hours or days later. Age is also a factor that contributes to the concentration of isoprene in breath with relatively low concentrations being identified in children's breath[61].

Lipid peroxidation of polyunsaturated fatty acids is a biological mechanism by which saturated hydrocarbons can be formed. Ethane and pentane are known to form in a ratio of 1:4 through this type of chain reaction. However, as pentane is more readily metabolised than ethane by cytochrome P450 enzymes, variations from this ratio can be due to changes in hepatic function in different individuals. Other factors such as increased mental and physical stress contribute to increases in levels of pentane and ethane[62]. Methylated hydrocarbons have also been reported in breath but their metabolic pathways have not been elucidated sufficiently for their diagnostic potential to be realised[1].

Anaerobic fermentation, which occurs through the action of bacteria in the gut gives rise to alcohols from methanol to heptanol[63]. These VOCs can transverse the intestinal barrier between the gut and the blood stream and although hepatic clearance is the primary mechanism alcohols can also find excretion routes through the lungs.

For a biomarker to have a physiological significance and diagnostic impact the biological pathways by which the substance is produced need to be identified. Sulfur-containing compounds can occur when methionine is not completely metabolised during transamination reactions and have been found to be significantly higher in patients with diminished hepatic function[62]. Kidney disorders and stomach ulcers can have an effect on the urea balance of the human body: If ammonia is not converted to urea and excreted in urine then it will be evacuated through the lungs. Patients with kidney failure have been known to have a distinctive breath malodour which can be attributed to increased concentrations of dimethylamine and triethylamine[1].

VOCs are generated in many parts of the body and depending on their origins they could potentially find an excretion route by the lungs. The extensive types and the geography of VOCs in healthy humans have been reviewed by de Lacy Costello *et al*[52]. This collection of VOCs reported from different biological samples (faeces, urine, breath, skin, milk, blood and saliva) show very little cross commonality of VOCs between the samples types (Figure 5). Out of the 872 individual VOCs found in breath it was reported that only 12 were present in all other six sample types[52]. This surprisingly low cross-compartmental correlation between the sample types in this review highlights the human body's ability to effectively and efficiently deal with VOC metabolites and eliminate them from the body. It must be considered that while this is a comprehensive view of what has been identified and reported in the literature the methods of collection and analysis of each of the compartmental samples will have been different. In such cases the overall extraction and suitability for multi-VOC metabolomic profiling methods may need to be brought into question in terms of the specific analyte's recovery or extractability form each matrix / sampling technique.



[#]Figure 5 A plot of the relative number of compounds in each class that have been detected in various biological samples. Figure taken without modification from[52]

[#] De Lacy Costello, B.; Amann, A.; Al-Kateb, H.; Flynn, C.; Filipiak, W.; Khalid, T.; Osborne, D.; Ratcliffe, N. M. *J. Breath Res.* **2014**, *8*, 014001 P4.

2.5 Disease diagnostics

The following section on disease diagnostics is taken from the publication: Thomas, P.; Turner, M. In *Global Metabolic Profiling: Clinical Applications*[64] and highlight two main areas of disease diagnostics where breath analysis has been applied over the last decade.

2.5.1 LUNG CANCER

Lung cancer can affect all areas of the respiratory system from the trachea down through to the alveoli and the surrounding lung tissue. Current clinical diagnostic methods for lung cancer use a combination of:

- Chest X-rays
- Computed Tomography
- Sputum Cytology
- Tissue Biopsy

The lesions have to be large enough to be discernible with these imaging/surgical approaches and the larger the lesion the more advanced the disease and the less likely a cure or extended period of remission. There is a need and demand for more sensitive and enhanced methods for the early detection of lung cancers.

One of the most influential and arguably the lead group in this area are the Menssana group (New Jersey, USA), led by Michael Philips, have published multiple studies on lung cancer. The first set of 22 VOCs that differentiated 60 participants with primary stage lung cancer from 50 healthy controls was published in 1999, with 100% sensitivity and 83% specificity[65]. The next study was larger with a panel of 280 participants with 4 disease classifications (healthy volunteers, participants with no histological evidence of cancer, participants with stage I disease, and participants with metastatic cancer) during which a total of 9 potential biomarkers were identified. This time the detection of disease through breath profiling had a sensitivity of 89 % and 83 % specificity [66]. These two studies employed the collection of exhaled breath on to sorbent tubes and subsequent analysis by thermal desorption gas chromatography mass spectrometry. Multi-variate data modelling based in forward stepwise discriminant analysis was used in both studies and random data selection approaches were used to test the data models. No individual compound was

identified as a discriminator of the disease, methylated hydrocarbons as a class of VOCs were however identified; and these are acknowledged markers of oxidative stress.

2.5.2 TUBERCULOSIS

Tuberculosis (TB) is the second leading cause of death from infectious disease worldwide. WHO reported 1.4 million tuberculosis deaths in 2011[67]. The statistics of infection are challenging: 40 million new infections each year; ca. 10 million new cases of the active disease each year; more than half a million new TB infections each year are resistant to Isoniazid and Rifampin; and strains resistant to practically all anti-TB drugs have been isolated in 55 countries. As a result of inadequate treatment and subsequent transmission there is an estimated global burden of 2 billion people infected with TB worldwide. Current diagnosis is based on the following panel of tests:

- chest x-ray
- sputum culture
- sputum smear
- nucleic acid amplification tests (NAAT)

While culture and microscopy remain the gold standard for detection it is unlikely that the clinical infrastructure will ever develop sufficiently to address the enormous need painted in stark colours by the statistics above. These approaches require aseptic microbiologic facilities and highly skilled technical staff and take days to complete. Such approaches have limited value in many in-community screening settings across the world where TB disease reservoirs have become well-established within the affected populations [68].

This case study also draws on the metabolic approach of the Menssana Research team. In 2007 a preliminary study was presented with 42 suspected cases of TB, 19 of whom would go on to provide positive sputum cultures, and 59 healthy controls[43]. A battery of multi-variate analysis tools were used on the data produced from controlled sampling of end-tidal breath onto adsorbent traps with thermal desorption GC-MS. The study was supported with in-vitro VOC samples obtained from M. tuberculosis cultures. The conclusion was a panel of 7 markers that coded for patients with suspected TB. The next study by the same team, published in 2010[69]

was larger with 226 participants sampled at four clinical centres: 38 at the University of California San Diego, California; 100 from The University of Santo Tomas, Manila, Philippines, and 66 from, De La Salle University Hospital also in the Philippines; finally 22 from The East London Tuberculosis Service, UK. The data processing approach had developed to use 900 8 s data bins with 4 s overlaps and the differences in VOC concentrations in the expired air and concordant environmental samples, termed alveolar gradients, in each of the data bins were ranked in terms of the resultant ROC curve. Monte Carlo simulations were employed to test the outcomes of the data modelling against randomly assigned diagnostic states of the participants. The participants were assessed by sputum culture, sputum smear, and chest x-ray. The combination of these tests was found to have area under the ROC curve of 0.65 vs. a value of 0.85 for the breath tests. In this study the candidate panel of markers was revised once again. Inflammatory oxidative stress was cited as one of the sources of these markers, as well as the infectious organism. The difficulty in establishing consistent marker identities in these studies may be attributed to several reasons. Different co-eluting contamination profiles introducing artefacts into the mass spectra, and the continued enhancement in the fidelity of the analytical systems are two prominent reasons. Consequently the use of library matches from mass spectrometry databases for compound identification without confirmatory standards is highlighted as a potentially important gap in assigning identities to candidate biomarkers. Nevertheless a double blind metabolic approach was found to diagnose TB faster and as reliably, if not more so, than the current accepted clinical protocols.

One of the most recent developments in this study was the translation from thermal desorption GC-MS to a point-of-care device; a thermal desorption GC-surface acoustic sensor based system that provided a diagnostic test in 6 min. This was achieved through the continued use of rigorous Monte Carlo based tests of the validity of the data drawn from multi-national and multi-centre studies with 251 participants (130 with active TB)[70].

These three papers addressing a global health challenge have probably not definitively described the VOCs associated with TB, however they have defined a bench-mark work-flow and process associated with a metabolic profiling approach to therapeutic marker discovery, and its translation to point of care diagnostics. The

author list in these publications confirms the statements by Risby of the importance of substantial multi-disciplinary teams and critical mass in breath research.[71]

2.6 Sampling exhaled breath

2.6.1 SINGLE END TIDAL BREATH SAMPLING

End tidal breath sampling is based on the principle that an adult expires approximately 4 litres of breath when exhaling deeply. The first fraction of the expired breath is made up of the dead space air, which has not been involved in gas exchange in the lungs. The middle fraction of breath is a mixture of both dead space and alveolar air. The final component of the single breath is the undiluted alveolar faction, which constitutes the last 1 to 2 litres of the breath[72]. To sample this portion of a single breath a syringe-based sampler is the most simple and convenient method. An advantage of this approach is that it requires little in the way of services such as compressed air, pumps or computer assisted sampling valves. Markes International produce a commercially available end tidal breath sampler called the Bio VOC™. This particular sampler has found applications in workplace VOC monitoring of benzene, xylenes and toluene (BTEX)[72] and has been found to effective for capture and analysis of over 20 VOCs metabolites[73].

The syringe style sampler incorporates a one way valve mechanism on the plunger[73]. The participant providing the sample exhales through a disposable cardboard mouthpiece into the barrel of the syringe until they can continue to exhale no further. Utilising this method the undiluted end fraction of alveolar breath passes through the syringe barrel. The disposable mouthpiece can then be removed and a plunger is connected to the syringe. Breath samples can then be transferred to an appropriate storage media or taken for direct analysis.

The disadvantage of single breath sampling is that only a small amount of alveolar breath is collected for analysis and thus applications are limited to measuring major constituents of breath or targeted analyte methods. An alternative approach would be to employ more sensitive detection systems however this incurs greater cost and makes analysis less viable if one day breath analysis is to be used in a routine clinical environment.

2.6.2 CO_2 AND PRESSURE CONTROLLED BREATH SAMPLERS

CO_2 and pressure controlled breath samplers employ a CO_2 sensor or pressure transducer which are used to monitor the breath profile obtained as a subject expires. The samplers use microprocessor controlled valves or pumps to collect a subject's breath at appropriate points during exhalation[11,14] The advantage of visualising a breath profile is that it allows the whole breath or an isolated particular phase of expiration to be collected (see Figure 3). The schematics of these systems vary from instrument to instrument but the overall principle of operation remains the same.

The main differences between systems occur in the techniques used for VOC capture and the inlet mechanism for the subject's breath. Miekisch *et al* described a CO_2 controlled system that employed a syringe to sample the alveolar faction of breath[74] This system is also able to incorporate a Tedlar ® bag for breath capture at the same sampling point. Other systems like those described by Basanta[11] and Phillips[75] employ thermal desorption tubes at the sampling point for pre-concentration of breath samples.

The second main difference between breath sampling apparatuses is the inlet system. Depending on the application the user has the option of supplying participants with filtered clean air or allowing them to breathe air from the immediate environment[74]. The use of a purified air supply minimises contributions from exogenous VOCs and provides a consistent field blank[11]. A further adaptation is the incorporation of a mask that is fitted over a subjects face and nose; this replaces the conventional mouthpiece and nose clip[74]. The mask provides an alternative for those studies that investigate lung diseases whose volunteers would be expected to have diminished lung function and may find it difficult or uncomfortable to breathe through a mouthpiece into a tube[11]. Whichever inlet system is used in the breath sampler apparatus, it has to be changed between volunteers. The tube inlet would require removal, disposal and replacement of a single use mouthpiece. When masks are employed they have to be cleaned, sterilized and purged of any VOCs from the cleaning process. The number of masks therefore becomes a rate-limiting step in the number of volunteers that can be sampled at any one time. A disadvantage to using this type of sampler is the service requirements, (compressed air supply, computers

for controlling the sampler and power supply) which make them more complex and less portable than other sampler systems.

2.6.3 WHOLE BREATH COLLECTION

Whole breath collection is the simplest and most straightforward approach to breath sampling. This technique involves the collection of breath into pre-cleaned, nitrogen purged Tedlar® style sample bags. There is very little required in the way of hardware, a simple breathing tube is all that is needed to fill the Tedlar® bag with expired air. Exhaled breath is collected directly with no distinction between the phases of breath and both dead space and alveolar air are collected simultaneously[76].

After collection the sample can be processed for analysis either through a gas injector system[34,77] or subjected to pre-concentration techniques such as solid phase micro extraction (SPME)[78,79] or cold trapping[34,77,38]. While this sampling mechanism requires no services, e.g. compressed air, there can be issues with condensation build up in the bags, and between samples the bags have to be purged and cleaned of any residual VOCs[76]. The drawback to sampling in this manner is the dilution effects occurring with the dead space air in respect to the alveolar air. As a result the responses obtained for endogenous VOCs in breath profiles when using this method of sampling have been reported to be two to three times less intense when compared to those samples taken from the alveolar fraction of breath alone[1]. Moreover, dead space air also has the largest concentration of contamination from exogenous sources of VOCs and therefore real care is required to ensure that a definite distinction is made between background contaminants and endogenous VOCs[1].

2.7 Methods of capture and analysis

There are two main approaches for the capture of exhaled breath for analysis, these maybe termed direct and archival. Within each of these two approaches there are a number of different techniques used to solve the challenge of producing robust reproducible measurements of exhaled breath.

2.7.1 DIRECT APPROACH

This approach is the simpler of the two sampling systems, where samples of exhaled breath are introduced directly into the instrumentation performing the analytical measurement. This has been achieved by incorporating a transfer-line, which participants blow through and the exhaled breath is transferred to the instrument where it undergoes analysis. This approach is generally used in applications where breath-by-breath and real-time analysis is performed. Applications of direct analysis of breath have been used in real-time, in-vivo monitoring and pharmacokinetics of valproic acid[80] and the real-time monitoring of fatty acids in breath[59,81–84]. The analytical systems have included extractive electrospray ionization-mass spectrometry (EESI-MS), atmospheric pressure chemical ionization-mass spectrometry (APCI-MS), Proton transfer reaction- mass spectrometry (PTR-MS) and metal oxide sensors, this is by no means an exhaustive list of the analytical instrumentation used in the direct analysis of breath and is just an example of the variety of techniques that have been used.

The direct approach has the advantage that real time data can be collected and monitored. Breath by breath analysis gives researchers the opportunity to examine changes in breath that can occur rapidly, such as isoprene concentrations during exercise[85], whereas collection methods would only offer the average over the multiple breaths collected. However there are two factors that affect the analysis of breath VOCs by the direct approach which are evident from the literature. The first is that sensitive instruments are needed because numerous analytes are present in the breath in low concentrations, whereas collection methods offer considerable enrichment of the exhaled breath samples. The second affliction is the requirement for the analytes to undergo protonization in order to be detected and identified. The direct detection methods are based around soft ionization mass spectrometric methods or sensors. PTR-MS, EESI-MS and APCI-MS all rely on breath analytes being able to undergo protonization. In PTR-MS for example as the proton transfer reaction requires the analyte in question to have a higher proton affinity than water to be successfully ionized and thus detected. In addition these ionization mechanisms can suffer from competitive ionization and ion suppression. These confounding factors should be considered as breath potentially contains upwards of 1000 individual components[86] that have broad concentration ranges (ppmv to pptv) with

little to no separation prior to the matrix of breath entering the ionization source, so there is a potential for matrix suppression. This particular confounding factor has yet to be addressed in the literature as it pertains to breath analysis and remains a known unknown. However all these detection systems have the advantage of incorporating a mass analyser which provides details on the mass to charge ratio (m/z) and the potential structure of an analyte. With the advent of time of flight systems the ability to discriminate analytes using accurate mass can also provide information on an analyte's elemental composition. In contrast the sensor systems are less discriminating, they can be tuned to specific analytes or functional groups[87] but for global profiling where changes may be seen between control and test groups, identification of analytes can be problematical. The sensor systems do have an advantage in their ability to be portable and more cost effective when compared to the mass spectrometer based instrumentation.

2.7.2 ARCHIVAL METHODS

Collection of exhaled breath has yielded a number of different approaches and the archival methods can be further subdivided into gas and adsorbent collection methods.

Gas collection methods

The two main gas collection methods used for breath are Tedlar® bags[88] and canisters[89].

Canisters are stainless steel cylinders. They require evacuation before sampling and draw in the sampled air or breath through a valve system. The inside surfaces of the canisters can be treated or passivated to provide a reasonably inert surface to minimize the interactions between the sampled VOCs and the inner walls of the canister. There can still remain some loss of particular classes of volatiles, for instance aldehydes and terpenes, after collection. The canister method is not a concentration method and may still require subsequent concentration of the collected sample before analysis to achieve sufficient limits of detection to study trace VOCs in the captured sample. The main advantage of canister sampling is that it allows for multiple replicate analyses of single collected samples. The down side of this technique is that it is costly and bulky making transport more problematic than sorbent based sampling.

Tedlar® bags are one of the simplest approaches to the collection of VOCs from breath. They have been applied in the capture of breath in a number of studies. The bags are constructed from sheets of a polyvinylfluoride polymer that are heat sealed. They are re-useable and are available in a variety of different sizes they are less expensive than the canisters but are very fragile in comparison and are at risk of punctures or tears. This method is also a non-concentration approach to collection of breath that can then be analysed directly or concentrated prior to analysis. Tedlar® bags require extensive conditioning prior to use as carry over can be an issue with this technique. Post sampling the bags suffer from losses through permeation and absorption/adsorption. One main drawback is that they are not suitable for use with humid sample matrices as recovery of VOCs have been shown to decline in such samples.

2.7.3 ADSORBENT BASED COLLECTION METHODS

Solid phase micro extraction (SPME) is a technique that was originally used for headspace analysis and liquid sample extractions. The SPME fibre is bonded to a stainless steel plunger and is a simple adsorption/desorption sampling technique. The fibre can be coated with a range of sorbent materials that are often either non-polar or highly polar. Typical sorbents used in SPME are polydimethylsiloxane (PDMS), polyacrylate or Carbowax. A divinylbenzene polymer may also be incorporated with the sorbent material, increasing the surface area of the sorbent to aid extraction efficiency. The fibres are usually 1 cm long and are very delicate. They are housed in a hollow needle, into which they are retracted when not being exposed to sample matrix. They have been used in the collection of breath directly from participants and as a pre-concentration technique for Tedlar® bag samples[79]. The sampling procedure requires the fibre to be exposed to the sample, usually for a fixed length of time until the fibre's sorbent reaches equilibrium with the analytes in the sample matrix. The fibre can then be transferred to a standard GC injection port and be thermally desorbed directly without the need for any additional instrumentation. While the fibres have been demonstrated to be applicable to breath analysis the analyte capture range is still limited due to limited number of sorbents bonded to the fibres that are commercially available and the small amount of sorbent

present on the fibre. Although re-useable, SPME fibres are relatively expensive compared to other sorbent capture methods like adsorbent tubes. Also, commercial fibres are designed for analysis directly after capture making them unsuitable for storage and transport of breath samples.

Adsorbent tubes and needle traps are different variants of the same methodology. The needle traps use a standard 22 gauge needle packed with sorbent material[90]. The adsorbent traps are much larger in comparison to the needle trap and can be constructed out of stainless steel or borosilicate glass. A standard adsorbent trap used in commercial thermal desorption systems are approximately 89mm in length with an internal diameter of 5 mm and have gastight sealing caps used for storage pre and post sampling. The combination of multiple sorbents within in the same tube enhances the range of analytes that can be captured for analysis. As with the SPME fibres the efficacy of the sorbent to capture VOCs is dependent on the analytes of interest and their intermolecular interactions with the sorbents in the tube or needle trap. The advantage of using this type of capture mechanism is that samples can be captured away from the laboratory environment and be transported with relative ease compared to the canisters and Tedlar® bags. Tubes and needle traps are robust[11,90] and providing they are sealed appropriately the chance of contamination or damage during transport is minimal. Adsorbent tubes have an additional advantage over the needle traps as the sorbent bed can be packed with much larger amounts of sorbents making sample breakthrough less likely and increases the amount of sample enrichment that can be achieved. While both sampling systems are re-useable the adsorbent tubes do require additional instrumentation for desorption of the sample analytes into the analytical system where as the needle traps utilize a standard GC injection port in combination with an autosampler add on for the desorption of the captured VOCs.

2.8 Ambient air correction in breath analysis

The composition and concentrations of VOCs in ambient air is a potential issue that requires consideration in the field of breath analysis. There have been numerous approaches, which have sought to deal with the potential confounding VOCs that are introduced from ambient air. One of the most common methods used is that of ambient air subtraction. Philips *et al*, described one such method which utilizes the technique of alveolar gradients[14]. In this approach those VOC responses (peak

areas of chromatographic responses) recorded from components detected in ambient air are subtracted directly from VOC responses detected in exhaled breath. The result is that the VOCs detected have either a positive or a negative alveolar gradient. Those VOCs that have a negative gradient are considered to be of exogenous origin while those with positive gradients are thought to be from endogenous origins. While this method is simple and logical, it has been shown that is does not reliably account for uptake of background VOCs in the air that the participant breathes at concentrations higher than 5 % of that found in breath[74,91].

An alternative to attempting to compensate ambient air is to provide clean purified air to the participants. The purified air supplied to participants can be monitored and controlled to give a higher degree of consistency in the composition and concentrations of VOCs between participants[11]. Studies reported in the literature have suggested that participants require as little as 4 minutes for a lung washout using purified air[92]. This approach may reduce the overall intensity of exogenous VOCs found in a breath sample but a total body washout out of exogenous VOCs may take hours or days to achieve[92,93]. While not an ideal solution to the problem of inhaled exogenous VOCs, it does offer potential for the reduction of inhaled VOCs and therefore where exogenous VOCs are reduced to less than 5% of those found in breath it should follow that subtracting VOCs from purified air after a lung washout could give more reliable results.

An additional factor that further complicates and confounds both these methods is the potential of a pre-exposure to VOCs that are ubiquitous in a participant's everyday environment. One such example may be found in the compound limonene. Limonene is a compound that has many potential sources which participants could be exposed to, for example limonene is found in household cleaning products and air fresheners, and therefore exposure from indoor air is one possible route. A second possible route of exposure is that limonene is used in both toothpastes and skincare / beauty products which expose the mouth directly and provide exposure through trans-dermal uptake respectively. Finally limonene is a natural occurring compound that is found in citrus fruits, herbs and spices, and fruit juices, so there is a high likelihood of exposure through ingestion[57]. As such subtraction of limonene from

background air measurements is likely to be insufficient to identify this VOC as being exogenous in origin based on the alveolar gradient method.

What is apparent from the literature is that there is this is a known problem with breath analysis and the subsequent identification of potential biomarkers. As there is currently no consensus for a standard method that deals with the elimination of confounding factors from ambient air, and in the absence of reliable toxicokinetic models for the uptake and washout of VOCs in ambient air, the subtraction of VOCs from breath must be treated with caution[92].

2.9 Methods of analysis

The detection systems used in the analysis of breath can be broken down into those instruments capable of direct analysis and those which offer off-line analysis. There are numerous systems that offer direct analysis of breath. These systems include PTR- MS[94–97], SIFT- MS[9,48,98], sensor-based[87,99] systems, EESI[100,84,101] and APCI-MS[100,102]. In addition to being able to be used for direct analysis many of these systems may also be used in conjunction with archival sampling systems and pre-concentration techniques for off-line analysis. While suitable for direct analysis of breath many of these systems, especially mass spectrometer based systems, are movable but far from portable. In addition these systems are expensive and are far from the stage where non-specialists may be able to use and maintain them effectively. Those instruments such as sensor-based systems that can be considered portable lack in their ability to detect the multitude of analytes found in breath, being limited to one (Breathalyzer) or only a few analytes or functional groups, compared to the mass spectrometer systems.

One of the most widely used and long established techniques for off-line analysis of VOCs in breath is the use of archival capture methods coupled with gas chromatography as a separation technique. This system of analysis was first demonstrated for breath analysis by Pauling in 1971[34], where a flame ionization detector was used and 250 breath components were observed. With modern GC instruments, mass spectrometers are commonly used for the detection of breath VOCs. While not portable or inexpensive this approach provides separation and detection of the largest range of VOCs in breath. Philips *et al* [86] used a GCxGC-MS system to analyse exhaled breath of 34 normal healthy participants resulting in the

identification of approximately 2000 individual VOC analytes and thus highlighting the technique as ideal platform for metabolomics style analysis of VOCs in breath.

2.9.1 THERMAL DESORPTION

The technique of thermal desorption for the introduction of VOCs to an analytical system such as a gas chromatograph utilizes archival methods of collection (previously described.) A description of the instrumentation and the thermal desorption process that pertains to the instrument platform used in this research are described in Chapter 3.5.

For thermal desorption to be used successfully as an introduction system there are two physical processes that need to occur.

- Adsorption
- Desorption

Adsorption

Adsorption is the process of adhesion of molecules onto the surface of an adsorbent. In the case of breath analysis we are concerned with the adsorption of molecules in the gas phase depositing onto the surface of a solid adsorbent[54].

The processes that are involved with molecules attaching themselves to the surface are physisorption and chemisorption. Physisorption describes the adsorption of molecules by Van der Waals interactions such as dipolar interactions. These types of interaction are weak but are long ranging. Chemisorption in contrast refers to molecules that adhere to the surface by forming chemical bonds such as covalent bonds. The distance between the surface and the molecule is much shorter when chemisorption is involved than when physisorption processes are involved[54].

The extent of adsorption of molecules on a sorbent is known as surface coverage (θ) and is often expressed as a fraction of the number of adsorption site occupied (X_O) over the number of adsorption sites available(X_A).

$$\theta = \frac{X_O}{X_A}$$

Equation 5

The rate of adsorption is the rate of surface coverage per unit time. The variation of θ with pressure at a given temperature is termed the adsorption isotherm.

If we consider the simplest isotherm where adsorption of a molecule on the surface would not proceed past a monolayer and that all adsorption sites were equivalent, and finally if we assume that the ability of a molecule to adsorb at a given site is independent of the occupation of any neighbouring sites then at a dynamic equilibrium: -

$$\theta = \frac{Kp}{1 + Kp}$$

Equation 6

This is the Langmuir isotherm. Where p is the partial pressure of the gas phase molecule and K is the rate of adsorption divided by the rate of desorption. Figure 6 illustrates the shapes of the Langmuir Isotherms at different rates of adsorption. The different shapes of the curves illustrate the different rates, K, observed at different temperatures, pressures or concentration of adsorbate. In Figure 6 the surface coverage increases with increasing concentration of adsorbate and only at very higher concentrations does the coverage approach 1, where almost all available sites for a molecule to adsorb on to the surface have been occupied.

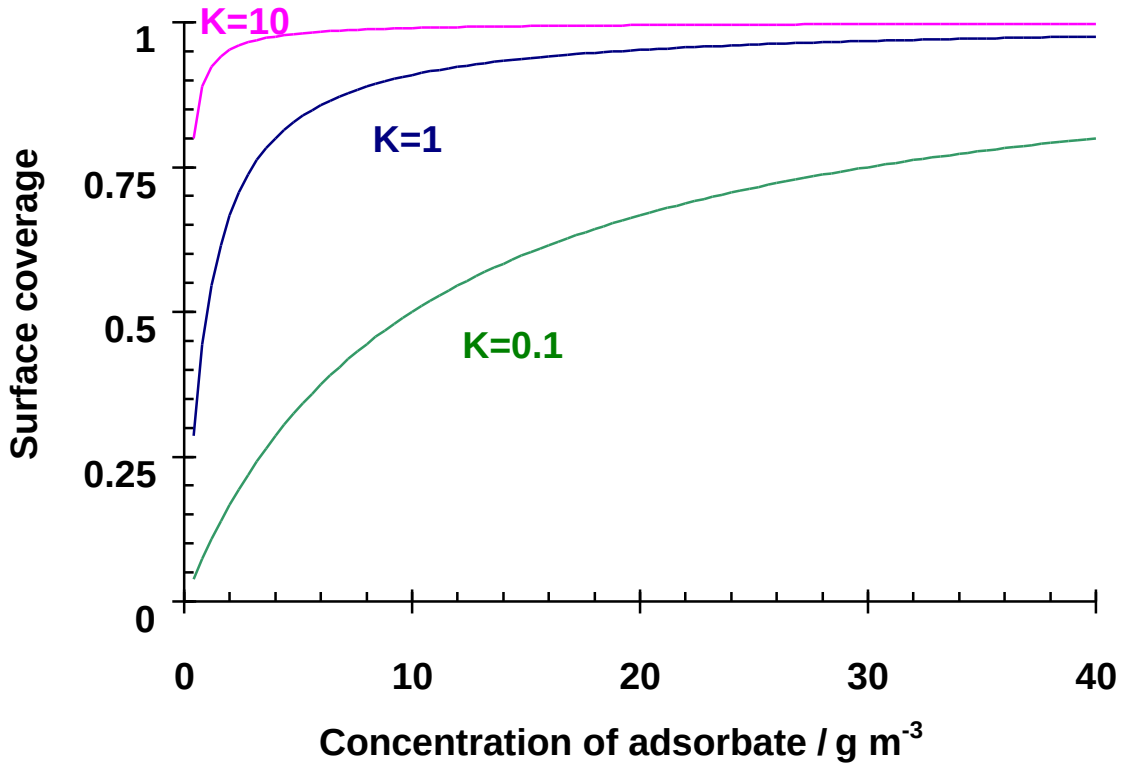


Figure 6 Example of the resultant Langmuir isotherms at different rates of adsorption(K)

Desorption is the process of departure of a molecule from the surface of an adsorbent where it was previously adsorbed. This process always has to be activated and is temperature dependant. The liberation of an adsorbate molecule from the surface of an adsorbent adopts a first order rate which can be described through an inversion of the Arrhenius equation.

$$k_d = Ae^{-E_d/RT}$$

Equation 7

Where: k_d is the rate constant of desorption

A is the pre-exponential factor

E_d is the activation energy

R is the gas constant

T is absolute temperature in degrees Kelvin

Therefore the mean lifetime ($\bar{\tau}$) of an adsorbed species may be expressed as follows:

$$\bar{\tau} \approx \tau_0 e\left(\frac{E_d}{RT}\right)$$

at *STP* $\bar{\tau} \approx 10^{-8} s$

$$\tau_0 = A^{-1}$$

$$\bar{\tau} = k_d^{-1}$$

Equation 8

Where: k_d is the rate constant of desorption

A is the pre-exponential factor

E_d is the activation energy

R is the gas constant

T is absolute temperature in degrees Kelvin

τ_0 is the period of vibration of a bond between the adsorbed molecule and adsorbent

$\bar{\tau}$ is the mean lifetime of an adsorbed species

On an adsorbent where many adsorbates or VOC molecules are present there will be a number of *Ed* that need to overcome to successfully recover the all of the adsorbed molecules present on the surface. The choice of sorbent in thermal desorption traps and tubes should be considered carefully to achieve the highest recoveries possible. In addition the storage conditions and length of storage should also be taken into account as the resonance time of a molecule on the surface of a sorbent is dependent on temperature.

2.9.2 BASICS OF GAS CHROMATOGRAPHY

The following section intends to give a basic overview of gas chromatography and is not intended as a fully comprehensive guide to the technique as it is already a long established analytical technique.

Chromatography is a physical separation process wherein a mixture of components is distributed between two phases. One phase is the stationary phase, which remains immobilised and the second is the mobile phase that moves in a definite direction. In the technique of gas chromatography the stationary phase can be a solid or a liquid, while the mobile phase is a gas (usually helium, nitrogen or hydrogen). GC is also an elution chromatography technique and as such the mobile phase is continually being passed along the chromatographic bed. Most modern gas chromatographs employ capillary columns which are made from fused silica and are open tubes, where a liquid stationary phase is coated on the inside walls of the column. These types of column vary in size and length (0.1mm - 0.53 mm ID and 15m to 100m) and have an array of stationary phase coatings that can be applied to the internal walls of the column at different thicknesses. The column is housed in the chromatograph inside a temperature-controlled oven, where the temperature can be kept constant (isothermal) or be varied throughout the method cycle (programmed temperature). The latter can be very useful when dealing with a mixture of compounds that have different boiling ranges[103,104].

The rate at which an analyte traverses the chromatography column is dependent on the thermodynamic equilibrium constant called the distribution coefficient Kc. For a given analyte (X), Kc is the ratio of the concentration of the analyte in the mobile phase verses that in the stationary phase Equation 9. From Equation 9 it can be seen that the higher the value of Kc the greater the concentration of analyte in the stationary phase. This leads to a longer the retention time of the analyte on the column. This interaction with the stationary phase and thus the retention time is dependent on three factors: the analyte, the stationary phase and temperature of the column.

$$K_c = \frac{[X]_S}{[X]_m}$$

Equation 9

As the column temperature increases the distribution coefficients decrease and thus the retention time of an analyte is reduced, this dependency is expressed in terms of the Clausius-Clapeyron equation[103,104].

$$\log P^0 = -\frac{\Delta H}{2.3\,RT} + constant$$

Equation 10

Where delta H is the enthalpy of vaporization at absolute temperature T. R is the gas constant and P^o is the compound's vapour pressure at T.

Although the term equilibrium constant is used there is in fact no static equilibrium and the movement of the analyte through the column may be thought of as a number of equilibrium steps or plates that occur through the length of the column. The number of these plates describes the efficiency of the column and may be calculated from the Equation 11.

$$N = \left(\frac{t_R}{\sigma}\right)^2 = 16\left(\frac{t_R}{W_b}\right)^2 = 5.54\left(\frac{t_R}{W_h}\right)^2$$

Equation 11

Where:*N* is the column efficiency or plate number

t_R is the retention of the analyte peak.

σ is the is the width associated with the Gaussian chromatographic peak at either σ 2.354 (W_h) or σ 4 (W_b)

W_b is the width at the base of the chromatographic peak

W_h is the width at half height of the chromatographic peak.

As an analyte moves through these equilibration steps a phenomenon known as band broadening occurs. The causes of band broadening that are inherent to chromatographic systems are explained by three effects (A) eddy diffusion, (B) longitudinal molecular diffusion and (C_s) mass transfer in the stationary phase. When band broadening is expressed in terms of plate height (H) as a function of the linear velocity of the mobile phase ($\bar{\mu}$) this is the van Deemter equation.

$$H = A + \frac{B}{\bar{\mu}} + C_s\bar{\mu}$$

Equation 12

However the van Deemter equation was originally envisaged for the band broadening effects in packed columns and as such when dealing with open tubular columns the A term of the original equation becomes redundant. Golay proposed a modified version of van Deemter equation in which the mass transfer term is further modified and expanded to account for the gas phase process in capillary columns. An additional mass transfer term was added to account for the mass transfer in the mobile phase (C_m).

$$H = \frac{B}{\bar{\mu}} + (C_s + C_m)\bar{\mu}$$

Equation 13

As a result when utilizing capillary columns in gas chromatography to achieve the maximum column, and in turn the highest resolution between analytes, one must seek to minimise the B and C terms. However this ideal is often not practical as it can lead to very long columns with lengthy run times and so in practice there often remains a degree of less than ideal resolution between analytes in highly complex mixtures such as breath. One way of tackling this problem is to introduce a mass spectrometric detector system[103,104].

2.9.3 AN INTRODUCTION TO QUADRUPOLE ION TRAP MASS SPECTROMETERS

The following section is designed to give an overview of the mode of operation and principles involved in quadrupole ion trap theory, it is not intended as a comprehensive guide to the mathematical theory of ion motion in an ion trap as this is considered to be outside the scope of this section.

Gas phase analytes that are eluted from the gas chromatograph enter the ion trap directly where they undergo electron ionisation. The analyte molecules interact with electrons that are produced from a heated rhenium filament and accelerated across the source at 70 eV. This process leads to the formation of an analyte radical cation. The technique is a “hard” ionization method and the resultant analyte radical cation undergoes extensive fragmentation, so that there is often little to none of the original cation observed. The ionization time in an ion trap is sufficient so that a predetermined number of ions are produced in a process known as automatic gain control. The resultant ions produced in the electron ionization process are gated into the ion trap[105].

The quadrupole ion trap consists of three shaped electrodes two of which are virtually identical. These two electrodes are known as the end cap electrodes and have a hyperboloidal geometry. The third electrode is hyperboloidal in shape and is called the ring electrode (Figure 7). The ring electrode is situated symmetrically between the two, end cap electrodes and is isolated from each end cap electrode by means of a quarts spacer[106–108].

