

Mass Spectrometric Analysis of Volatile Organic Compounds in Oesophago-Gastric Cancer

Sacheen Kumar BSc (Hons) MBBS MRCS (Eng)

Department of Surgery and Cancer, Imperial College London,
St Mary's Hospital, London, UK

2014

A thesis submitted to Imperial College London for the degree of
Doctorate of Philosophy

Supervisors

Professor George B Hanna PhD FRCS

Department of Surgery and Cancer, Imperial College London,
St Mary's Hospital, London, UK

Professor David Smith DSc FRS

Institute for Science & Technology in Medicine, Keele University,
Stoke-on-Trent, UK

THESIS DECLARATION

I hereby declare that I am the sole author of this thesis and that all work within it is my own. Any individuals who carried out any collaboration work are appropriately credited.

Signed:

Dated:

Sacheen Kumar

COPYRIGHT DECLARATION

‘The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work’

ABSTRACT

In the UK, oesophago-gastric (OG) cancer remains a disease with poor patient outcomes and only 35% of patients are suitable for potentially curative treatment at first presentation. Early detection of these cancers is necessary with Volatile Organic Compound (VOC) analysis being a promising area to explore for new endogenous biomarkers.

In this research, the application of Selected-Ion Flow Tube Mass Spectrometry (SIFT-MS), a real-time technique, for the analysis of VOCs in biofluids and exhaled breath has been investigated. The primary aim of this research was to identify potential VOC biomarkers in oesophago-gastric cancer; the long-term application being the translational development of VOCs as non-invasive diagnostic tests. A total of 13 VOCs emitted from gastric content were investigated in the first study; acetone, formaldehyde, acetaldehyde, hexanoic acid, hydrogen sulphide, acetic acid, hydrogen cyanide and methyl phenol were found to be significantly different between oesophago-gastric cancer and healthy groups. In a study on the headspace of urine, acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulphide, methanol and phenol were found to be at significantly different concentrations between oesophago-gastric cancer, positive control and healthy groups. Methodology experiments including the optimisation of VOC sampling techniques with evaluation of exogenous contamination and variability measures have also been conducted.

Investigation of reference ranges for C3-C10 aldehydes in healthy subjects using SIFT-MS have been undertaken for the first time; C4-C10 aldehydes were present in the exhaled breath of healthy persons at concentrations < 3 ppbv. In a study on 210 consecutive patients, increased concentrations of fatty acids, phenols and aldehydes were observed in the exhaled breath of patients with oesophageal and gastric adenocarcinoma. The potential effects of clinical confounding factors have been also assessed and possible explanations for the observed differences have been discussed. The results of the pilot study on acute infection demonstrated minimal influence on the cancer-linked VOCs in OG cancer, except for butanal. The analysis of gastric content, urine and exhaled breath have proven that trace compounds are more influential than abundant metabolites in upper gastro-intestinal malignancy. It is also demonstrated that VOC profiles rather than individual compounds are more accurate in the molecular-orientated diagnosis of oesophago-gastric cancer.

PEER REVIEWED PUBLICATIONS & PRESENTATIONS

Publications

- Mass Spectrometric Analysis of Exhaled Breath for the Identification of Volatile Organic Compound Biomarkers in Esophageal and Gastric Adenocarcinoma

Kumar S & Huang J, Abbassi-Ghadi N, Mackenzie HA, Veselkov KA, Hoare JM, Lovat LB, Španěl P, Smith D, Hanna GB – accepted for publication in *Annals of Surgery* (2015 Jan 8. [Epub ahead of print]. PMID: 25575255)

- Investigation of C3-C10 aldehydes in the exhaled breath of healthy subjects using Selected Ion Flow Tube-Mass Spectrometry (SIFT-MS)

Huang J & Kumar S, Hanna GB – published in *Journal of Breath Research* (2014 Sep; 8(3):037104. PMID: 25190002)

- Selected ion flow tube mass spectrometry analysis of exhaled breath for volatile organic compound profiling in esophago-gastric cancer

Kumar S & Huang J, Abbassi-Ghadi N, Španěl P, Smith D Hanna GB – published in *Analytical Chemistry* (2013 Jun;18;85(12):6121-8. PMID: 23659180)

- Selected ion flow tube mass spectrometry analysis of volatile metabolites in urine headspace for the profiling of gastro-esophageal cancer

Huang J & Kumar S, Abbassi-Ghadi N, Španěl P, Smith D Hanna GB – published in *Analytical Chemistry* (2013 Mar 19;85(6):3409-16 PMID: 23421902)

- SIFT-MS analysis of headspace vapour from gastric content for the diagnosis of gastro-oesophageal cancer.

Kumar S & Huang J, Cushnir JR, Španěl P, Smith D Hanna GB – published in *Analytical Chemistry* (2012 Nov 6;84(21):9550-7 PMID: 23035898)

Presentations

- ‘Sniffing Out Cancer’

Kumar S, Hanna GB – invited lecture at 6th MATTU Oesophago-Gastric Symposium in the Postgraduate Medical School, University of Surrey, Guildford (December 2014)

- SIFT-MS Analysis of Exhaled Breath for Volatile Organic Profiling of Oesophago-Gastric Cancer

Kumar S, Huang J, Abbassi-Ghadi N, Mackenzie HA, Hoare JA, Španěl P, Smith D, Hanna GB – oral presentation at the Society of Academic & Research Surgery (SARS) meeting in Robinson College, Cambridge (January 2014)

- SIFT-MS Analysis of Headspace Vapour from Gastric Content for the diagnosis of Gastro-Oesophageal Cancer

Kumar S, Huang J, Hanna GB (Academic Surgical Unit, Imperial College, London) – oral presentation shortlisted for the Patey Prize at the Society of Academic & Research Surgery (SARS) meeting in London (January 2013)

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to the many people who have helped me during the course of my research. The first person I would like to thank is my mentor - Professor George B Hanna. His patience, constant guidance and scientific training have been instrumental in helping me to complete this research. Professor Hanna has been an inspirational role model over the last three years and I have learnt many important lessons which will remain with me throughout my career. I am indebted to Professor David Smith and Professor Patrik Španěl, both of whom provided me with invaluable advice and their expert insights during my research. They have taught me the principles of Selected Ion Flow Tube Mass Spectrometry and their continuous support was essential to the completion of this thesis. I would like to thank Dr. Juzheng (Frank) Huang who was an everpresent laboratory partner. I would like to thank Mr. Nima Abbassi-Ghadi for his help with patient recruitment and in the collecting of clinical data. I am grateful to Mr. Hugh Mackenzie and Dr. Kirill Veselkov for their help with statistical analysis. I would like to thank Mr. Krishna Moorthy, Dr. Alison Knaggs and the St. Mary's Hospital Endoscopy Staff for all their help with patient recruitment. I am also thankful and indebted to Miss. Anisha Perera, who was with me from the beginning till the end of my research time on the 10th floor at St. Mary's.

My final words of gratitude and appreciation are for my family and in particular my parents. I am and always will be eternally grateful for the unwavering support you have given me throughout my career. This thesis would not have come to fruition without your patience, understanding and words of encouragement, and therefore belongs to you as much as me.

TABLE OF CONTENTS

ABSTRACT	iii
PEER REVIEWED PUBLICATIONS & PRESENTATIONS	iv
TABLE OF CONTENTS	vii
LIST OF FIGURES	xii
LIST OF TABLES	xiv
ABBREVIATIONS	xv
Chapter 1 – OESOPHAGEAL AND GASTRIC CANCER	1
1.1 Introduction	2
1.2 Oesophageal and Gastric Cancer Trends	2
1.2.1 Incidence	2
1.2.2 Survival	3
1.2.3 Diagnosis and National Guidelines	3
1.2.4 Diagnostic Biomarkers in Oesophageal and Gastric Cancer	5
Chapter 2 – VOLATILE ORGANIC COMPOUNDS AND CANCER	7
2.1 Introduction	8
2.2 Objective	10
2.3 Methods	10
2.3.1 Search Strategy	10
2.3.2 Study Quality Assessment	11
2.4 Results	11
2.4.1 Search Strategy Results and Selected Studies	11
2.4.2 Quality assessment of Studies	13
2.5 Discussion	14
2.5.1 Exhaled Breath VOCs and Cancer	14
2.5.2 Urine VOCs and Cancer	16
2.5.3 Sensor Technologies	17
2.5.4 Blood VOCs and Cancer	18
2.5.5 VOCs and Cancer Cell lines	18
2.5.6 SIFT-MS and PTR-MS	19
2.6 Conclusions	20
2.6.1 Aldehydes	21
2.6.2 VOCs as potential diagnostic tools	22