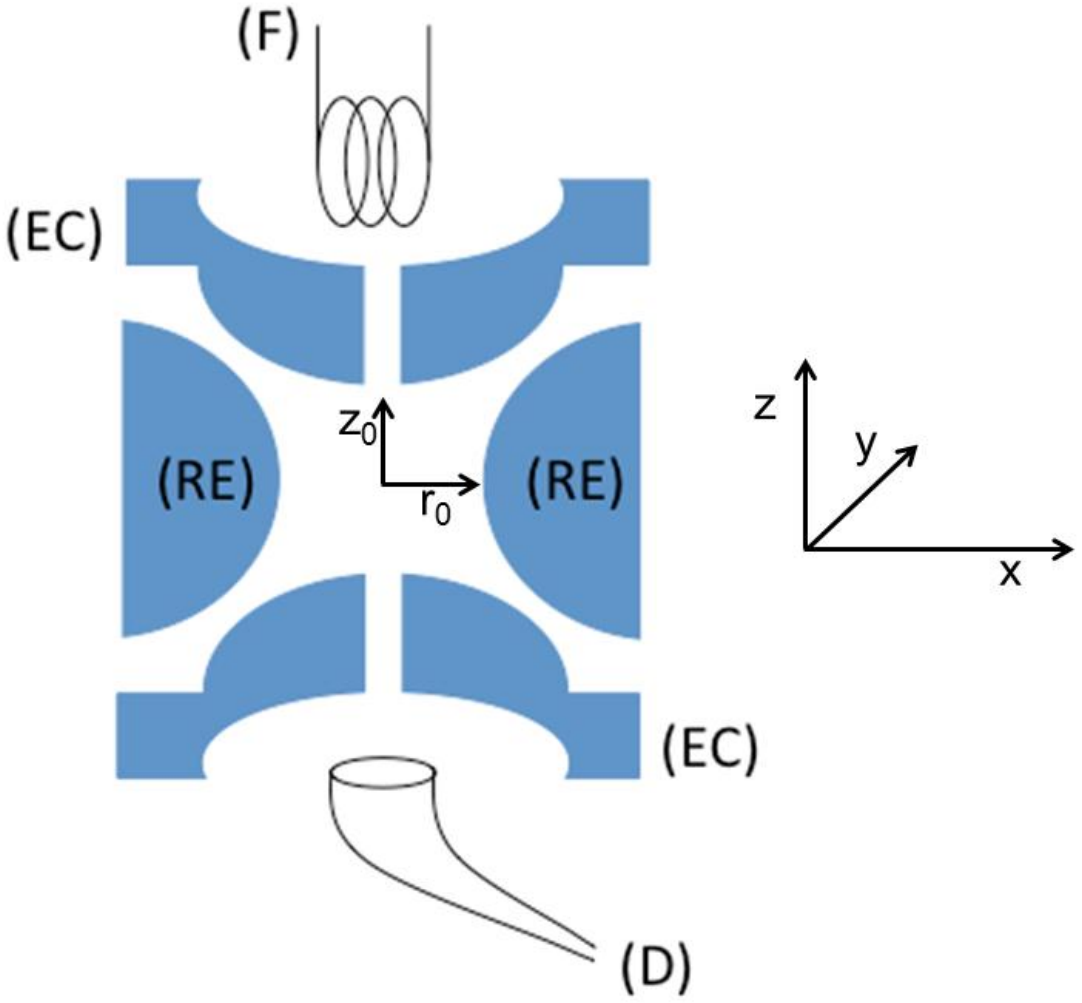


Figure 7 Cross section of a typical quadrupole ion trap mass analyser illustrating the geometry. F is the filament of the ion source and D is the electron multiplier detector. EC are the end-cap electrode and RE is the centre ring - electrode

As the ions are gated in to the analyser they immediately come under the influence of the trapping potential in the ion trap, this is achieved by the ring electrode having an applied initial RF voltage. This causes ions of a particular m/*z* to come under the influence of the imposed quadrupole field thus trapping them. The motion of the ions within the quadrupole field can be described mathematically by the Mathieu equation[105].

$$\frac{d^2u}{d^2\xi} + (a_u - 2q_u cos2\xi)u = 0$$

Equation 14

Where: u represents the coordinates of x, y and z. a and q are the trapping parameters and are dimensionless. ξ is defined as follows:-

$$\xi = \frac{\omega t}{2} = \frac{2\pi f t}{2}$$

Equation 15

Where

f is the fundamental RF frequency of the ion trap (ca. 1MHz)

ω is angular frequency applied to the RE electrode

t is time

In a conventional quadrupole mass analyser the ion motion that results from the applied potentials to the 4 rods of the quadrupole occurs in two directions (x and y) with the motion on the z axis accruing from the acceleration of the ions from the ion source, optics as they enter the quadrupole. With the ion trap however the motion of the ions in the quadrupole field occur in three dimensions x, y and z. When the geometry of the ion trap is taken in to account the resulting cylindrical symmetry results in $X^2+Y^2 = r^2$ and so the motion of ions in an ion trap may be expressed in terms of r and z coordinates[106].

In terms of the Mathieu equation if *u* can be either r or z then as a result the coefficients a and q may be summarised using Equations 16 and 17

$$a_u = -2a_r = \frac{-16zeU}{m(r_0^2 + 2z_0^2)\omega^2}$$

Equation 16

$$q_u = -2q_r = \frac{8zeV}{m(r_0^2 + 2z_0^2)\omega^2}$$

Equation 17

Where m = mass of an ion

e = electronic charge

z = number of charges on the ion

V = is the RF frequency or AC voltage

U = the DC potential and is proportional to a_u and is equal to zero

For an ion to have a stable trajectory within the mass analyser, the ions must have simultaneous stability in both the r and z axes. By plotting q_u against a_u it is possible to illustrate where the regions of stability residue. This type of plot is known as a Mathieu stability diagram (Figure 8). Those regions which overlap closest to the origin provide the stable trajectory necessary for an ion to remain caged in the quadrupole field. The stability of the ion trajectory is dependent on the stability parameters β_r and β_z which are in turn dependent on a and q. The limits of this first stability region are defined by $0 < \beta_r, \beta_z < 1$

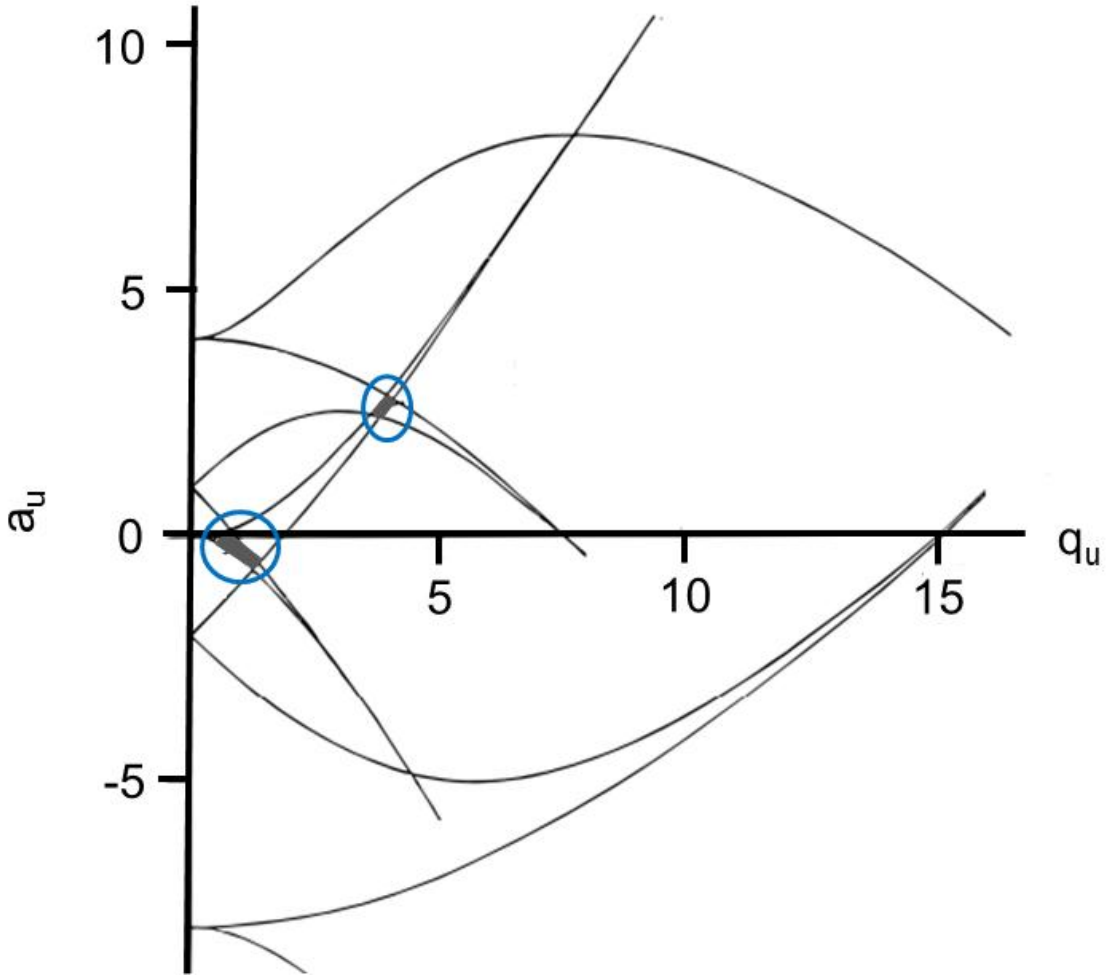


Figure 8 Mathieu stability diagram, the shaded indicates the regions of a/q space that overlap and provide stable trajectories for an ion to remain in the quadrupole field.

As U=0 (i.e. no DC voltage is applied) the ion trap will operate along the q_u axis. At a given U/V ratio ions of varying m/z can be found to reside in a line that crosses the stability region. Ions with the highest m/z are located closer to the origin than those with lower m/z ratios. The regions of stability found in a/q space can then be plotted as envelopes that have a characteristic shape. If V is increased q_z is also increased for each ion. At the point where $q_z = 0.908$ ($a_z = 0$), $\beta = 1$ and the ion has reached its stability limit. If V increases past this point the ion's trajectory becomes unstable and the ion is then ejected from the ion trap in the z axis. This mode of operation is known as mass-selective instability mode and as the RF or V is continually increased ions of increasing m/z are ejected from the mass analyser and detected resulting in a full scan mode spectrum[106]

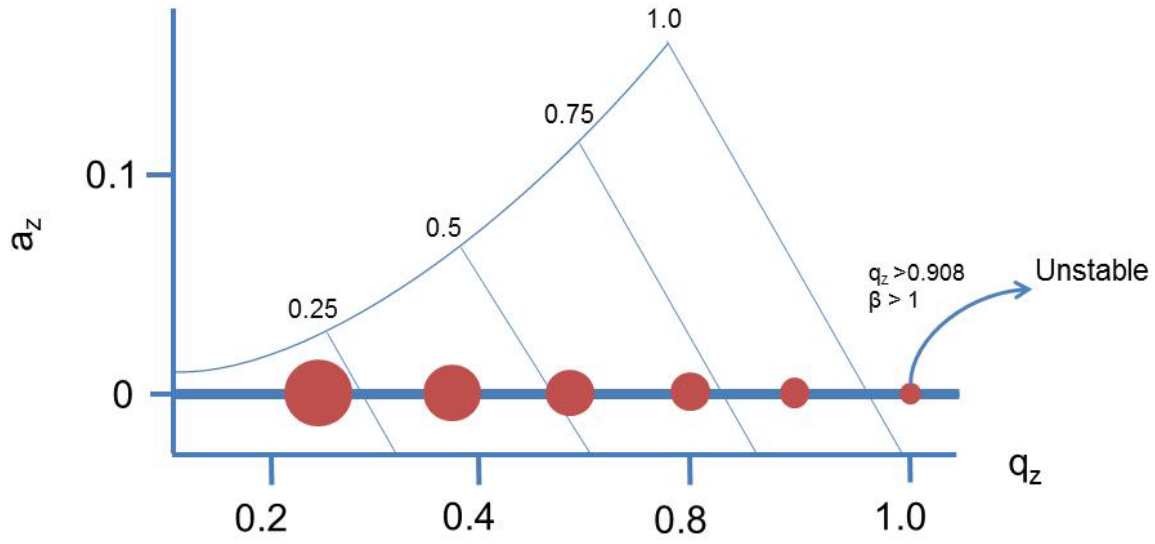


Figure 9 Example of ion ejection at the stability limit.

2.10 Data processing

2.10.1 MULTIVARIATE ANALYSIS

The multitude of data that can be generated in "-omic" style mass spectrometry approaches is overwhelming. The interpretation of such complex data sets is challenging as they typically contain hundreds of variables and many observations. Simple uni-variate or bi-variate statistical approaches are simply not up to dealing with such extensive data sets. Two approaches used to tackle such vast data sets are principal component analysis (PCA) and partial least squared discriminate analysis (PLS-DA)[109].

PCA is an unsupervised multivariate analysis method and does not involve the identification of particular groups within the data sets by the analyst.

The PCA process aims to reduce the dimensions of observational space in which a given set of observations reside. To accomplish this the PCA model creates new linear combinations of variables that characterize the observations studied, these new variables are termed principal components.

The first step in the PCA process is to subtract the mean observed value of each of the variables; then a covariance matrix can be constructed[110,111]. The number of different covariance values for a data set with n- variables can be summarized by Equation 18

$$\frac{n}{(n-2)x2}$$

Equation 18

The covariance matrix for n-variables may be defined using Equation 19.

$$C^{n\times n} = (C_{i,j}, C_{i,j} = cov(Dim_i, Dim_j))$$

Equation 19

where $C^{n\times n}$ is the matrix with n- variables and n- observations and Dim_x is the xth variable. Therefore if the data set has n-variables then the covariance matrix will have n-rows and n-columns, with each matrix entry being the calculated covariance between two variables. Using the covariance matrix, eigenvectors and eigenvalues can be calculated. As the covariance matrix contains n x n rows and columns it is possible to calculate n- eigenvectors each that are all perpendicular to each other[110,111].

Once the eigenvectors are found the next step is to order them in descending eigenvalue. The eigenvectors with the highest values have the most significance. These eigenvectors can then be combined and transposed into a matrix, which is then combined with the original matrix to produce the principal components[109]. This can be summarized in Equation 20.

$PC1 = (a_{11}X_1 + a_{12}X_2 + \ldots a_{1n}X_n)$

$PC2 = (a_{21}X_1 + a_{22}X_2 + \ldots a_{2n}X_n)$

etc.

Equation 20

These new principle components can then be represented graphically in terms of a score plot. The score plot shows how the observations trend or cluster compared to each other and acts like map of the observations. The projection of a score plot from two principal components to a 2D surface is shown in Figure 10. Typically to highlight any potential outliers in the data set a 95 % tolerance is applied and is represented by an ellipse drawn on the score plot. This ellipse is based on Hotelling's T^2 distribution[110,111]. The variables responsible for any separation or clustering apparent in the score plot can be elucidated by means of interrogation of a loading plot. A loading plot is the plot of all the variables measured plotted on the same axis diagram as the principal components. As a result any variables that are subsequently overlaid in the corresponding quadrant to the apparent clusters from the score plot can be isolated and elucidated.

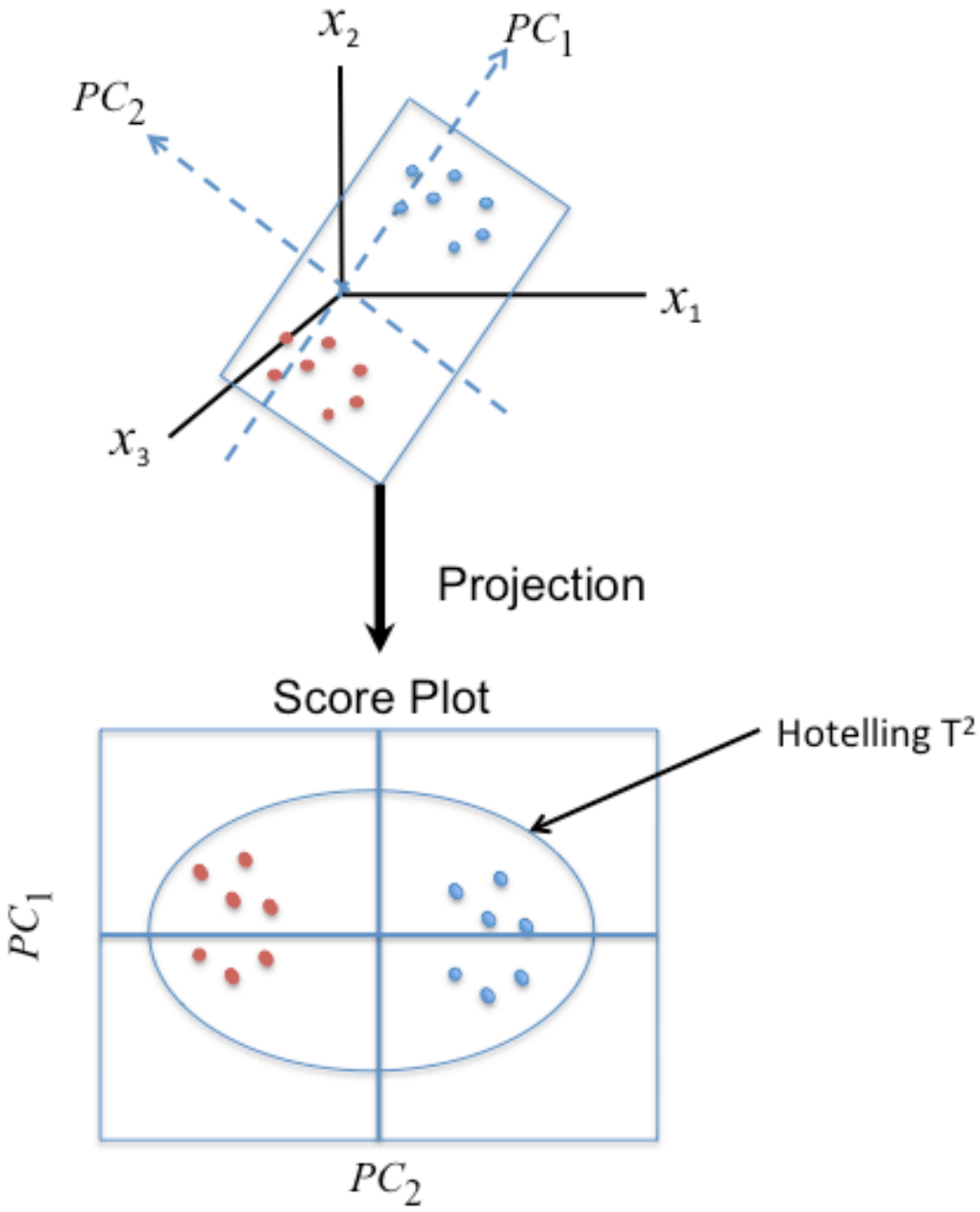


Figure 10 Projection of a score plot from calculated principle components.

The second approach to multivariate analysis that has been used in “omic” style studies is partial least squares discriminate analysis. The aim of partial least squares discriminate analysis, like that of PCA, is to reduce the dimensions in the data sets and to highlight those variables that contribute significantly to any separation observed in the dataset. The approach of PLS-DA differs from PCA, in that this multivariate analysis procedure is a supervised method where the analyst will indicate to the model the class membership of the observations in the data set. These classifications may be two groups, for example healthy vs. disease

participants. The model uses this information to build a learning or training set of data which generates a set of rules to enable the fit of new unknown observations to the correct groups or classes. The application of these rules to unknown observations enables the assessment of the predictive ability of the model. The results then provide the maximum amount of discrimination available between the two classifications of observations.

The partial least squares model attempts to do this by using linear combinations of the original X variables in order to obtain a linear discriminant function. An example of this is given in Equation 21.

$$Y_1 = a_1X_1 + a_2X_2 + \cdots a_nX_n$$

Equation 21

As in the PCA model the coefficients a1, a2…etc. are assigned so that the Y exhibits the maximum variation between the classifications of the observations. This method of multivariate analysis has been used extensively in metabolomic studies, where the number of variables vastly out weights the number of observations.

The vast discrepancy between the number of variables and the amount of observations in this type of data means that only part of the data contains any useful information. Beyond a certain point the useful information starts to diminish and the principal components will then only be modelling noise. As a result cross validation of the model needs to be undertaken to ensure that spurious relationships in the data are avoided. A widely used approach of cross validation is the “leave one out” method. This approach works by keeping a fraction of the observations of the data out of the model; these observations are then predicted by the model and compared with the actual values obtained. This process is repeated until all observations have been kept out once and only once[112].

2.10.2 POTENTIAL PITFALLS OF MULTIVARIATE ANALYSIS METHODS WHEN USED IN CONJUNCTION WITH METABOLOMIC DATA SETS.

The application of multivariate statistics enables researchers to interrogate the multitude of data acquired using modern chromatography based separation methods coupled with mass spectrometric detection. These large datasets containing many variables make multivariate analysis susceptible to both confounding variables and voodoo correlations.

Confounding variables are variables that have real statistical correlation with a particular class or group of observations[113]. As a result they can be highlighted using multivariate approaches as potential principal components. It is worth noting that these variables are not fictitious eigenvectors generated through the mathematical modelling process, they are in fact real artefacts present within the samples that have undergone analysis. These artefacts may be introduced through improper sampling, poor storage or be produced from the analytical system itself. Other sources of confounding variables may arise through participants' interactions with their environments and in the case of clinical studies potential confounding factors may arise from the medication the study volunteers are taking to manage their medical condition. The identification and elimination of confounding variables from data sets should be standard practice, as inclusion of these components increases the number of variables, which further contributes to the potential rise of voodoo correlations.

Voodoo correlation was a term that was first introduced in the field of social science[113–115]. These correlations are seemingly real correlations between variables and observations. However they appear coincidently in sets of data where there are a disproportionately high number of variables compared to the number of observations. This scenario then must apply to the types of data sets examined in "-omic" style studies for biomarker correlations with disease. Miekisch *et al* demonstrated an example of voodoo correlations by using 128 coin tosses to generate random numbers for a population size of n = 8 (four controls subjects and four test subjects). Miekish reported that on average they only needed 16 coins before finding one that was able to correctly predict 7 out of 8 subjects correctly[113].

Instruments such as GCxGC–TOF–MS used for the analysis of exhaled breath yield upto 3000 VOC components in an individual's breath[86]. The profiles obtained from such systems have highlighted the complexity of breath and show how many variables are on offer to the analytical or clinical scientist when it comes to finding potential voodoo correlations within the resultant data sets.

The question must then arise: how can these confounding factors and correlations be eliminated from data sets used in the interpretation of breath studies? The first and most obvious way is to ensure that there are a sufficient number of independent observations being produced in studies. However this solution poses a problem as Miekish *et al* recommends that 5 to 10 times more observations should be used than there are variables[113]. As 3000 VOC components have thus far been shown to be present in breath[86], this would mean that for any study to statistically acceptable it would require somewhere between 15000 to 30000 observations to be made. The author has yet to discover a published study that even comes close to this number of observations within the breath analysis community. In many cases this first approach is practically unfeasible and as a result there are a number of precautions that can be used to ensure the validity of the results obtained from "-omic" style studies. These include but are not limited to:-

- Standardisation of sampling and analysis procedures
- Rigorous quality control procedures for the instrumentation used
- Acquisition of blank samples from air supplies or room air
- Removal of contributing VOCs found to be present in blanks to aid in reducing the number of measured variables
- Relevantly matched participant control cohorts
- Cross validation of statistical models
- Voodoo testing of statistical software by generation of random data sets with equivalent numbers of variables and observation as the data set under analysis.
- The use of metadata, such as body mass index, age, sex, medication, alcohol intake etc.
- Investigation of the proposed biomarkers origins in the human body.
- Biomarker validation and conformation through independent studies at different clinical centres and different participant cohorts.

These suggested precautions should be considered throughout the study design, participant selection, sample collection and analysis and statistical analysis. At each stage there is the possibility to introduce a multitude of confounding factors that can lead to erroneous results when data mining for potential biomarkers[113,116].

Chapter 3 Method and systems

3.1 Overview

Figure 11 is a summary of the processes for the collection and analysis of exhaled breath used in this research. Further, the data modelling and statistical approaches used in this work are also described.

The collection of expired breath has been approached through the application of the adaptive breath sampler described[11]. This procedure uses an instrumented non-vented RESMED Mirage Series 2, full-face mask (RESMED (UK) Ltd, Oxfordshire, and UK.) to collect distal breath, which includes the alveolar fraction, onto a mixed-bed thermal-desorption tube. The masses of analyte associated with these procedures are typically in the pg to ng region and the potential for contamination, artefacts and interference is always present. The procedures, described below, for conditioning and preparing the materials and systems required for the collection of breath samples needs to be followed carefully to ensure that samples and analytical results are valid.

3.2 Ethics

This research was undertaken in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The breath samples were collected using a protocol approved by the Local Ethics Advisory Committee (Ref: Loughborough University G09-P5) from volunteer participants who gave informed consent.

Each participant was given an overview of the study and what their part in the study would involve. This document was known as the participant information sheet and an example of the participant information sheets can be found in. This information was provided to each participant prior to any visit to give a breath sample and on arrival participants were given the opportunity to ask any questions pertaining to the study. Following this the participants were given a demonstration of the sampler and shown what would happen when a breath sample was taken. If the participant then wished to continue with the study they were presented with two further documents, firstly an informed consent document and a health screen questionnaire. Following this all

participants were then given a study specific identifier and the consent forms were sealed, this process pseudo-anonymised the participants.

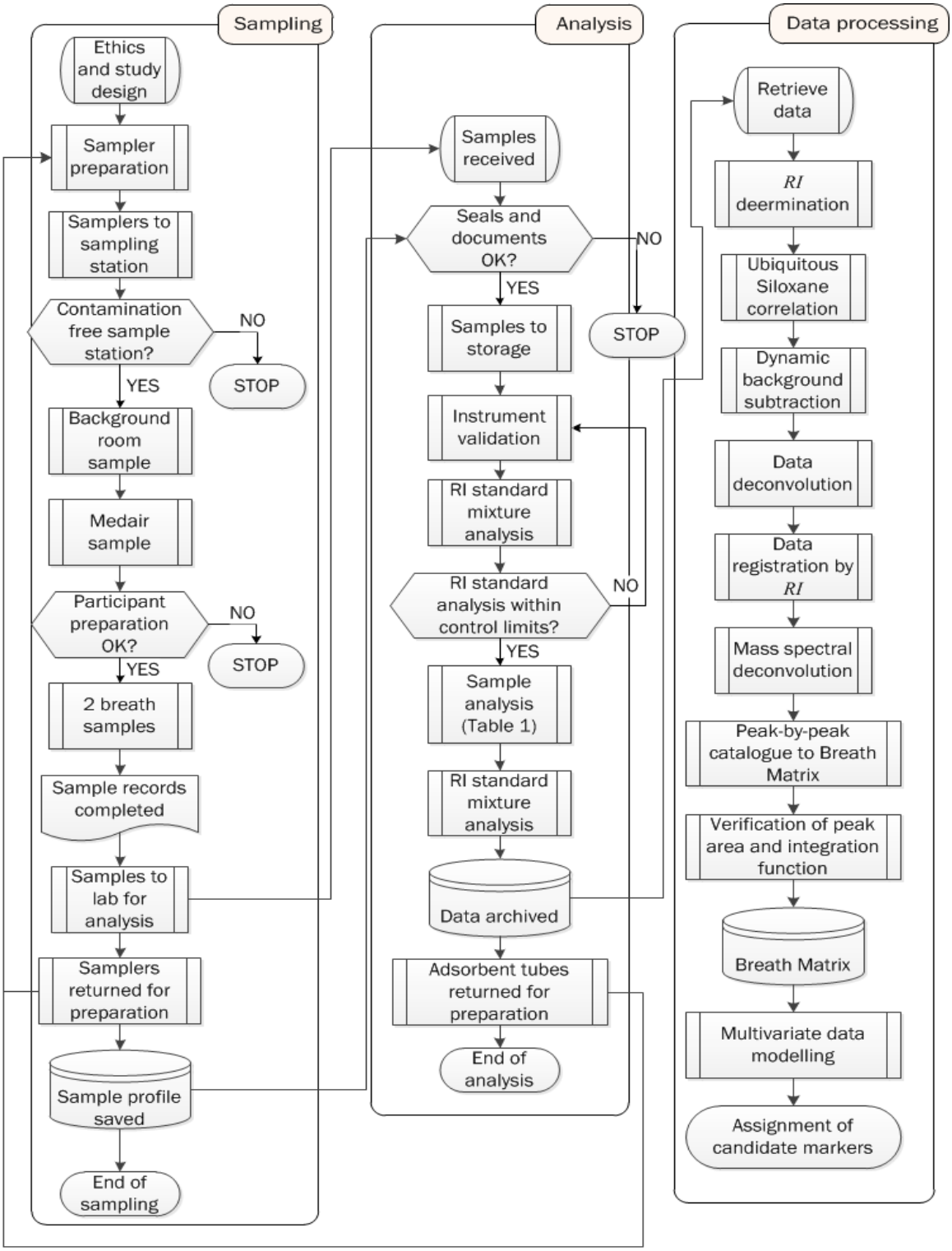


Figure 11 An overview of the breath analysis work-flow adopted for this research showing the three main stages of the process. Taken from [117].

3.3 Sampler preparation

3.3.1 ASSEMBLING AND DISASSEMBLING OF FULL-FACE MASKS

The RESMED full face masks comprise 5 main parts that are completely detachable, a diagram of a participant wearing the full-face mask can be found in Figure 12. To disassemble the mask, firstly remove the forehead support and forehead pad located at the top of the mask. Once separated the forehead pad may be removed from the support. The mask is now ready to have the mask cushion removed from the mask support, this is achieved by removing the cushion clip from the mask support and then gently lifting the cushion pad from the locating rim of the mask support. Spreading and lifting the locating clip on the interior of the mask support will allow the removal the elbow fitting. To reassemble the mask this procedure should be performed in reverse order.

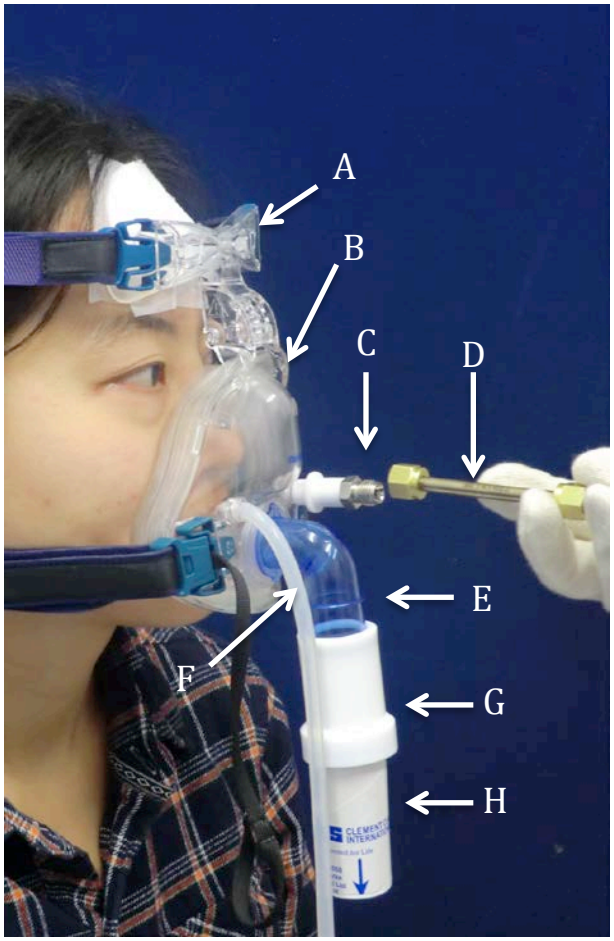


Figure 12 Illustration of a participant wearing a full face mask (A) forehead support and pad, (B) face mask, (C) adaptive sampler sampling capillary, (D) Thermal desorption tube (E) elbow fitting, (F) silicon tube connecting the mask to adaptive sampler pressure transducer (G) T-piece and air inlet and (H) one way disposable valve.

3.3.2 STERILIZATION AND CLEANING OF THE FULL-FACE MASKS

To reduce cross-contamination of samples and eliminate transmission of communicable disease masks must only be used once with an individual participant and then cleaned and sterilized before reuse.

A sufficient amount of the disinfectant Cidex® OPA (Johnson & Johnson UK) solution was transferred to a plastic container such that the component parts of the facemask could be fully submerged. The Cidex® OPA solution contains 0.55 % *ortho*-phthalaldehyde as the active ingredient and therefore all operations involving the solution were carried out in a fume cupboard and appropriate PPE was worn at all times. The facemask was disassembled, see above, and all parts of the mask were submerged in the Cidex ® OPA solution before a bottlebrush was used to ensure that all areas of the mask were cleaned. Once clean the components of the mask were left completely submerged in the solution for a further 10 min before removal and immediate and rapid transfer into a bucket containing approximately 5 dm^3 of potable water, where they were rinsed for a minimum of 1 min. Each mask component was then transferred to a second bucket containing a further 5 dm^3 of potable water and rinsed thoroughly. The mask-components were rinsed a third time to ensure that all the disinfectant was removed before they were placed to air-dry on clean paper towelling. Once dry the components were inspected and any solid residue from the water was removed by polishing with a paper towel before all component parts were placed in a vacuum oven (set at 50^oC and -1 atm) for a minimum of 8 hr to vacuum polish the components to ensure that potential VOC contaminants from the cleaning procedure were supressed. After this process the masks were reassembled and placed into sealed clean air-tight containers ready for use. Note that the air-tight containers were cleaned sterilized and prepared using the same procedure.

3.3.3 CONDITIONING AND EVALUATION OF MIXED BED THERMAL DESORPTION TUBES

The thermal desorption sampling tubes used to trap VOCs from exhaled breath were industry standard 89 mm long 6.4 mm id stainless steel tubes. Each tube was packed with Tenax™ /Carbograph TD1 (Part No C2-AAXX-5032, from Markes International, Llantrisant, UK). Figure 13 is an engineering drawing of the adsorbent sampling-tubes used in this research.

Figure 13 An image of a thermal desorption tube showing the sampling and conditioning flow directions flow (green and red arrows respectively)

The adsorbent media in the sampling-tubes was conditioned before use and in-between reuse of the sampling tube, to remove VOC contamination, using a conditioning manifold (Figure 14) placed inside a conventional GC oven, designed to condition eight sampling tubes simultaneously. Purified nitrogen was supplied at a minimum flow of 100 $cm^3.min^{-1}$ to all eight thermal desorption tubes. The thermal desorption tubes were connected with 1/4" Swagelok fittings (PTFE ferrules) such that the direction of the gas flow was in the conditioning flow direction to back-flush sampled materials out of the adsorbent bed (see Figure 13). The manifold system was tested for the absence of gas leaks and the nitrogen-flow was measured at each tube to ensure that the minimum 100 $cm^3.min^{-1}$ was flowing through each tube. The tubes were then purged with nitrogen for a minimum of 5 mins, to remove residual oxygen and prevent oxidative damage of the adsorbent, before they were heated to 310^oC for a minimum of 30 min using the GC oven. After heating the tubes were allowed to cool to ambient temperature and then removed from the manifold and sealed with a long-term storage cap. The conditioned adsorbent tubes were then analysed by TD-GC-MS to verify they were free from any potential VOC contaminants.

Conditioned tubes were evaluated by TD-GC-MS desorbed with a rapid (10 min) high temperature (310^oC) isothermal GC method. The resultant chromatogram and spectrum of the desorption profiles was inspected, and any response over 1M counts

in the TIC trace above baseline indicated potential contamination or degradation of the adsorbent in the tube. When such a situation arose the tube was desorbed again and if the elevated response persisted the tube was removed from service. After blanking the adsorbent tubes were recapped using the long-term storage caps and placed in a laboratory refrigerator at 4^{o}C until required or 1 week whichever the sooner.

Figure 14 The adsorbent tube conditioning manifold was installed into a GC oven where up to eight adsorbent sample tubes were processed at the same time.

3.4 Sampling exhaled breath

3.4.1 EXHALED BREATH COLLECTION

A conditioned full-face mask was equipped with a PTFE t-piece fitted with a disposable one-way non-rebreathing cardboard valve. The t-piece was then connected to a filtered air supply. A flow rate of > 20 $dm^3.min^{-1}$ of air was supplied to the mask through a triple filter system (Parker 2000 Balson series with 100-12-BX, 100-12-DX and CI100-12-000 2002-series grade filters). The mask was then fitted to the participant's face using the head straps supplied by the manufacture and adjusted so that a seal was formed between the mask cushion and the participant's face. The participant was then asked to breath normally for a minimum of 5 mins so that a normal breathing rhythm could be established at which point the adaptive breath sampler was connected to the mask.

3.4.2 ADAPTIVE BREATH SAMPLER

The adaptive breath sampler used to collect VOC breath samples has previously been described in detail[11] (Figure 15) therefore only a brief overview will be given here. The sampler unit was controlled by a laptop computer (DELL Inspiron 1100 with Windows XP operating system) equipped with a 6024E PCMI data acquisition card (National Instruments) running Labview™ Virtual Instrument (VI) software. A 25-pin RS232 plug connects the sampler unit to the laptop computer. The sampler employs a pressure transducer with a programmable gain amplified circuit to monitor the breathing profile of the participants via the Labview VI software. During the exhalation of the distal breath the VI software automatically switches the sampler's two electro fluidic micro-switch valves from the vent to the sampling position. An Escort Elf® Sampling Pump (Part No 497701, MSA, UK) was used to accurately sample 0.8 dm^3 min^{-1} through the valves when in the sampling and vent positions.

A silicon tube was connected from the sampler unit's pressure transducer to the first luer fitting on the full-face mask and acted as the pressure measurement line for the sampler.

The second luer fitting housed the VOC sampling manifold and comprised a PTFE mask connector with male luer fitting which housed a 1/4'' to 1/16" reducing union that was fitted with a 5 cm length of 0.53 mm id MXT passivated stainless steel column

Figure 16. The column protruded into the mask through the luer fitting and seated above the participants upper lip. Held in position by a nut and graphite/vespel ferrule connected to the 1/16" end of the reducing union the column acted as the sampling outlet of the mask. A conditioned and blanked thermal desorption tube was connected to the 1/4'' side of the reducing union using the brass nut and PTFE ferrule of the brass sealing cap supplied with the adsorbent tube from the manufacturer. Silicone tubing connected the outlet of the adsorbent tube to the sampler unit's electro fluidic micro-switch valves and sampling pump. Figure 17 is an example of the pressure transducer's output with the switching signal superimposed on top of the breath profile.

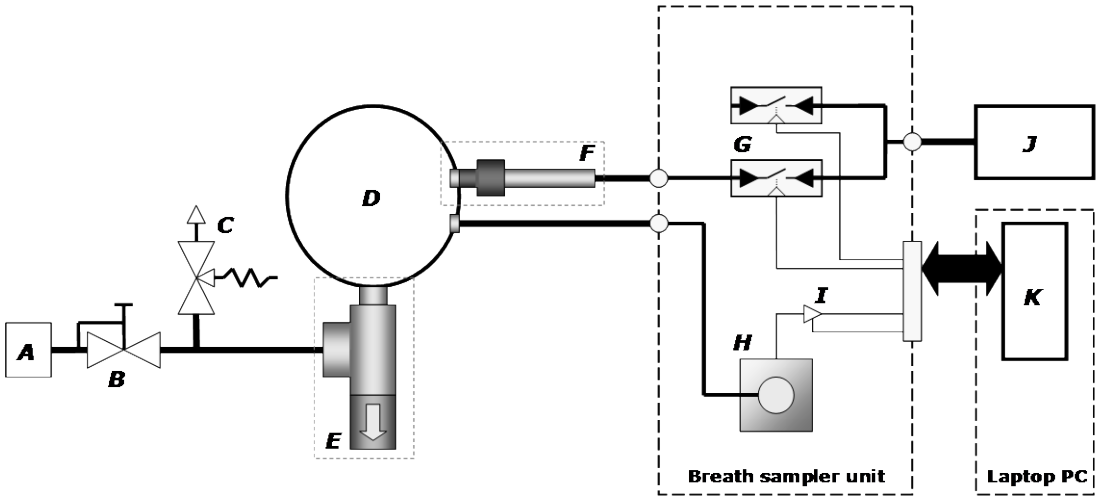


Figure 15 Schematic diagram of the adaptive breath sampler[11] A)Purified air supply; B) pressure regulator; C) pressure relief valve; D) Resmed full face mask; E) air supply assembly; F) adsorbent sampler assembly (exploded diagram see Figure 16); G) micro-control valves; H) pressure transducer; I) electronics interface; J) precision air sampling pump; J) Laptop with DAC card and Labview software.

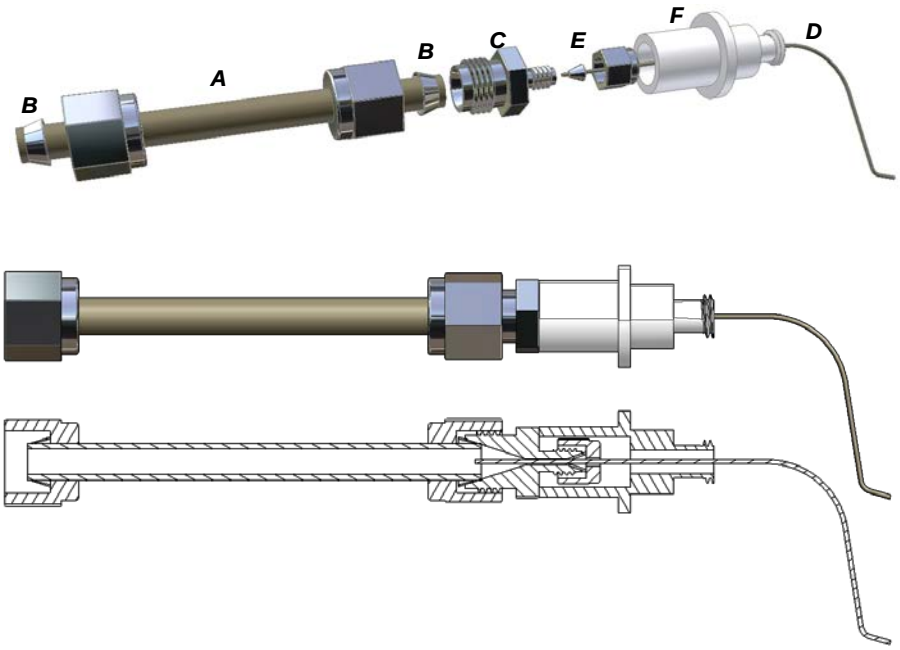


Figure 16 Adsorbent sampler assembly. The adsorbent trap (A) was connected to the sampling line through a 1/4" push fit connector sealed with a PTFE ferrule, (B). The other end of the trap was fastened into a 1/4" to 1/16" reducing union, (C), also with a PTFE ferrule (B). A shaped 5cm MXT passivated stainless steel column (D) was passed through the mask connector and reducing union so it was inserted into the top of the adsorbent trap. A 1/16" nut with a 0.53mm vespel ferrule sealed the capillary in place (E). The mask connector (F) was machined from PTFE with a male Luer fitting. The reducing union was held firmly in place while the capillary was threaded and positioned carefully inside the mask, before the Leur fitting secured the whole assembly (Figure reproduced from Theo Koimtzis original artwork)[11].

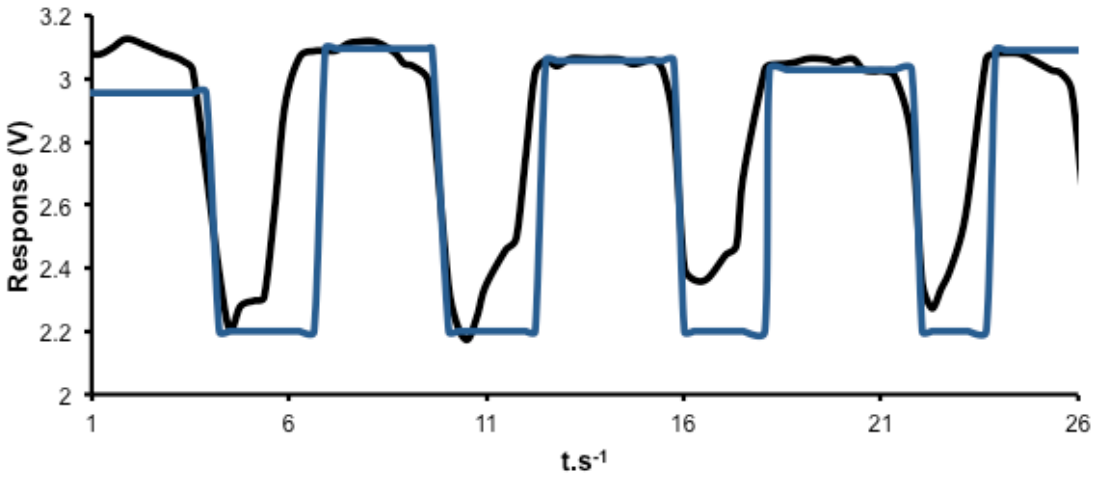


Figure 17 Example of a typical breath profile obtained from the adaptive breath sampler. Black line illustrates the inhalation and exhalation of the profile of the participant over 26 seconds. Blue line highlights the switching of the sampling values on (high blue plateau) and off (low blue plateau) indicating the fraction of the expired breath collected on to the thermal desorption tube

3.5 Breath sample analysis by thermal desorption gas chromatography mass spectrometry (TD-GC-MS)

3.5.1 THERMAL DESORPTION SAMPLER

Sample introduction to the GC-MS is performed using a Series 1 Unity thermal desorption platform (Markes International, UK). The Unity system employs a two-stage thermal desorption process which comprises a glass refocusing cold-trap that has an internal diameter of 2 mm and is packed with identical adsorbent to that of the sample tube used for the breath sample collection described in Section 3.3.3. The cold-trap is commercially available from Marks International and is a general-purpose hydrophobic cold-trap. The cold-trap refocusing system is electrically cooled by a Peltier system that is capable of generating temperatures down to -10°C. **Figure 18** illustrates the thermal desorption inlet process. When a sample is placed into the sample oven the Unity system performs a leak-test to ensure that there will be no loss of sample during the desorption process. On successful completion of the leak-test the system purges the sample tube with the carrier gas (helium) before any heat is applied to the tube, this process removes any air in the sample tube that may cause damage to the sorbents through oxidation from residual oxygen when the tube is heated to high temperatures. The primary desorption flow through the sample tube

(45 cm^3 min^{-1}) is higher than the flow of the GC capillary column (2 cm^3 min^{-1}). The sample tube is then heated rapidly to a high-temperature (300°C) for 5 mins. The primary desorption-flow is directed through the cold-trap (-10°C) where the VOC analytes freeze and concentrate into a small volume in the cold-trap adsorbent. The gas-flow to the cold-trap is then reversed and directed into the capillary column of the GC-MS system at a rate of 2 cm^3min^{-1}. The cold-trap is then heated rapidly to 300°C causing desorption of the concentrated VOCs onto the head of the analytical capillary column. The operating parameters for the thermal desorption process used in this work are summarised in **Table 1**.

Table 1 Thermal desorption parameters for the Series 1 Unity Thermal Desorption

Parameter	Setting
Tube desorption	300 ^oC for 5 minutes
Desorption flow	45 ml min^{-1} Splitless
Cold trap	General purpose hydrophobic
Trapping Temperature	-10 ^oC
Trap desorption	300 ^oC for 5 minutes
Trap desorption flow	2.0 ml min^{-1} Splitless
Split ratio	splitless
Flow path temperature	180 ^oC

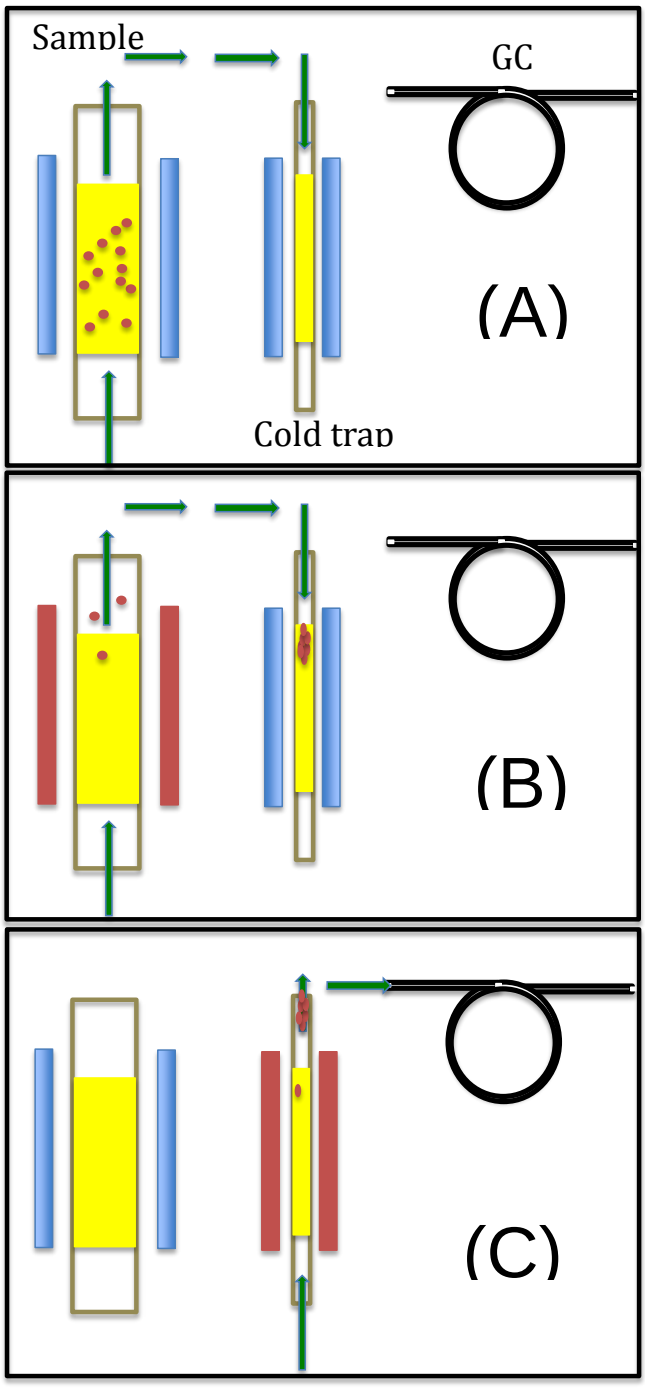


Figure 18 Overview of the thermal desorption process in the Unity thermal desorption platform. (A) Sample pre purge, He flows through the sample tubes at ambient temperature to purge the sample tube of any oxygen. (B) Sample cold trapping, the sample tube is heated desorbing the analytes off the surface of the adsorbent in the tube in to the Helium flow and is refocused on to the cold trap. (C) Sample introduction, the helium flow supplied to the cold-trap is reversed and diverted to the analytical column. The cold-trap is then heated rapidly desorbing the sample analytes.

3.5.2 GAS CHROMATOGRAPHY MASS SPECTROMETRY

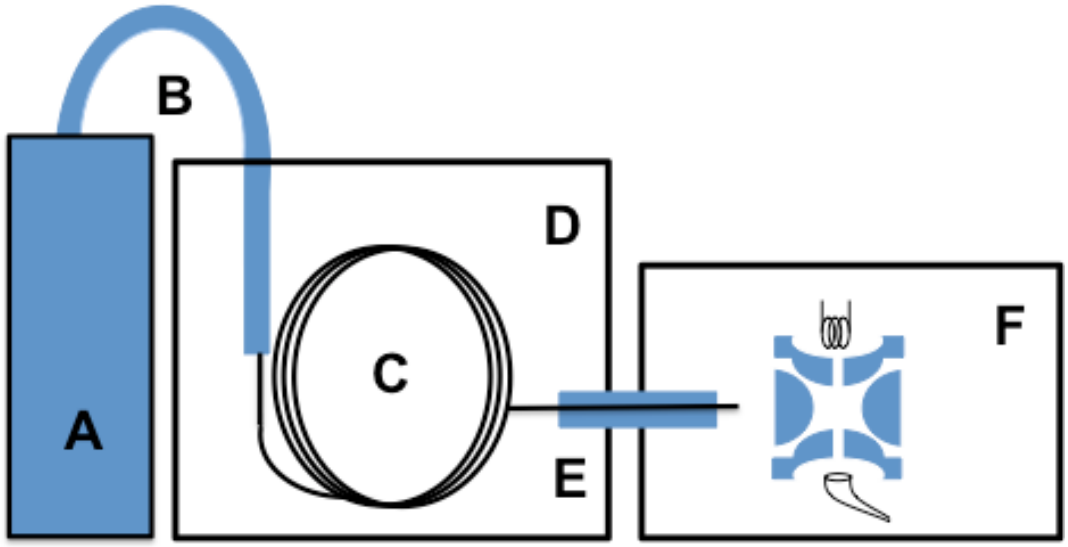


Figure 19 Schematic diagram of TD-GC-MS, (A) thermal desorption sampler, (B) Heated transfer line, (C) capillary column, (D) Gas Chromatograph, (E) Mass spectrometer heated transfer line and (F) Ion trap mass spectrometer.

Chromatographic separation of VOC analytes from exhaled breath samples was performed with a Varian 3800 GC equipped a 60 m x 0.25mm x 0.25µm capillary column with a 5% phenyl 95% dimethyl arylene siloxane stationary phase (DB5-MS, Agilent technologies). The analytical column was connected directly to the outlet transfer-line of the thermal desorber and was supplied with a constant flow (2 cm^3 min^{-1}) of helium carrier gas through the instrument's pneumatic flow controller. The oven's initial temperature was 40 ^oC with no hold-time, the oven temperature was then increased at a rate of 5 ^oC min^{-1} to 300^oC. The final hold-time was 8 min bringing the total run-time to 60 min.

Detection of VOC analytes was performed with a Varian 4000 ion-trap mass spectrometer. The spectrometer uses a 3-D ion-trap mass analyzer with an electron ionization (EI) source. The system was configured in internal ionization mode, which allows both EI and positive chemical ionization (PCI) to be performed in separate time segment in the same analytical run. The column effluent-flow enters the ionization source though a heated transfer line set to 300^oC, located between the spectrometer and GC oven. In addition to the column effluent flow the spectrometer also employs a helium buffer gas flow at 4 cm^3 min^{-1}, which is directed into the trap and causes the ions to adopt a more compact orbit. This buffering produces higher

resolution when the ions are ejected from the mass analyzer. After ejection, the ions strike a conversion dynode, which begins the signal multiplication process performed by an electron multiplier. A comprehensive list of instrumental conditions is given in Table 2.

Table 2 Gas chromatography conditions (Varian 3800)

Parameter	Setting
Column	DB 5 MS 60 m x 0.25 mm id x 0.25 µm
Initial oven temperature	40 ^{o}C Hold for 0 min
Oven temperature ramp	300 ^{o}C at 5 ^{o}C min^{-1} Hold for 8 min
Total analysis time	60 min
Carrier gas	Helium
Carrier gas flow	2.0 ml min^{-1}

Mass spectrometer conditions (Varian 4000)

Parameter	Setting
Trap temperature	150 ^{o}C
Manifold temperature	50 ^{o}C
Transfer line temperature	300 ^{o}C
EI scan mode	EI Auto
EI scan frequency	2 Hz
EI scan range	m/z 40 – 45

3.6 Method development.

At the outset of this research the GC-MS methodology used a 30 M analytical column with a DB5-MS stationary phase and a flow rate of 4 ml min-1. The GC temperature program incorporated ramps at different rates and temperature holds. The Varian 4000 mass spectrometer's ion trap analyzer was factory standard. Due to the nature of the volatiles contained in breath and the high water content of both the breath samples and contributions from the Unity thermal desorption platform this initial methodology was improved and re-optimized specifically for the analysis of the breath samples that were collected in the studies presented in this research. The conditions of analysis were optimized to those presented in Table 2.

Gas chromatography method development

In order to ensure that chromatographic data was aligned over the course of the studies presented in this thesis a chromatographic method was required that could be used in conjunction with retention indexing. Previous methods presented in the literature[11,118] that have used the adaptive breath sampler have employed GC methods that incorporated multiple temperature gradients and hold steps throughout the oven program. While retention indexing can be performed on these types of methods successfully, the application of retention indexing for data alignment is less reliable than that of linear temperature programing over long periods of time due to changes in the column performance. Changes in the column performance may be brought about through general maintenance or through stationary phase degradation due to age and use. Calculations of component retention indices when using temperature programing must be carried out under constant flow and a constant linear temperature program[119] in order to establish alignment of chromatographic data over a longitudinal time frame. It was found that a temperature gradient of 5oC/min from and initial oven temperature of 40oC to a final oven temperature of 300oC afforded adequate separation of the VOCs with lower boiling points. The final temperature of 300oC ensured that all potential analytes captured using the Tenax / Carbotrap thermal desorption tubes were eluted from the analytical column, eliminating any potential carryover effects. In addition to the oven program the inlet temperature (Unity thermal desorption platform transfer line) was increased from 140oC to 200oC. This improved the transfer of higher molecular weight components such as eicosane (part of the retention index standard) from the thermal desorber to

the analytical column. Finally a 60M x 0.25mm x 0.25u DB5 MS capillary column replaced the original 30 M x 0.25mm x 0.25u column and the constant flow rate was reduced to 2ml / min which provided increased resolution enabling improved de-convolution of the chromatographic data.

Instrumentation development

The Varian 4000 mass spectrometer is equipped as standard with 3 analyzer electrodes (Figure 7) that have been silica-coated[120]. This protective layer is around 1 µm thick is strongly bonded but is susceptible to damage when exposed to acids and bases or abrasive cleaning chemicals such as aluminium oxide. As such the manufacturer's recommended cleaning methods for the parts that are silica-coated are limited to sonication with solvents such as methanol or dichloromethane and non-abrasive cleaners such as mild detergent (pH 6 to 7.5) and water.

The high water content of the breath samples and the inherent water associated with the Markes Unity thermal desorption platform causes a high degree of column bleed in this combination of instrument hardware for the analysis of VOCs in exhaled breath samples. As a result the mass analyser becomes dirty and the ionization time falls below the manufacturer's specification of 20,000 to 25,000 us in full scan mode. The reduction in the ionization time occurs extremely rapidly (over 6 sample injections) reducing the ionization time below 8,000 us and making it inconsistent between samples. This reduction in ionization time occurs due to the automatic gain control (AGC) attempting to maintain a total ion count of 20,000 cps in the ion trap in order to reduce the potential of space charging effects. The resulting high proportion of background ions to analyte ions in this dirty system means that the majority of molecules being ionized are background ions. This intern results in less analytes molecules becoming ionized and therefore negatively impacts on the mass spectrometers sensitivity, overall performance and sample-to -sample consistency. The combination of high system bleed and limitations on the cleaning protocols leads to incomplete cleaning of the analyzer electrodes resulting in a permanent loss of sensitivity and performance. To address this issue the manufacturer (Varian. UK) was approached for the production of non-standard stainless steel DC and RF electrodes that would enable abrasive cleaning methods to be adopted this

adaptation over the standard electrodes supplied allowed completed cleaning of the electrodes, this resulted in mass spectrometer was able to meet the minimum manufacture specification for the ionization time consistently between runs ensuring system performance was maintained for longer periods of between maintenance intervals.

3.6.1 QUALITY CONTROL PROCEDURES

The management of day-to-day system compliance was performed by a series of checks and standards. To ensure the TD-GC-MS instrument was performing consistently the mass spectrometer operation was evaluated through a number of checks

- Air and Water check.
- Background and Ionisation time check.
- Calibration and ionisation response check
- Auto tune

This series of checks are recommended by the manufacturer and were evaluated before performing any analysis to ensure that the mass spectrometer was working correctly and was free from any interferences. Results of these checks were recorded in a daily log and the following specifications were applied to each test.

Air and water

The results and testing procedure for the air water check was performed by the Varian Saturn software and returned a result of OK, High or Massive. The system evaluated the air by comparing the ratio of m/z 28 to m/z 32 ions. The water content of the system was evaluated on comparison of the response of m/z 19 to m/ z 18 ions. Due to the nature of the thermal desorption platform the evaluation of the air and water was deemed to be acceptable if the air result was OK and the water result was High.

Background and ionization time check.

This check was performed under a system file provided by the instrument manufacturer (Varian). The system was turned on and a range of m/z 50 to 750 was

evaluated. This check highlighted any background-ions present in the system and provided a read back on the ionization time achieved by the spectrometer for a specified number of ions entering the detector. The specification was a maximum ionization time of 25 ms or a total of 20000 ions whichever was the fastest. The specification for this check, set by the manufacturer, was an ionization time of greater than 20 ms before the target ions were reached.

Calibration and ionization response check.

To evaluate the mass spectrometer sensitivity and mass tune, an ionization response check was performed on a daily basis. This involved the introduction of a perfluorotributylamine (PFTBA) standard into the mass spectrometer through the actuation of a solenoid valve connected to the Vacuum system of the mass spectrometer. The characteristic ions monitored by this calibration standard were m/z 69, 131, 264, 416 and 502. The specification for this system test was all ions had to be present and m/z 69 was the most abundant producing a response of 38000 to 43000 counts per second. The ion m/z 502 should also be present at an abundance of at least 1.5 % of that of the response of the ion m/z 69. Calibration of all ions should not exceed m/z $\pm$ 0.2 of the masses indicated above.

Auto tune

If the calibration and response check did not meet the specification described above then the system underwent the manufacturer's auto tune. If the system did not meet specification after an auto tune was performed then maintenance/repair was performed to restore the system to an acceptable standard.

Retention index standard

In addition to the daily checks performed on the mass spectrometer, a retention index standard was also analysed to evaluate the total system performance. The retention index standard was prepared in-house and was mix of n-alkane hydrocarbons. The standard mixture contained 50 µg cm^{-3} of each of octane, nonane, decane, undecane, dodecane, tetradecane, heptadecane and eicosane in methanol. A volume of 0.2 µl of this solution was injected onto a conditioned and blanked thermal desorption tube and analysed under the conditions specified in Table 2 with the addition of a 1 in 10 split applied on the trap desorption. This yielded

an on-column mass of 1 ng for each alkane. The response and retention times of the standards were monitored to ensure that the chromatography and thermal desorption systems were performing satisfactorily.

3.7 Data processing

3.7.1 THE NEED FOR DATA PROCESSING

The data obtained from the analysis of breath samples by TD–GC–MS provides an analyte rich environment containing upwards of 300 individual chromatographic entities. The majority of these components are from exhaled breath and the filtered air being supplied to the mask. A small contribution comes from the facial skin of the participant and artefacts from the analytical system. Many of the sampled VOCs co-elute in the analysed breath samples. As a result a single chromatographic separation alone is not sufficient to distinguish all potential analytes in the sample chromatograms. This complexity in the data was addressed by the following data processing steps:

- deconvolution of the a mass spectral data,
- registration of the deconvoluted analytes by retention indexing
- Integration of potential analyte peaks areas.

This sequence of procedures produced a catalogue of analytes in the breath samples and the formation of a breath matrix, which may then be applied to further statistical analysis.

3.7.2 DECONVOLUTION OF MASS SPECTRAL DATA

Spectrum deconvolution of the chromatographic peaks was performed using Analyser Pro (Spectral Works, Cheshire, UK) and the list of the settings elucidated for the de-convolution of the exhaled breath sample data with Analyser Pro is given in Table 3 . On average Analyser Pro isolated 325 individual components within a typical breath sample chromatogram. Resultant data was exported in the form of a table and provided the chromatographic peak start retention time, retention time and m/z of the ions of a particular compound. An illustration of this process may be seen in Figure 20.

Table 3 Deconvolution parameters

Detection parameter	Settings	
Area threshold	500	
Signal to noise	3	
Width threshold	0.06	min
Resolution	low	
Minimum ions for a component	2	

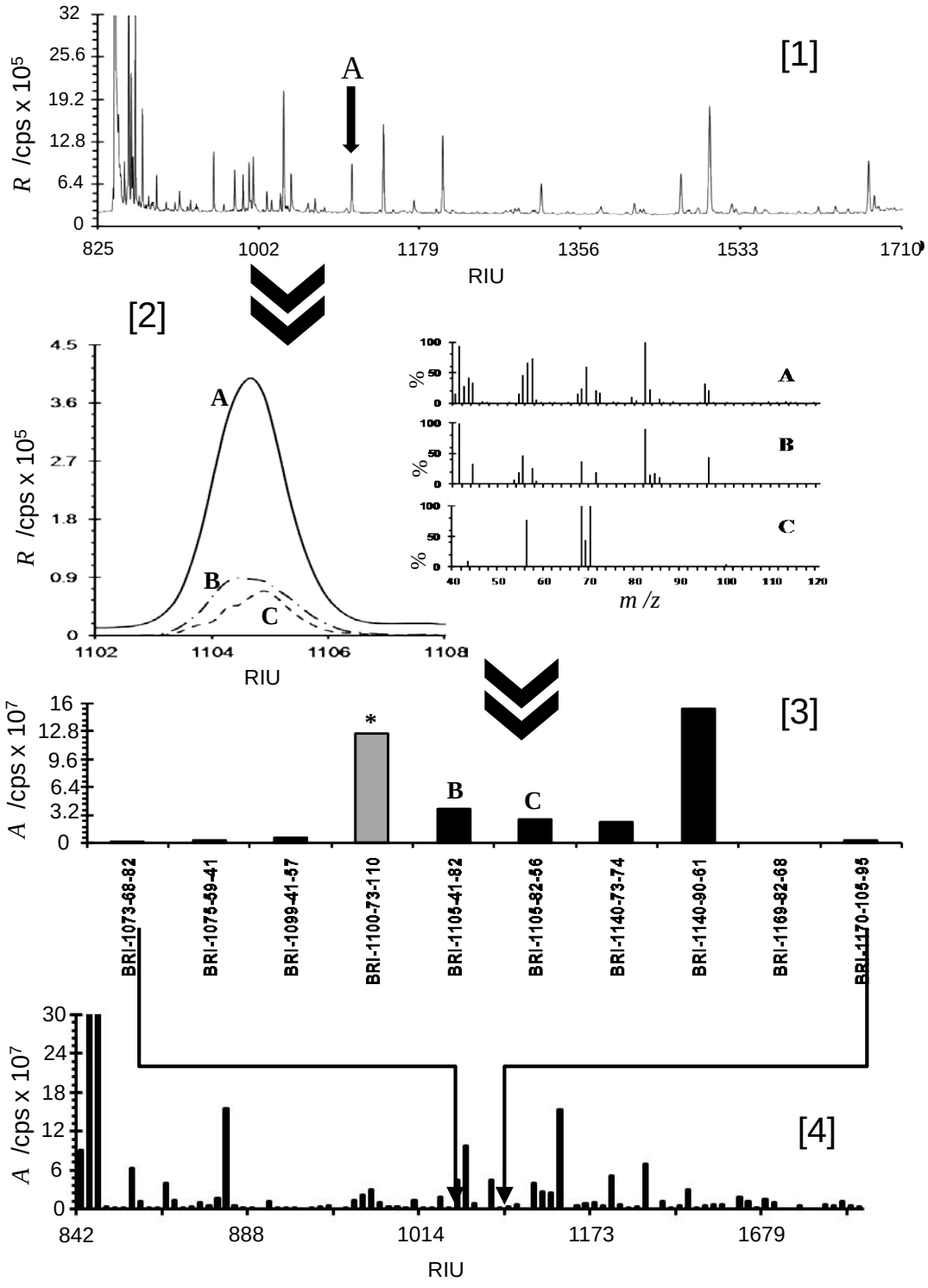


Figure 20. The deconvolution process. [1] shows the TIC TD-GC-MS spectrum with a co-eluting peak at A identified. Following peak deconvolution in [2] three peaks A and minor peaks B and C were isolated and the selected ion chromatographic peak is shown in [2]. Integration of these peaks against their retention index [3] results in the retention index registered histogram [4] required for the multi-variate analysis[117].

3.7.3 ALIGNMENT OF DATA BY RETENTION INDEXING

To compensate for variability in the analytical method over extended operations, especially variation associated with gradual column degradation, retention indexing was employed (see Appendix 2). With each batch of samples analysed, a retention index standard was also run. The standard provided a primary retention index ladder based on the n-alkanes within the mixture. In each case the theoretical retention index was plotted against the observed retention time of each alkane. From this primary retention index ladder, a secondary retention index ladder was constructed. The secondary retention index ladder used five ubiquitous siloxane components present in the chromatograms of all the exhaled breath samples[1]. The secondary retention index ladder enabled the timescale of the chromatogram to be converted from time to retention index units, using the equation of the graph produced from the retention time obtained from the sample of the ubiquitous siloxanes versus their calculated retention index.

3.7.4 INITIAL COMPONENT CATALOGUING

The combined application of retention index registration and de-convoluted mass spectra for each individual analyte yielded a unique identifier that was assigned to each isolated breath component which resulted in the creation of a catalogue of breath VOCs. This unique identifier uses a numeric method for describing the component and comprises retention index and then the five most abundant ions in the deconvoluted spectrum. To indicate that each component had originated from an exhaled breath sample the prefix BRI was added to the identifier code for example BRI-1701-57-43-85-112.

3.7.5 COMPONENT INTEGRATION AND BREATH-MATRIX FORMATION

Extracted Ion chromatograms (XIC) were used for the integration of the individual components identified through spectral deconvolution. Integration of each component was performed using Varian Saturn workstation software.

---

[1] The siloxane artefacts are produced through the interaction between the moisture in the breath sample which causes hydrolysis of the sorbents of the thermal desorption tubes and cold trap, and the GC columns' stationary phase.

The construction of a matrix was performed by listing all samples from a particular study (observations) against the total catalogue of identifier entries (variables) from that study. The resultant grid was then filled in with the integrated peak areas of each component from each sample. Where a component was not present in a sample the value of zero was used. This matrix could then be represented as a heat map of observations against variables and could be directly inserted into statistical analysis software. An example of a heat map obtained from one of these matrices is given in Chapter 5.

3.8 Statistical treatment of data

Multivariate analysis was performed on all studies using a matrix constructed as described in Section 3.7.1 and the software program SIMCA-P+ (V12.0, Umetrics UK). The theoretical aspects of the statistical methods employed using the SIMCA-P+ software have already been discussed in Section 2.10.1. The data matrix can be imported directly into the SIMCA-P+ Software from a standard Microsoft Excel spreadsheet. An example of the matrix format is given in Figure 21

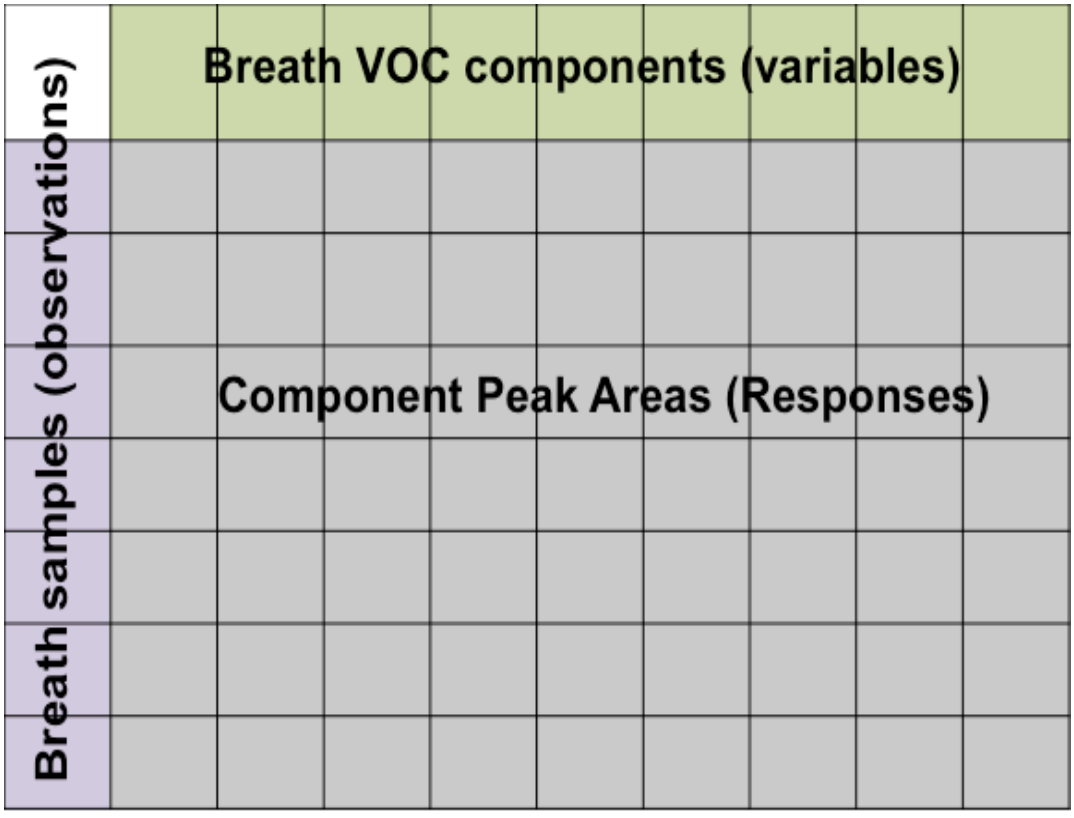


Figure 21 The breath matrix summarising the observations recovered from a breath study.

SIMCA-P+ software uses the original data from the matrix to form a work-set that is used to define which observation to include or exclude from the MVA. The observations may also be grouped into classes, for example pre- and post- an intervention. In each of the three studies presented in Chapter 4,Chapter 5 and Chapter 6 all observations were included, initially, in the work-sets and were classified according to the sample type for each particular study.

The variables in the work-set may also be included or excluded from the model at this stage. The SIMCA-P+ software has an automated evaluation tool that will include or exclude variables based on the number of responses in each variable set. The value is inputted as a percentage and has a default setting of 50%. To evaluate the data sets used in these studies the variables were considered by the software and included in the data analysis if the variable had values for responses above zero for more than or equal to 25% of the total number of observations for that particular variable.

The responses for each of the variables may then be scaled to one of four scaling types (None, Unit variance, (UV), Pareto (par) or control (Ctrl)). The final option to address in the work-set was to select the required model, in the research presented in Chapter 4,Chapter 5, and Chapter 6 partial least squared discriminant analysis PLS-DA and principal component analysis (PCA) models only, were used to classify the data obtained in each study.

The complete work-set could now have the model applied and generate component variables using the Autofit function in the SIMCA-P+ software. This function generates component variables until the maximum number of permissible variables for a particular module had been generated or until the predictive ability (Q2) of the model begins to degrade and the resulting Q2 value becomes lower than that of the previous component variable. In addition to the Q2 value for each component variable the SIMCA-P+ software also generates a Fit value (R2) for each variable.

Cross validation of the models generated in the SIMCA-P+ software is performed automatically and is accomplished using the “leave one out” method. The default conditions for the validation require the data to be divided into to 7 groups. A model is then generated for the data, which is devoid of one of these several groups and is repeated until all groups and therefore all variables have been left out. In each

instance a predictive residual sum of squares (PRESS) value is generated and if a new component has an elevated PRESS value compared with the previous component then the new component is retained.

Transformation of the model in to a visual format is achieved by generating a Score-plot where potential trends in the separation of the observation types can be assessed by hierarchical clustering. The separation of the different clusters of observations in the score-plots can then be attributed to the differences in the variable responses through loading plots or contribution plots.

When the Initial approach did not yield any separation of the observation types using the PCA approach the model type was modified to a supervised PLS-DA model. The model was refitted using the Autofit function and score plots for PLS-DA model were then generated and assessed. In addition to the score and loading plots the PLS-DA model also allows the generation of an S-Plot for each component variable. The S-plot enables the identification of the most reliable and most intense variables in the data and is a useful tool to identifying putative markers within groups of observations. The identified variables from the S-plot could then be extracted and re-modelled using the unsupervised PCA approach described above to both confirm and enhance the separation of the observations.

Chapter 4 It's all in the mind and sometimes in the breath

4.1 An introduction to psychological stress

Psychological stress (stress) is experienced by everyone and it is a healthy response to a challenging or threatening environment, evoking what is known as the general adaptation syndrome (the fight or flight response). It is conceivable that participants providing breath samples for diagnostic research during a visit to a clinical facility might be experiencing differing degrees of stress. The UK Health and Safety Executive (HSE) reported over 400,000 cases of psychological stress related complaints in the UK from 2009 to 2010[121]. Psychological stress disorders such as anxiety and depression have also been shown to have a higher prevalence in patients suffering from COPD[122] and lung cancer compared to healthy volunteers[123] Stress and anxiety may be a common concurrent factor in many breath samples.