2.6.3 Thesis Aims	23
Chapter 3 – SELECTED ION FLOW TUBE MASS SPECTROMETRY	24
3.1 The Selected Ion Flow Tube Mass Spectrometry Technique	25
3.2 Ion Chemistry in SIFT-MS	26
3.3 The development of the Profile 3 SIFT-MS instrument	26
3.4 Modes of acquisition of SIFT-MS data.....	28
3.5 Sensitivity and limit of detection of the Profile 3 SIFT-MS instrument.....	28
3.6 Direct analysis of ambient gaseous samples	29
3.7 The SIFT-MS Kinetics Library.....	33
3.8 Headspace of Biofluids and Exhaled Breath Sampling	34
Chapter 4 – VOCs in the headspace of gastric content.....	35
4.1 VOCs within gastric content in OG Cancer	36
4.1.1 Objective	36
4.1.2 Introduction.....	36
4.1.3 Methods.....	37
4.1.3.1 Patient selection	37
4.1.3.2 Gastric content sampling.....	38
4.1.3.3 SIFT-MS analysis	39
4.1.3.4 Statistical analysis.....	40
4.1.4 Results.....	41
4.1.4.1 Analysis of full scan spectra	41
4.1.4.2 Reproducibility and validity of the sampling method.....	42
4.1.4.3 MIM analysis of selected VOCs in the headspace of gastric content	43
4.1.4.4 VOCs within gastric content headspace to differentiate cancer from healthy controls	47
4.1.5 Discussion.....	49
4.1.5.1 Aldehydes and Ketones.....	49
4.1.5.2 Acids	50
4.1.5.3 Alcohol and Phenol derivatives	51
4.1.5.4 Sulphur and nitrogen-containing compounds	52
4.1.6 Conclusions.....	53
4.2 The effect of pH on VOCs in gastric content headspace	54
4.2.1 Objective	54
4.2.2 Methods.....	54
4.2.2.1 Gastric content sampling and storage	54
4.2.2.2 Gastric content sample preparation & SIFT-MS analysis.....	54

4.2.3 Results.....	55
4.2.4 Discussion.....	58
4.2.4.1 Aldehydes	58
4.2.4.2 Acids	58
4.2.4.3 Acetone & Methyl Phenol.....	59
4.2.4.4 Sulphur and nitrogen-containing compounds	59
4.2.5 Conclusions.....	60
Chapter 5 – VOCs in the headspace of urine	61
5.1 Objective.....	62
5.2 Introduction.....	62
5.3 Methods.....	65
5.3.1 Patient selection	65
5.3.2 Urine sampling procedures	65
5.3.3 SIFT-MS analysis of urine headspace.....	66
5.3.4 Statistical analysis	67
5.4 Results.....	68
5.4.1 Analysis of full scan spectra using H_3O^+ and NO^+	68
5.4.2 Reproducibility and validity of the sampling method	69
5.4.3 MIM analysis of selected VOCs in the headspace of urine	70
5.4.4 pH measurements of urine	76
5.4.5 VOCs within urine headspace to differentiate OG cancer from non-cancer controls.....	77
5.5 Discussion.....	79
5.5.1 Aldehydes and Ketones.....	79
5.5.2 Sulphur compounds.....	80
5.5.3 Acids	81
5.5.4 Alcohols and Phenol derivatives.....	82
5.6 Conclusions.....	84
Chapter 6 – VOCs and Exhaled Breath.....	85
6.1 Exhaled Breath VOCs in OG Cancer	86
6.1.1 Objective	88
6.1.2 Introduction.....	88
6.1.3 Methods.....	89
6.1.3.1 Patient selection	89
6.1.3.2 Breath Sample Collection	90
6.1.3.3 SIFT-MS analysis of exhaled breath samples.....	92

6.1.3.4 Statistical analysis	93
6.1.4 Results.....	94
6.1.4.1 Water vapour level of breath samples	94
6.1.4.2 Reproducibility and validity of the sampling method	95
6.1.4.3 Surgical Sciences Laboratory room air and Nalophan bag measurements	95
6.1.4.4 Analysis of selected VOCs in exhaled breath	96
6.1.4.5 Exhaled Breath VOCs to discriminate cancer from non-cancer controls	101
6.1.5 Discussion	103
6.1.5.1 Abundant breath VOCs	103
6.1.5.2 Hexanoic Acid.....	105
6.1.5.3 Phenol and phenol derivatives	107
6.1.6 Conclusions.....	109
6.2 Exhaled breath Aldehydes in a Healthy population	110
6.2.1 Objective	110
6.2.2 Introduction.....	110
6.2.3 Methods.....	111
6.2.3.1 Subjects.....	111
6.2.3.2 Breath sample collection.....	111
6.2.3.3 SIFT-MS analysis	112
6.2.3.4 Ion chemistry	112
6.2.4 Results.....	114
6.2.4.1 Aldehydes quantification and variations within healthy subjects	114
6.2.4.2 C3 to C10 aldehyde concentrations in exhaled breath of healthy subjects	116
6.2.5 Discussion	119
6.2.6 Conclusions.....	121
6.3 Exhaled breath VOC risk prediction model in oesophageal and gastric adenocarcinoma	122
6.3.1 Objective	122
6.3.2 Introduction.....	122
6.3.3 Methods.....	123
6.3.3.1 Study Population.....	123
6.3.3.2 Procedures.....	123
6.3.3.3 SIFT-MS Analysis	124
6.3.3.4 Statistical analysis	126
6.3.3.5 Risk Prediction Model	127
6.3.4 Results.....	129

6.3.4.1 Patient Characteristics	129
6.3.4.2 Volatile Organic Compound Analysis	135
6.3.4.3 Risk Prediction Model	148
6.3.5 Discussion	150
6.3.6 Conclusions	153
6.4 The effect of acute infection on cancer-related VOCs	154
6.4.1 Objective	154
6.4.2 Introduction	154
6.4.3 Methods	155
6.4.4 Results	156
6.4.5 Discussion	162
6.4.6 Conclusions	163
Chapter 7 – Summary, Conclusions and Future work	164
REFERENCES	171
APPENDICES	183

LIST OF FIGURES

Figure 2.1: Systematic Review Search Strategy	12
Figure 2.2: QUADAS-2 Assessments of Study Quality	13
Figure 2.3: Bio-samples analysis strategy in oesophageal and gastric cancer.....	23
Figure 3.1 Schematic diagram of the <i>Profile 3</i> SIFT-MS instrument.....	31
Figure 3.2: Water vapour and acetone measurements using the MIM mode	32
Figure 4.1 Suction Trap for the collection of gastric content	38
Figure 4.2: Full scan spectra of SIFT-MS analysis of the headspace of gastric content	41
Figure 4.3: Box-Whisker plots of acetaldehyde, formaldehyde, acetone, acetic acid, hexanoic acid and pH measurements of the gastric content of cancer, positive control and normal upper GI groups	45
Figure 4.4: Box-Whisker plots of methanol, methyl phenol, hydrogen sulphide and hydrogen cyanide from the gastric content of cancer, positive control and normal upper GI groups.....	46
Figure 4.5: ROC curve for oesophago-gastric cancer diagnosis using gastric content VOCs	48
Figure 4.6: Bar graphs of acetaldehyde, formaldehyde, hexanoic acid and acetic acid concentrations from the headspace of gastric content at pH 1, 2, 4, 6 & 8	56
Figure 4.7: Bar graphs of acetone, hydrogen sulphide, hydrogen cyanide and methyl phenol concentrations from the headspace of gastric content at pH 1, 2, 4, 6 & 8	57
Figure 5.1: Full scan spectra of the headspace of urine using H_3O^+ and NO^+ precursor ions	68
Figure 5.2: Box-Whisker plots of acetaldehyde and acetone concentrations in the headspace of urine of OG cancer patients, positive controls and healthy groups	72
Figure 5.3: Box-Whisker plots of hydrogen sulphide in the headspace of urine of OG cancer patients, positive controls and healthy groups	73
Figure 5.4: Box-Whisker plots of acetic acid and hexanoic acid in the headspace of urine of OG cancer patients, positive controls and healthy groups	74
Figure 5.5: Box-Whisker plots of methanol and phenol in the headspace of urine of OG cancer, positive control and healthy groups	75
Figure 5.6: Bar graphs of the pH measurement of urine samples across the study groups	76
Figure 5.7: ROC curve for the diagnosis of oesophago-gastric cancer from non-cancer controls using urine VOCs.....	78
Figure 6.1: Nalophan [®] bags employed for the collection of exhaled breath samples	91
Figure 6.2: Box-Whisker plots of acetone, methanol and isoprene from the exhaled breath of oesophago-gastric cancer patients, positive controls and the healthy upper GI tract cohort ...	99
Figure 6.3: Box-Whisker plots of hexanoic acid, phenol, methyl phenol and ethyl phenol from the exhaled breath of oesophago-gastric cancer patients, positive controls and healthy controls	100
Figure 6.4: ROC curve of discrimination results for the diagnosis of oesophago-gastric cancer from positive controls using exhaled breath VOCs	102
Figure 6.5: Full scan spectrum of C3-C10 aldehydes within exhaled breath	115
Figure 6.6: Box-Whisker plots of exhaled breath C3 to C10 aldehydes in healthy subjects.	117
Figure 6.7: Correlation of the concentrations of hexanal with heptanal in exhaled breath ...	118

Figure 6.8: Box-Whisker plots of isoprene, acetic acid and acetone in exhaled breath of healthy subjects.....	118
Figure 6.9: Heat map of the investigated VOCs from all patients across the study groups ..	137
Figure 6.10: Box-Whisker plots of exhaled breath butyric, pentanoic and hexanoic acids...	138
Figure 6.11: Box-Whisker plots of exhaled breath phenol, methyl phenol and ethyl phenol	139
Figure 6.12: Box-Whisker plots of exhaled breath heptanal, nonanal and decanal.....	140
Figure 6.13: Box-Whisker plots of the abundant breath metabolites - methanol, acetone, isoprene and ammonia	141
Figure 6.14: ROC curves for gastric and oesophageal cancer versus normal upper GI tract	143
Figure 6.15: ROC curves for gastric and oesophageal cancer versus all non-cancer controls.....	144
Figure 6.16: Frequency histogram of VOCs used in the diagnostic model to discriminate OG adenocarcinoma from non-cancer controls	148
Figure 6.17: ROC curves for Model subset and Validation subset of the diagnostic prediction model to discriminate upper gastro-intestinal adenocarcinoma from non-cancer controls ...	149
Figure 6.18: Box-Whisker plots of pentanoic acid and hexanoic acid in acute infection	158
Figure 6.19: Box-Whisker plots of phenolic compounds in acute infection	159
Figure 6.20: Box-Whisker plots of butanal, pentanal, hexanal and heptanal in acute infection	160
Figure 6.21: Box-Whisker plots of octanal, nonanal and decanal	161

LIST OF TABLES

Table 3.1: Sample Kinetics Library entries for acetone using all three precursor ions	34
Table 4.1: Summary of the analytical information of VOCs analysed in gastric content	43
Table 4.2: Summary of VOCs that were significantly different in the headspace of gastric content between cancer and healthy groups.....	44
Table 5.1: Summary of the analytical information of VOCs detected within the headspace of urine	70
Table 5.2: Summary of the seven VOCs that were significantly different between oesophago-gastric cancer and non-cancer controls	71
Table 6.1: Median water vapour levels in Nalophan® bags at 30, 60 120 and 240 minutes: ..	94
Table 6.2: Summary of the characteristic ions of the VOCs analysed by the MIM mode in exhaled breath	97
Table 6.3: Summary of the most abundant VOCs present in exhaled breath and the 4 VOCs that were significantly different between the oesophago-gastric cancer cohort and controls .	98
Table 6.4: Individual % signals for m/z 117, m/z 135 & m/z 153 for the measurement of hexanoic acid in exhaled breath.....	106
Table 6.5: Characteristic analyte and m/z of acetic acid, acetone, isoprene and the C3-C10 aldehydes analysed using SIFT-MS	112
Table 6.6: Median concentrations and intra-individual CV(%) of C3 to C10 aldehydes in exhaled breath	117
Table 6.7: The running order for VOC analysis of exhaled breath using SIFT-MS	125
Table 6.8: Patient characteristics across study groups.....	130
Table 6.9: Clinical staging of the oesophageal and gastric adenocarcinoma groups.....	131
Table 6.10: Assessment of medication use across the study groups.....	132-134
Table 6.11: Analytical information of the 29 VOCs analysed within exhaled breath.....	136
Table 6.12: The statistically significant VOCs and abundant breath metabolites across the study groups.....	142
Table 6.13: Summary of ROC analysis, sensitivity and specificity of discriminating cancer groups from control groups.....	145
Table 6.14: Room air and Nalophan® bags measurements of each VOC.....	147
Table 6.15: Concentrations of fatty acids, phenols, aldehydes and abundant breath VOCs in community acquired pneumonia and urinary tract infection	157

ABBREVIATIONS

COPD - Chronic Obstructive Pulmonary Disease
DNA - deoxyribonucleic acid
GC-MS - Gas-Chromatography Mass Spectrometry
GI - gastro-intestinal
GNP - gold nanoparticles
HCC - hepatocellular carcinoma
H.pylori - Helicobacter pylori
MPM - malignant pleural mesothelioma
Nuclear Magnetic Resonance (NMR)
NSCLC - non-small cell lung cancer
OG Cancer - Oesophago-Gastric Cancer
OGD - Oesophago-Gastro-Duodenoscopy
ppbv - parts-per-billion by volume
PTR-MS - Proton Transfer Reaction Mass Spectrometry
RNA - ribonucleic acid
SIFT-MS - Selected Ion Flow Tube Mass Spectrometry
SPME – Solid-phase microextraction
VOC - Volatile Organic Compound

Chapter 1

OESOPHAGEAL AND GASTRIC CANCER

1.1 Introduction

Cancer remains the leading cause of death worldwide: it accounted for 7.6 million deaths in 2008¹. Despite an increasing incidence, mortality from cancer has been decreasing over the last four decades². Between 1979 and 2008, the European age-standardised mortality rates for all malignant neoplasms fell by 20% from 219 to 176 per 100,000 population³. Screening programmes and disease awareness coupled with treatment advances have contributed to this decrease in cancer mortality rates. However, it is estimated that a further 30-40% of these deaths can be prevented¹. Early disease detection remains one of the most important aspects in further improving cancer-related outcomes and survival rates. The majority of screening processes involve invasive procedures to appropriately diagnose those with cancer and suitably confirm the true negative cohort. Although these screening procedures are invariably necessary, there always remains a small but non-negligible morbidity risk associated with invasive procedures. Moreover, one of the main factors in deciding whether to institute a screening programme (population-based or selective) for a specific cancer is dependent on the prevalence of that cancer within the target population⁴; although this is justifiable in terms of healthcare costs and resource allocation, this results in the less common cancers being detected in their earliest stages and hence these cancers tend to have poor outcomes.

1.2 Oesophageal and Gastric Cancer Trends

1.2.1 Incidence

Oesophageal cancer is the currently the sixth leading cause of cancer death worldwide; in 2008, there were 482,000 new cases of oesophageal cancer diagnosed and 407,000 oesophageal cancer deaths recorded globally⁵. There is also significant geographical variation in the prevalence of the different histological subtypes of oesophageal cancer, with squamous cell carcinoma far more common in south-central Asia and south-eastern Africa^{5,6}. There has been a steady decrease in the oesophageal squamous cell carcinoma incidence in the West, with one of the main factors being the associated decline in smoking rates. However, the incidence of adenocarcinoma of the gastro-oesophageal junction has significantly increased in Europe and the USA in the past thirty years^{5,7}; the evidence shows that both reflux oesophagitis and obesity are primary risk factors contributing this increase^{7,8,9,10,11,12}. A significant decline in the incidence of gastric cancer has been observed over the past seventy years, which has been attributed to the control of *H.pylori* infection and dietary changes¹³. Nonetheless, it remains the

second leading cause of cancer-related death worldwide^{5,14}. In 2008, there were 989,000 new cases of gastric cancer diagnosed and 738,000 gastric cancer deaths recorded globally⁵. The highest incidence of gastric cancer is seen in East Asia, Eastern Europe and South America^{15,16,17}. Although the survival rates are slightly better in comparison to oesophageal cancer (with age standardised 5 year relative survival of 27.7% in the USA), the incidence-mortality ratio is higher than for the more common cancers such as breast, prostate and colorectal^{5,15}.

1.2.2 Survival

Oesophageal and gastric malignancies account for 15% of cancer-related deaths globally¹⁵. The poor prognosis of these cancers is all too evident with age-standardised 5-year relative survival rates, from 2005-2009, in the UK of 13% and 17.6% for oesophageal and gastric cancer, respectively¹⁸. The statistics are similar in North America, where the five-year survival rates, from 2002-2008, were 20% and 3% for regional oesophageal cancer and distant oesophageal cancer, respectively¹⁹.

1.2.3 Diagnosis and National Guidelines

Oesophago-gastric (OG) cancers often present to clinicians with a spectra of non-specific symptoms/signs. These can include nausea, vomiting, abdominal pain, dysphagia, odynophagia and/or weight loss amongst others. The variable nature of the early signs of OG cancer and lack of ‘alarm’ symptoms until the disease is at a more advanced stage contribute to the aforementioned poor survival statistics. However, the guidance from the National Institute of Clinical Excellence (NICE) for suspected OG Cancer focusses on late symptoms in its referral criteria for investigation²⁰. The guidelines state that ‘An urgent referral for endoscopy or to a specialist with expertise in upper gastrointestinal cancer should be made for patients of any age with dyspepsia who present with any of the following’:

- chronic gastrointestinal bleeding
- dysphagia
- progressive unintentional weight loss
- persistent vomiting
- iron deficiency anaemia

- epigastric mass
- suspicious barium meal result.