Stress causes a cascade of hormonal releases. The most well known is the secretion of cortisol, but this is only one aspect of the body's response. Gluconeogenesis, glucogenolysis and lipolysis are also stimulated and increased and levels of renin and angeotensin II enzyme are produced causing elevation in blood-pressure (BP) and heart-rate (HR)[23]. Research into the exposure to stress prior to influenza vaccination has led to claimed increases in the level of antibody response in women[124] Long-term exposure to psychological stress has been reported to cause headaches[125], back pain[126] and reduced cognitive function[127]

Such observations lead to the research questions:

- Does stress change breath VOC profiles?
- and, if so, are putative markers of stress potential confounding factors for the diagnosis of disease by breath analysis.

At the outset of this study the observation and description of the effect of psychological stress on breath VOCs was unreported. This study, therefore, sought to engender a stress response in human breath using a Paced Auditory Serial Addition Task (PASAT) intervention[128,129].

An adjunct to this study was a tentative examination of the potential of objective stress monitoring through exhaled VOC markers.

4.2 Study objectives

- Evaluate the effect of a PASAT on the VOC composition of exhaled breath.
- Identify any VOCs in the breath samples that indicate a change in homeostasis after a stressful event.
- Develop an understanding of the potential origins of those VOCs present in which changes are observed and how this may potentially confound future studies into breath research.

4.3 Overview of the study

4.3.1 VOLUNTEER PARTICIPANTS

The ethical framework of this study has been reviewed in Section 3.2, and the protocols were reviewed and granted a favourable opinion the Local Ethics Advisory Committee (Ref: Loughborough University; G08-P7 and G09-P5). All participants taking part in the study gave written informed consent.

Male (n=10) and female (n=13) volunteers aged between 19 and 26 years participated in the study, and 22 complete sets of breath samples were obtained. The study used a randomised cross-over design with the participants attending two sampling sessions. In one session they were asked to sit comfortably and listen to relaxing music (neutral intervention), in the other they were asked to undertake cognitive and mental arithmetic tests (PASAT intervention). The mental arithmetic test (PASAT) was augmented with enhanced audio/visual alarms to intensify their perception of errors, and to mislead the participants with false-positive error alarms.

The sampling/test sessions ran in the morning. Before the sampling/test sessions the participants had fasted for a minimum of 9 hr and had drunk only water during this time. Further, the participants had refrained from washing or using oral hygiene products on the morning of their visit. In each session, two VOC breath samples were collected. The first breath sample was taken prior to beginning the

experimental interventions (neutral /PASAT intervention) and the second directly after completion of the intervention.

4.3.2 HEART-RATE AND BLOOD-PRESSURE MEASUREMENTS

Heart-rate and blood-pressure measurements were recorded using a Datex Ohmeda S/5 patient monitoring system. A finger clip was attached to the index finger of the participant's left hand in order to monitor the heart rate. A Critikon 23 -33 cm Dura-Cuf® blood-pressure cuff was applied to the participant's upper left arm. Three baseline blood-pressure and heart--rate readings were taken at 2 minute intervals. A further 12 readings were taken over the course of the experimental session at 12, 16, 19, 23, 30, 42, 45, 48, 49, 57, 60, 65 and 71 minutes. Heart rate, systolic blood-pressure, diastolic blood-pressure and the mean arterial pressure were recorded at each time point.

4.3.3 COGNITIVE FUNCTION TESTS

During each of the two visits the participants also undertook three cognitive function tests before and after the intervention period. The tests were a Stroop test, a Flanker test, and a Rapid Visual Information Processing test (RVIP). Each test lasted approximately 5 minutes and they were always given in the same order. In brief, the Stroop test, involved presenting the participants with words for a colour e.g. blue, red, green etc. The text colour for these words was different to the written colour presented. The aim was to discern the written word and not the colour it was presented in. The Flanker task required the participant to view an image of three arrows in a line, each pointing randomly left or right. The aim was identify correctly the direction of the centre arrow while ignoring the flanking arrows. Finally, the RVIP test presented the participants with a series of random numbers from 1 to 9 and required the participant to indicate when they observed three consecutive odd numbers. These tests were developed at the School of Sport, Exercise and Health Science at Loughborough University.

4.3.4 PACED AUDITORY SERIAL ADDITION TASK (PASAT)

The PASAT, originally developed to assess neurological injury[129], was used to induce stress in a reproducible manner in the participants[128]. The PASAT requires participants to perform mental arithmetic in which they listened to a recording of a

series of random single digit numbers (1 to 9) where each number was followed by a pause. Participants were instructed to take the first number they heard and add it to the second number. They were then required to state their result. If they were incorrect, a loud buzzer sounded to indicate a wrong answer while the next number in the series was presented for the participant to add to the last number they had been given and for them to once more state the result. The task ran for ten minutes with the frequency at which the numbers were presented increasing. At pre-defined points over the task, regardless of the accuracy of an answer the error-buzzer sounded; causing frustration, disorientation and feelings of self-doubt.

The control for the PASAT intervention session required participants to listen to ten minutes of calm classical music comprising four instrumental pieces; Truman Sleeps composed by Burkhard Dallwitz, Guitar Flute & String composed by Moby, Gabriel's Oboe composed by Ennio Morricone and Suite Bergamasque: Clair De Lune performed by Dame Moura Lympany.

The collected samples were then analysed by GC-MS under the conditions in Table 2.

4.4 Results and discussion

4.4.1 BLOOD PRESSURE AND HEART RATE

Individual heart rates, systolic blood pressures and diastolic blood pressures were collated and the average observations show the physiological responses to the PASAT and neutral interventions (Figure 22). The period of PASAT intervention (23 to 60 minutes) elicited a significant increase in these physiological responses, compared to the relaxed intervention where no discernible changes were observed. The baseline and cognitive function test portions of the two experimental sessions also showed no significant difference prior to the experimental intervention (0 to 23 minutes).

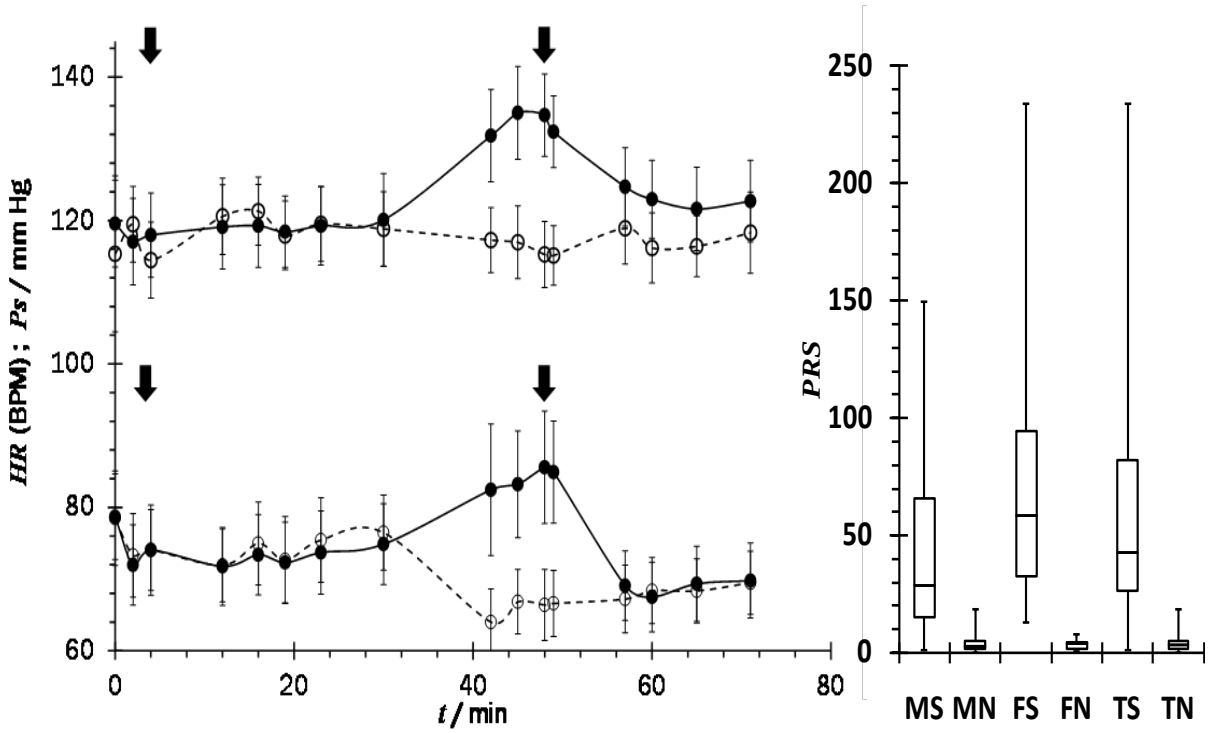


Figure 22 Heart rate and blood pressure responses in PASAT and neutral sessions. Left Averaged responses of 22 participants for systolic blood pressure (top), and heart rate (bottom) over the 70 minute PASAT (solid line) and relaxing music (dashed line) interventions. 0-23 minutes 1st cognitive function session, 24-48 minutes experimental intervention and 49-70 minutes 2nd cognitive function session. The error bars designate the 95 % confidence limit on the mean. Right Box-whisker summaries of the physiological stress response (PSR) scores classified by male stress (MS), male neutral (MN), Female stress (FS), Female neutral (FN) and Total stress (TS) and Total neutral (TN).

4.4.2 PASAT RESPONSE SCORE

In order to evaluate the extent at which a participant had been effected by the two interventions a PASAT Stress response score (PSR) was constructed. The participants' individual physiological responses to the PASAT and neutral interventions were scored on the basis of observed changes in their systolic blood pressure and heart rate. The average of the 3 initial baseline observations were used to give blood pressure and heart rate start values (P_S^{start} and HR^{start}). The maximum observed value of the four readings taken during the intervention periods (P_S^{max} and HR^{max}) were assigned the points of maximum "stress". Equation 22 defines a PASAT stress response score (*PSR*) derived in-house for the PASAT (PSR_S) and neutral interventions (PSR_N).

$$PSR = \frac{\Delta P_S}{P_S^{start}} \times \frac{\Delta HR}{HR^{start}} \times 1000$$ **Equation 22**

The *PSR* sought to emphasise the change in systolic blood pressure that occurred during the intervention (ΔP_S) and relate this back to the baseline observations(P_S^{Start}). Heart rate was also included in the same manner providing a ratio between the overall change (ΔHR) and the baseline values. The *PSR* score in Equation 22 was designed to give higher score for those individuals showing the largest physiological responses under the interventions. The *PSR* scores for the participants are summarised in Table 4 and Figure 22. Of the 22 participants 21 showed significant physiological responses to the PASAT intervention consistent with a stress response and this is reflected in their *PSR* score. The median *PSR* score during the neutral intervention was 3.5 compared to 428 for the PASAT intervention. Male and female responses were not statistically different with male/female participants' median *PSR* scores for neutral and PASAT interventions being 2.9/3.6 and 28.9/58.6 respectively. Female *PSR* scores showed a greater range of responses during the PASAT intervention (*PSR* = 12 to 233), while it was the male participants who had the greatest variability (*PSR* = 0.1 to 18.6) in the neutral state. Overall, the physiological responses as summarised by the *PSR* score indicated a significant acute response to the PASAT intervention that was consistent with the individuals experiencing psychological stress over the intervention period.

4.4.3 CHROMATOGRAPHIC DATA

The gas chromatographic-mass spectrometric (GC-MS) data were processed following a work-flow that has been described previously in Chapter 3. Retention times were correlated to a retention index scale constructed from a primary retention ladder generated from a quality assurance standard for the analysis, and a secondary retention index scale based on the retention times of five ubiquitous siloxane components present in all breath samples as a result of hydrolysis of the active phases within the analytical system. Figure 23 illustrates a typical example of the total ion current (TIC) chromatograms with both retention time and retention index values (RIU) obtained from the filter air blanks and breath samples collected under the two experimental sessions.

The GC-MS data were deconvolved and a typical breath sample would yield between 300 and 400 individual components. A breath matrix was constructed that contained 154 breath VOCs that were unique to breath and not found in any background sample. Each component was assigned a unique reference number. Saturn MS workstation software (Varian Inc, UK) was used to integrate peak areas for each breath component's extracted ion chromatograms (XIC) and both the ID assignments and the peak areas made up the breath matrix as described previously in Chapter 3.

Table 4 PASAT response scores of participants during the neutral (PSR_N) and PASAT intervention sessions (PSR_S).

Ref	1*	2	3*	4	5*	6	7*	8	9*	10	11*	12	13*	14	15*	16	17*	18	19*	20	21*	22
M/F	M	F	M	M	M	F	F	M	F	M	M	F	F	F	M	M	F	F	F	F	F	M
	PASAT intervention																					
HR^{max}	90	119	70	92	80	77	105	91	90	90	96	100	82	134	95	84	146	66	97	95	105	85
HR^{start}	78	94	54	90	69	68	81	76	66	73	59	66	72	84	75	76	88	42	82	61	65	54
P_S^{start}	141	121	112	136	121	113	105	120	129	141	113	119	102	117	107	138	112	112	116	126	110	144
P_S^{max}	152	148	122	142	144	124	125	136	155	168	129	138	117	146	121	153	152	120	133	154	132	162
PSR_S	11.6	60.6	27.5	0.9	30.3	12.9	56.7	26.1	74.2	44.8	87.6	84.5	19.7	149	34	11.1	233	40.8	27.5	125	125	72.6
	Neutral intervention																					
HR^{max}	81	102	65	72	74	83	98	85	69	101	71	89	101	73	82	95	113	54	79	94	71	70
HR^{start}	75	90	58	65	68	79	85	82	61	81	60	72	84	56	73	78	81	49	60	79	68	56
P_S^{start}	126	120	125	118	120	107	102	125	120	135	113	113	111	116	107	133	118	108	120	119	103	147
P_S^{max}	127	121	125	124	122	122	101	118	123	138	122	111	109	123	114	131	117	107	122	116	105	148
PSR_N	1.0	0.8	0.1	5.3	1.6	7.4	1.9	2.0	3.2	4.7	15.5	5.0	4.2	18.6	7.7	3.9	2.7	1.2	4.5	4.5	1.0	1.4

Note: The suffix * indicates that the participants were subjected to the PASAT session on their first visit.

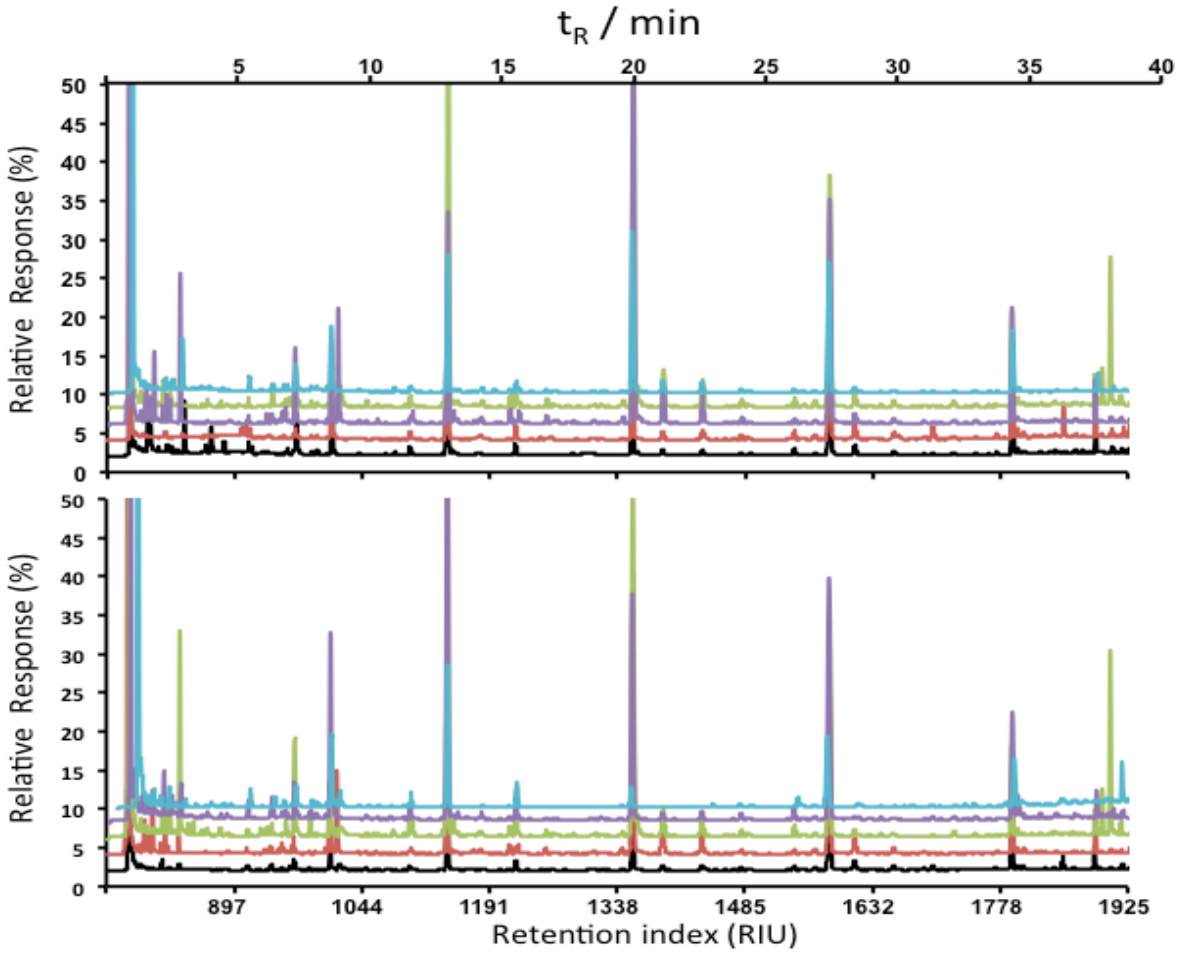


Figure 23 Examples of total ion current (TIC) chromatograms of five participants [Top] relaxed session and [Bottom] stressed session breath samples. Lower x-axis shows retention index units (RIU) and upper x-axis illustrates the equivalent retention time (tR), overlaid chromatograms are offset by 2%

4.4.4 MULTI-VARIATE ANALYSIS (MVA)

The breath data matrix was subject to MVA using SIMCA-P+ software package. Those compounds that were present in less than 30% of the breath samples collected were excluded from the data modelling, this value ensured that potential present / absence compounds would not be excluded. This inclusion criterion yielded a resultant data set of 50 breath components that were used in the subsequent multi-variate analysis (MVA).

The breath data matrix consisted of the peak areas of the 50 VOC breath components (variables) against the participants (observations). MVA was conducted using SIMCA-P+ software (Version 12, Umetrics, UK). Partial least squared discriminate analysis (PLS-DA) was initially performed on the post-intervention

samples from both PASAT and neutral experimental sessions. All variables were assigned to a single block and a weighting of $1/\sqrt{Block}$ was applied making the total variance of all variables equal to 1. In addition Pareto Variance (ParN) was selected for base scaling of the data set; this scaling method takes each variable, subtracts it from the mean of that variable set and divides it by the square root of the standard deviation for the variable set[130]. The S-plot generated from this statistical model (Figure 24) indicated six possible VOC variables (α1 to α6) that changed in response to interventions experienced by the participants.

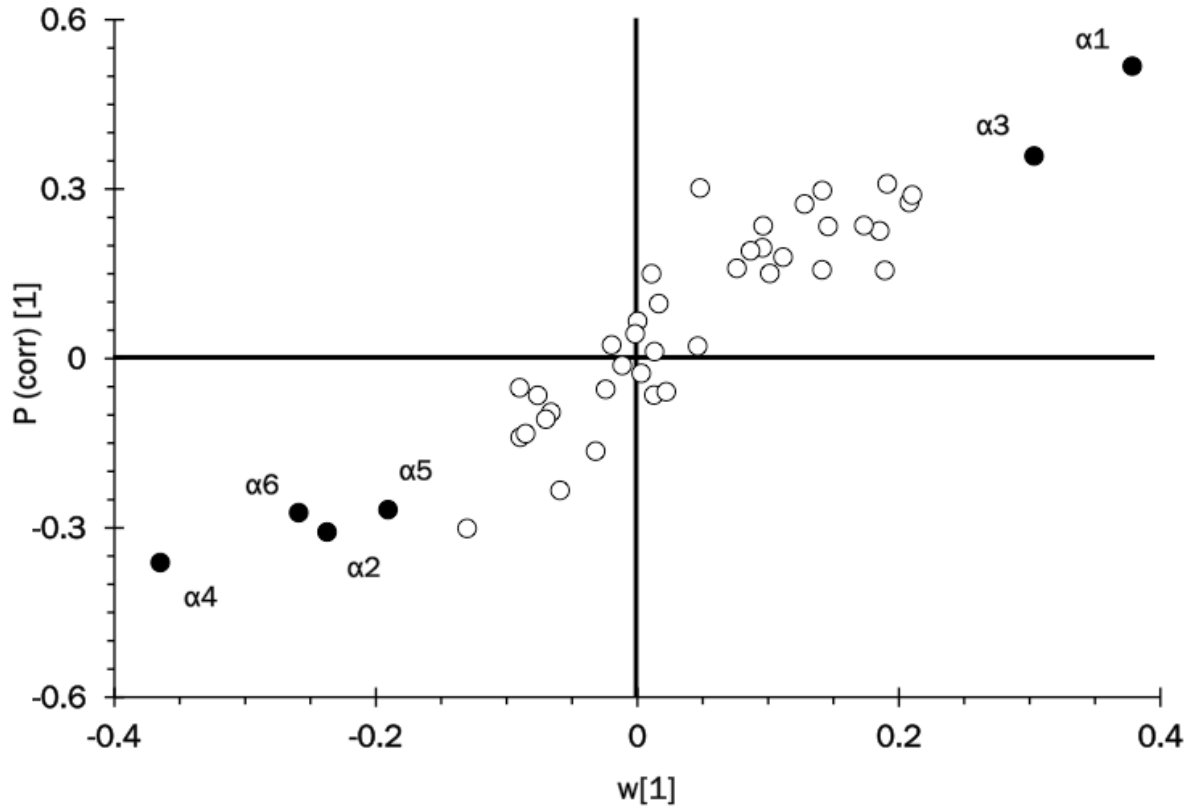


Figure 24 S-Plot of variables used for discrimination of stressed and neutral states S-plot generated from the modelled covariance (P(corr)[1]) and modelled correlation (w[1])] of the partial least square discriminate analysis (PLS-DA) from the intensities (peak areas) of 50 breath components (variables) from 44 breath samples (observations). The six variables highlighted (black dots) were found to provide the highest magnitude (intensity) of the variables in the data set and greatest reliability (P(corr)[1] axis) of the variables in the data set; indicating potential putative markers sensitive to PASAT intervention

Principle component analysis (PCA) of these six VOC variables enabled the separation between the PASAT and neutral observations to be evaluated. The score plots generated from these data are given in Figure 25. In each plot responses from the PASAT and neutral interventions are seen to cluster and a distinction between breath profiles obtained under the two experimental sessions can be observed. The principle components identified exhibited 44.1% [PC1] and 15.7% [PC2] of the total explained variance of the data set for male participants yielding a 100% sensitivity (% of correctly assigned true positive observations) for the male PASAT observations and 90% selectivity (% of correctly assigned true negative observation) for the neutral observations. In contrast to this the female observations produced a

weaker model and yielded a sensitivity of 83.3% and a selectivity of 91.6%. The total variance modelled by each component was 34.5% and 28.4% for [PC1] and [PC2] respectively. Further when the multivariate results were cross-correlated to the *PSR* scores the outliers were found to have anomalous *PSR* scores. Participants 4 and 6 had low *PSR* scores during the PASAT intervention and Participants 11 and 14 scored higher *PSR* score in the neutral intervention than any of the other participants (Table 4) finally participant 17 had the largest *PSR* score during the PASAT intervention and was classed as an outliner by the multivariate analysis.

4.4.5 PRELIMINARY IDENTIFICATION OF PASAT SENSITIVE BREATH COMPONENTS

Elucidation of the VOC components α1 to α6 was based on three tests. The first was a satisfactory match of the de-convoluted spectrum generated from Analyzer Pro to the respective entry in the NIST MS library (NIST05). Satisfactory is defined in this instance as forward (direct matching of spectrum peaks to library reference spectrum) and reverse (matching by omitting spectrum peak not present in the library spectrum) match factor with a value greater than, or equal to, 650[131], see Table 5.

The second test was a satisfactory match between the observed retention index value (IU) and the estimated retention index value obtained from NIST05 MS library. Satisfactory is defined as a difference not greater than ±50 IU. The recorded retention index value for α2, see Table 5, of 1063 IU for BRI1063-105-77-52-106-51 (indicated by mass spectrometric library matching to be 2-hydroxy-1-phenylethanone) did not match the estimated NIST retention index value of 1272 IU. No record of an experimental retention index for this compound for a DB5 (5% phenyl 95 % methyl) stationary phase appears to be currently available.

The third and final test was based on matching the preliminary assignments in Table 5 to standards analysed under the same conditions (chapter 3). An adsorbent sampling tube was spiked with 0.1 µl of a 50 µg ml-1 methanol solution of: indole (α1), 2-hydroxy-1-phenylethanone (α2), 1,4-cyclohexadiene, 1-methyl-4-(1-methylethyl) (α4), benzaldehyde (α5), 2-ethylhexan-1-ol (α6); all reagents were purchased from Aldrich. The on-column mass of each of the components was 500

pg. 2-methylpentadecane (α3) was not available commercially and was not included in this test.

These three tests confirmed the assignments of indole, 2-hydroxy-1-phenylethanone, benzaldehyde, and 2-ethylhexan-1-ol. α4 was not confirmed as 1,4-cyclohexadiene, 1-methyl-4-(1-methylethyl) however, the mass spectrum indicates that α4 was likely to be a terpene. 2-methylpentadecane (α3) remains putatively annotated and Table 5 summarises the PASAT sensitive components.

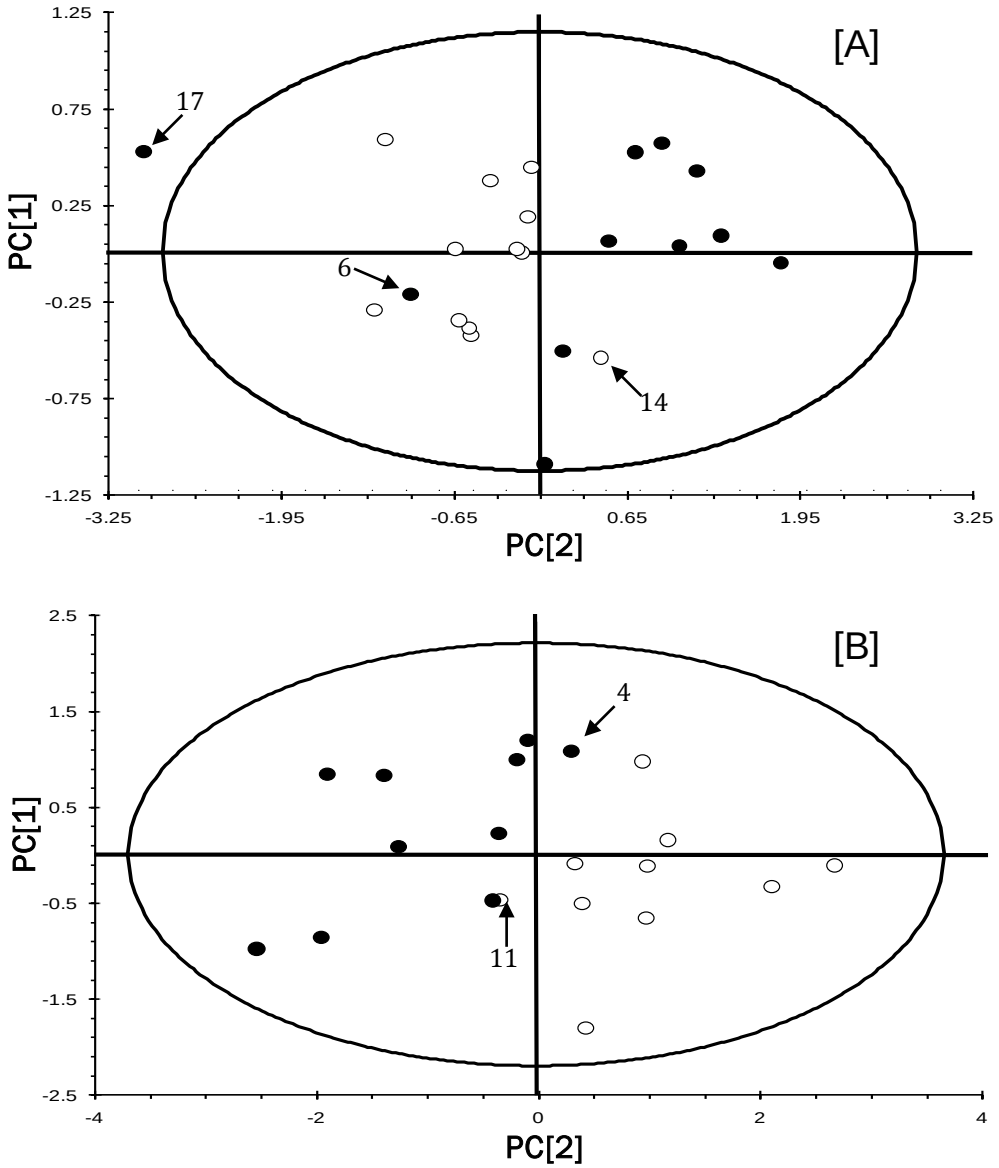


Figure 25 Unsupervised principle component analysis (PCA) of six stress sensitive breath components identified from PLS-DA (Figure 24) model on [A] 12 observations from the stressed (black dots) experimental session and 12 observations from relaxed (grey squares) experimental session from female participants. [B] 10 observations from the stressed experimental session and 10 observations from the relaxed experimental session from male participants. 100 % sensitivity was obtained for the male stressed observations and 90% selectivity was obtained for the relaxed observations. In contrast to this the female observations yielded a sensitivity of 83.3% and a selectivity of 91.6% was observed. Outliers are annotated with participant references given in Table 4

Table 5 Identification of PASAT sensitive breath components C1 to C6 with reverse and forward spectral match and de-convoluted spectrum lists.

Ref	Name and library entry	Deconvolved mass spectra	MS F/R	RI O/E	CAS
α1	Indole BRI-1286-117-90-89-63-73	117 (999); 90 (411); 89 (310); 63 (151); 73 (130)	850/889	1286/1276	120-72-9
α2	2-hydroxy-1-phenylethanone BRI-1063-105-77-52-106-51	105 (999); 77 (311); 52 (172); 106 (73); 51 (66)	889/885	1063/1272	118-93-4
#α3	2-methylpentadecane BRI-1527-41-57-43-85	41 (999); 57 (786); 71 (776); 43 (619); 85 (412)	650/658	1527/1548	1560-93-6
*α4	unknown terpene compound BRI-951-93-91-92-79-121	93 999); 91 (435); 92 (430); 79(264);121 (217)	-	951/ -	-
α5	Benzaldehyde BRI-976-105-77-52-106-95	105 (999); 77 (239); 52 (222); 106 (103); 95 (72)	813/813	976/982	100-52-7
α6	2-ethylhexan-1-ol BRI-1019-41-57-70-85-55	41 (999); 57 (756); 70 (620); 85 (543); 55 (389)	732/830	1019/995	106-76-7

Note: MS F/R Forward (F) and reverse (R) mass spectral match of deconvolved mass spectrum to NIST05 library search.

RI O/E Observed (O) and estimated (E) retention index value for deconvolved peak.

 # Putatively annotated by spectral match with NIST05 Library only as reference standard is not commercially available.

*** BRI-1527-41-57-43-85 NIST match indicates 1,4-Cyclohexadiene, 1-methyl-4-(1-methylethyl), this was not confirmed with a reference standard.**

The origins of these VOCs in breath are not well defined and may be generated through a combination of independent mechanisms. Benzaldehyde is a common food component and occurs naturally in fruits like apples and honey[132,133] likewise terpenes occur naturally in foodstuffs[134]. These components may be due to dietary intake and may not have an endogenous origin. Indole is associated with the production of the essential amino acid tryptophan, which in turn is part of the pathway that produces serotonin; previously reported to be involved in cardiovascular and psychological response to acute stress[135] Indole may also be generated from tryptophan metabolism by indole positive bacteria in the gastro intestinal system. 2-Ethylhexan-1-ol has been associated with polyvinylchloride (PVC)[136] and also as a potential VOC marker of lung cancer[137] In addition 2-ethylhexan-1-ol has also been observed in the headspace of in-vitro cultures of lung carcinoma A549 cells[138]. Methylated hydrocarbons have been reported as the endogenous products of oxidative stress and numerous VOCs in this class have been identified as potential biomarkers of various cancers[40,138]

Example responses from Participant 14, for these six compounds are presented in Figure 26. Indole and 2-methylpentadecane exhibit an increased or up-regulated response during the PASAT intervention. Indole signals increased from +140% to +360% in response to the PASAT intervention, compared to the Neutral intervention samples. 2-methylpentadecane increased to a lesser extent, between +30% to +80% increases compared to the Neutral intervention. The remaining four components show a reduced or down-regulated response compared to the neutral Intervention.

The collective scores for the six VOC markers were combined by taking their Euclidian distance, based on the peak intensities (I_R) to generate a marker score (*CMS*), and the box whisker plots in Figure 27. The median *CMS* for the neutral intervention was 5.42×10^4 compared to 4.62×10^4 for the PASAT intervention. Female participants exhibited the greatest range and variability during the PASAT intervention. In contrast male participants showed the widest range and variability during the neutral intervention. In comparison to the *PSR* the *CMS* results show equivalent trends in the female participants with an increase in range and variability in breath samples collected under

PASAT intervention when compared to the results from the neutral intervention. This correlation was not reflected in the male participants, one reason for this may be due to the lower intensities observed for the down regulated breath components indicating a higher rate of clearance.

$$CMS = \sqrt{\sum_{\alpha=1}^{6} I_R^2}$$

Equation 23

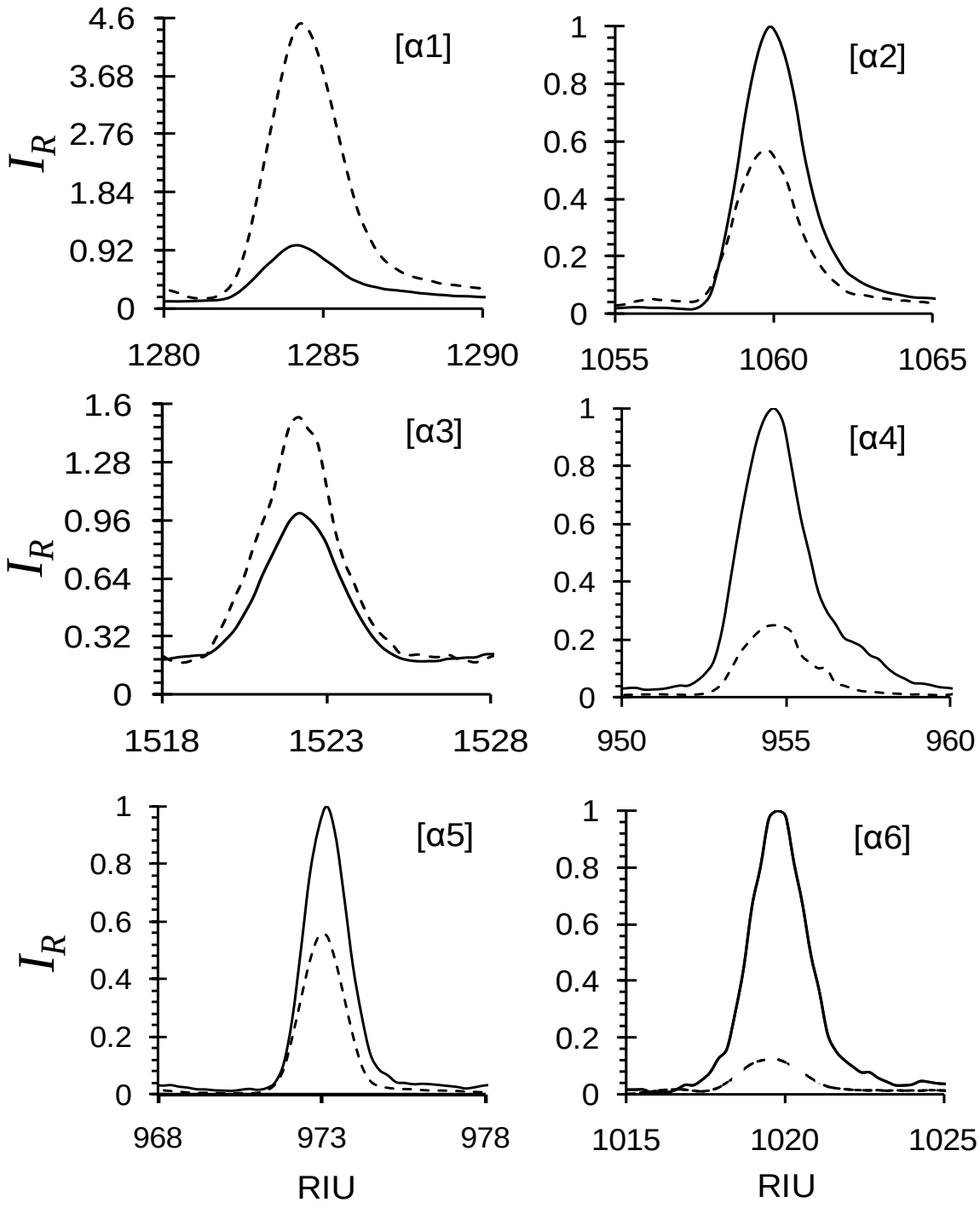


Figure 26 Overlaid extracted ion chromatogram (XIC) responses for stressed experimental session (dashed line) and relaxed experimental session (solid line) for stress sensitive breath components, Indole (α1), 2-hydroxy-1-phenylethanone (α2), 2-methylpentadecane (α3), unknown terpene (α4), benzaldehyde (α5) and 2-ethylhexan-1-ol (α6) from a single male participant. The intensities (IR) displayed have been normalised with respect to the relaxed experimental session to indicate those components that have been up-regulated and down-regulated as a result of undertaking the PASAT.