The guidance also states that for patients less than 55 years old, endoscopic investigation of dyspepsia is not necessary in the absence of alarm symptoms. The Health department of the Scottish executive has also published guidelines for urgent referral in suspected cases of upper GI cancer²¹; their recommendations are similar to NICE and include the following:

- Dysphagia – food sticking on swallowing (any age)
- Dyspepsia at any age combined with one or more of the following ‘alarm’ symptoms:
 - weight loss
 - proven anaemia
 - vomiting
- Dyspepsia in a patient aged 55 years or more with at least one of the following ‘high risk’ features:
 - onset of dyspepsia less than one year ago
 - continuous symptoms since onset
- Dyspepsia combined with at least one of the following known risk factors:
 - family history of Upper GI cancer in more than 2 first degree relatives
 - family history of colorectal cancer (familial adenomatous polyposis, hereditary non-polyposis colorectal cancer)
 - Barrett’s oesophagus
 - pernicious anaemia
 - peptic ulcer surgery over 20 years ago
 - known dysplasia, atrophic gastritis, intestinal metaplasia
- Jaundice
- Upper abdominal mass
- Back pain and weight loss

Due to the expansile nature of the oesophageal and gastric muscular layers, there is often a masking effect of a growing tumour. When patients do present to a doctor with alarm symptoms, the tumour has often invaded deep into the muscularis propria and spread to the surrounding lymph nodes or distant sites; this results in only one-third of patients diagnosed

with OG cancer being considered suitable for treatment with curative intent²². However, diagnosis of these cancers at a treatable stage is associated with a significant survival benefit. Treatment of early oesophageal adenocarcinoma, either endoscopically or surgically, has a reported overall 5-year survival rate of 83-96%²³. There are currently no established screening programmes for oesophageal or gastric cancer in Western countries. Only Japan and South Korea have population-based screening programmes targeting gastric cancer^{24,25}. However, in Western countries, these endoscopy-based screening approaches are not justified given the low incidence of significant pathology and associated healthcare costs (estimated at ~£55,000 for detecting each upper gastro-intestinal malignancy)²⁶.

Alarm symptoms such as dysphagia and weight loss have also been shown to have low predictive value for underlying malignancy, despite forming a major part of the 2-week wait primary care referral guidance for suspected upper gastro-intestinal cancer²⁷. A recent study on inequality in oesophago-gastric cancer outcomes reported that OG cancer patients belonging to general practices with low rates of gastroscopy referral are at increased risk of poor outcomes²⁸. However, a retrospective analysis of 2000 patients undergoing endoscopy for upper gastrointestinal symptoms demonstrated that no abnormal findings were identified in 53% of patients and only 1% of patients undergoing endoscopy were diagnosed with cancer²⁹. This creates a diagnostic dilemma for general practitioners; they have no objective methods other than endoscopy for identifying those at highest risk of an upper gastrointestinal malignancy whilst they cannot refer every patient with non-specific upper gastrointestinal symptoms for an Oesophago-Gastro-Duodenoscopy (OGD). Thus, there remains an important clinical need to develop novel diagnostic techniques for early disease detection in both oesophageal and gastric cancer.

1.2.4 Diagnostic Biomarkers in Oesophageal and Gastric Cancer

Research investigating potential non-endoscopic diagnostic biomarkers in oesophageal and gastric cancer has primarily concentrated on molecular targets. In a study on cell-free DNA (cfDNA), increased levels of cfDNA95 were observed in patients with oesophageal adenocarcinoma compared with healthy controls³⁰. Mutations in chromatin-remodelling genes including *ARID1A* and *SMARCA4* have been reported in 25% of patients with oesophageal adenocarcinoma^{31,32}; dysfunction of chromatin regulatory mechanisms appear to be crucial in oesophageal adenocarcinoma tumourigenesis^{33,34,35}. Jin *et al.* observed AKAP12 promoter

hypermethylation in 53% of patients with oesophageal adenocarcinoma and dysplastic Barrett's oesophagus however the same mutation was absent in patients with a normal oesophagus³⁶. Reprimo and TAC1 promoter hypermethylation have also been shown to be increased in patients with oesophageal adenocarcinoma compared to control subjects^{37,38}. However, these studies have also confirmed the marked heterogeneity that exists between oesophageal adenocarcinomas; *TP53* is the only identified tumour suppressor gene to be mutated in more than half of oesophageal cancers³¹.

Serum biomarker panels including a combination of Epidermal Growth Factor Receptor (EGFR), transthyretin (TTR), regulated upon activation, normally T-expressed and presumably secreted (RANTES) and vitronectin (VN) demonstrated a diagnostic accuracy of 88% for gastric adenocarcinoma³⁹. RUNX3 methylation was observed in the serum of 65% of patients with gastric carcinoma compared with 8.1% of patients with chronic gastritis⁴⁰; RUNX3 methylation was also higher in gastric cancer tissue than in healthy gastric mucosa⁴¹.

RASSF1A methylation and CDKN2A methylation has been observed in 34% and 44% of patients with gastric adenocarcinoma whilst being absent in healthy controls^{42,43}. Several non-invasive technologies have also been trialled in the diagnosis of oesophageal and gastric cancer. A non-invasive occult blood detector was tested in 200,000 asymptomatic individuals however the sensitivity for oesophageal and gastric cancer was only 3.4%⁴⁴. Cytological techniques utilising balloon and sponge capture devices have also been developed for oesophageal cancer⁴⁵. However, these cytological methods have demonstrated poor sensitivity for diagnosing both dysplasia and cancer (24-47%)⁴⁵. Hence, alternative diagnostic strategies for oesophageal and gastric cancer need to be investigated.

Chapter 2

VOLATILE ORGANIC COMPOUNDS AND CANCER

2.1 Introduction

Halitosis and systemic disease states have been associated since the days of Hippocrates. Physical breath analysis during examination of a patient has often provided the clinician with important clues in reaching the correct diagnosis. Foetor oris associated with appendicitis, foetor hepaticus linked to chronic liver failure and the fruity odour associated with diabetic ketoacidosis have all been described over the years. However, the roots of modern breath gas analysis can be traced back to the Nobel Prize winner Linus Pauling. In 1971, he demonstrated by gas liquid partition chromatography that human breath is a complex mixture containing over 250 volatile organic compounds (VOCs)⁴⁶; he also demonstrated that there were over 280 VOCs within urine vapour⁴⁶. The term volatile organic compound (VOC) encompasses organic chemicals that have a high vapour pressure at room temperature. VOCs have routinely been measured in the assessment of environmental gases and within the flavour & fragrance industry for many years; Gas Chromatography-Mass Spectrometry, GC-MS, is considered a 'gold-standard' method for forensic substance identification and has been extensively employed for this purpose. However, many of the biochemical pathways relating to the origin of these VOCs within humans are as yet unknown; it is postulated that these endogenous VOCs potentially hold the key to identification of new biomarkers for underlying disease processes. Since the original discovery of Pauling, research in this field has focussed on identifying VOCs associated with certain disease states; the ultimate translational application for this research being the quantification of VOCs as a non-invasive diagnostic and therapeutic monitoring tool in medicine. In 1983, Manolis published a seminal review on the diagnostic potential of breath analysis, including identifying disease-specific VOCs⁴⁷. Although the potential of such non-invasive techniques is readily acknowledged, there remain limited applications of human breath analysis. The most well-known example of exhaled breath analysis remains the quantification of blood alcohol levels; this test is almost exclusively employed by law enforcement agencies for the purposes of roadside breathalyser testing. In the field of medicine, the ¹³C urea breath testing for *H. pylori* bacterium, hydrogen breath test for small bowel bacterial overgrowth and the measurement of exhaled nitric oxide as a biomarker of airways inflammation in asthma represent the few routinely employed clinical breath tests^{48,49,50}. Hence, there remains much opportunity to develop VOCs as non-invasive diagnostic tools.

The detection of VOCs in cancer has been investigated using both innovative and conventional methods. Sonoda *et al.* made global headlines after they trained a Labrador retriever dog in scent detection of colorectal cancer⁵¹. After completion of 4 years of training using a reward-

based approach, the sniffer dog was then expected to identify those with colorectal cancer from breath and watery stool samples. On each occasion, a cancer sample was placed in a room with 4 control samples⁵¹. Detection of colorectal cancer in exhaled breath using this method had a sensitivity of 91% and a specificity of 99%⁵¹. The sensitivity of canine scent detection for watery stools was 97% with a specificity of 99%⁵¹. The study demonstrated that canine olfaction of VOCs could discriminate patients with colorectal cancer from controls with good diagnostic accuracy. Ehmann *et al.* wanted to investigate the robustness of a ‘so far unknown’ VOC fingerprint within the exhaled breath of lung cancer patients⁵². To test this hypothesis, they trained four dogs (two German shepherd dogs, one Australian shepherd dog and one Labrador retriever) of both sexes (two females, two males) aged 2.5–3 yrs in scent detection for lung cancer⁵². They collected breath samples from 220 participants; lung cancer was identified by the dogs with a sensitivity of 71% and a specificity of 93%⁵². Lung cancer detection was independent of Chronic Obstructive Pulmonary Disease (COPD) status, smoking and food odours⁵². Although the results of both studies are clearly impressive, it is impractical to implement canine olfaction for clinical screening purposes; the cost and length of time required in dog training as well as the practicalities of maintaining an appropriate dog facility do not make it justifiable. More importantly, both studies further support the theory that cancer-specific VOCs exist, which could potentially be identified through chemical analytical techniques.

However, the identification, characterisation and accurate quantification of VOCs associated with certain diseases continue to pose a major scientific challenge. In recent years, there have been technological advances in gas phase analytical techniques, allowing one to measure VOCs emitted from exhaled breath, headspace of bodily fluids and histological specimens with accuracy at levels of parts-per-billion by volume (*ppbv*), and even down to parts-per-trillion by volume (*pptv*)⁵³. The mass spectrometry tools available have increased the analytical chemist’s capabilities in detecting the smallest differences between biological samples and several encouraging VOC-disease associations have been described. However, the use of the technology remains confined to the laboratory and robust scientific protocols are essential as both dilution and contamination of gaseous samples remain a source of potential error. There has been no methodology standardization thus far and it has been postulated that sampling errors may be one of the major factors contributing to variation in VOC concentrations between studies⁵⁴. Since Pauling’s original study on VOCs in exhaled breath and urine vapour, there has been much research on VOCs and disease; the literature pertaining to cancer has been assessed in the form of a systematic review.

2.2 Objective

The aim of this systematic review was to identify volatile organic compounds (VOCs) linked to cancer disease states.

2.3 Methods

2.3.1 Search Strategy

Methodology for this review was adapted from that for compiling systematic reviews of clinical studies of healthcare interventions⁵⁵. Bibliographic searches using PubMed (1951–2011), EMBASE (1947–2011), Web of Science (1900–2011) and Google Scholar (1980–2011) electronic databases were conducted up to and including 19th November 2011 for studies on Volatile Organic Compounds and Cancer. Search strategy incorporated 2 different domains of MeSH-terms and key words combined by “AND”. The first domain contained terms on volatile organic compounds/volatile compounds and the second had terms on cancer/malignancy/neoplasm. Selection was not restricted by language. Two reviewers (S Kumar and J Cushnir) independently screened titles and abstracts of studies identified through the searches. For those fulfilling inclusion criteria, full texts of potentially relevant articles were retrieved and assessed for eligibility in final review. Further potentially relevant articles were identified through the searching of reference lists of included studies. Inclusion criteria were studies employing mass spectrometry or VOC sensor technology and investigating human biological samples or cell lines only. Exclusion criteria encompassed all reviews, method papers, letters, comments, perspectives and papers without original experimental data. Studies that reported the same patient population were also excluded, except for the most recent publication. Both reviewers independently extracted data from the selected studies including primary author; year of publication; number and types of specimen; mass spectrometry platform; and identified cancer-related VOCs. The primary outcome measure was the identification of VOCs observed at varying concentrations within cancer samples.

2.3.2 Study Quality Assessment

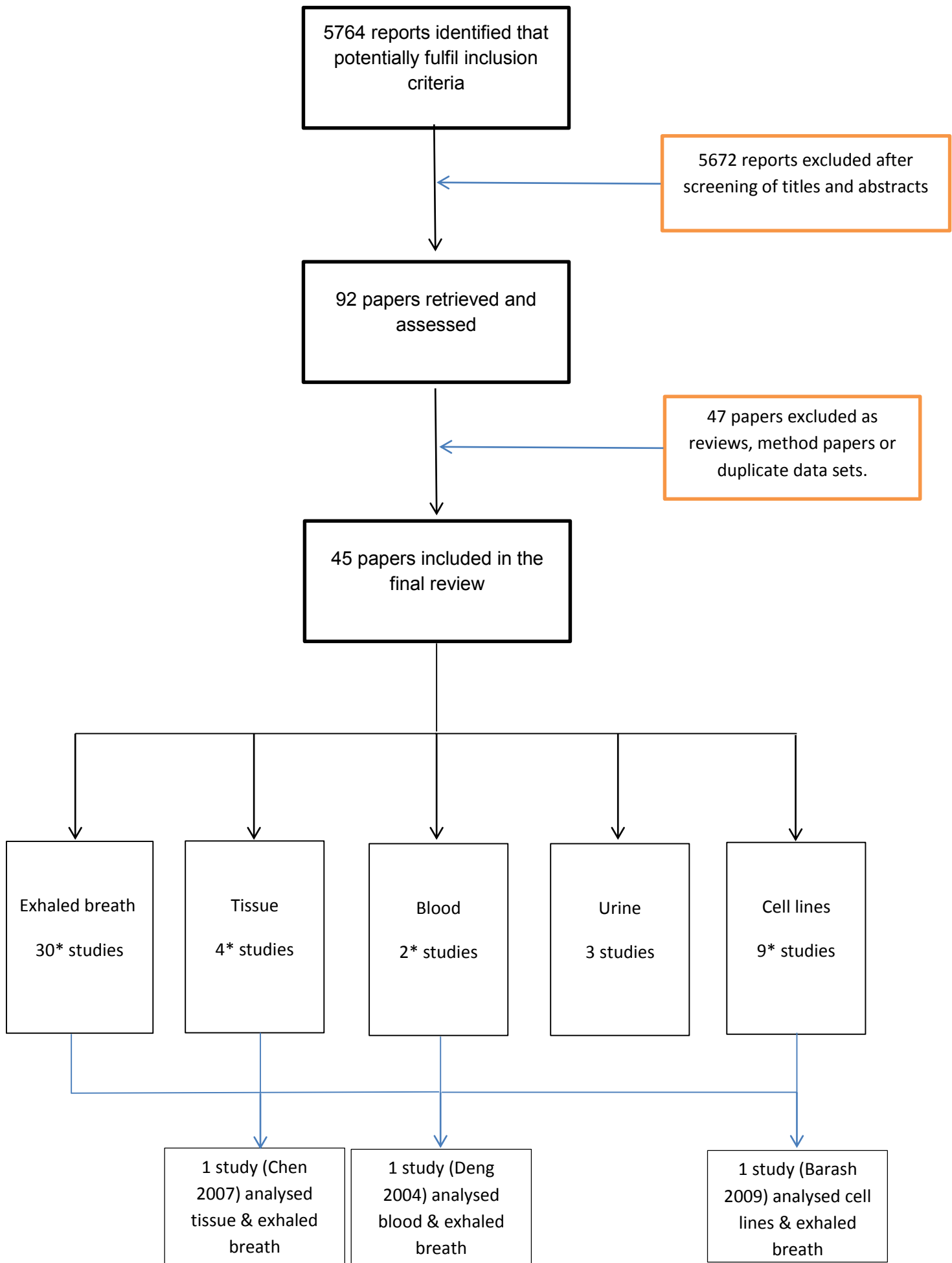
The QUADAS-2 tool was used to assess the quality of the studies which involved recruitment of patients or analysed patient samples⁵⁶; this quality evaluation tool consists of 4 domains including patient selection, index test, reference standard, and flow of patients through the study⁵⁶. The reference standard is defined as the best available method to establish the presence or absence of the target condition⁵⁷. In this review, the reference standard was histological confirmation of the cancer diagnosis. Each domain is assessed in terms of the risk of bias and the first three are also assessed in terms of concerns regarding applicability.

2.4 Results

2.4.1 Search Strategy Results and Selected Studies

In total, 5764 potential articles were identified from the literature search (Figure 2.1 - Google Scholar 2907, PubMed 1171, Web of Science 1169, EMBASE 517). Ninety-two studies potentially fulfilled the inclusion criteria. A total of 45 studies (Exhaled breath - 30, Tissue - 4, Blood - 2, Urine - 3, Cell lines - 9) were included. The tables in Appendix A include the primary author, year of publication, number and types of specimen and mass spectrometry platform employed for all the relevant studies included in the review.