4.5 Study summary

Changes in the exhaled breath VOC profiles of healthy volunteers were observed in participants who underwent a PASAT intervention. Four of the VOCs have been identified through mass spectral, retention index and chemical standard matches. A combination of all six components enabled the population to be classified as PASAT Responsive or Neutral with a sensitivity of 90.9% and a selectivity of 90.9%. These data correlated with their physiological scoring; the *PSR* score, and the individuals who were not correctly classified exhibited possible stress, induced by the unfamiliar surroundings of the test rooms and laboratory environments, or who had low *PSR* scores; indicative of not experiencing stress during the PASAT intervention.

The biological origin and role of these VOCs in breath has yet to be completely described. The four down-regulated VOCs may have been depleted through enhanced ventilation from elevated heart-rate, and respiratory rate. In contrast 6-methyl-pentadecane and indole, which were up–regulated, may be indicative of biochemical response to the PASAT intervention.

The PASAT intervention has been developed specifically to induce signs and symptoms associated with stress in the participants. These preliminary findings are indicative of the existence of a stress-related response in the VOC profiles of humans. The concentrations involved may be estimated at this stage to range from 20 ng m^{-3} for indole to 4.8 µg m^{-3} for benzaldehyde and targeted studies with labelled standards would be a logical follow-up to this study.

Demonstrations of the utility of breath analysis in providing next generation in-community diagnostic capability through monitoring biochemical changes in the human body continue to be reported[70]. This study indicates that the human VOC profile may also be sensitive to non-physical stimuli. Physical disease may not be the only contributing factor in breath profiling and the acquisition of samples and data from people undergoing traumatic and

emotionally challenging diagnoses associated with serious diseases may need to consider this factor.

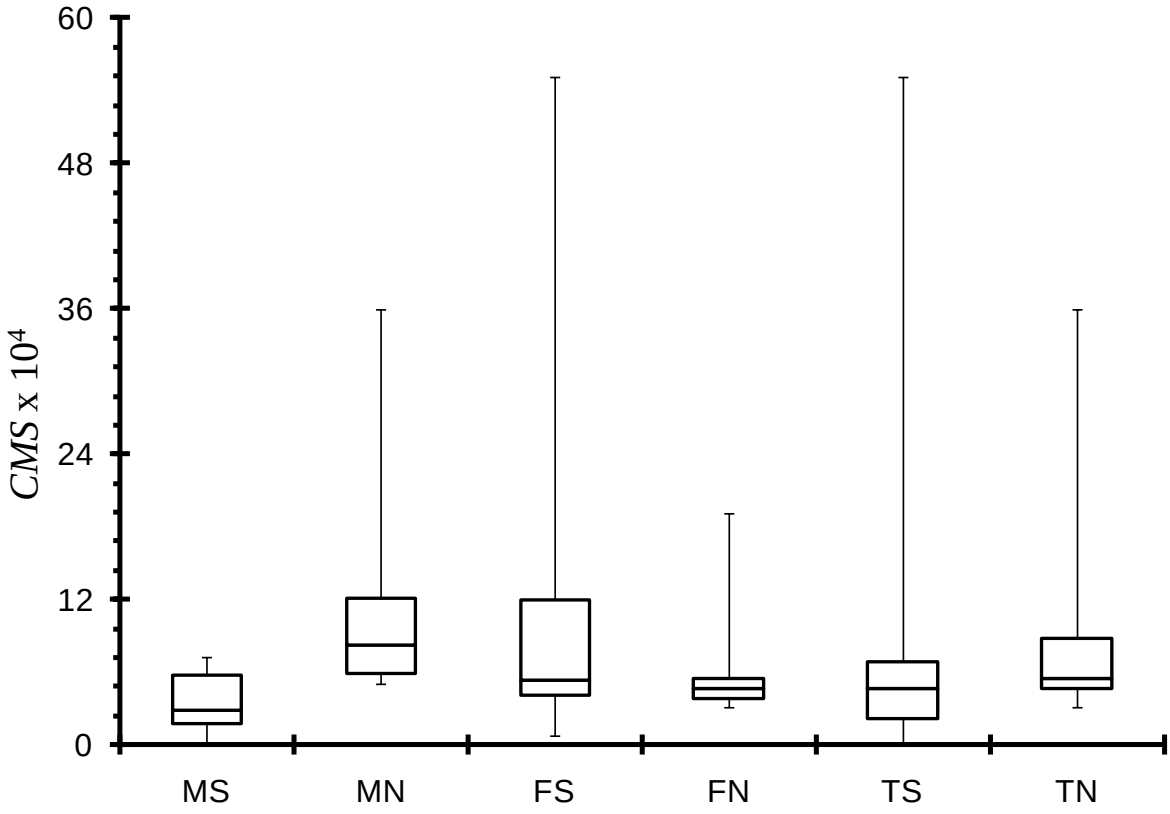


Figure 27 Box-whisker summaries of the combined marker scores (CMS) for stressed and neutral participants, classified by male stressed and neutral interventions (MS and MN), female stressed and neutral (FS and FN) and the total group stressed and neutral (TS and TN) results.

Chapter 5 Changes in VOC profiles of exhaled breath before and after induced physiological stress

5.1 Introduction to VOCs and exercise

During the onset of exercise the human body responds by increasing gas-exchange in the lungs and the independent regulation of organ blood flow. As a result of this new internal environment the biochemical reactions in the body change to accommodate the increased workload and to keep the human body functioning. At rest the average cardiac output is approximately 5 dm^3min^{-1}, and this increases rapidly during moderate exercise to somewhere in the region of 25 dm^3 min^{-1} Figure 28 highlights this[23].

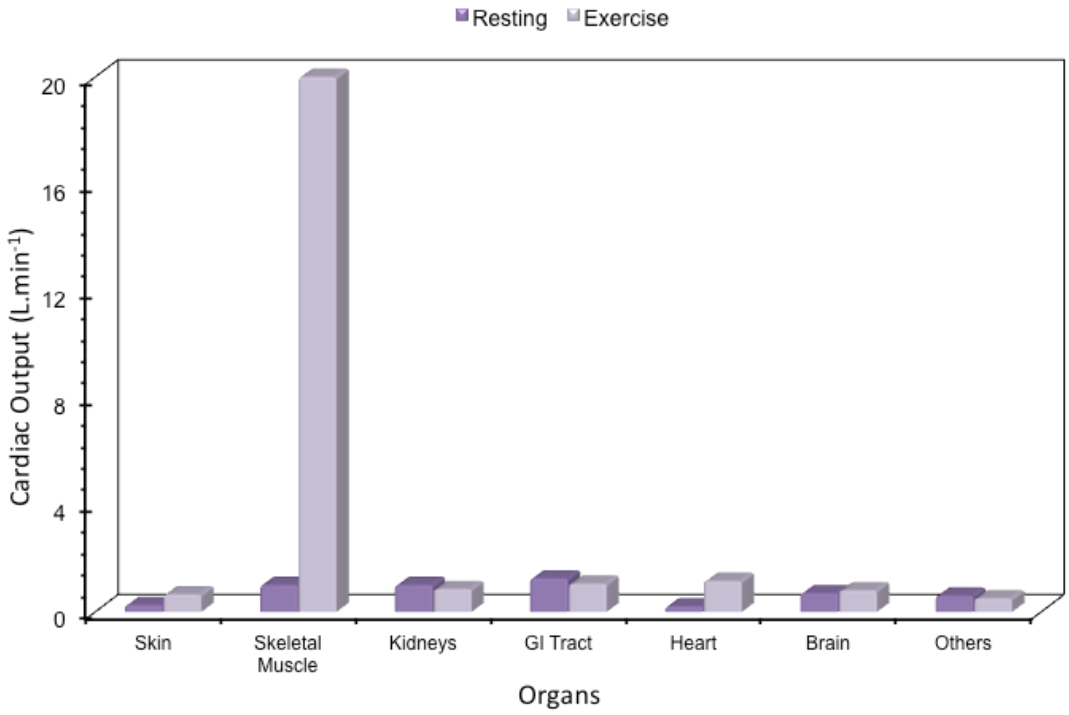


Figure 28 Bar Chart illustrating the approximate proportion blood of the total cardiac output at rest (5 dm3 min-1) and during exercise (25 dm3 min-1) to organs of the body (graph drawn from data presented in Stanfield, C. L. Principles of Human Physiology; 4th ed.; Pearson, 2009)

At rest skeletal muscle receives approximately 20% of the total cardiac output and this increases to approximately 80% when the body is engaged in

exercise. It is helpful to note that both the skin and heart show an increase in blood flow during exercise. Blood-flow in the gastro intestinal (GI) tract and kidneys, decreases as the metabolic demands in these areas is low during exercise[23].

Gas exchange in the lungs increases under physical exercise with the intake of O_2 and output of CO_2 increasing to over 10 times the resting-rate. This is characterised by the respiratory exchange ratio (RER); the ratio of the volume of CO_2 expired per unit time (V_{CO_2}) to the volume of O_2 inspired per unit time (V_{O_2}) inspired (equation 22). At rest the RER ≈ 0.8, and during exercise when there is a greater reliance on carbohydrates for energy production the RER ≈ 1.0[139].

During exercise the O_2 uptake or V_{O_2} increases linearly with the workload until the V_{O_2} becomes constant; known as the $V_{O_2max.}$ Workloads above the V_{O_2max} require the body to use anaerobic glycolysis for the production of energy.

$$RER = \frac{V_{CO_2}}{V_{O_2}}$$

Equation 24

Monitoring of breath VOCs during exercise has been approached by a number of different analytical techniques and include PTR-MS[82,140], sensors[141] API- MS[142] and SPME-GC-MS[143]. The most common technique used to monitor VOCs in breath during exercise has been PTR-MS as continual monitoring during the exercise regime can be achieved.

The exercise protocols presented in the literature used to induce physical stress though exercise have incorporated both resting and active exercise phases. The workloads on the exercise phases of the protocols ranged between 50 W and 100 W over time periods between 5 to 15 min. The resting phases, which are dispersed throughout the exercise protocol, lasted between 3 to 12 min. Further studies presented in the literature have negated any resting phases and precede the exercise phase by a warm up period. Other exercise phases used have included a stepped workload. In these scenarios

the literature reports the workload was increased in steps of 20 W every 3 min over a period of 24 mins[144]. Conversely the exercise protocols have also been tailored to the individual participants with a study using a 20 $W.min^{-1}$ incremental workload up to a sub maximum of 80 % of the participants' maximum workload[142]

The cohorts reported in the literature for studies involving breath VOCs during exercise have been on small pilot study scale with many cohort sizes in single figures or low teens[82,142,144]. Participant panels have included populations of both sexes and have incorporated both smokers and non-smokers. The age range reported has also been broad, ranging from early 20's to mid 50's[142]

The changes in acetone and isoprene profiles under different exercise protocols using an ergometer have been reported[82,142,144]. An increase in both acetone and isoprene during the exercise phases has been shown, however they did not follow similar profiles. Acetone increases in response rapidly until it reaches a plateau, the plateau then continued until exercise stopped and then reduces rapidly[82]. This trend with acetone appears to happen consistently throughout each exercise phase with responses of similar concentration in each subsequent exercise phase when rest periods are incorporated in the protocol. The isoprene response differs markedly to that of acetone: in isoprene there is an abrupt increase in the response that is followed by a rapid decline. This response profile is repeated over each phase when rest periods are included however in each subsequent exercise phase the initial response observed is lower than the previous phase. The isoprene and acetone profiles have been confirmed across different research groups, however differences in the profiles of acetone were observed when the exercise protocol was based on continuous exercise. In contrast to the protocols using rest periods the acetone also follows initially the same raise and plateau, however when an extended exercise protocol was used, acetone exhibits a reduction in response after approximately 20 min of continual exercise. King *et al* concluded the observed profiles could be ascribed to different exhalation patterns during exercise in the case of acetone, while changes observed for isoprene were due mainly to changes in pulmonary gas exchange[82].

In addition to PTR-MS, VOCs have been monitored simultaneously using SPME-GC-MS for four VOCs in expired breath[144]. Acetone, isoprene, acetaldehyde and hexanal have been monitored as they are produced through different metabolic pathways: lipolysis, cholesterol biosynthesis, a metabolic product of ethanol, and oxidative stress, respectively. Only the PTR-MS data was reported from this comparison study, however it was reported that the data from PTR-MS and SPME-GC-MS follow similar trends with the exception of the initial warm up phase where it is commented that the two data sets diverge.

One of the few approaches to the analysis of VOCs in breath during exercise that does not utilize PTR-MS is the use of APCI-MS for real time analysis of VOCs. The expired breath is passed in to an APCI source set in negative ionization mode. This study reported numerous responses of VOCs that show variation during exercise, however no VOC identities were elucidated and only the m/z ratios were reported.

At the outset of this study there was little research into the VOCs in breath associated with exercise and recovery using a global metabolic approach with TD-GC-MS.

5.2 Study objectives

- Evaluate the effect of moderate exercise on the VOC composition of exhaled breath.
- Identify possible VOC markers of recovery to homeostasis following exercise.
- Develop an understanding of the potential origins of those VOCs in which changes are observed and how this may potentially confound future studies into breath research.

5.3 Study overview

The study panel consisted of 10 male participants aged between 18 and 35. All participants were healthy non-smokers and gave informed consent prior to taking part. This study used a combined sampling approach to collect breath,

skin VOCs, saliva, urine and blood. The local ethical committee at Loughborough University (G09-P5) approved the sampling protocols used in the study.

The experimental design comprised four Phases; Phases A to D. Each phase lasted approximately 30 min. Phases A, C and D were resting phases where the participants were required to sit upright in a chair. Phase B was the exercise fraction of the study and during this phase participants were relocated onto an ergometer and required to peddle for 30 min at a resistance based on the participant's body weight. The resistance was set at 2 $W.Kg^{-1}$. A comprehensive timeline for the experiment if given in Figure 29

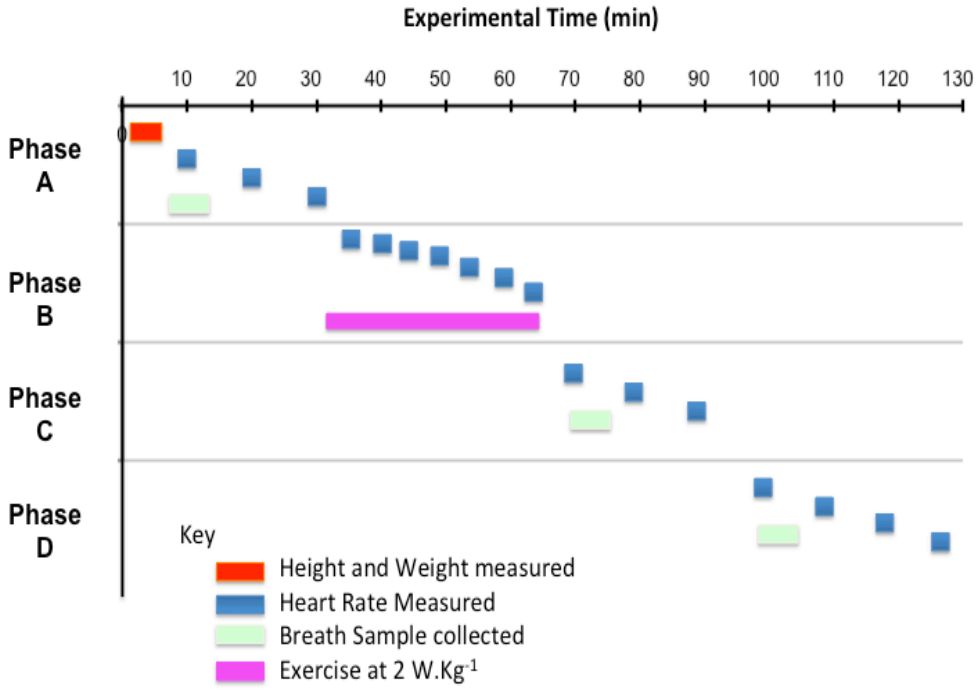


Figure 29 Experimental timeline of samples in each phase of the experiment

Participants undertaking the study were asked to arrive without exerting themselves (i.e. Cycling or running) at the sampling station in one of two slots (7:30 am or 10:30 am), having not brushed their teeth or washed. The participants were further briefed and shown the exercise facilities, they then had the chance to ask any questions and be shown how to adjust and use the ergometer. Each participant's weight and height was recorded (Table 6) and they were fitted with a heart rate monitor. The participant's heart rate was

recorded every 10 min throughout Phase A, Phase C and Phase D and every 5 min of Phase B of the experiment. These data are summarized in Table 7.

Table 6 Summary of participant height, weight and calculated body mass index (BMI).

Participant ID	Weight /Kg	Height / cm	Calculated BMI
CZ01	69.64	176.9	22.3
CZ02	73.46	173.5	24.4
CZ03	82.92	189.3	23.1
CZ04	83.52	196.0	21.7
CZ05	72.20	174.2	23.8
CZ06	74.02	181.8	22.4
CZ07	92.08	191.5	25.1
CZ08	88.12	177.9	27.8
CZ09	100.04	176.5	32.1
CZ10	88.66	168.0	31.4

The participants provided three breath samples, see chapter 3. The first sample was collected at the start of Phase A. The second sample was collected at the start of Phase C and the final sample was collected at the beginning of Phase D. Breath samples were not collected during Phase B due to the impracticality of using the adaptive breath sampler on the ergometer. A volume of 2.0 dm^3 of exhaled breath was collected at each sampling point.

In addition to the breath samples a 2.0 dm^3 air samples was acquired of the filtered air used to supply the masks. All samples were stored at 4^oC for up to 14 days until analysis by TD-GC-MS was undertaken. The analysis conditions and quality procedures are outlined in detail in chapter 3.

Table 7 Summary of recorded heart rate for the ten participants throughout each phase of the experiment.

Phase	t.min^{-1}	Participant Heart Rate (BPM)									
		CZ01	CZ02	CZ03	CZ04	CZ05	CZ06	CZ07	CZ08	CZ09	CZ10
A	0	52	63	72	60	76	56	78	76	90	57
A	10	59	61	72	58	75	64	76	55	86	62
A	20	52	59	80	63	60	66	73	75	81	57
A	30	67	71	80	59	70	57	74	71	79	55
B	35	101	92	69	74	96	109	109	102	74	131
B	40	120	126	130	131	123	103	133	139	126	147
B	45	129	148	138	141	149	108	149	156	197	154
B	50	137	157	151	149	162	111	160	165	210	141
B	55	144	167	152	153	169	111	167	173	218	148
B	60	149	168	157	165	173	114	164	175	209	149
B	65	152	173	159	165	183	113	168	176	215	149
C	70	90	171	N/R	169	188	114	182	122	213	85
C	80	73	102	93	90	135	76	101	114	120	67
C	90	80	82	85	73	105	69	97	97	110	73
D	100	78	78	67	73	91	65	93	98	111	71
D	110	66	68	75	70	83	67	94	89	96	63
D	120	69	76	72	73	80	67	87	82	89	66
D	130	62	61	67	64	78	67	84	77	88	61

5.4 Results

5.4.1 HEART RATE RESULTS

The lowest resting heart rate recorded in phase A was 52 beats per min (BPM) and the highest recorded heart rate in phase A was 90 BPM. As the exercise portion of the experiment began a rise in heart rate was observed with a maximum heart rate of 218 BPM being recorded from a single participant during Phase B. The maximum change observed between resting and exercise was 144 BPM and plot of the percentage (%) change in heart rate above the initial resting value is presented in Figure 30. Using the exercise protocol described previously in this chapter the trend shown indicates that a significant increase in heart rate was observed for all

participants over the course of Phase B. In the post exercise portion of the study (Phase C) a rapid decline in heart rate was observed over 30 mins and by the end of the second recovery phase (Phase D) the heart rate had returned to a rate that was not significantly different from the initial heart rate readings.

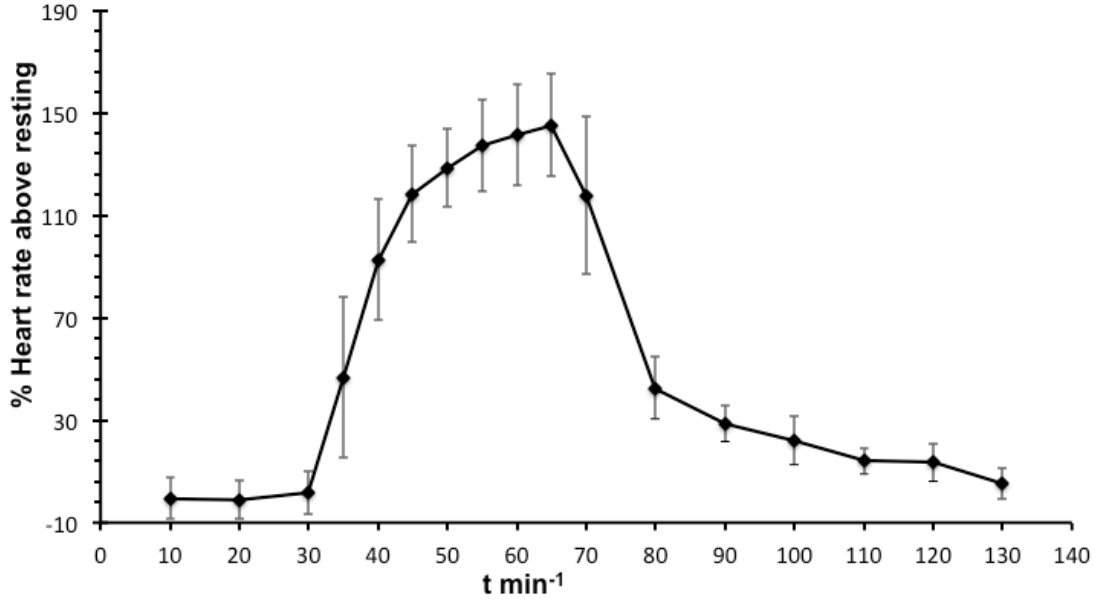


Figure 30 Plot of the mean % heart rate change above the initial recorded heart rate reading verses the experimental time in min. Error bar shown are 95% confidence intervals.

5.4.2 ACETONE AND ISOPRENE RESULTS

The assessment of the TD-GC-MS data generated from this study was performed initially using the acetone and isoprene responses obtained from phases A C and D. The chromatographic responses for each VOC were integrated and the peak areas interrogated. These data have been graphically summarized as box and whisker plots in Figure 31 and Figure 32. This approach was adopted to qualitatively assess and contrast how the observed data compares to those presented in current literature. In addition the heart rate data from phases A, C and D have also been summarized in box whisker plots (Figure 33) to aid comparison with the acetone and isoprene data.

The observed responses for isoprene exhibit similar interquartile ranges for each phase of the experimental study with phase C showing the greatest

variability overall. The interquartile ranges of phase A and C show similar variation and show similar distribution of responses above and below the median line on each plot. In contrast phase D shows a similar interquartile variation as A and C, however the majority of responses in the lower 50% of the interquartile range exhibit more variation that the upper 50% above the median which is the opposite of the distributions observed in phases A and C.

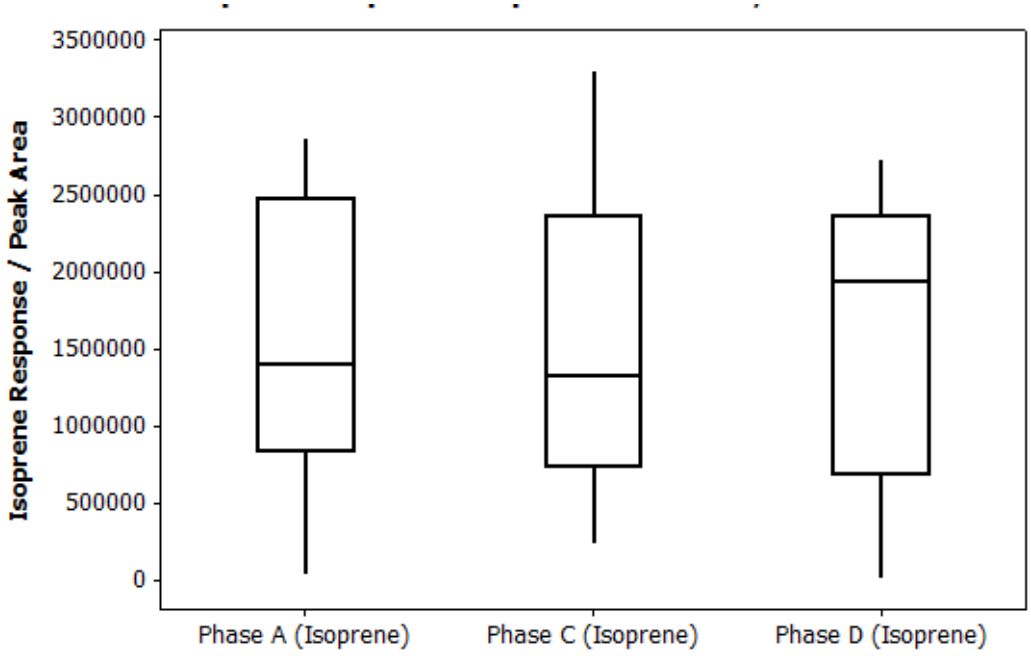


Figure 31 Boxplot of observed isoprene responses in Phases A, C and D.

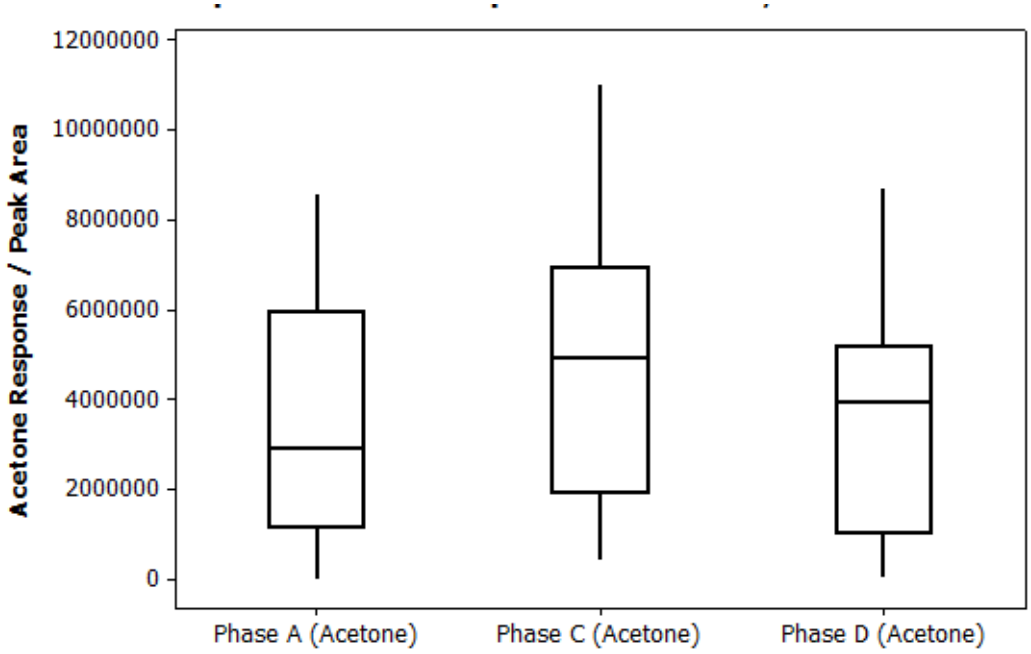


Figure 32 Boxplot of observed acetone responses in Phases A, C and D.

Study of the acetone responses reveals differences from the isoprene. The box whisker plots show similar interquartile variability between the three experimental phases and with the greatest range of intensities being observed in phase C. The median trend also differs from that of isoprene showing a higher median in phase C, which then decreases in the second rest phase (phase D). Again with Phase D the lower 50% of responses exhibit a higher degree of variation.

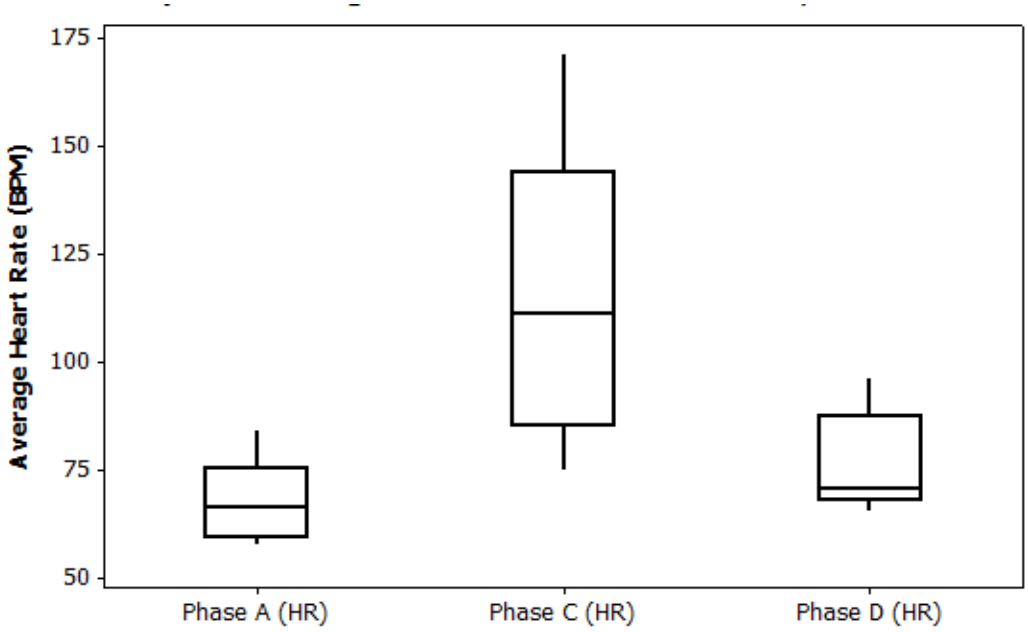


Figure 33 Boxplot of observed heart rate responses in Phases A, C and D

The results of the heart rates, graphically represented in the box whisker plots (Figure 32) for the 3 phases; indicates how the participants' heart rates varied in each phase. Phase A illustrates the heart rate of the participants at rest with a little variation and an interquartile range that is evenly distributed around the median with a variation of less than 25 BPM. This has significantly changed directly after exercise with minimal overlap of the phase A and C plots. The Phase C plots as with the Acetone and isoprene shows the largest variation overall. Phase C plot exhibits partial overlap of with phase D with phase D showing a trend that mostly aligns with Phase A, with the exception of the upper 50% exhibiting an increased variation above the median when compared to the data observed for Phase A.

5.4.3. CORRELATION OF HEART RATE WITH ACETONE AND isoprene responses.

The mean peak area responses for both acetone and isoprene were plotted against the average heart rate observed in each phase. The resultant graphs are given in Figure 34. The data yielded three points for each graph, the graph produced for isoprene responses and heart rate exhibited no significant correlation. The resultant graph for acetone shows the general trend between the mean acetone response and the mean heart rate.

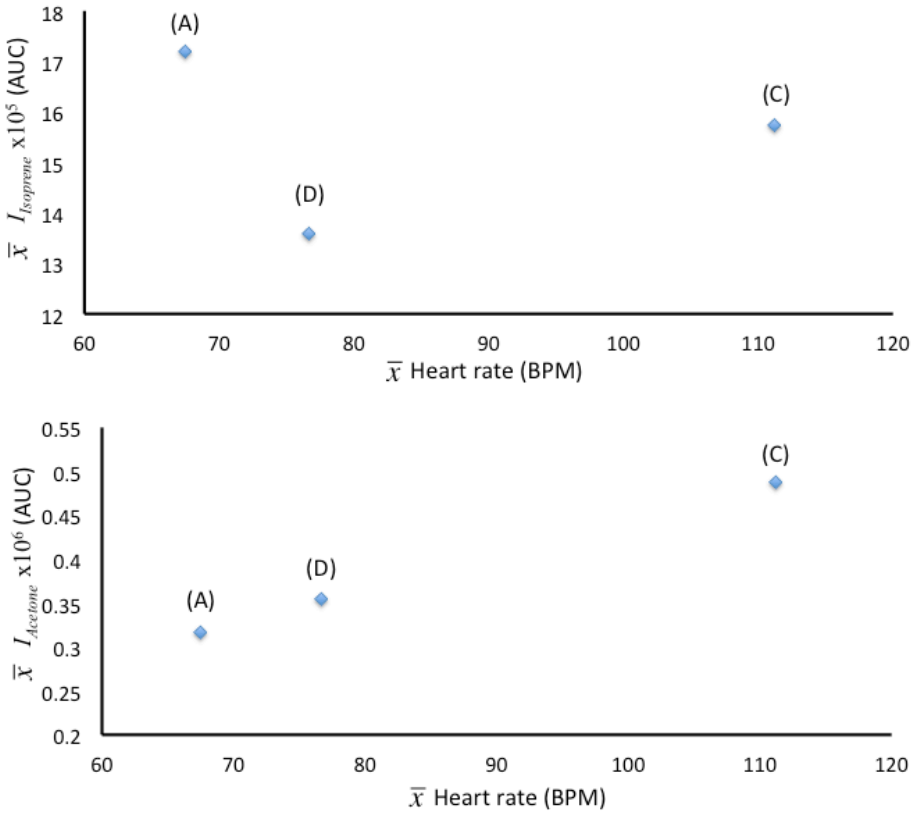


Figure 34 Correlation between average HR and average observed intensity of isoprene (top) and acetone (bottom) signals.

5.4.3. MULTIVARIATE ANALYSIS OF TD-GC-MS

A total of 346 individual VOCs were observed in the 40 samples collected and subsequently analysed by TD-GC-MS. The sample data was processed using the methodology described in chapter 3. The resultant heat map of the integrated chromatographic peak areas for each of the 10 participants in Phases A, C and D and 10 filter air blanks (BLK) is given in Figure 35. The heat map or matrix was then transferred to SIMCA-P+ software, (Umetrics, UK, version 12.0) where multivariate analysis was performed.

The matrix was checked for variables that had responses in less than 25% of observations. This value was chosen as there were four classes of observation (Phase A, Phase C, Phase D and Filtered Air Blank) in the data set. This approach ensured that any VOC components of the breath samples

that appear in only one class of observations; effectively presence / absence based approach, would not be excluded. The result of this test was that no variables were excluded from the matrix. Further to this the data set was scaled using Pareto scaling. The model type selected was PCA-X this was an unsupervised principle component analysis (PCA).

The model produced initially investigated the responses yielded in the matrix from the observations of the pre-exercise phase (phase A) and post exercise phase (phase C) of the study. The model generated two principle components that were classed as significant by the SIMCA-P+ software. The first principle component [t1] explained 12.7% of the variance of the total variance of X (R^2) and the second principle component [t2] explained 18.9 % of the total variance. The model yielded Q2 values of 0.832 and 0.202 for [t1] and [t2] respectively. The score plot produced from this model is given in Figure 36.

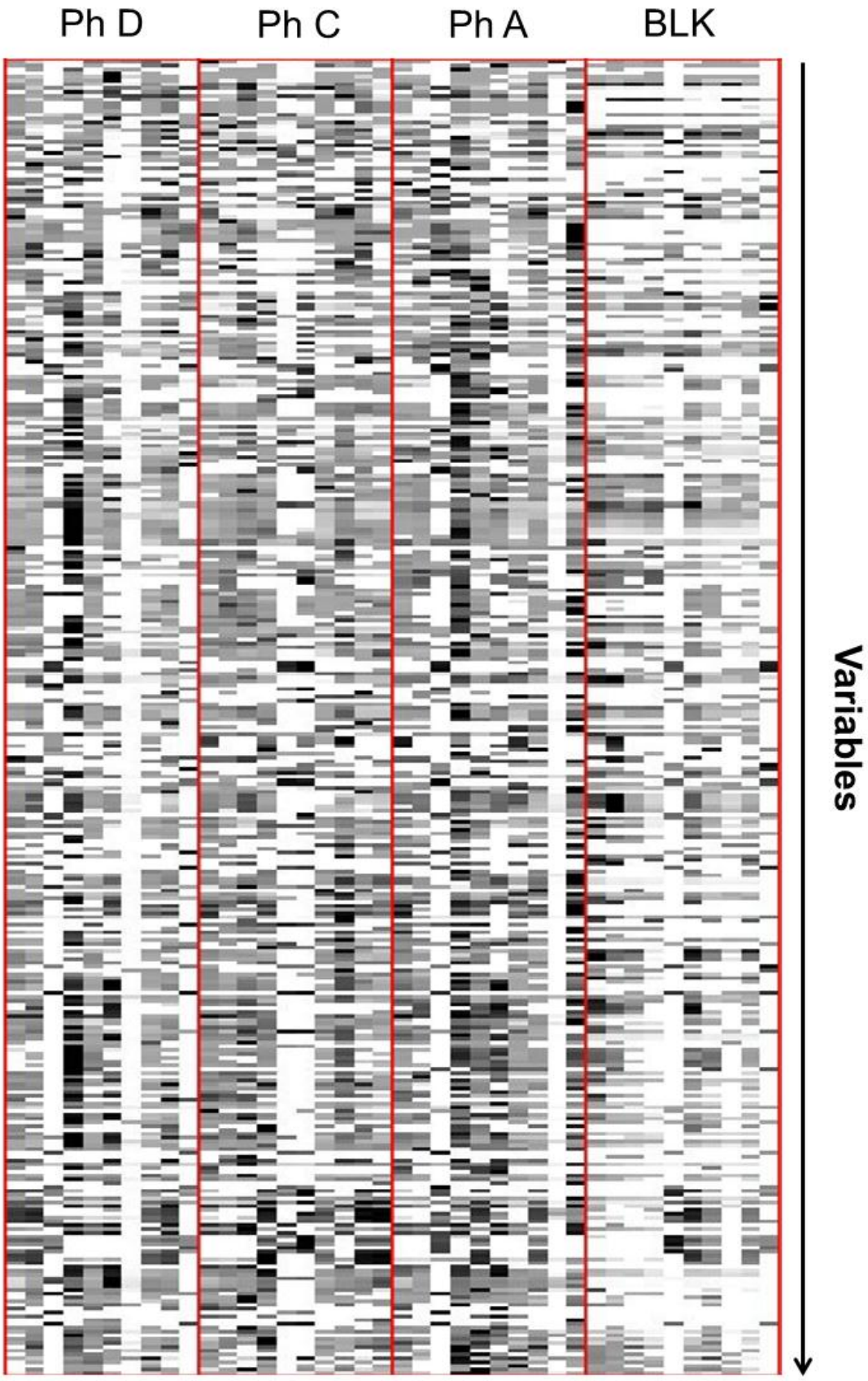


Figure 35. Heat maps of the breath data-matrices for air blanks, and sample collected in experimental phases A, C and D.

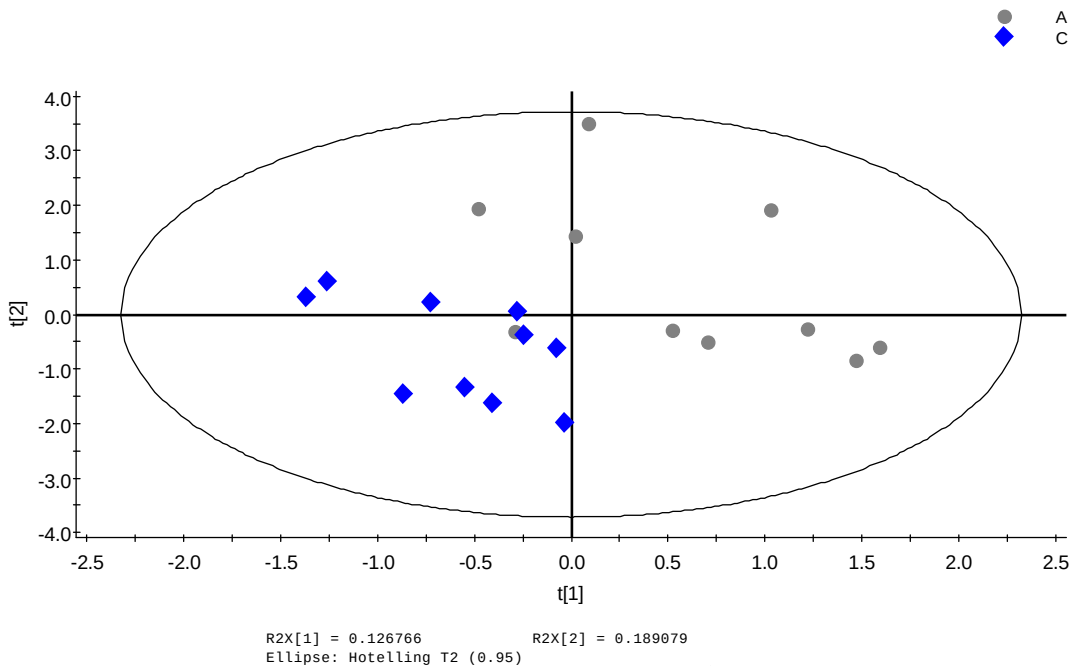


Figure 36 PCA score plot of Phase A and C, Phase A observations indicated by grey circles and Phase C observations illustrated by the blue diamonds. The two principle components generated by the PCA model explain a total of 31.5 % of the variation observed.

A cluster analysis performed on the model showed there were two distinct clusters between the responses for observations in phase A and phase C. The model exhibited a sensitivity and specificity of 100.0% and 90.0% respectively. The false positive observation from this data set was generated from participant CZ03.

There were a total of 11 variables that were found to be responsible for the separation observed in the PCA, a list of these variables are given in Table 8.

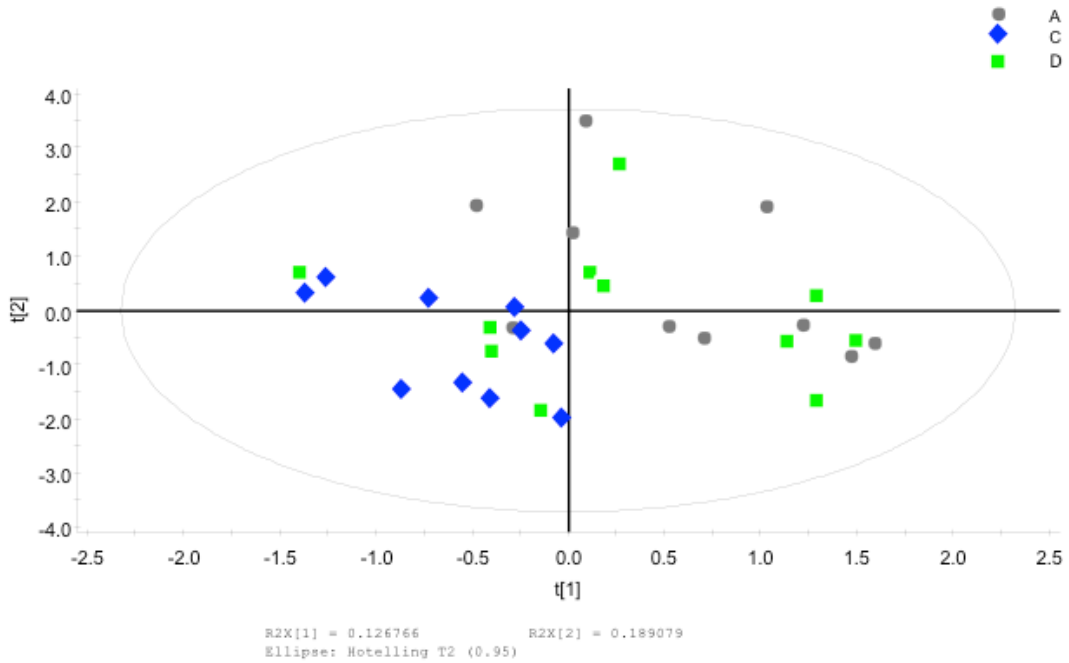


Figure 37 PCA Score plot of phases A, C and D. Phase A observations indicated by grey circles, Phase C observations illustrated by the blue diamonds and D observations are shown by the green squares. Notice only two clusters of observations are still present compared with Figure 36 as the Phase D observations are spread over the previous two experimental phase observations.

The PCA analysis was shown including phase D observations. The resultant score plot is given in Figure 37. The results of this analysis and the subsequent cluster analysis show that the three classes of observations yield two clusters. However, with this data set the score plot shows that the recovery phase samples (Phase D) exhibit a 60/40 split between the pre-exercise phase and the post exercise phase. The variables that were found to be responsible for the separations observed in the score plots are listed in Table 8. To elucidate these variables, the de-convoluted spectrum of each compound was compared with NIST 11 mass spectral database. The reverse fit matches for each compound are listed in Table 8. In addition to the spectral matching the retention index values obtained experimentally from the construction of a primary and secondary retention index ladders were also compared to the NIST library matches. Where a tentative identification was

made using both the retention index and spectral matching these components have been named. When the spectral matching returned reverse fit results below 700 and / or there was no substantial correlation to the experimental retention index values obtained when compared to those in the NIST database then the assignment has been given Unknown VOC.

Table 8. Candidate exercise response markers.

Ref	Library Index	NIST REV match	NIST RIU	Putative annotation
1	VOC 617.8_59_43_0_0_0	950	481	Acetone
2	VOC 622.6_67_68_53_42_0	890	502	Isoprene
3	VOC 622.7_95_97_49_93_77	-	-	Unknown VOC # 1
4	VOC 631.2_81_67_53_93_95	-	-	Unknown VOC # 2
5	VOC 636.9_95_94_97_92_45	936	514	Dimethyl selenide
6	VOC 637_94_97_45_93_91	-	-	Unknown VOC #3
7	VOC 685.1_43_86_41_71_58	811	673	3-Methyl-2-butanone
8	VOC 702.7_61_90_48_41_47	762	707	Propane, 1-(methylthio)-
9	VOC 717.8_88_73_45_71_47	839	726	1-Propene, 1-(methylthio)-, (E)-
10	VOC 726.7_94_95_45_79_97	933	724	Disulfide, dimethyl
11	VOC 756.3_97_98_69_45_53	903	779	Thiophene, 3-methyl-

5.4.4 TRENDS IN VOC RESPONSES IN BREATH POST EXERCISE.

The responses for each of the VOC components in breath responsible for the separation observed in the PCA for phase A and C are summarised in Figures 37 to 39. The trends observed for (3) VOC 622.7_95_97_49_93_77 show a reduced inter-subject variability in phase C compared to phase A or D. In addition Phase C also exhibited a down regulated response with respect to the component peak area response that reduced from a maximum of 31 x 10^4 cps in Phase A to 13 x 10^4 cps in Phase C. After the 30 minute resting period the inter-subject variability observed increased. The peak response observed in Phase D also showed an up-regulated response with respect to Phase C

with maximum of 232 x 10^4 cps being recorded for one participant. The general trend of the responses of VOC 622.7_95_97_49_93_77 are contrary to those observed for acetone and heart rate. This trend of down regulated responses in the VOCs listed in Table 8, occurs for four additional compounds, (6) VOCs 637_94_97_45_93_91, (8) VOC 702.7_61_90_48_41_47, (10) VOC 726.7_94_95_45_79_97 and (11) VOC 756_97_98_69_45_53. The tentative assignment of compounds 8, 10 and 11 reveals that they are sulfur-containing compounds and include dimethyl disulphide. Dimethyl disulphide and dimethyl selenide (compound 5) have been reported previously as markers of reduced liver function in exhaled breath[145]. A potential source of sulfur-containing compounds is the incomplete metabolism of sulfur-containing amino-acids through the transamination pathway[145]. A second potential source of sulfur containing compounds in the body is the GI tract as our diet contains a number of foods such as onions, garlic and eggs[146]. The reduction in the observed responses of these VOCs directly after exercise may potentially indicate their metabolic origins lie in the organs that receive reduced blood flow as the metabolic demands on them are low during the physiological stressful conditions associated with exercise.

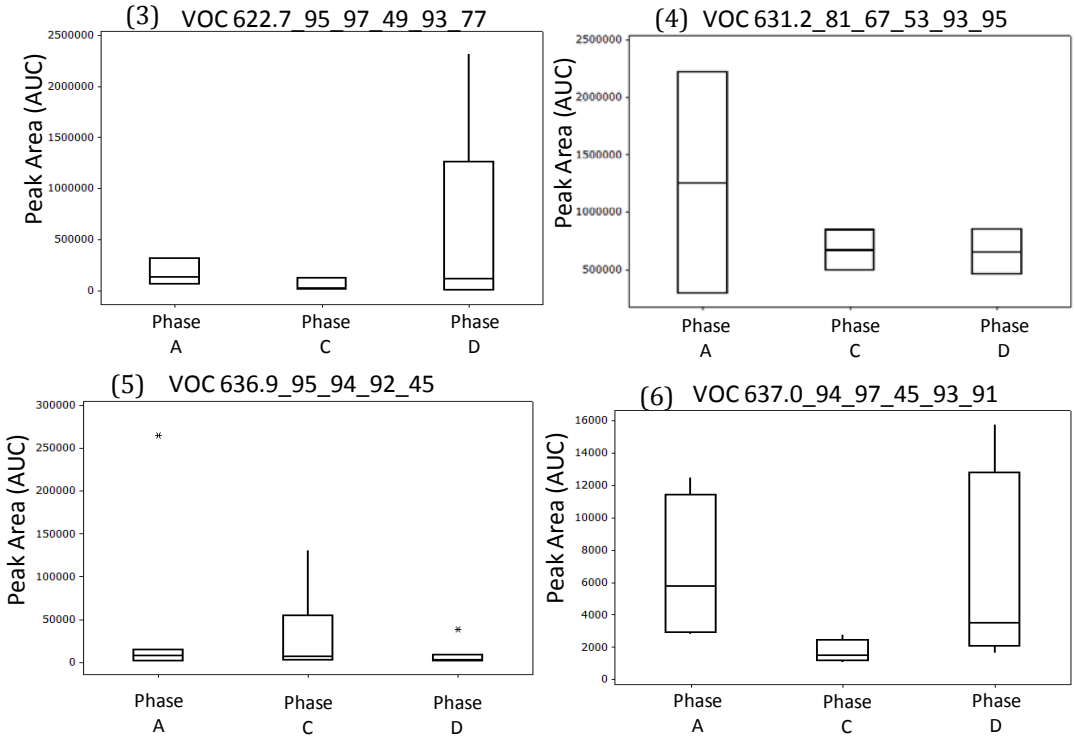


Figure 38 Box whisker plots of the peak areas observed for VOC breath components responsible for the separation in the score plots (Figure 36) between phases A and C.

VOC compounds (5) VOC 636.9_95_94_97_92_45 and (7) VOC 685.1_43_86_41_71_58 exhibit the same overall trend as that of acetone. These compounds tentatively identified as dimethyl selenide and 3-methyl-2-butanone are compound that have been reported in breath and have been used to classify reduced liver function[145] while 3-methyl-2-butanone has been reported in breath and urine[147,52] and as a ketone it is a reasonable assumption that it follows the same elimination pathways as acetone, which would potentially explain why is shows the same general trend in response following exercise. Dimethyl selenide also appears to follow this trend, selenium is an essential mineral and finds its major elimination routes from the human body via breath and urine like that of ketones.

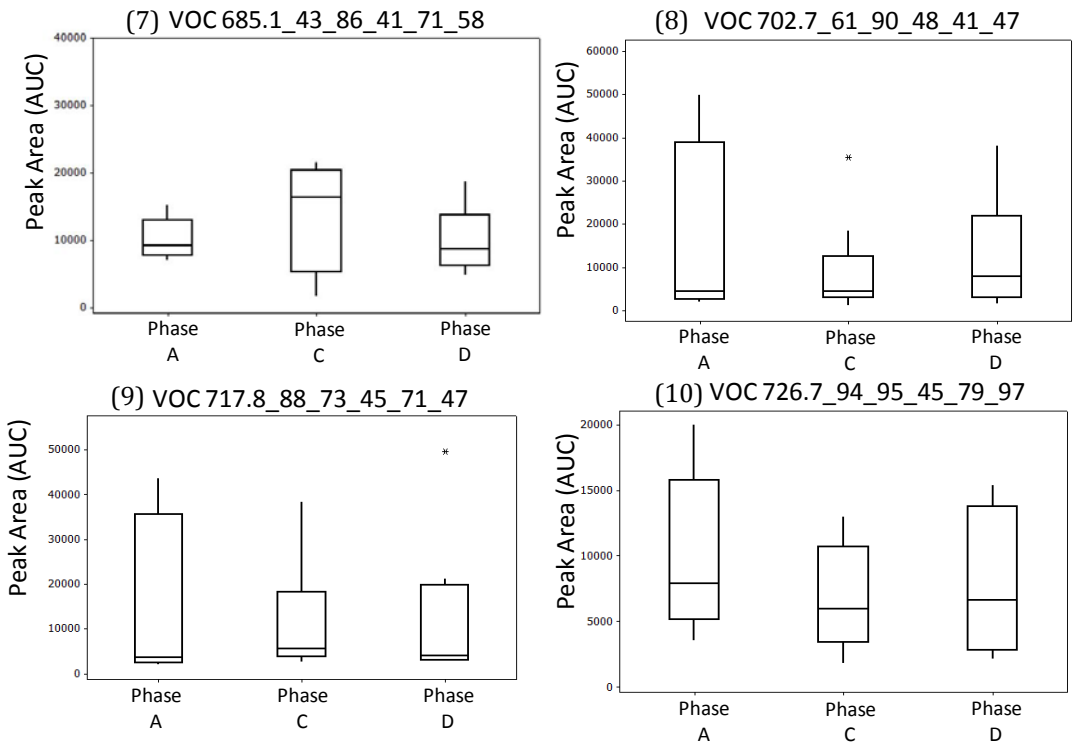


Figure 39 Box whisker plots of the peak areas observed for VOC breath components responsible for the separation in the score plots (Figure 36) between phase A and C

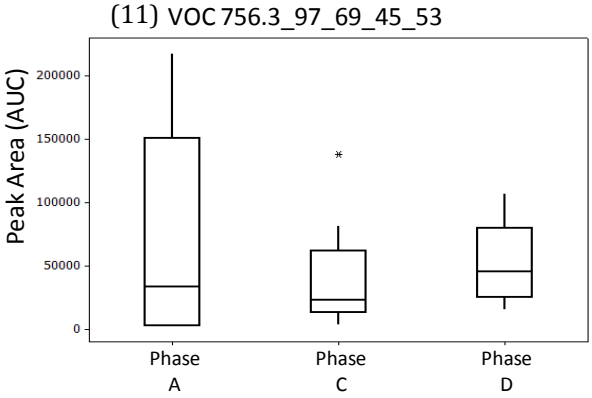


Figure 40 Box whisker plot of the peak areas observed for VOC breath component responsible for the separation in the score plots (Figure 36) between phases A and C

5.5 Study summary

The protocol was successful in causing a physiological response to exercise this is indicated by the elevated heart rate (Figure 30).

Due to the hardware limitations of the adaptive sampler breath samples for TD-GC-MS were only able to be collected for experimental phases A, C and D.

The samples from the experimental phases A and C were able to successfully distinguish between breath collected pre and directly post exercise with a sensitivity and specificity of 100.0% and 90.0% respectively.

A total of 11 VOCs were found to be responsible for the separation observed in the PCA Score plots. These included acetone and isoprene, which agrees with previously published observations in the literature. These 11 components also showed that when phase D sample observations were incorporated into the statistical model they were distributed between the pre and post sets of samples indicating that there are differing rates at which the participants recovered from the exercise protocol. Within the suite of VOC markers found to discriminate physiological stress there were four sulphur-containing VOCs that are suggestive that the liver metabolism is sensitive to exercise. One selenide compound was also identified which have been reported to discriminate healthy and liver disease patients.

Chapter 6 The role of the lower intestinal tract in breath variability

6.1 Introduction

There are numerous VOCs that are present in samples of human faeces and faecal VOC profiles in breath contain putative biomarkers of bowel disease[40,13,52,148,149] . Many of the VOCs isolated from faecal samples have also been found in breath and this may be explained, in part, by the venous system around the anus and the colon has veins that are connected to both to the hepatic system and the inferior vena cava (Figure 42) and how they integrate with the rest of the cardiovascular system (Figure 41). The inferior vena cava is the large portal vein that returns the de-oxygenated blood from the lower body to the heart. The returning blood enters the left ventricle of the heart and is pumped to the lungs where it then undergoes gas exchange, before being transported back to the heart and re-circulated around the body. There is therefore the potential for VOCs that are present in the gastrointestinal system to be transported directly to the lungs and excreted as part of normal gas exchange without encountering first-pass metabolism in the liver.

6.2 Study objectives

- Observe the effect of defecation on breath VOC profiles
- Establish if defecation is a factor that needs controlling in breath bio-marker prospecting studies.

6.3 Overview of study

The study panel consisted of 10 participants (6 male and 4 female) aged between 23 and 33. All participants were healthy non-smokers and gave informed consent prior to taking part. Participants undertaking the study were asked to school their bowel movements so they could undergo a normal evacuation in the morning between 09:00 and 11:00 hours. The participants were required give two breath samples collected in the manner described in chapter 3. The first of the two samples was collected at the point at which the participant started to feel the need to defecate. The second breath sample

was collected 10 minutes after defecation. In both instances the sample collection volume was 2.0 dm^3. In addition to the breath samples two additional air samples were acquired; (i) a 2.0 dm^3 of the filtered air used to supply the mask and (ii) a passive VOC sample; which the participant was required to wear at the start of the sampling procedure. The average passive sampling time was 30 minutes and this measurement was used as an environmental control to account for exposure to any immoderate sources of VOCs from the building or toilet facilities. The acquired samples were stored at 4°C for up to 72 hours until analysis by TD-GC-MS was undertaken. The analysis conditions and quality procedures are outline in detail in Chapter 3.

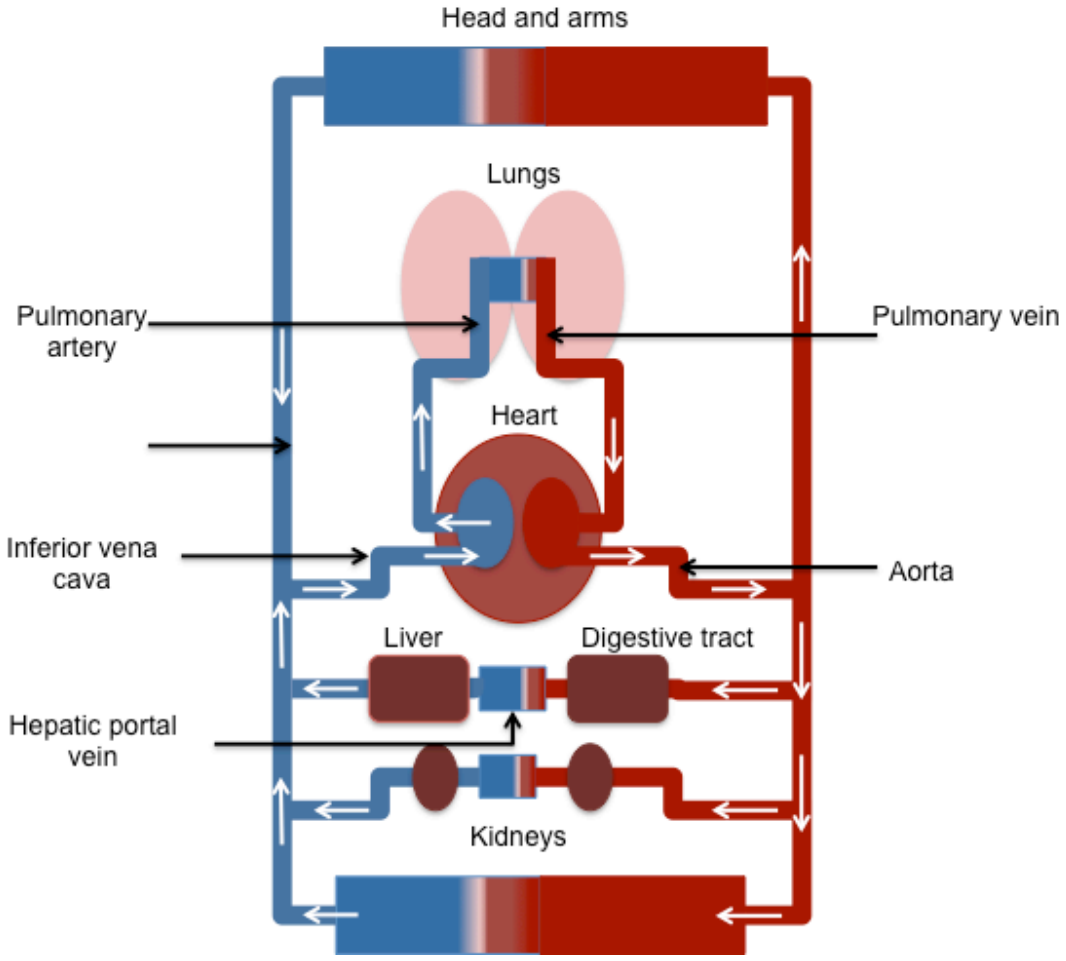


Figure 41 Schematic diagram of the human cardiovascular system

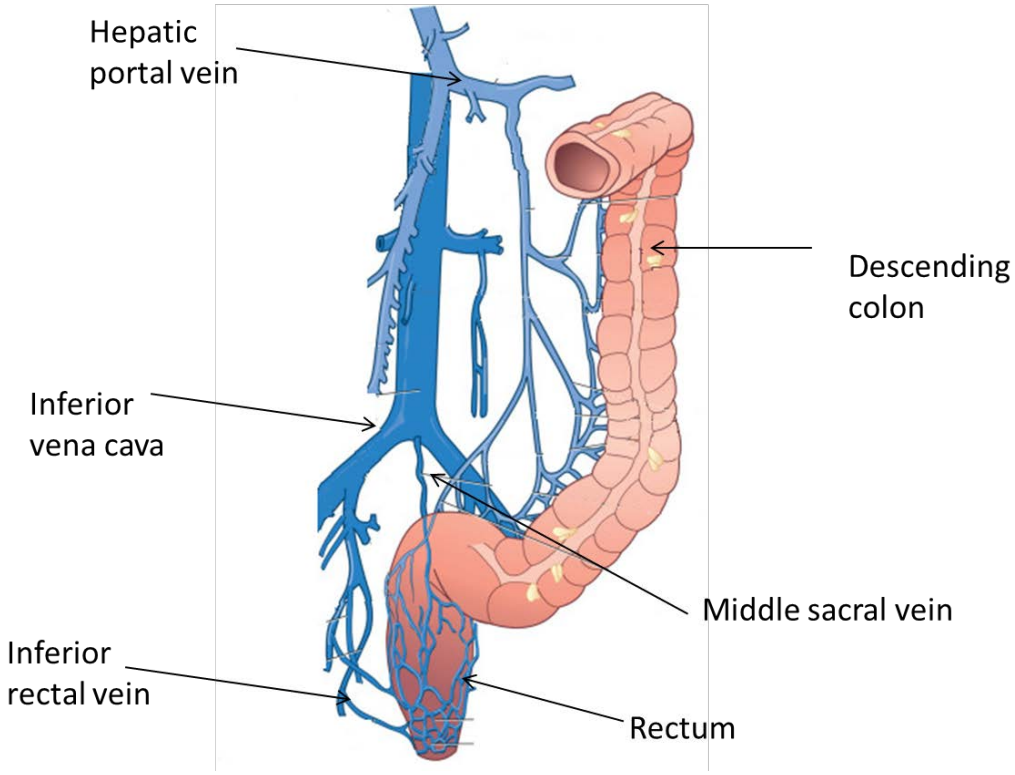


Figure 42 Anatomical drawing showing the veins associated with lower intestine and rectum[150]

6.4 Results

6.4.1 STATISTICAL EVALUATION OF VOC BREATH PROFILES OBSERVED IMMEDIATELY BEFORE AND AFTER DEFECATION

The TD–GC–MS data sets were processed following the work-flow described in chapter 3 and employed in the studies covered in chapters 4 and 5. Samples were aligned by retention index, spectra were deconvoluted and the resultant component peak areas calculated and entered into a breath data-

matrix with a total of 263 variables (spectral and time resolved distinct components) and 20 observations (10 x pre and 10 x post defecation breath samples). They were modelled with multivariate analysis using SIMCA-P+ (version 12, Umetrics, UK). The model was examined for any observation or variables that had less than 50 % of the total responses and none of the observational responses were found to have less than 50% of the total responses in the entire dataset. The dataset was then scaled with Pareto scaling and each observation was assigned as either Pre or Post. An OPLSDA model was then applied to the data and in order to classify possible discriminator VOCs within that provided a separation between the two classes of observations.

The resultant model provided a separation of the two classes of observations, however no significant principle components were isolated. The RX2 and Q2 results generated are given Table . The S-plot generated from the model identified a total of 6 variables that were responsible for the separation observed in the OPLSDA model.

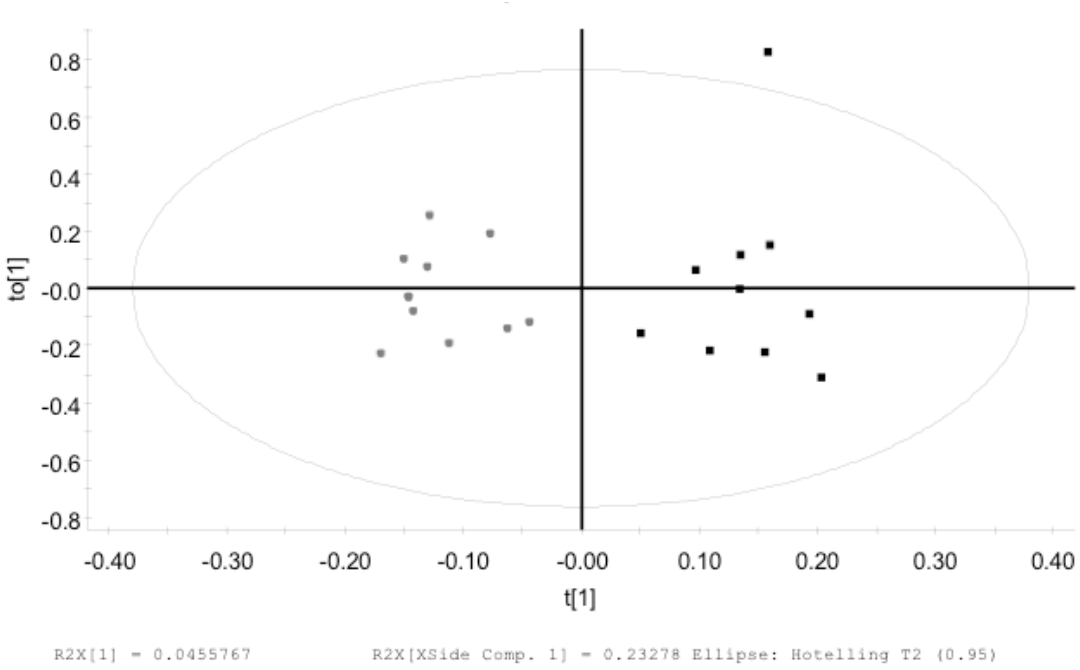


Figure 43 OPLS-DA Score plot of 10 pre bowel evacuation (grey circles) breath samples and 10 post bowel evacuation (black square) breath samples

Table Results of the OPLS-DA model generated by Umetrics SIMCA-P statistical package

	R2X	R2X (cum)	Eigenvalue	R2Y	R2Y (cum)	Q2	Q2(cum)	significance
Model								
1	0.0456	0.631	0.684	0.885	0.885	-0.527	-0.527	NS
Orthogonal		0.585			0			
1	0.233	0.233	3.49	0	0			NS
2	0.223	0.456	3.34	0	0			NS

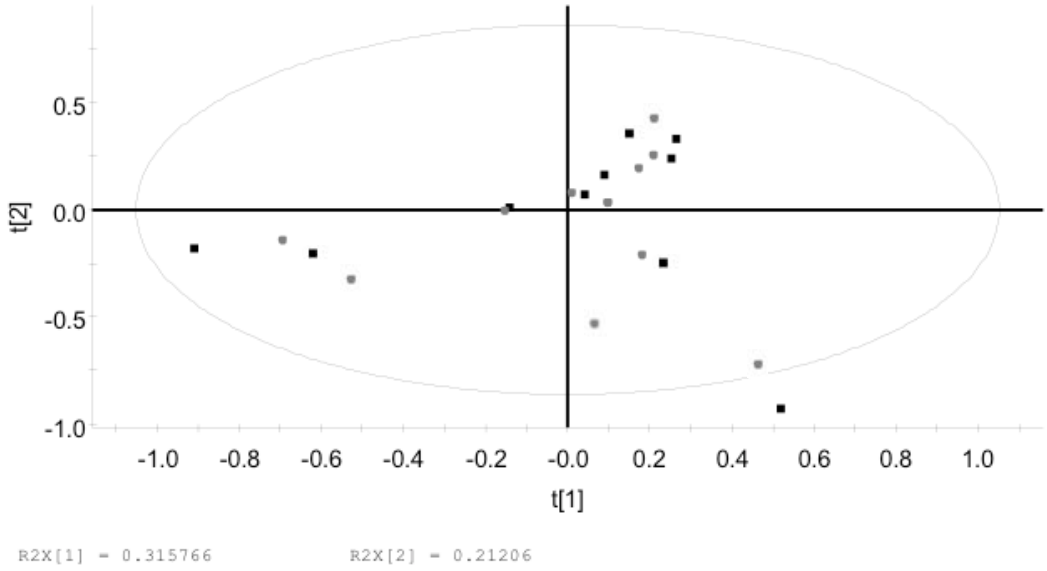


Figure 44 PCA Score plot of 10 pre bowel evacuation (grey circles) breath samples and 10 post bowel evacuation (black square) breath samples using unsupervised model generated from the variables identified from the OPLSDA analysis

The responses for these variables were isolated and a principle component analysis (PCA) was performed to establish whether discrimination of the two sets of observations could be retained using an unsupervised model. The PCA model generated three principle components. The principle components generated in the unsupervised PCA were all significant according to either rule 1 (R1) or rule 2 (R2) of the cross validation rules[130], however their predictive ability (Q2) was poor. The resultant score plot of components t1 and t2 from the model is given in Figure 44.

Evaluation of the score plot revealed the there was insufficient separation between the two sets of observations to discriminate with any certainty between samples taken before and after a bowel evacuation. Further scrutiny of the score plot revealed that while there was no common discrimination between the pre and post samples, the individual sample sets from each participant were all separated to various degrees on the score plot.

6.4.2 OBSERVATIONS OF VOC BREATH PROFILES BEFORE AND AFTER DEFECATION

In the absence of a meaningful statistical model to identify potential VOCs that change after bowel evacuation a further examination of the individual participant samples identified differences in the responses of a number of components in the total ion current chromatograms (Table 11). The extracted ion chromatograms presented in Figure 44, 45 and 46 have been normalised to the response of the pre bowel movement samples for each VOC component. The VOC components presented in these figures show reductions in response between 30 to 100% in breath after defecation had occurred. A summary of the VOCs observed to change and their tentative identifications based on spectral searching with NIST 11 spectral database are given in Table 9.

Table 9 VOCs observed to reduce post bowel evacuation

NIST database assignment	BRI Reference	Match (Rev)
5-hepta-2-one-(6-methyl)	BRI_978-93-108-43-41-67	783
unknown Alcohol	BRI_871-41-56-55-69-70	-
unknown aldehyde	BRI_998-41-56-82-81-67	-
2 propylpentanol	BRI_1026-57-41-70-83-56	799
Nonanal	BRI_1112-81-67-41-70-69	779
Decanal	BRI_1223-81-67-83-41-82	804
Unknown VOC #4	BRI_1198-81-95-67-41-71	-
methylsalicylate	BRI_1213-81-95-67-41-71	802
Benzene-1-methyl-2-(1methylethyl)	BRI_1023-119-91-117-115-77	891
Anethole	BRI_1311-148-147-117-105-133	889
Indole	BRI-1286-117-90-89-63-73	902
Unknown Terpene	BRI_1028-93-79-67-94-68	-

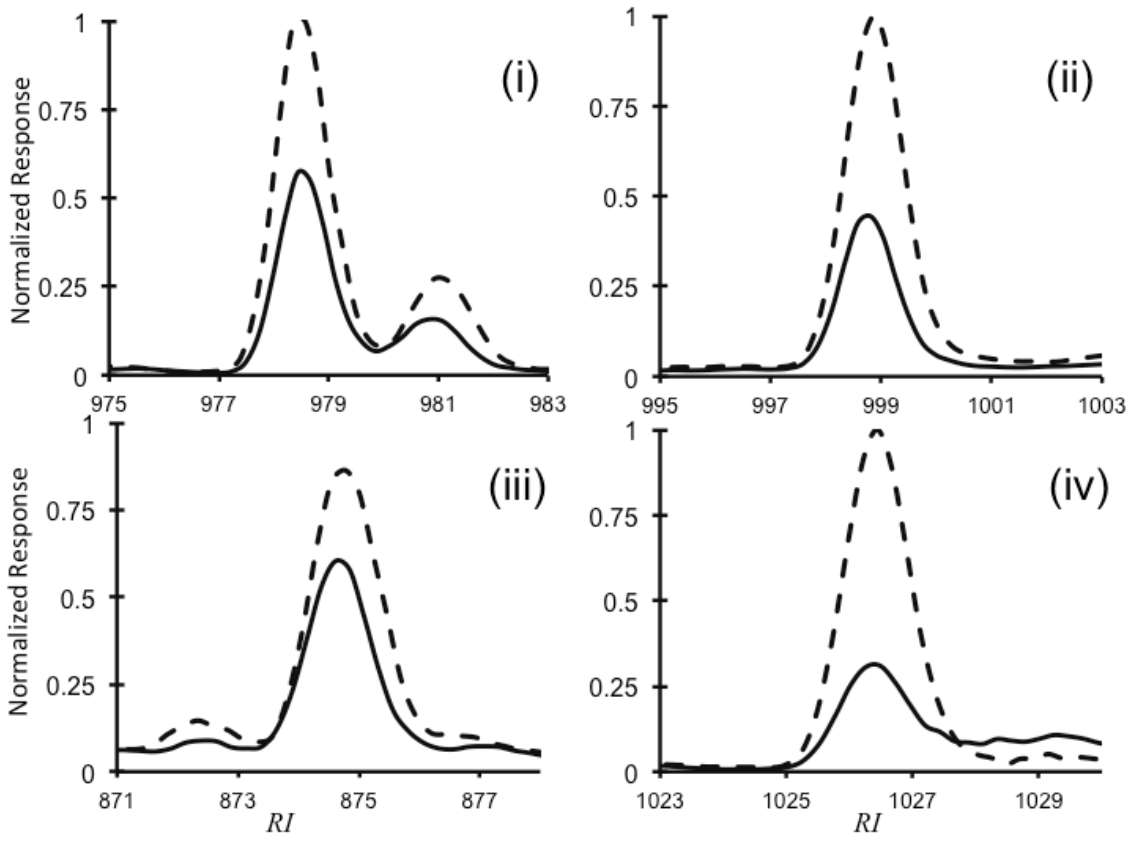


Figure 45 **Example of extracted ion response plotted against calculated retention index of detected analytes in pre (dashed line) and post (solid line) defecation breath samples observed to reduce in intensity after a bowel evacuation. Responses have been normalized to pre sample analyte intensities (i) BRI_978-93-108-43-41-67, (ii) BRI_998-41-56-82-81-67, (iii) BRI_871-41-56-55-69-70 and (iv) BRI_871-41-56-55-69-70**

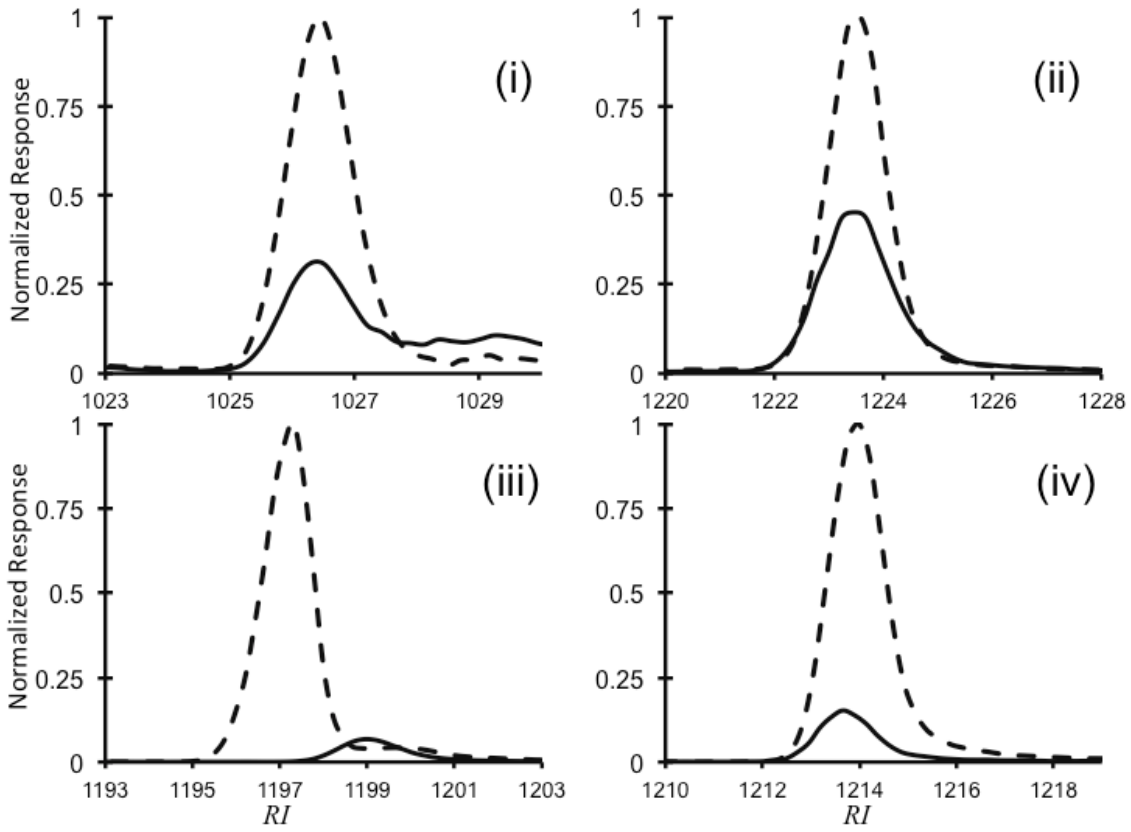


Figure 46 Example of extracted ion response plotted against calculated retention index of detected analytes in pre (dashed line) and post (solid line) defecation breath samples observed to reduce in intensity after a bowel evacuation. Responses have been normalized to pre sample analyte intensities (i) BRI_1112-81-67-41-70-69, (ii) BRI_1223-81-67-83-41-82 (iii) BRI_1198-81-95-67-41-71and (iv)BRI_1213-81-95-67-41-71.