Figure 2.1: Systematic Review Search Strategy

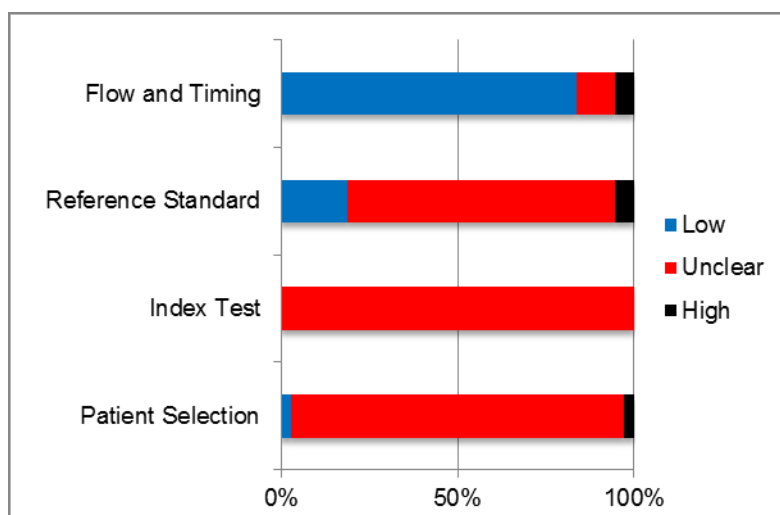


2.4.2 Quality assessment of Studies

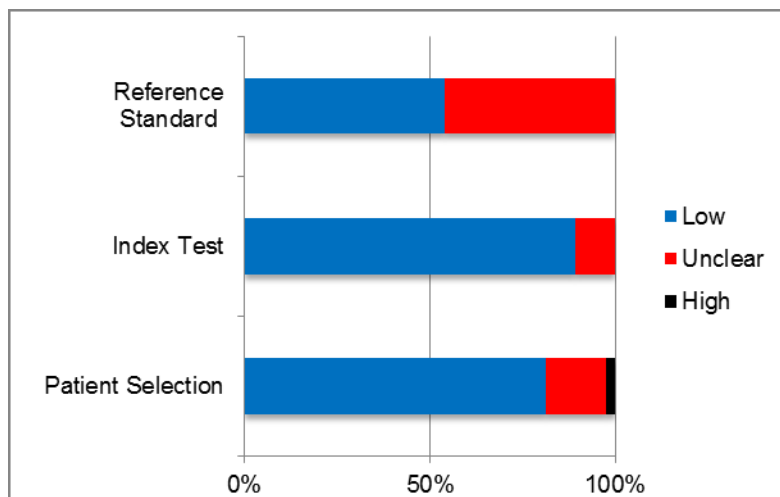
The QUADAS-2 study quality assessment results are shown in Figures 2.2A & 2.2B. There was an unclear risk of bias for patient selection in the majority of studies due to the lack of clarity regarding patient recruitment. In particular, it was unclear whether consecutive patient data had been included. The risk of bias for the index test and reference standard were unclear in most studies, due to a lack of information regarding blinding of assessors. Forty-five percent of the studies had an unclear risk of applicability for the reference standard, mostly due to the limited information regarding histological assessment of control samples.

Figure 2.2: QUADAS-2 Assessments of Study Quality (A) Bar chart demonstrating an assessment of risk of bias and (B) assessment of applicability concerns

(A)



(B)



2.5 Discussion

The results of this review demonstrated that Gas Chromatography Mass-Spectrometry (GC-MS) is the most frequently employed mass spectrometry platform. GC-MS coupled with sample collection and pre-concentration techniques, in particular solid-phase microextraction (SPME), have further enhanced the isolation and identification of VOCs from mixed samples. Pawliszyn *et al.* originally developed the SPME technique, which is simple to use, relatively inexpensive and negates the need for solvents⁵⁸. It allows the collected sample to undergo pre-concentration prior to GC-MS analysis. SPME uses fibres that are coated with different materials to allow for optimal adsorption of compounds of interest. Deng *et al.* investigated 5 different fibre coatings to optimise recovery of VOCs from the headspace of blood in patients with lung cancer⁵⁹. This study highlighted the significant method development necessary when employing the GC-MS technique for analysis of VOCs emitted by biological samples.

2.5.1 Exhaled Breath VOCs and Cancer

One of the earliest studies on exhaled breath VOCs by Phillips *et al.* investigated 108 patients with abnormal radiographs (of which 60 patients had histologically confirmed cancer)⁶⁰. They trapped a range of volatile organic compounds from samples from patients with and without lung cancer on a sorbent trap. These VOCs were later desorbed, concentrated by 2-stage cryo-focusing and analysed using GC-MS. Twenty two breath VOCs, including alkanes, alkane derivatives and benzene derivatives were used to distinguish between patients with and without lung cancer (100% sensitivity, 81.3% specificity for stage 1 lung cancer)⁶⁰. Poli *et al.* employed SPME/GC-MS to analyse the VOCs in the breath of patients with non-small cell lung cancer (NSCLC), patients with COPD, smokers and healthy controls⁶¹. The VOC profile in the NSCLC group was also analysed 3 weeks after these patients had undergone pulmonary surgery. None of the VOCs alone allowed for discrimination between the study groups. However, it was observed that both isoprene (2-methyl-1,3-butadiene) and decane levels significantly decreased after surgery⁶¹. More recently, Poli *et al.* used an on-fiber-derivatisation SPME sampling technique with GC-MS to investigate C₃-C₉ aldehydes in the exhaled breath of 40 NSCLC lung cancer patients and 38 healthy controls⁶². In this study, they employed a PDMS/DVB fibre (Supelco, Bellefonte, BA, USA) specifically for extraction of the aldehydes; they excluded formaldehyde and acetaldehyde from analysis due to the presence of these compounds in both indoor and outdoor environments as well as being combustion products of

tobacco smoke. The concentrations of all the C₃-C₉ aldehydes were significantly increased in the patients with NSCLC in comparison with controls⁶². Their results also demonstrated that the selected aldehydes had good discriminatory power in being able to distinguish between asymptomatic non-smokers and patients with NSCLC (92.1% of controls vs. 90% of NSCLC were correctly classified).

Fuchs *et al.* investigated volatile aldehydes in the exhaled breath of 12 lung cancer patients (8 NSCLC, 4 small cell lung cancer), 12 smokers and 12 healthy volunteers⁶³. In this study, the reactive aldehydes were derivatised to their corresponding stable oximes using on-fibre derivatisation (SPME-OFD) and subsequently analysed with GC-MS⁶³. Exhaled pentanal, hexanal, octanal and nonanal concentrations were found to be significantly higher in lung cancer patients than in smokers and healthy controls. However, there were no statistically significant differences in the concentrations of acetaldehyde, propanal, butanal, heptanal and decanal between the three groups. Fuchs *et al.* also reported that the sensitivity and specificity of their identified exhaled breath VOCs for cancer discrimination were comparable to conventional serum markers and CT imaging; however it is noted that their study was pilot in nature and had a relatively small sample size⁶³. Song *et al.* employed SPME/GC-MS to assess mixed expiratory samples from 43 patients with NSCLC and 41 healthy volunteers⁶⁴. Their study reported that the peak areas (indication of concentration) of 1-butanol (sensitivity of 95.3% and a specificity of 85.4%) and 3-hydroxy-2-butanone (a sensitivity of 93.0% and a specificity of 92.7%) were significantly higher in the exhaled breath of patients with NSCLC than those in the control subjects⁶⁴.

De Gennaro *et al.* designed a cross-sectional, case control study in which they recruited 39 patients, divided equally into 3 groups⁶⁵. They analysed exhaled breath VOCs from patients with histologically confirmed malignant pleural mesothelioma (MPM), those with similar occupational asbestos exposure and healthy controls. The exhaled breath VOCs were adsorbed on sorbent cartridges, then thermally desorbed and analysed by GC-MS. Through univariate statistical analysis, it was found that cyclohexane was the only compound that could distinguish between MPM and the other 2 groups⁶⁵. Cyclopentane was reported as the dominant compound to discriminate between the occupational asbestos exposure group from the mesothelioma patients and healthy controls; it was proposed by De Gennaro *et al.* that cyclopentane could potentially be a good indicator of long term asbestos exposure.

The studies in this review indicate that it is highly unlikely that a single VOC biomarker would be able to distinguish cancer patients from controls. This has led to several authors exploring

the concept of VOC profiling and discriminative analysis of the resultant VOC patterns from biological samples. Phillips *et al.* collected exhaled breath samples from 51 women with abnormal mammograms & biopsy-proven breast cancer as well as 42 age-matched controls⁶⁶. After analysis with GC-MS, the results were analysed using fuzzy logic software. The women were randomly assigned to either the training or prediction set in a 2:1 ratio. Five breath biomarkers of breast cancer were identified – 2-propanol, 2,3 dihydro-1-phenyl-4(1H)-quinazolinone, 1-phenyl-ethanone, heptanal and isopropyl myristate⁶⁶. The fuzzy logic model incorporating the aforementioned breath biomarkers predicted breast cancer with a 93.8% sensitivity and 84.6% specificity⁶⁶. Phillips *et al.* also applied a similar predictive model to a cohort of lung cancer patients⁶⁷; 66% of patients were randomly assigned to the training set and the remaining 33% of patients were put in the prediction set. Thirty five candidate VOCs were identified in the training set and 16 of these VOCs were selected for inclusion for fuzzy logic analysis. In this study, the model predicted primary lung cancer with 84.6% sensitivity and 80.0% specificity⁶⁷. Qin *et al.* investigated exhaled breath VOCs in hepatocellular carcinoma (HCC)⁶⁸. In this study, they recruited 30 patients with histologically confirmed HCC, 27 patients with hepatitis B-induced cirrhosis and 36 healthy volunteers⁶⁸. They employed a SPME/GC-MS method to analyse the collected breath samples. 3-hydroxy-2-butanone was only compound found to be statistically significant between all three groups. 3-Hydroxy-2-butanone had a sensitivity of 83.3% and specificity of 91.7% to discriminate healthy volunteers and HCC patients; there was a sensitivity of 70.0% and a specificity of 70.4% between hepatitis B-induced cirrhosis group and patients with HCC⁶⁸.