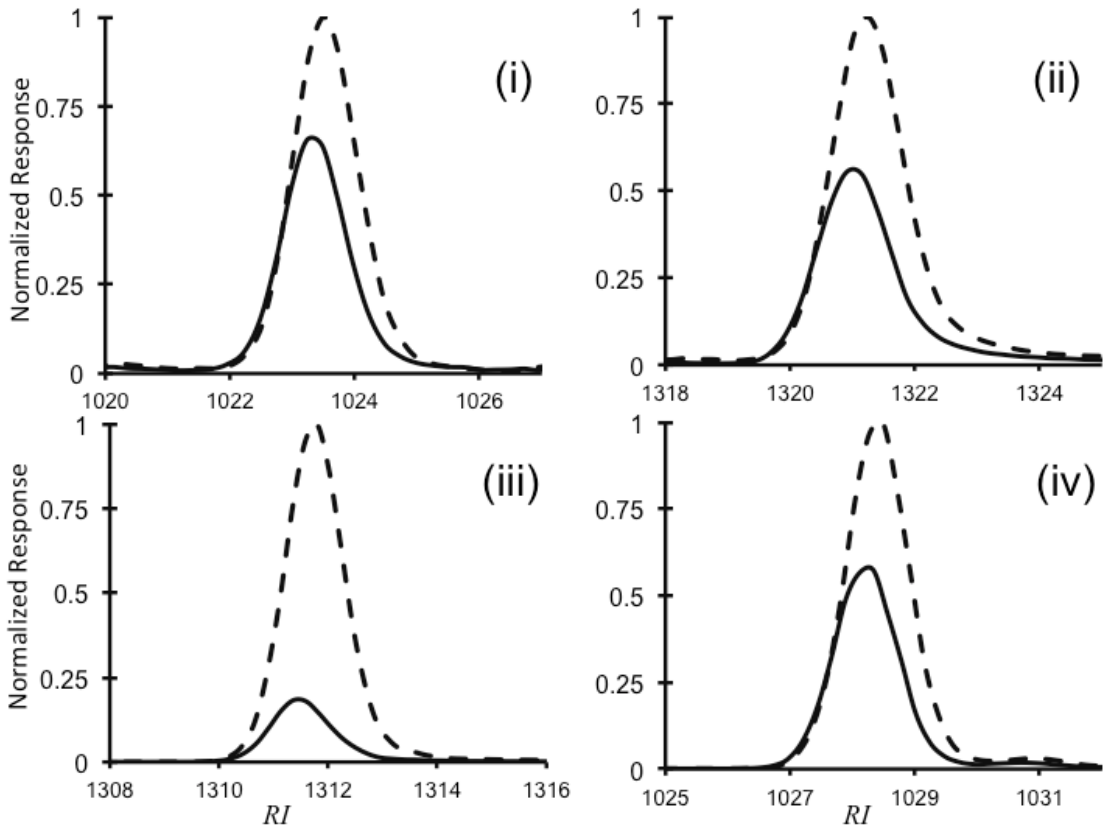


Figure 47 Example of extracted ion response plotted against calculated retention index of detected analytes in pre (dashed line) and post (solid line) defecation breath samples observed to reduce in intensity after a bowel evacuation. Responses have been normalized to pre sample analyte intensities. (i) BRI_1023-119-91-117-115-77, (iii) BRI_1311-148-147-117-105-133, (ii) BRI-1286-117-90-89-63-73 and (iv) BRI_1028-93-79-67-94-68

6.5 Study summary

The multivariate analysis of these data generated from pre- and post-defecation breath samples yielded statistical models that generated significant principle components but had poor predictive ability. As a result there was no discernible discrimination possible between the two observation sets in this study. There was however a degree of separation within the participants' individual sample sets, which could be seen from both the score plots generated from the statistical analysis and the visual observation made on the chromatographic data obtained. A total of 12 VOCs were observed to reduce over the entire data set after defecation.

The observed inter-subject variability suggests that there is a faecal component in breath. It is important to note that the reduction of the VOCs concentration in breath was in all but one case not completely depleted which would indicate that faeces in the rectum was not the only source of these VOCs in the breath samples.

The VOCs identified in this study have been observed previously in faecal samples and while present in faecal samples their origins have yet to be fully elucidated. Potential origins of the VOCs could be:

- The intake of different foodstuffs (diet), and their subsequent metabolism, catabolism, conversion, oxidation/reduction and hydrolysis.
- The microbiological metabolism within the gastrointestinal tract

One or both of these sources could account for the contribution of the VOCs excreted in breath and this study frames the complexity of seeking to account for gastrointestinal factors in designing breath samples.

At this stage it appears reasonable to propose that the defecation cycle has a significant effect on the VOCs present, or absent, in an individual's breath. It is also apparent that microbiological, lifestyle, diet, phenotypic and environmental factors combine to create individualised profiles and that significantly higher powered experimental design will be needed to establish if there are generic faecal discriminator markers.

The study presented here was undertaken as a pilot study with a small population. The wealth of variation observed between the profiles in this study, where no controlled stimulus was applied, highlights two main factors that should be taken into consideration when attempting to follow up on this study.

- No common VOC artefact in breath was found across all participants; with a lack of any commonality among the VOCs associated with defecation a significantly larger, single-gender participant cohort will be needed in follow up studies.

- The participants were not subject to dietary controls, and this study indicates that this might be helpful in participant preparation protocols for biomarker discovery.

Chapter 7 Discussion, conclusions and future work.

7.1 Overview of thesis

This research utilises optimized conditions for the chromatographic, instrumentation and data processing methods for the detection of VOCs in exhaled breath. The thesis describes new putatively annotated biomarkers associated with psychological and physiological stress. It further investigates how the everyday occurrences of stress and defecation can impact and confound the prospecting of breath biomarkers when using an untargeted approach. With the premise that small everyday tasks and situations impact on VOCs breath profiles established, it highlights the need for pre-sampling participant protocols when embarking on breath analysis studies.

7.1.1 CHAPTERS 1 AND 2

The first two chapters set the proposition of this thesis and established the research objectives. A Gedankan experiment is used to explore the complexity of the VOC profiles obtained from exhaled breath. The fundamental theory of gas exchange in the lungs is discussed along with an introduction to thermal desorption, GC-MS and the multivariate statistical methods commonly used in breath analysis. Finally the different techniques of breath capture and analysis that have been reported in the literature are reviewed and the applications of breath analysis are reviewed.

7.1.2 CHAPTER 3

This chapter outlines the methodology used in this research and describes in detail how breath was collected and the analytical methods employed with an outline of the instrumentation used. The development of the methodology and the final optimized conditions are given. The untargeted global metabolomic approach adopted for the data handling and discovery are discussed in detail.

7.2 Overview of the studies

7.2.1 BREATH VOCS AND THE PACED AUDITORY SERIAL ADDITION TASK

The administration of the PASAT caused an elevated amount of anxiety and frustration in the study participants. This was confirmed by the increased heart rate and blood pressure data obtained before, during and after the task.

The study reported the first example of breath VOCs being affected by a PASAT and indicated the existence of psychological stress-related response in the VOC profiles of healthy humans.

There were six candidate VOC markers that were shown to be responsible for the separation observed in the multi-variate analysis. The analysis of the data also indicated that gender has a role in the extent of the stress response observed for these marker compounds. The resultant score plots yielded 100 % sensitivity and 90% specificity for the male participants and 83.3% sensitivity and specificity of 91.6% for female participants.

Indole and 2-methylpentadecane; two of the candidate markers were found to increase in intensity (up regulated) in samples collected after the PASAT was administered. The remaining four showed a decrease in intensity (down regulated).

Consequently this highlights the fact that non-physical stimuli can have a bearing of the VOCs observed in breath. These findings illustrate that physical disease may not be the only contributing factor in breath profiling of people undergoing traumatic and emotionally challenging diagnoses associated with serious diseases: The participants' stress and anxiety levels may need to be considered as a confounding factor in the biomarker discovery process.

7.2.2 CHANGES TO BREATH VOCS FROM INDUCED PHYSIOLOGICAL STRESS.

The breath samples collected show that the acetone intensities reflect the profiles obtained from the heart rate data observed. These results are consistent with the published literature.

The metabolomics approach to the VOC profiles illustrated that pre and post exercise breath samples differed significantly when interrogated using the data processing protocol previously described in chapter 3.

There was 100% sensitivity and 90% specificity shown in the separation observed from the breath samples collected pre (Phase A) and post (Phase C) exercise.

When the recovery phase samples were incorporated in to the model they showed that 60 % were classified as pre exercise samples and 40 % were classified as post exercise samples. This potentially indicates differences in the fitness levels and therefore the recovery rates for each participant. This finding suggesting that 30 minutes was insufficient for 40 % of participants to recovery from the exercise protocol to a metabolic state resembling that prior to exercise.

There were 11 VOCs that were found to be responsible for the separation observed in the score plots. In these VOC compounds there were four that were found to be sulfur-containing compounds, which have previously been associated with liver function.

The results of this study highlighted new candidate markers that have not been previously reported and are possibly associated with physical stress through exercise. These VOC markers are also potentially indicative of physical recovery after exercise. This finding supports the fact that human breath is not constant and external stimulus can cause a deviation in homeostasis. Furthermore after the removal of the external stimulus the human body attempts to recover homeostatic environment of breath.

These findings if investigated further could help in the personalised training regimes of athletes or patients recovering from physical trauma through physiotherapy.

The study also highlights potential confounding factors in the markers associated with liver function tests though breath VOC analysis.

7.2.3 INFLUENCE OF THE LOWER INTESTINAL TRACT ON BREATH VOC PROFILES

The breath VOC profiles pre and post bowel evacuation have shown that inter-individual variability was not consistent in the cohort of this study. As a result no common markers that correspond to bowel evacuation were identified.

The intra-individual variability of breath VOCs found pre and post sample were shown to be random, but significant changes in VOCs were found in breath. While multivariate analysis did not establish any common VOCs that were able to separate the pre and post samples, the score plots showed that there was separation between the individual participants' profiles. This highlights three important conclusions:

Significant contributions of a number of VOCs in breath find their origins in the bowel.

The time of when a participant's last bowel movement occurred can influence presence and or concentration of VOCs found in breath.

Potentially the effects of diet not only provide an immediate effect on breath profiles directly after consumption such as coffee, but also can be much longer lived providing a source of VOCs in the bowel long after their initial consumption. The ramifications of this is that diet and the bowel are potentially a large contributor to the VOCs found in breath and that there is a need to identify what VOCs are generated in the bowel to remove potential confounding factors when considering metabolomics workflows. Also anxiety and general fitness levels of participants need to be considered when building study panels and planning clinical studies. At the very least this data should be collected and used when considering controls and test groups that are being scrutinised by statistical methods. A list of the VOCs identified in the studies described in this thesis are listed in Table 11.

Table 10 List of VOCs highlighted in the studies presented in this thesis.

Breath Library Ref	NIST assignment	Match (Rev)	Potentially associated with	RIU
BRI_617-59-43-0-0-0	Acetone	950	Physiological stress	617
BRI_622-67-68-53-42-0	Isoprene	890	Physiological stress	622
BRI_623-95-97-49-93-77	Unknown VOC # 1	-	Physiological stress	623
BRI_631-81-67-53-93-95	Unknown VOC # 2	-	Physiological stress	631
BRI_637-95-94-97-92-45	Dimethyl selenide	936	Physiological stress	637
BRI_638-94-97-45-93-91	Unknown VOC #3	-	Physiological stress	638
BRI_685-43-86-41-71-58	3-Methyl-2-butanone	811	Physiological stress	685
BRI_703-61-90-48-41-47	Propane,1(methylthio)-	762	Physiological stress	703
BRI_718-88-73-45-71-47	1-Propene,1(methylthio)	839	Physiological stress	718
BRI_727-94-95-45-79-97	Disulfide, dimethyl	933	Physiological stress	727
BRI_756-97-98-69-45-53	Thiophene, 3-methyl-	903	Physiological stress	756
BRI_871-41-56-55-69-70	Unknown Alcohol	-	Bowel	871
BRI_951-93-91-92-79-121	Unknown terpene compound		Psychological stress	951
BRI_976-105-77-52-106-95	Benzaldehyde	813	Psychological stress	976
Breath Library Ref	NIST assignment	Match (Rev)	Potentially associated with	RIU
BRI_978-93-108-43-41-67	5-hepta-2-one- (6-methyl)	783	Bowel	978

BRI_998-41-56-82-81-67	Unknown aldehyde	-	Bowel	998
BRI_1019-41-57-70-85-55	2-ethylhexan-1-ol	830	Psychological stress	1019
BRI_1023-119-91-117-115-77	Benzene-1-methyl-2- (1methylethyl)	891	Bowel	1023
BRI_1026-57-41-70-83-56	2-propylpentanol	799	Bowel	1026
BRI_1028-93-79-67-94-68	Unknown terpene	-	Bowel	1028
BRI_1063-105-77-52-106-51	2-hydroxy-1-phenylethanone	885	Psychological stress	1063
BRI_1112-81-67-41-70-69	Nonanal	779	Bowel	1112
BRI_1198-81-95-67-41-71	Unknown VOC #4	-	Bowel	1198
BRI_1213-81-95-67-41-71	Methyl salicylate	802	Bowel	1213
BRI_1223-81-67-83-41-82	Decanal	804	Bowel	1223
BRI_1286-117-90-89-63-73	Indole	889	Psychological stress	1286
BRI_1286-117-90-89-63-73	Indole	902	Bowel	1286
BRI_1311-148-147-117-105-133	Anethole	889	Bowel	1311
BRI_1527-41-57-43-85	2-methylpentadecane	658	Psychological stress	1527

7.3 Future work

7.3.1 VALIDATION OF POTENTIAL CANDIDATE MARKERS FOR PSYCHOLOGICAL AND PHYSIOLOGICAL STRESS.

The studies in this thesis have been of a pilot study size with the cohorts of participants being a relatively small number when compared to more conventional large-scale metabolic studies. The approach of this research was not undertake large scale participant studies but rather to investigate potential causes of changes in breath profiles from a baseline using healthy control participants. As a result the potential candidate markers highlighted in the physical and psychological stress studies have yet to be validated in large-scale targeted studies.

In addition there also remain a number of unconfirmed VOCs that would benefit from confirmation using commercially available standards as this would potentially help to elucidate their biological origins. In some instances where there is not a commercial standard available, further elucidation could potentially be achieved using MS^n and or chemical ionization modes using instrumentation described in chapter 3. This could also be accomplished using high-resolution mass analysers to provide higher mass accuracy and elemental compositions.

The study described in chapter 6 would also benefit from a larger scale investigation with stricter dietary control of the cohort to attempt to elucidate common VOCs, which may allow for separation when scrutinized using multivariate statistics.

In addition to the larger scale studies it is proposed that quantification of the candidate markers should also be undertaken using matrix matched standards and potentially the incorporation of an internal standard into the breath sampling procedure. The latter being a difficult challenge as there are hundreds of VOCs in breath with a variety of functional groups, polarities and molecular weights, so multiple internal standards may be required. Ideally

such compounds would be introduced as early as possible in the sampling procedure or even administered to the participant prior to sample collection.

7.3.2 DEVELOPING FUTURE STUDIES BASED ON THE CURRENT FINDINGS

The following sections provide an outline of potential future studies that are based on and further the research presented in this thesis.

The application of stress has only been evaluated using healthy participants that have undergone a single acute stressful episode. A valuable study would include participants that undergo chronic stress. These cohorts could be made up of participants from stressful professions such as air traffic controllers, soldiers or fire fighters.

A diurnal study targeting the potential candidate markers for physical and psychological stress identified in this research would allow for the monitoring of the inherent variability of these breath components. The study would aim to collect samples over the course of hours, days and weeks and could be aimed at competing athletes as they train and prepare for competitions. This approach would aim to establish potential links between physical and mental fitness in athletes.

7.4 Concluding remarks

How we think, how we move and what we eat, the very environment we find ourselves in defines who we are. Breath VOC analysis has a long history, but our understanding of how rich, complex and variable the matrix of breath can be, has only started to be realized in the last 15 years. The factors examined in this research effect our body's biochemistry and this is in turn reflected in our breath. The studies here however, were not concerned with infections or diseases but sort to explore how day to day activities such as exercise and toileting, as well as psychological stress impacts on VOC profiles. In the course of this research new putative VOC markers have been presented that can successfully discriminated participants that had been exposed to psychological stressful situation from those who were in a relaxed state. In

addition to this group of markers a second panel of putative annotated VOCs were elucidated that were associated with the physical stress of exercise. These two studies highlight potential confounding factors that could impact on future breath VOC studies. They show that if a panel participants or part of that panel are exposed to physical stress or acute psychological stress prior to the collection of breath samples then it would be conceivable when using MVA methods that potential false positive could be identified. The investigation into the how defecation changes the composition of breath highlighted how a bowel movement can cause major changes in the concentrations and occurrence of VOCs observed in the breath of individuals.

From these findings we ideally need to consider and be aware of the management of care and hospital facilities in terms of toileting. We should also be assessing people's state of physical activity as it has an effect on breath. This is yet to be fully quantified but is observed in this research. Finally, that a participant's emotional state may need to be noted and assessed prior to sample collection. However attempting to control all of these aspects is highly impractical and may well never be achieved. In the same way though by not considering these factors, which can impact significantly on the breath volatalome and instigate a deviation from homeostasis in human breath, we may introduce more complexity which could either mask or confound small metabolite signals in future studies. One potential compromise to this would be the introduction of a detailed participant diaries or the development of using wearable personal technology to gather a range of meta-data, such as sleeping patterns, heart rate and exercise patterns. What is apparent is that further investigation is needed to fully understand what a normal everyday breath VOC profile should look like. The research presented here is only a small fraction of the types of everyday activities that could potentially cause significant deviations in the baseline of breath. Understanding what constitutes normal breath would undoubtedly enable more reliable and robust interpretation of breath VOCs as markers in disease diagnostics.

Chapter 8 References

(1) Miekisch, W.; Schubert, J. K.; Noeldge-Schomburg, G. F. E. *Clin. Chim. Acta.* **2004**, *347*, 25–39.

(2) Turner, M.; Guallar-Hoyas, C.; Kent, A.; Wilson, I.; Thomas, C. *Bioanalysis* **2011**, *3*, 2731–2738.

(3) Pleil, J. D.; Lindstrom, a B. *Clin. Chem.* **1997**, *43*, 723–730.

(4) Mueller, W.; Schubert, J.; Benzing, A.; Geiger, K. **1998**, *716*, 27–38.

(5) Grote, C.; Pawliszyn, J. *Anal. Chem.* **1997**, *69*, 587–596.

(6) Lord, H.; Yu, Y.; Segal, A.; Pawliszyn, J. *Anal. Chem.* **2002**, *74*, 5650–5657.

(7) Phillips, M.; Greenberg, J. *J. Chromatogr. B Biomed. Sci. Appl.* **1991**, *564*, 242–249.

(8) Karl, T.; Prazeller, P.; Mayr, D.; Jordan, A.; Rieder, J.; Lindinger, W.; Schaefer, M. L.; Wongravee, K.; Holmboe, M. E.; Heinrich, N. M.; Dixon, S. J.; Zeskind, J. E.; Kulaga, H. M.; Brereton, R. G.; Reed, R. R.; Trevejo, J. M.; Cope, K. a; Watson, M. T.; Foster, W. M.; Sehnert, S. S.; Terence, H.; Fall, R. a Y.; Fall, R. **2014**, 762–770.

(9) Diskin, A. M.; Spanel, P.; Smith, D. *Physiol. Meas.* **2003**, *24*, 107–119.

(10) Di Natale, C.; Macagnano, A.; Martinelli, E.; Paolesse, R.; D'Arcangelo, G.; Roscioni, C.; Finazzi-Agrò, A.; D'Amico, A. *Biosens. Bioelectron.* **2003**, *18*, 1209–1218.

(11) Basanta, M.; Koimtzis, T.; Singh, D.; Wilson, I.; Thomas, C. L. P. *Analyst* **2007**, *132*, 153–163.

(12) Basanta, M.; Jarvis, R. M.; Xu, Y.; Blackburn, G.; Tal-Singer, R.; Woodcock, A.; Singh, D.; Goodacre, R.; Thomas, C. L. P.; Fowler, S. J. *Analyst* **2010**, *135*, 315–320.

(13) Probert, C. S. J.; Ahmed, I.; Khalid, T.; Johnson, E.; Smith, S.; Ratcliffe, N. *J. Gastrointestin. Liver Dis.* **2009**, *18*, 337–343.

(14) Phillips, M.; Greenberg, J.; Awad, J. *J. Clin. Pathol.* **1994**, *47*, 1052–1053.

(15) Jones, A. W. *Forensic Sci. Int.* **1985**, *28*, 157–165.

(16) Alving, K.; Weitzberg, E.; Lundberg, J. M. *Eur. Respir. J.* **1993**, *6*, 1368–1370.

(17) Atherton, J. C.; Spiller, R. C. *Gut* **1994**, *35*, 723–725.

(18) Phillips, M.; Herrera, J.; Krishnan, S.; Zain, M.; Greenberg, J.; Cataneo, R. N. *J. Chromatogr. B Biomed. Sci. Appl.* **1999**, *729*, 75–88.

(19) Silkoff, P. E. *Am. J. Respir. Crit. Care Med.* **1999**, *160*, 2104–2117.

(20) Smith, G. P. *Appetite* **2008**, *51*, 428–432.

(21) Naish, J. Syndercombe Court, D. *Medical Sciences E-Book*; 2nd Editio.; Elsevier Health Science, 2009.

(22) Woods, S. C.; Ramsay, D. S. Homeostasis: Beyond Curt Richter. *Appetite*, 2007, *49*, 388–398.

(23) Stanfield, C. L. *Principles of Human Physiology*; 4th ed.; Pearson, 2009.

(24) Paschke, K. M.; Mashir, A.; Dweik, R. a. *F1000 Med. Rep.* **2010**, *2*, 56.

(25) Liddell, K. *Postgrad. Med. J* **1976**, *52*, 136–138.

(26) Phillips, M. In *Disease markers in exhaled breath*; Marczin, N.; Kharitonov, S.; Yacoub, M.; Barnes, P., Eds.; Taylor & francis group, 2002; pp. 219–231.

(27) Smith, M. *Lancet* **1982**, *2*, 1452–1453.

(28) DUVEEN, D. I.; KLICKSTEIN, H. S. *Bull. Hist. Med.* **1955**, *29*, 164–179.

(29) Anstie, F. E.; Baldwin, A. D. *public Heal. then now* **1977**, *67*, 679–681.

(30) Hubbard, R. S. *J. Biol. Chem.* **1920**, *XLIII*, 57–65.

(31) Haldane, J. S.; Priestley, J. G. *J. Physiol.* **1905**, *32*, 225–266.

(32) Borkenstein, R. F.; Smith, H. W. *Med. Sci. Law* **1961**, *2*, 13–22.

(33) Metz, G.; Gassull, M. A.; Leeds, A. R.; Blendis, L. M.; Jenkins, D. J. *Clin. Sci. Mol. Med.* **1976**, *50*, 237–240.

(34) Pauling, L.; Robinson, a B.; Teranishi, R.; Cary, P. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 2374–2376.

(35) Kaji, H.; Hisamura, M.; Saito, N.; Murao, M. *Clin. Chim. Acta.* **1978**, *85*,

279–284.

(36) Wallace, L. A.; Pellizzari, E.; Hartwell, T.; Rosenzweig, M.; Erickson, M.; Sparacino, C.; Zelon, H. *Environ. Res.* **1984**, *35*, 293–319.

(37) Gordon, S.; Szidon, J. *Clin. Chem.* **1985**, *31*, 1278–1282.

(38) Phillips, M.; Greenberg, J. *Clin. Chem.* **1992**, *38*, 60–65.

(39) Phillips, M.; Cataneo, R. N.; Saunders, C.; Hope, P.; Schmitt, P.; Wai, J. *J. Breath Res.* **2010**, *4*, 026003.

(40) Peng, G.; Hakim, M.; Broza, Y. Y.; Billan, S.; Abdah-Bortnyak, R.; Kuten, a; Tisch, U.; Haick, H. *Br. J. Cancer* **2010**, *103*, 542–551.

(41) Dallinga, J. W.; Robroeks, C. M. H. H. T.; van Berkel, J. J. B. N.; Moonen, E. J. C.; Godschalk, R. W. L.; Jöbsis, Q.; Dompeling, E.; Wouters, E. F. M.; van Schooten, F. J. *Clin. Exp. Allergy* **2010**, *40*, 68–76.

(42) Van Berkel, J. J. B. N.; Dallinga, J. W.; Möller, G. M.; Godschalk, R. W. L.; Moonen, E. J.; Wouters, E. F. M.; Van Schooten, F. J. *Respir. Med.* **2010**, *104*, 557–563.

(43) Phillips, M.; Cataneo, R. N.; Condos, R.; Ring Erickson, G. a; Greenberg, J.; La Bombardi, V.; Munawar, M. I.; Tietje, O. *Tuberculosis (Edinb).* **2007**, *87*, 44–52.

(44) Ruzsanyi, V.; Baumbach, J. I. *Int. J. Ion Mobil. Spectrom.* **2005**, *8*, 5–7.

(45) Turner, M. a; Bandelow, S.; Edwards, L.; Patel, P.; Martin, H. J.; Wilson, I. D.; Thomas, C. L. P. *J. Breath Res.* **2013**, *7*, 017102.

(46) Fiehn, O. *Trends Analyt. Chem.* **2008**, *27*, 261–269.

(47) O'Hara, M. E.; Clutton-Brock, T. H.; Green, S.; Mayhew, C. a. *J. Breath Res.* **2009**, *3*, 027005.

(48) Turner, C.; Spanel, P.; Smith, D. *Physiol. Meas.* **2006**, *27*, 13–22.

(49) Phillips, M.; Herrera, J.; Krishnan, S.; Zain, M.; Greenberg, J.; Cataneo, R. N. *J. Chromatogr. B. Biomed. Sci. Appl.* **1999**, *729*, 75–88.

(50) Council directive 1999/13/EC. *Off. J. Eur. Communities* **2000**, *L 269*, 1–15.

(51) Miekisch, W.; Schubert, J. *TrAC Trends Anal. Chem.* **2006**, *25*, 665–673.

(52) de Lacy Costello, B.; Amann, A.; Al-Kateb, H.; Flynn, C.; Filipiak, W.; Khalid, T.; Osborne, D.; Ratcliffe, N. M. *J. Breath Res.* **2014**, *8*, 014001.

(53) West, J. B. *Respiratory Physiology, The Essentials*; 8th ed.; Wolter Kluwer, 2008.

(54) Atkins, P.; De Paula, J. *Physical chemistry Akins*; 6th ed.; Oxford University Press, 2006.

(55) D'Souza, J. C.; Jia, C.; Mukherjee, B.; Batterman, S. *Atmos. Environ.* **2009**, *43*, 2884–2892.

(56) Edwards, R.; Schweizer, C.; Llacqua, V.; Lai, H.; Jantunen, M.; Bayeroglesby, L.; Kunzli, N. *Atmos. Environ.* **2006**, *40*, 5685–5700.

(57) Beauchamp, J. *J. Breath Res.* **2011**, *5*, 037103.

(58) Kalapos, M. P. *Biochim. Biophys. Acta - Gen. Subj.* **2003**, *1621*, 122–139.

(59) Thekedar, B.; Szymczak, W.; Höllriegl, V.; Hoeschen, C.; Oeh, U. *J. Breath Res.* **2009**, *3*, 027007.

(60) Stone, B. G.; Besse, T. J.; Duane, W. C.; Evans, C. D.; DeMaster, E. G. *Lipids* **1993**, *28*, 705–708.

(61) Lechner, M.; Moser, B.; Niederseer, D.; Karlseder, A.; Holzknecht, B.; Fuchs, M.; Colvin, S.; Tilg, H.; Rieder, J. *Respir. Physiol. Neurobiol.* **2006**, *154*, 478–483.

(62) Buszewski, B.; Kesy, M.; Ligor, T.; Amann, A. *Biomed. Chromatogr.* **2007**, *566*, 553–566.

(63) Lindinger, W.; Taucher, J.; Jordan, a; Hansel, a; Vogel, W. *Alcohol. Clin. Exp. Res.* **1997**, *21*, 939–943.

(64) Thomas, P.; Turner, M. In *Global Metabolic Profiling: Clinical Applications*; Future Science Book Series; Future Science Ltd, 2014; pp. 140–161.

(65) Phillips, M.; Gleeson, K.; Hughes, J.; Greenberg, J.; Cataneo, R.; Baker, L.; Mcvay, W. *Lancet* **1999**, *353*, 1930–1933.

(66) Phillips, M. *Chest* **2003**, *123*, 2115–2123.

(67) WHO. *Global tuberculosis report 2014*; 2014; Vol. 39.

(68) Parida, S. K.; Kaufmann, S. H. E. *Drug Discov. Today* **2010**, *15*, 148–157.

(69) Phillips, M.; Basa-Dalay, V.; Bothamley, G.; Cataneo, R. N.; Lam, P. K.; Natividad, M. P. R.; Schmitt, P.; Wai, J. *Tuberculosis (Edinb).* **2010**, *90*, 145–151.

(70) Phillips, M.; Basa-Dalay, V.; Blais, J.; Bothamley, G.; Chaturvedi, A.; Modi, K. D.; Pandya, M.; Natividad, M. P. R.; Patel, U.; Ramraje, N. N.; Schmitt, P.; Udwadia, Z. F. *Tuberculosis* **2012**, *92*, 314–320.

(71) Solga, S. F.; Risby, T. H. In *Volatile Biomarkers: Non-Invasive Diagnosis in Physiology and Medicine*; Amann, A.; Smith, D., Eds.; Elsevier Ltd, 2013; pp. 19–24.

(72) Markes International. *The BIO VOCTM Sampler: Award-Winning Device for Non-Invasive Measurement of the Body Burden of Volatile Chemical Pollutants*.

(73) Kwak, J.; Fan, M.; Harshman, S. W.; Garrison, C. E.; Dershem, V. L.; Phillips, J. B.; Grigsby, C. C.; Ott, D. K. *Metabolites* **2014**, *4*, 879–888.

(74) Miekisch, W.; Kischkel, S.; Sawacki, A.; Liebau, T.; Mieth, M.; Schubert, J. K. *J. Breath Res.* **2008**, *2*, 026007.

(75) Phillips, M. *Anal. Biochem.* **1997**, *247*, 272–278.

(76) Amann, a; Poupart, G.; Telser, S.; Ledochowski, M.; Schmid, a; Mechtcheriakov, S. *Int. J. Mass Spectrom.* **2004**, *239*, 227–233.

(77) Robinson, A. B.; Cary, P.; Pauling, L. *Time* **1972**, *XI*, 1971–1973.

(78) Ligor, T.; Ligor, M.; Amann, a; Ager, C.; Bachler, M.; Dzien, a; Buszewski, B. *J. Breath Res.* **2008**, *2*, 046006.

(79) Buszewski, B.; Ulanowska, A.; Ligor, T.; Denderz, N.; Amann, A. *Biomed. Chromatogr.* **2009**, *23*, 551–556.

(80) Gamez, G.; Zhu, L.; Disko, A.; Chen, H.; Azov, V.; Chingin, K.; Krämer, G.; Zenobi, R. *Chem. Commun. (Camb).* **2011**, *47*, 4884–4886.

(81) Martı, P. **2008**, *80*, 8210–8215.

(82) King, J.; Kupferthaler, a; Unterkofler, K.; Koc, H.; Teschl, S.; Teschl, G.; Miekisch, W.; Schubert, J.; Hinterhuber, H.; Amann, a. *J. Breath Res.* **2009**, *3*, 027006.

(83) Toyooka, T.; Hiyama, S.; Yamada, Y. *J. Breath Res.* **2013**, *7*, 036005.

(84) Chen, H.; Wortmann, A.; Zhang, W.; Zenobi, R. *Angew. Chemie - Int. Ed.* **2007**, *46*, 580–583.

(85) King, J.; Kupferthaler, a; Unterkofler, K.; Koc, H.; Teschl, S.; Teschl, G.; Miekisch, W.; Schubert, J.; Hinterhuber, H.; Amann, a. *J. Breath Res.* **2009**, *3*, 027006.

(86) Phillips, M.; Cataneo, R. N.; Chaturvedi, A.; Kaplan, P. D.; Libardoni, M.; Mundada, M.; Patel, U.; Zhang, X. *PLoS One* **2013**, *8*, 1–8.

(87) Di Natale, C.; Paolesse, R.; Martinelli, E.; Capuano, R. *Anal. Chim. Acta* **2014**, *824*, 1–7.

(88) Beauchamp, J.; Herbig, J.; Gutmann, R.; Hansel, A. *J. Breath Res.* **2008**, *2*, 046001.

(89) Barker, M.; Hengst, M.; Schmid, J.; Buers, H. J.; Mittermaier, B.; Klemp, D.; Koppmann, R. *Eur. Respir. J.* **2006**, *27*, 929–936.

(90) Mleth, M.; Schubert, J. K.; Gröger, T.; Sabei, B.; Kischkel, S.; Fuchs, P.; Hein, D.; Zimmermann, R.; Miekisch, W. *Anal. Chem.* **2010**, *82*, 2541–2551.

(91) Spaněl, P.; Dryahina, K.; Smith, D. *J. Breath Res.* **2013**, *7*, 017106.

(92) Risby, T. H.; Solga, S. F. *Appl. Phys. B* **2006**, *85*, 421–426.

(93) Lourenço, C.; Turner, C. *Metabolites* **2014**, *4*, 465–498.

(94) King, J.; Kupferthaler, A.; Frauscher, B.; Hackner, H.; Unterkofler, K.; Teschl, G.; Hinterhuber, H.; Amann, A.; Högl, B. *Physiol. Meas.* **2012**.

(95) Thekedar, B.; Szymczak, W.; Hoeschen, C.; Oeh, U. *Ger. Res.* **2009**, *3*.

(96) Schwarz, K.; Filipiak, W.; Amann, a. *J. Breath Res.* **2009**, *3*, 027002.

(97) Wehinger, a; Schmid, a; Mechtcheriakov, S.; Ledochowski, M.; Grabmer, C.; Gastl, G.; Amann, a. *Int. J. Mass Spectrom.* **2007**, *265*, 49–59.

(98) Smith, D. *Int. J. Mass Spectrom.* **2001**, *209*, 81–97.

(99) Machado, R. F.; Laskowski, D.; Deffenderfer, O.; Burch, T.; Zheng, S.; Mazzone, P. J.; Mekhail, T.; Jennings, C.; Stoller, J. K.; Pyle, J.; Duncan, J.; Dweik, R. a; Erzurum, S. C. *Am. J. Respir. Crit. Care Med.* **2005**, *171*, 1286–1291.

(100) Berchtold, C.; Meier, L.; Zenobi, R. *Int. J. Mass Spectrom.* **2011**, *299*, 145–150.

(101) Gamez, G.; Zhu, L.; Disko, A.; Chen, H.; Azov, V.; Chingin, K.; Krämer, G.; Zenobi, R. *Chem. Commun. (Camb).* **2011**, *47*, 4884–4886.

(102) Borsdorf, H.; Nazarov, E.; Miller, R. *Anal. Chim. Acta* **2006**, *575*, 76–88.

(103) McNair, H. M.; Miller, J. M. *Basic gas chromatography*; Techniques in analytical chemistry series; Wiley, 1998.

(104) Grob, R. L.; Barry, E. F. *Modern Practice of Gas Chromatography*; 4th ed.; Wiley, 2004.

(105) de Hoffman, E.; Stroobant, V. *Mass Spectrometry Principles and Applications*; 3rd ed.; John Wiley & Sons, Ltd, 2007.

(106) March, R. E. **1997**, *32*, 351–369.

(107) Stafford, G. C. J. .; Kelly, P. E. .; Syka, J. E. P. .; Reynolds, W. E. .; Todd, J. F. J. Recent improvements in and analytical applications of advanced ion trap technology. *Int. J. Mass Spectrom. Ion Processes*, 1984, *60*, 85–98.

(108) Jonscher, K. R.; Yates III, J. R. *Anal. Biochem.* **1997**, *244*, 1–15.

(109) Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*; 5th ed.; Pearson/Prentice Hall, 2005.

(110) Smith, L. I. A Tutorial on Principal Components Analysis, http://www.iro.umontreal.ca/~pift6080/H09/documents/papers/pca_tutorial.pdf (accessed Dec 18, 2014).

(111) Shlens, J. *arXiv Prepr. arXiv1404.1100* **2014**.

(112) Easrment, H.; Krzsnowski, W. *Technometrics* **1982**, *24*, 73–77.

(113) Miekisch, W.; Herbig, J.; Schubert, J. K. *J. Breath Res.* **2012**, *6*, 036007.

(114) Vul, E.; Harris, C. R.; Winkielman, P.; Pashler, H.; Vul. *Perspect. Psychol. Sci.* **2009**, *4*, 274–290.

(115) Schagen, I. *Educ. J.* 33–36.

(116) Goodacre, R.; Broadhurst, D.; Smilde, A. *Metabolomics* **2007**.

(117) Guallar-Hoyas, C.; Turner, M. a.; Blackburn, G. J.; Wilson, I. D.; Thomas, C. P. *Bioanalysis* **2012**, *4*, 2227–2237.

(118) Gahleitner, F.; Guallar-Hoyas, C.; Beardsmore, C. S.; Pandya, H. C.; Thomas, C. P. *Bioanalysis* **2013**, *5*, 2239–2247.

(119) Ettre, L. S.; Words, K. *Chromatographia* **2003**, 491–494.

(120) Creek, W. 4000 GC / MS Hardware Operational Manual http://www.agilent.com/cs/library/usermanuals/public/914998.pdf (accessed Nov 20, 2015).

(121) Health and Safety Executive. *Stress and Psychological Disorders in Great Britain 2013*; 2013.

(122) Hynninen, K. M. J.; Breitve, M. H.; Wiborg, A. B.; Pallesen, S.; Nordhus, I. H. *J. Psychosom. Res.* **2005**, *59*, 429–443.

(123) Heffner, K. L.; Loving, T. J.; Robles, T. F.; Kiecolt-Glaser, J. K. *Brain. Behav. Immun.* **2003**, *17 Suppl 1*, S109–S111.

(124) Edwards, K. M.; Burns, V. E.; Reynolds, T.; Carroll, D.; Drayson, M.; Ring, C. *Brain. Behav. Immun.* **2006**, *20*, 159–168.

(125) Houle, T.; Nash, J. M. *Headache* **2008**, *48*, 40–44.

(126) Dahl, J.; Wilson, K.; Nilsson, a. *Behav. Ther.* **2004**, *35*, 785–801.

(127) Marin, M.-F.; Lord, C.; Andrews, J.; Juster, R.-P.; Sindi, S.; Arsenault-Lapierre, G.; Fiocco, A. J.; Lupien, S. J. *Neurobiol. Learn. Mem.* **2011**, *96*, 583–595.

(128) De Boer, D.; Ring, C.; Wood, M.; Ford, C.; Jessney, N.; McIntyre, D.; Carroll, D. *Psychophysiology* **2007**, *44*, 639–649.

(129) Gronwall, D. M. *Percept. Mot. Skills* **1977**, *44*, 367–373.

(130) Umetrics, A. *User's Guide to SIMPCA-P, SIMPCA-P+" Version 11.*; Umetrics, Sweden, 2005.

(131) Epa, N.; Mass, N. I. H.; Library, S.; Ei, V. Z.; Sparkman, J. a. *NIST 2008 User Guide*; 2004.

(132) Jerković, I.; Marijanović, Z.; Kezić, J.; Gugić, M. *Molecules* **2009**, *14*, 2717–2728.

(133) Mattheis, J. P.; Buchanan, D. a.; Fellman, J. K. *J. Agric. Food Chem.* **1991**, *39*, 1602–1605.

(134) Vikram, a.; Lui, L. H.; Hossain, a.; Kushalappa, a. C. *Ann. Appl. Biol.* **2006**,

148, 17–26.

(135) Davies, S. J. C.; Hood, S. D.; Argyropoulos, S. V; Morris, K.; Bell, C.; Witchel, H. J.; Jackson, P. R.; Nutt, D. J.; Potokar, J. P. *J. Clin. Psychopharmacol.* 2006, *26*, 414–418.

(136) Tuomainen, A.; Stark, H.; Seuri, M.; Hirvonen, M. R.; Linnainmaa, M.; Sieppi, A.; Tukiainen, H. *Environ. Health Perspect.* 2006, *114*, 1409–1413.

(137) Haick, H.; Gang, P.; Tisch, U.; Zilberman, Y.; Pisula, W.; Feng, X.; Müllen, K. *Proc. SPIE* 2009, *7418*, 741809–741809 – 15.

(138) Filipiak, W.; Sponring, A.; Filipiak, A.; Ager, C.; Schubert, J.; Miekisch, W.; Amann, A.; Troppmair, J. *Cancer Epidemiol. Biomarkers Prev.* 2010, *19*, 182–195.

(139) Tanabe, Y. *Chest* 2001, *119*, 811–817.

(140) Schubert, R.; Schwoebel, H.; Mau-Moeller, A.; Behrens, M.; Fuchs, P.; Sklorz, M.; Schubert, J. K.; Bruhn, S.; Miekisch, W. *Metabolomics* 2012, *8*, 1069–1080.

(141) Bikov, A.; Lazar, Z.; Schandl, K.; Antus, B. M.; Losonczy, G.; Horvath, I. *Acta Physiol. Hung.* 2011, *98*, 321–328.

(142) Shimouchi, A.; Nose, K.; Shirai, M. *IEEE Sens. J.* 2010, *10*, 85–91.

(143) King, J.; Mochalski, P.; Kupferthaler, a; Unterkofler, K.; Koc, H.; Filipiak, W.; Teschl, S.; Hinterhuber, H.; Amann, a. *Physiol. Meas.* 2010, *31*, 1169–1184.

(144) Schwoebel, H.; Schubert, R.; Sklorz, M.; Kischkel, S.; Zimmermann, R.; Schubert, J. K.; Miekisch, W. *Anal. Bioanal. Chem.* 2011, *401*, 2079–2091.

(145) Van den Velde, S.; Nevens, F.; Van Hee, P.; van Steenberghe, D.; Quirynen, M. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 2008, *875*, 344–348.

(146) Taucher, J.; Hansel, A.; Jordan, A.; Lindinger, W. *J. Agric. Food Chem.* 1996, *44*, 3778–3782.

(147) Arasaradnam, R. P.; McFarlane, M. J.; Ryan-Fisher, C.; Westenbrink, E.; Hodges, P.; Thomas, M. G.; Chambers, S.; O'Connell, N.; Bailey, C.; Harmston, C.; Nwokolo, C. U.; Bardhan, K. D.; Covington, J. a. *PLoS One*

2014, *9*, e108750.

(148) Garner, C. E.; Smith, S.; de Lacy Costello, B.; White, P.; Spencer, R.; Probert, C. S. J.; Ratcliffe, N. M. *FASEB J.* **2007**, *21*, 1675–1688.

(149) Walton, C.; Fowler, D. P.; Turner, C.; Jia, W.; Whitehead, R. N.; Griffiths, L.; Dawson, C.; Waring, R. H.; Ramsden, D. B.; Cole, J. A.; Cauchi, M.; Bessant, C.; Hunter, J. O. *Inflamm. Bowel Dis.* **2013**, *19*, 2069–2078.

(150) ScienceOK. Inferior vena cava http://www.scienceok.com/human-body-systems/circulatory/inferior-vena-cava/ (accessed Dec 18, 2014).

APPENDICES

Appendix 1. Generic Breath Sampling Ethics

A1.1 Generic participant information sheet

Breath analysis in <Study Title>>

Part 1

Invitation to participate

We would like to invite you to take part in a research study. Before you decide you need to understand why the research is being done and what it would involve for you. Please take time to read the following information carefully. Talk to others about the study if you wish. This part (Part 1) tells you the purpose of this study and what will happen to you if you take part. Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear, or if you would like more information. Please take time to decide whether or not you wish to take part.

What is the purpose of this study?

We are testing a new way of detecting changes in the human body, through measurement of tiny amounts of chemical substances in the breath. These substances arise from all the processes and reactions that occur in the body. We think that analysis of these compounds may be able to give us information about health, diet and diseases. This research may lead to a test that can detect serious diseases earlier in the future.

Why have I been invited?

You will have been invited because you have << specific inclusion criteria goes here>> and this makes you part of the group of people we are particularly interested in seeing.

Do I have to take part?

It is up to you to decide.

We will describe the study and go through this information sheet, which we will then give to you. We will then ask you to sign a consent form to show you have agreed to take part. You are free to withdraw at any time, and you do not give us a reason.

What will happen to me if I take part?

If you agree to take part we will arrange for you to visit our <<laboratory>> for a single visit lasting about 45 minutes. During this visit we will first ask you a series of questions relating to factors that can influence the chemical substances in the breath. These will include questions about food, drink, and medications recently consumed, and smoking. We will also ask you some questions about personal care including toiletries, perfumes and personal care products. We will also ask you for information about diet and digestion, and finally we will ask for some information about <Inclusion criteria specified here>.

For the same reasons, you will be asked to refrain from eating, drinking (other than unflavoured water) and smoking for two hours prior to the study visit.

After we have collected this information, we will collect a breath sample. This involves wearing a face mask that is connected to an air supply. You will be asked to breathe normally throughout the test, whilst sitting comfortably in a chair. A sampling tube is connected to the mask on which we will collect the sample. Up to 5 samples will be collected in total and the process will take about half an hour.

What will I have to do?

We need you to not smoke, eat or drink (although you are permitted to drink unflavoured water) for two hours prior to the study visit.

We also ask you not to use perfumed products (which include not only perfume, but also deodorants, after-shaves), or any facial products (such as make-up or creams (scented or un-scented)) on the day of the study visit.

What are the possible disadvantages and risks of taking part?

We do not anticipate any risk from taking part in this study. You may experience some initial discomfort wearing the mask. We have tested the mask with volunteers with chronic bronchitis and emphysema without any difficulties. Some patients experienced slight claustrophobia when first wearing the mask but this was in very few cases and was not enough to cause anxiety. We will always be with you during this process and we will how to release the mask quickly should you wish to take it off.

What are the possible benefits of taking part?

There will be no direct benefits to you from taking part in this research. The aim is to develop new tests that may be used in future to detect diseases in people or to provide information that will help changes in lifestyle to be accurately assessed.

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. You may withdraw from the study at any time; either before or during the study. The detailed information on this is given in Part 2.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. The details are included in Part 2.

If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2

What will happen if I don't want to carry on with the study?

You are under no obligation to participate in this study and you may withdraw from the whole study, or to contribute only to a part of the study (for example you may not wish to give a breath sample, or answer the questionnaire, or *vice versa*), in which case we will use only the sample(s) and answers you agreed to provide.

We will ensure that you are able to stop the sampling procedure at any time quickly and easily should you want to.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions (contact number 0150 922 2549). You may ask a researcher to clarify anything about the study at any time, both before and during sampling. If you remain unhappy and wish to complain formally, you can do this through Loughborough University Complaints Procedure. Details can be obtained from the University.

In the event that something does go wrong and you are harmed during the research due to someone's negligence then you may have grounds for a legal action for compensation against Loughborough University but you may have to pay your legal costs.

Will my taking part in this study be kept confidential?

All information which is collected about you during the course of the research will be kept strictly confidential, and any information about you which leaves the Laboratory will have your name and address removed so that you cannot be recognised.

What will happen to the samples I give?

The breath samples are collected in small tubes that will be processed at Department of Chemistry at Loughborough University. Your name will be removed, and the samples identified only with a code. No genetic testing will be performed on any of your samples.

What will happen to the results of the research study?

It is intended to publish the results of this study in a scientific journal. Data generated will be kept at the Department of Chemistry at Loughborough University. Your name and address will not be kept with the data although a record of your participation will be maintained by the principle investigator. The measurement data will only be made available to people who are directly involved with the study.

Who is organising and funding the research?

This study is a collaboration between <<details here>>

Who has reviewed the study?

This study has been reviewed and been given a favourable opinion by Loughborough University Ethics Committee to protect your safety, rights, wellbeing and dignity.

Who should I contact for further information?

Prof. C. L. Paul Thomas

Professor of Analytical Science

Department of Chemistry

Loughborough University

LE11 3TU

Tel: 0150 922 2549

Thank you for taking time to read this information sheet and considering helping us with our research. Please retain a copy for your reference.

A1.2 Generic informed consent form

CONSENT FORM Breath analysis in <Study Title>> .

Mr/Mrs/Miss/Ms/Dr.....

Name.....

Age.....

Contact details

Trial Supervisor

Prof. Paul Thomas, Professor Analytical Science, Department of Chemistry, Loughborough University, LE11 3TU

Tel. No. 0150 922 2549

Before participating in our study we need you to complete this consent form.

Please tick the appropriate box.

I have been given a copy of the information sheet **Breath analysis in** <Study Title>>	Yes	No
I understand that my participation in this study is voluntary and that I am under no obligation to take part.	Yes	No
I have read and understood the information sheet	Yes	No
I have had the opportunity to ask questions and I understand the answers I have been give.	Yes	No
I understand that I may withdraw from the study at any time	Yes	No
I have been shown how to stop the sampling procedure if I want to.	Yes	No
I consent to participate in this study	Yes	No

PLEASE SIGN HERE

Date

Many thanks for your help.

Researcher Use:

Participant identifier:

A.1.3 Example of Generic Participant health screen questionnaire.

Health Screen Questionnaire for Breath analysis in <Study Title>> .

Mr/Mrs/Miss/Ms/Dr.....

Name.....

Age.....

Contact details

Trial Supervisor

Prof. Paul Thomas, Professor Analytical Science, Department of Chemistry, Loughborough University, LE11 3TU

Tel. No. 0150 922 2549

Before participating in our study we need you to complete this health screen questionnaire.

Please tick the appropriate box.

Do you have any chest or breathing problems?	Yes	No
Are you taking or using any medicines, pills, tablets, ointments, injections or any other drug, either from your doctor or on your own accord?	Yes	No
Do you have or have you ever had any heart or blood pressure	Yes	No
Have you experienced nausea, sickness or dizziness in the last week	Yes	No
Is there anything else concerning your health, you think we should know about?	Yes	No

PLEASE SIGN HERE

Date

Many thanks for your help.

Researcher use

Participant identifier:

Appendix 2. Retention Indexing

Retention Index Method Protocol

Scope and Introduction

Incorporation of a retention index ladder into the analysis of complex samples allows for data from inter-laboratory analyses to be aligned and compared directly, independently of the type of instrument manufacturer, column length or column flow.

Principle

A primary retention index is generated by using a retention index standard and a modified version of Kovat's retention index equation which allows for temperature programming of the gas chromatography system (Equation 1 shows this modified form).

$$I = \frac{t_{R(Unknown)} - t_{R\mu}}{t_{RN} - t_{R\mu}} \cdot 100 \cdot z \cdot \mu \qquad (1)$$

Where: $t_{R(Unknown)}$ = Retention time of component of interest (min).

$t_{R\mu}$ = Retention time of previous lower molecular weight known component (min)

t_{RN} = Retention time of next higher molecular weight known component (min)

z = Difference in carbon number between μ and N

μ = Carbon number of the lower molecular weight known component.

Using the known straight chain hydrocarbons from the retention index standard, retention index values are assigned based on the carbon number of the component. The values assigned to each of the components are then plotted against their respective retention times to produce a linear retention ladder. The equation generated from the linear trend line produced is then used to assign retention index values to the known components that are

present in all sample chromatograms to form a secondary retention index. The secondary retention index is then used to align the all sample data

Safety

Equipment	**Nature of Hazard**	**Precautions to be taken**
GC	Heat	When the GC is in use the oven, injectors and detectors are all heated. Do not touch these areas. Do not operate the GC with any covers or parts removed
Gas supply pipe lines	High pressure	Never apply oil or grease to joints Wear safety glasses when working with compressed gas Close valves when apparatus is not in use
Mass Spec Transfer Line	Heat Vacuum	When the GC oven door is open the transfer line is exposed and is heated even when the GC oven is off
Chemical	**Classification**	**First Aid Measures**
Hydrocarbons Standards	Xn, N, R65, R65 - R67, R20 - R36/37/38 - R51/53	**General advice** Consult a physician. Show this safety data sheet to the doctor in attendance. **If inhaled** If breathed in, move person into fresh air. If not breathing give artificial respiration Consult a physician. **In case of skin contact** Wash off with soap and plenty of water. Consult a physician. **In case of eye contact** Rinse thoroughly with plenty of water for at least 15 minutes and consult a physician. **If swallowed**

		Do NOT induce vomiting. Never give anything by mouth to an unconscious person. Rinse mouth with water. Consult a physician. **Handling** Avoid inhalation of vapour or mist. Normal measures for preventive fire protection.
Dichloromethane	Xn, Carc.Cat.3, R40	**General advice** Consult a physician. Show this safety data sheet to the doctor in attendance. **If inhaled** If breathed in, move person into fresh air. If not breathing give artificial respiration Consult a physician. **In case of skin contact** Wash off with soap and plenty of water. Consult a physician. **In case of eye contact** Rinse thoroughly with plenty of water for at least 15 minutes and consult a physician. Aldrich - 239763 www.sigma-aldrich.com Page 2 of 5 **If swallowed** Never give anything by mouth to an unconscious person. Rinse mouth with water. Consult a physician. **Handling** Avoid inhalation of vapour or mist. Normal measures for preventive fire protection.

Chemicals

1- Chlorodecane	CAS: 1002-69-3	Available from Sigma (24205)
1- Chlorohexane	CAS: 544-10-5	Available from Sigma (24771)
1-Chlorononane	CAS: 2473-01-0	Available from Sigma (238457)
1- Chlorooctane	CAS: 111-85-3	Available from Sigma (25660)
1- Decanol	CAS: 112-30-1	Available from Sigma (30605)
1-Hexanol	CAS: 111-27-3	Available from Sigma (73117)
2- Heptanol	CAS: 543-49-7	Available from Sigma (51800)
2- Octanol	CAS: 126-96-6	Available from Sigma (74858)
5- Nonanol	CAS: 623-93-8	Available from Sigma (74308)
Decane	CAS: 124-18-5	Available from Sigma (30540)
Dichloromethane	CAS: 79-05-2	Fisher Scientific (D/1859/17)
Dodecane	CAS: 112-40-3	Available from Sigma (44010)
Eicosane	CAS: 112-95-8	Available from Sigma (44818)
Heptadecane	CAS: 629-78-7	Available from Sigma (51578)
Nonane	CAS: 111-84-2	Available from Sigma(74250)
Octane	CAS: 111-65-9	Available from Sigma (74821)
Tetradecane	CAS: 629-59-4	Available from Sigma (87139)
Undecane	CAS: 1120-21-4	Available from Sigma (94000)

Apparatus and Glassware

An analytical balance capable of weighing to 50 ± 1 mg

3 x 10 ml volumetric flasks (grade A) with stoppers

Auto pipette capable of dispensing in the range of 20 to 100 ul

Weighing boat

Tenax® / Carbotrap Mixed mode thermal desorption tube

1 µl glass sample syringe

Reagents and Solutions

Mixed Stock Retention Index Solution.

Using a calibrated auto pipette transfer 70 ± 5 µl of each of the components 1 to 16 in Table 4 into a clean 10 ml volumetric flask containing approximately 4 ml of dichloromethane.

Make the volumetric flask up to volume using dichloromethane. Stopper the volumetric flask and invert several times to ensure the content is mixed.

(This solution contains approximately 7 µ / ml of components 1 to 16 listed in Table 4.)

Heptadecane and Eicosane Stock Standard Solution.

Using a five place analytical balance weigh 50 ± 1 mg of Eicosane in to a suitable weighing boat.

Quantitatively transfer the contents of the weighing boat to a clean, dry 10 ml volumetric flask using not more than 8 ml of dichloromethane

With the aid of a calibrated auto pipette transfer 50 ± 5 µl of Heptadecane to the 10 ml volumetric flask

Make the volumetric flask up to volume with dichloromethane, stopper the volumetric flask and invert several times to ensure the content is mixed.

(This solution contains approximately 5 mg / ml and 5 µl / ml of heptadecane and eicosane.)

Retention Index Solution

Into a clean 10ml volumetric flask containing approximately 8 ml of dichloromethane transfer 20 ± 2 all of Mixed Stock Retention Index Solution using a calibrated auto pipette

Into the same 10 ml volumetric flask transfer 20 + 2 all of Heptadecane and Eicosane Stock Standard Solution using a calibrated auto pipette.

Make the volumetric flask up to volume with dichloromethane, stopper the volumetric flask and invert several times to ensure the content is mixed.

(This solution contains approximately 50 ng/µl of each of the compounds in Table 1 and heptadecane and eicosane).

Instrumentation

Markes International Unity Thermal desorber, equipped with a general purpose hydrophobic cold trap.

Gas Chromatograph with an oven capable of heating at a rate of 40^{o}C

Mass spectrometer capable of scanning 40 to 450 m/z

Operating conditions

Table A2.1. Summary of Marks International thermal desorption sampler parameters

Parameter	**Setting**
Primary desorption flow	50 cm^3 min^{-1}
Primary Split	10:1 Ratio
Primary desorption temperature	300 ^{o}C
Primary desorption time	5 min
Cold trap volume	0.019 cm^3
Cold trap temperature	-10 ^{o}C
Cold trap packing	U-T2GPH (General purpose hydrophobic)
Secondary desorption flow	2 cm^3 min^{-1}
Secondary desorption temperature	300 ^{o}C

Secondary desorption time	5 min
Trap heating rate	100 ^{o}C min^{-1}
Transfer line temperature	140 ^{o}C

Table A2.2. Summary of Varian 3800 gas chromatograph conditions

Parameter	**Setting**
Column	60 m x 0.25 mm x 0.25 µm DB 5
He carrier gas flow	2.0 cm^3 min^{-1}
Initial oven temperature	40 ^{o}C
Initial hold time	0 min
Oven temperature program	3.3 ^{o}C min^{-1} to 90 ^{o}C
	2.5 ^{o}C min^{-1} to 140 ^{o}C
	10 ^{o}C min^{-1} to 300 ^{o}C hold for 8.85 min
Total run time	60 min

Table A2.3. Summary of Varian 4000 ion trap mass spectrometer conditions

Parameter	**Setting**
Scan type	Full
Mass range	40 to 445 m/z
Tune type	Auto
Ionization type	EI
Target TIC	20000 counts
Max ion time	25000 µ Seconds
Emission current	10 µ A

Total run time	60 min
Scan time	0.82 Seconds
Transfer line temperature	300 ^{o}C
Trap temperature	150 ^{o}C
Manifold temperature	50 ^{o}C
Solvent delay time	5.0 min

Retention Index Analysis

A retention Index standard should be run at least once a day when breath samples are being analysed. The retention index provides not only a primary retention ladder, should it be required, but also quality control to ensure that the instrument is responding in an appropriate way day to day.

Remove the Swagelok fitting from the sampling side of the thermal desorption tube.

With the aid of a 1 µl syringe inject 0.1 µl of Retention Index Solution onto a mixed sorbent bed thermal desorption tube into

Quickly replace the Swagelok fitting to the sampling side of the thermal desorption tube.

Insert the desorption tube into the Unity Thermal desorber.

Analyse the retention index sample under the conditions described in section 8.

Retention Index Assignment

Retention Index assignment is achieved by designating a retention index value (RI) to each of the straight chain hydrocarbon components in the retention index standard. The RI for each of the 8 hydrocarbons is based on the carbon number of the component. The assignments used in this method to align data are given in Table A2.4.

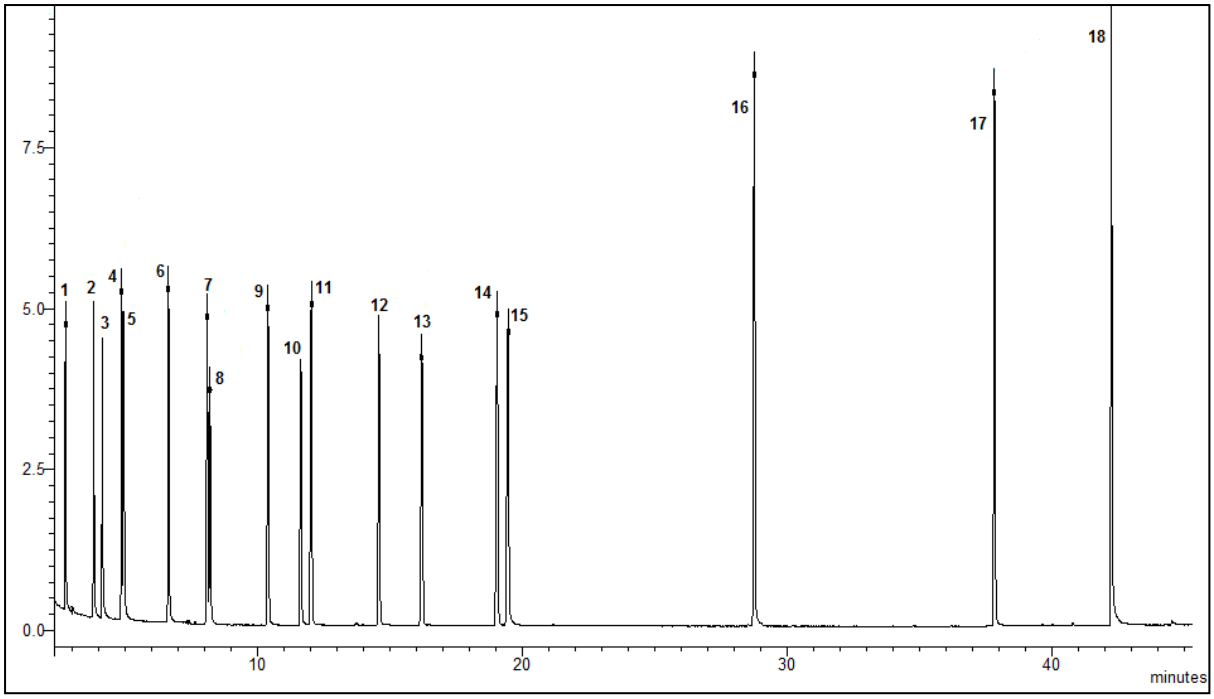


Figure A2.1 Example of a Total Ion Current (TIC) of a Retention Index Standard.

Table A2.4. Component list of identified peaks in the retention index standard.

Peak Assignment	Analyte Name	Assigned RI		Peak Assignment	Analyte Name	Assigned RI
1	Octane	800		10	Undecane	1100
2	1-Chlorohexane			11	Nonanol	
3	1-Hexanol			12	1-Chlorononane	
4	Nonane	900		13	Dodecane	1200
5	2-Heptanol			14	Decanol	
6	1-Chloroheptane			15	1-chlorodecane	
7	Decane	1000		16	Tetradecane	1400
8	2-Octanol			17	Heptadecane	1700
9	1-Chlorooctane			18	Eicosane	2000

Construction of Primary Retention Index Ladder.

Plot a graph of the retention times and the assigned RI values (Table A2.4) for the Octane, Nonane, Decane, Undecane, Dodecane, Tetradecane, Heptadecane and Eicosane.

Use the graph to obtain and record the intercept (C) and gradient (M) of the trend line produced.

Secondary Retention Index

It is not always practical to add numerous retention index compounds to a sample at point of collection for the purpose of aligning samples and to aid analyte prospecting. In such instances a secondary retention index can be employed. Under the specified method conditions in section 8, a number of endogenous siloxane components that result from the interactions of moisture found in breath and the sorbent beds in the thermal desorption tube have been selected. These siloxane components are continually present in all the sample chromatograms and have been selected to build such a retention index.

With the aid of a breath sample chromatogram, identify and record the retention times for the 5 endogenous siloxane peaks (S1 to S5) indicated in Figure A2.2.

Using the retention times from the identified siloxane components and equation (2) calculate the RI value for each of the siloxane compounds.

$$RI = \frac{t_{R(\text{component})}}{\text{M}} - \text{C} \qquad (2)$$

Where: $t_{R(\text{component})}$ = Retention time of Identified component

M = Gradient of the graph recorded in 10.1.2

C = Intercept of graph recorded in 10.1.2

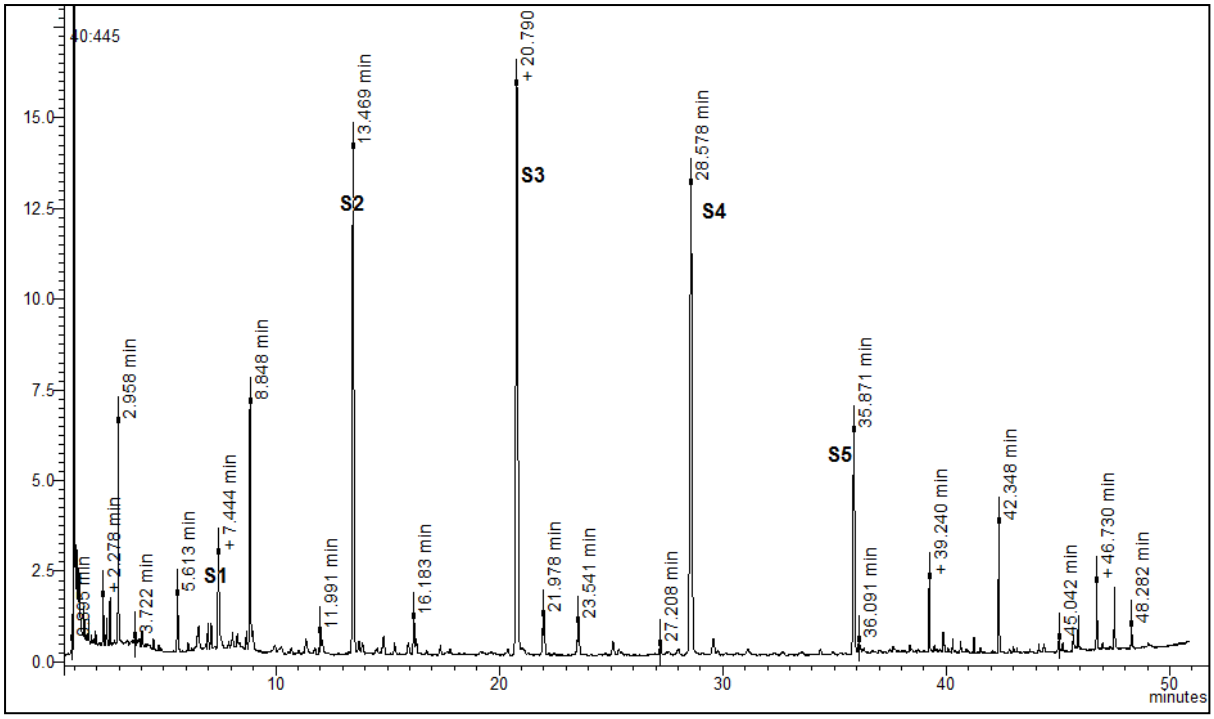


Figure A2.2 Example of breath sample chromatogram with endogenous siloxane peaks (S1 to S5) highlighted

Sample Alignment

Sample alignment is achieved by assigning a static RI value to the known siloxane peaks (S1 to S5) that occur in a particular sample chromatogram. The retention times for each of the siloxane and the assigned RI values are plotted to produce a calibration curve, from which the remaining analyte peaks for that particular sample may be assigned an RI value, therefore enabling the direct comparison with other breath sample chromatograms.

Using the static RI values in Table A2.5 and the retention times for the siloxane components in a breath sample construct a linear calibration graph.

Using the gradient (M) and the intercept (C) obtained from the constructed calibration graph (12.1), calculate the RI for the remaining components in the breath sample chromatogram using the component's retention time and equation 2.

For each breath sample chromatogram follow steps 12.1 and 12.2

The resultant list of retention indices and chromatogram components will now be directly comparable to each other.

Table A2.5 Static RI values and approximate retention times for siloxane peaks used in secondary retention index

Peak Assignment	Static RI Value	Approximate Retention Time / minutes
S1	998	7.5
S2	1142	13.3
S3	1303	20.6
S4	1516	28.5
S5	1664	36.8

Method Transfer Assessment

For the comparison of different data obtained through analysis on alternative equipment or in a different laboratory a method transfer assessment must be performed to ensure correct alignment of data. To accomplish this, the results of two assessments need to be compared.

Assessment 1

A Retention Index standard must be prepared as described in section 6.

The new and supplied retention index standards must be run under the conditions in section 8. A minimum of six repeats for each standard should be performed.

The newly prepared retention index standard is compared directly with that of a supplied retention index standard to ensure that all components are present when analysing the retention index under the conditions in section 8.

Each of the hydrocarbons in each chromatogram should be integrated and the peak areas recorded.

From the peak areas for each set of repeats the mean response and the relative standard deviation for each component should be calculated.

Assessment 2

Two duplicate breath samples are required to be analysed by each system / laboratory.

The component peaks are then required to be aligned using the secondary retention index (Section 11) and sample alignment (Section 12) procedures.

The within lab duplicate samples should then be compared to ensure that alignment between the samples has been achieved.

The mean Retention Index values of the five most intense analyte peaks (not including siloxane peaks used to construct the secondary retention index) should then be obtained for each set of duplicate samples and compared to ensure that the samples have been correctly aligned.

Appendix 3. Peer Reviewed Publications

Turner, M. A. (2014). Metabolic profiling of lung disease. In I. Wilson (Ed.), *Global Metabolic Profiling: Clinical Applications* (pp. 140-161). Future Science Ltd. doi:10.4155/9781909453616

Ratiu, I. -A., Bocos-Bintintan, V., Turner, M., Moll, V. -H., & Thomas, C. L. P. (2014). Discrimination of Chemical Profiles of Some Bacterial Species by Analyzing Culture Headspace Air Samples Using TD-GC/MS. *CURRENT ANALYTICAL CHEMISTRY*, *10*(4), 488-497.

Malkar, A., Devenport, N. A., Martin, H. J., Patel, P., Turner, M. A., Watson, P., . . . Creaser, C. S. (2013). Metabolic profiling of human saliva before and after induced physiological stress by ultra-high performance liquid chromatography-ion mobility-mass spectrometry. *Metabolomics*, *9*(6), 1192-1201.

doi:10.1007/s11306-013-0541-x

Turner, M. A., Bandelow, S., Edwards, L., Patel, P., Martin, H. J., Wilson, I. D., . . . Thomas, C. L. P. (2013). The effect of a paced auditory serial addition test (PASAT) intervention on the profile of volatile organic compounds in human breath: a pilot study. *J. Breath Res.*, *7*(1), 1-11.

doi:10.1088/1752-7155/7/1/017102

Guallar-Hoyas, C., Turner, M. A., Blackburn, G. J., Wilson, I. D., & Thomas, C. L. P. (2012). A workflow for the metabolomic/metabonomic investigation of exhaled breath using thermal desorption GC-MS. *BIOANALYSIS*, *4*(18), 2227-2237. doi:10.4155/BIO.12.193

Turner, M. A., Guallar-Hoyas, C., Kent, A. L., Wilson, I. D., & Thomas, C. L. P. (2011). Comparison of metabolomic profiles obtained using chemical ionization and electron ionization MS in exhaled breath. *BIOANALYSIS*, *3*(24), 2731-2738. doi:10.4155/BIO.11.284

Huo, R., Agapiou, A., Bocos-Bintintan, V., Brown, L. J., Burns, C., Turner M.A., . . . Thomas, C. L. P. (2011). The trapped human experiment. *JOURNAL OF BREATH RESEARCH*, *5*(4). doi:10.1088/1752-7155/5/4/046006