2.5.2 Urine VOCs and Cancer

Guadagni *et al.* employed SPME/GC-MS to investigate hexanal and heptanal urinary excretion in 10 patients with lung cancer and 25 healthy subjects⁶⁹. Hexanal levels were slightly higher in the cancer cohort; there were no differences observed in urinary heptanal concentrations between the 2 groups⁶⁹. Silva *et al.* investigated urinary volatile metabolites as potential cancer biomarkers using urine samples from 33 cancer patients (14 leukaemia, 12 colorectal and 7 lymphoma) and 21 healthy controls⁷⁰. Through analysis with SPME/GC-MS, benzene derivatives, terpenoids and phenols were the most common compounds observed in the cancer patients; ketones and sulphur compounds were the main classes that were isolated from the urine headspace of the healthy controls. Significant increases were observed in the peak area of 2-methyl-3-phenyl-2-propenal, p-cymene, anisole, 4-methyl-phenol and 1,2-dihydro-1,1,6-

trimethyl-naphthalene in the cancer cohort, whereas statistically significantly lower abundances of dimethyl disulphide were noted⁷⁰.

2.5.3 Sensor Technologies

Electronic nose technology using an array of VOCs has also been applied in the prediction of cancer. The majority of electronic nose studies employed an array of composite nano-sensors; the data generated from the electronic nose is analysed with learning algorithms such as artificial neural networks which characterise 'patterns of VOCs' rather than allow for formal identification of individual VOCs. Dragonieri *et al.* investigated electronic nose technology to distinguish VOC-patterns in the exhaled breath of lung cancer patients from those with COPD and healthy controls⁷¹. Canonical discrimination analysis with the optimal number of vectors showed the cross-validation results were 85% accurate for distinguishing lung cancer vs. COPD and 90% accurate for distinguishing lung cancer vs. healthy controls⁷¹. A potential limitation in this study was the small sample size (n=10 for each group); there was also no discrimination possible between different stages of lung cancer. Peng *et al.* used a combination of gold nanoparticles (GNP) sensor arrays and GC-MS to analyse exhaled breath samples from 96 subjects with a variety of cancers (lung, colon, breast and prostate) and 81 healthy volunteers⁷². The nanosensors were able to suitably discriminate between patients with lung, colon and breast cancer and healthy controls⁷². There was a slight overlap between the data for the prostate cancer and healthy controls; it was hypothesised that this overlap was due to a higher proportion of the prostate cancer patients having stage 1 disease compared to the other cancer patients. By combining all the cancers together, it was possible to distinguish between all four cancer groups and healthy controls, with minimal overlap between the prostate cancer and healthy controls⁷². The same breath samples were also analysed using SPME/GC-MS. A specific SPME fibre cannot extract all VOCs in a sample and hence it was not possible to identify all the cancer-specific VOCs to which the nanosensor array responded to. The results demonstrated that it was possible to achieve reasonable discrimination between healthy controls and patients with a specific type of cancer. Mazzone *et al.* analysed exhaled breath from 143 individuals (49 with non-small cell lung cancer, 18 with chronic obstructive pulmonary disease 15 with idiopathic pulmonary fibrosis, 20 with pulmonary hypertension, 20 with sarcoidosis and 21 controls) using a colorimetric sensor array⁷³. The colorimetric sensor array contained 36 spots, each with different sensitivities to a range of VOCs, which changed colour depending on the analytes they come into contact with⁷³. In this study, they reported colorimetric spots to be generally responsive to different VOCs but not specific. Applying data

from the colometric spots, a predictive model for lung cancer was created which demonstrated a sensitivity of 73.3% and specificity of 72.4% for lung cancer⁷³.

2.5.4 Blood VOCs and Cancer

Deng *et al.* investigated exhaled breath and blood samples from 10 patients with NSCLC and 10 healthy controls⁵⁹. They reported a range of VOCs including aromatic and aliphatic alcohols, ketones and aldehydes from both bio-samples. Specifically, hexanal and heptanal were found to be at significantly higher concentrations in the blood of lung cancer patients compared to healthy controls. These 2 compounds were also detected in the exhaled breath of lung cancer patients but not in the controls, these results being consistent with a previous study on VOCs in lung cancer⁶⁰. Xue *et al.* employed SPME/GC-MS to investigate the VOCs within the blood of liver cancer patients⁷⁴. Comparative analysis of the liver cancer blood group (n=19) and the controls (n=18) revealed 19 VOCs to be significantly different between the 2 groups⁷⁴. Of the compounds that were statistically significant, 3 peaks were observed at increased concentrations in liver cancer blood – hexanal (sensitivity 94.7%, specificity 100%), 1-octen-3-ol (sensitivity 84.2%, specificity 100%) and octane (sensitivity 89.5%, specificity 100%)⁷⁴.

2.5.5 VOCs and Cancer Cell lines

The majority of the studies on cell lines have primarily investigated lung cancer. Filipiak *et al.* conducted VOC analysis on the non-small cell lung cancer (NSCLC) cell line CALU-1 (which expresses the mutant K-ras gene)⁷⁵. The cultivated cell lines were incubated and 200mL of the headspace gas was retrieved for analysis by thermal desorption GC-MS. The results demonstrated 2,3,3-trimethylpentane, 2,3,5-trimethylhexane, 2,4-dimethylheptane and 4-methyloctane were released from the headspace of the cell line CALU-1⁷⁵. Sponring *et al.* investigated the headspace of NCI-H2087 lung cancer cell line using GC-MS⁷⁶; there was a significant increase in 2-ethyl-1-hexanol and 2-methylpentane in the headspace of NCI-H2087 lung cancer cells and statistically significant lower concentrations of acetaldehyde were present when compared to medium only⁷⁶. Filipiak *et al.* employed the same analytical methodology to investigate the VOCs released by the lung cancer cell line A549 and used primary human bronchial epithelial cells (HBEpC) & human fibroblasts (hFB) as controls⁷⁷. The compounds 2-pentanone and 2,4-dimethyl-1-heptene were significantly elevated within the headspace of the A549 cell line⁷⁷. However, several aldehydes (including 2-ethylacrolein and 2-methyl-2-butenal) were found exclusively in the headspace of the medium controls and not within the

headspace of the cell lines; it was postulated that the cancer cell lines may be consuming these aldehydes⁷⁷. Sponring *et al.* conducted one further study of volatile gas analysis in the headspace of NCI-H1666 lung cancer cells, but in contrast to the aforementioned studies, no unequivocal release of VOCs was observed⁷⁸.

2.5.6 SIFT-MS and PTR-MS

GC-MS has been the most frequently employed analytical method in the investigation of VOCs to date. However, there are several important technical considerations for GC-MS analysis of biological samples; it remains an offline technique requiring appropriate GC-column selection and pre-analysis sample preparation. Analysis processing times are lengthy and VOCs with low molecular weight (some of which have been proposed as potential biomarkers) can be difficult to accurately quantify with GC-MS. The advent of newer ionization techniques such as atmospheric pressure chemical ionization (APCI), proton transfer reaction mass spectrometry (PTR-MS), and selected ion flow tube mass spectrometry (SIFT-MS) have allowed one to perform real-time, online sample analysis^{79,80,81}. In comparison to GC-MS techniques, SIFT-MS has the advantage of real-time VOC quantification without the need for any pre-concentration steps. SIFT-MS has been employed to detect and quantify VOCs in several disease states within the headspace of cancer cell lines and biofluids as well as exhaled breath. Smith *et al.* employed SIFT-MS to investigate acetaldehyde in the non-small cell lung cancer cell lines - SK-MES and CALU-1⁸². Acetaldehyde was produced by both lung cancer cell lines and its concentration was found to be proportional to the number of cancer cells in the medium⁸². Španěl *et al.* investigated formaldehyde within the headspace of urine using SIFT-MS⁸³. Fourteen patients with bladder cancer, 24 patients with prostate cancer and 14 healthy volunteers were recruited into this seminal study. Formaldehyde was clearly increased in the headspace of the urine from the cancer patients compared with the healthy controls; it was proposed that the analysis of the headspace of urine could be a valuable non-invasive diagnostic tool for the future⁸³.

Proton transfer reaction mass spectrometry (PTR-MS) is another online analytical method which has been employed in the analysis of VOCs. Wehinger *et al.* assessed exhaled breath VOCs in 17 patients with primary lung cancer and 170 healthy individuals⁸⁴. The results of this study indicated that VOC-31 a product ion of mass-to-charge ratio, m/z , of 31 (tentatively protonated formaldehyde) and VOC-43 a product ion of m/z 43 (tentatively a fragment of protonated iso-propanol) were present at significantly higher concentrations in the exhaled

breath of lung cancer patients compared to controls⁸⁴. A disadvantage of PTR-MS is that one can only tentatively identify the compounds; the authors did propose that PTR-MS should be used in conjunction with GC-MS to confirm the identity of peaks of interest⁸⁴. Brunner *et al.* employed PTR-MS to investigate the VOCs in the headspace of cultured human cells (2 cancerous and 2 non-cancerous human cell lines)⁸⁵. The cell lines used were (i) lung epithelial tumour cells A-549 and retinal pigment epithelium cells hTERT-RPE1 and (ii) squamous lung carcinoma cells EPLC and immortalised human bronchial epithelial cells BEAS2B. A compound with a mass-to-charge ratio (m/z) of 45 was significantly consumed by the cancerous cell lines A-549 and EPLC but not by the controls⁸⁵. They tentatively assigned the VOC detected at $m/z=45$ to be acetaldehyde and indicated its potential as a cancer biomarker; however, the authors stated that the reasons for the variation in acetaldehyde metabolism remain unclear.

2.6 Conclusions

The results of this review demonstrate that VOC analysis for cancer diagnostics is a promising area for further development. However, in terms of the phases of screening biomarker development, VOC analysis for cancer remains in phase 1 (pre-clinical exploratory studies) and there remains much research to be done before VOC biomarkers can be utilised in routine clinical practice⁸⁶. GC-MS has been the most widely employed technique for the analysis of VOCs and it can be considered as a ‘gold standard’ analytical method, although there is much sample preparation necessary when analysing biological samples with this platform. SPME has been a critically important tool in VOC extraction from a variety of biological matrices. However, it is accepted that a SPME filter is unlikely to extract all the VOCs in a biological sample to the same degree as a real-time, online mass spectrometric technique could⁸⁷. The constitution of a biological sample is also altered with the use of necessary solvents and this may affect adsorption/desorption efficiencies of VOCs. This is an important limitation to consider for VOC analysis in cancer with GC-MS and there remains a need to develop real-time, online technologies such as PTR-MS & SIFT-MS.

2.6.1 Aldehydes

The majority of the studies performed global VOC analyses of their biological samples and it was evident that aldehydes were repeatedly linked with several cancer types. An aldehyde is defined as an organic compound containing a carbonyl centre (-CHO) bonded to a hydrogen and an R group. Aldehydes have multiple industrial and commercial applications, in particular formaldehyde (e.g. textile resins, automobile manufacturing) and acetaldehyde (e.g. chemical intermediate for acetic acid and ester production). Aldehydes are produced within the human body but they are not common in the natural building blocks such as amino acids, nucleic acids and lipids. The catalytic dehydrogenation of alcohols results in the production of aldehydes and ketones. When these aldehydes are produced within the human body, they are most commonly oxidised to carboxylic acids by aldehyde dehydrogenase enzymes. In recent years, there has been much epidemiological and laboratory research examining the potential mutagenic and carcinogenic properties of aldehydes. Zhang *et al.* performed a meta-analysis of epidemiological studies based on the hypothesis that high formaldehyde exposure can induce leukaemia⁸⁸. Their results suggested that formaldehyde exposure can increase the risk of leukaemia by 54%. Teng *et al.* studied the effects of formaldehyde in rat hepatocytes and found that there was a marked increase in formaldehyde cytotoxicity when either alcohol dehydrogenase or aldehyde dehydrogenase were inhibited⁸⁹. Wang *et al.* investigated the genotoxic mechanisms of formaldehyde and their group showed that formaldehyde induces both DNA damage via DNA adducts and DNA protein cross-links formation as well as through chromosomal aberrations in exposed rats⁹⁰. Rager *et al.* studied formaldehyde exposure in human lung cells and found that formaldehyde alters miRNA patterns that regulate gene expression⁹¹. The evolving evidence regarding formaldehyde has meant that as of June 10th 2011, the US National Toxicology Programme described formaldehyde as ‘known to be a human carcinogen’.

Significant evidence linking acetaldehyde and carcinogenesis has also been reported. Acetaldehyde is the first metabolite of ethanol oxidation and there is a known increased risk of liver and aerodigestive cancers in those with heavy alcohol consumption⁹². Several epidemiological studies have demonstrated that the risk of upper digestive tract cancers, specifically squamous cell oesophageal cancer, is markedly increased among ALDH2-deficient alcohol users compared to those with the normal enzyme^{93,94,95}. Fanelli *et al.* examined the effects of acetaldehyde accumulation in rat mammary tissue and concluded that this may be a

critical event in the development of breast cancer⁹⁶. The World Health Organization (WHO) has concluded that acetaldehyde derived from alcohol is a Group 1 carcinogen to humans⁹⁷. The mutagenic potential of several other aldehydes, including malondialdehyde, acrolein and glyceraldehyde, have also been implicated in oxidative stress/lipid peroxidation pathways⁹⁸. However, VOCs in oesophageal or gastric cancer have not been thus far investigated.

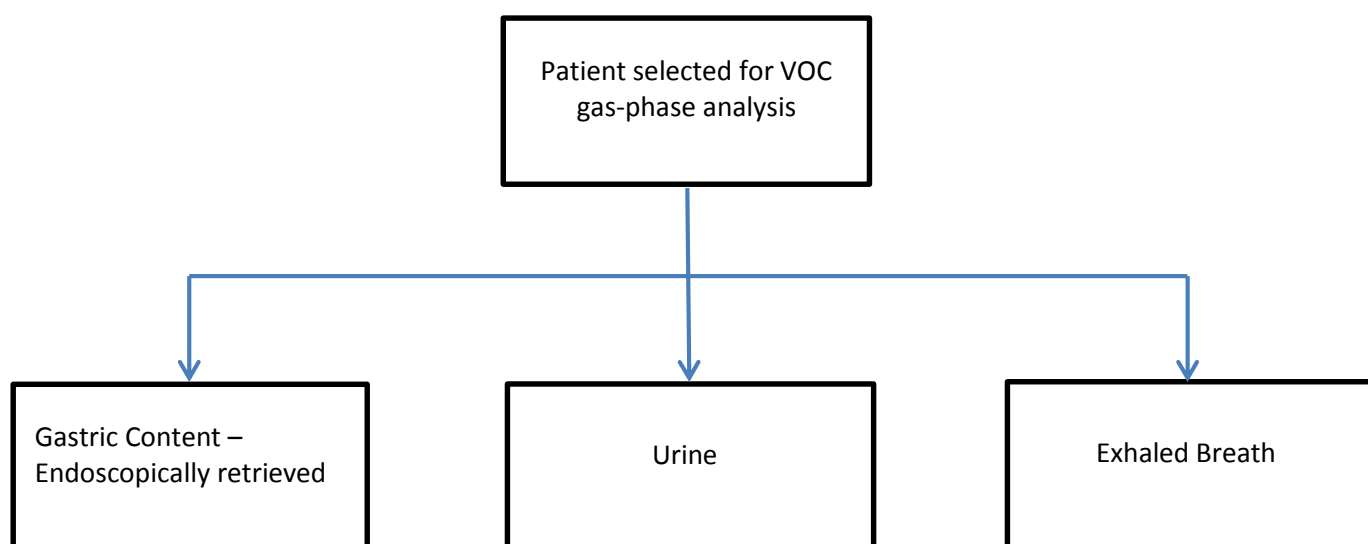
2.6.2 VOCs as potential diagnostic tools

Several published papers have suggested that the VOCs linked to a particular type of cancer could be used as biomarkers. However, any diagnostic test for cancer screening needs to have an acceptable level of sensitivity (i.e., true positive rate) and specificity (i.e., true negative rate). In the UK, the incidence of oesophageal and gastric cancer is 13.6 per 100,000 persons and 11.7 per 100,000 persons, respectively⁹⁹; this incidence is relatively low when compared to other cancers. For example, breast cancer and prostate cancer have an incidence of 156.8 per 100,000 women and 133.7 per 100,000 men, respectively⁹⁹. Hence, a screening tool for a disease with a low incidence must have excellent sensitivity and an acceptable specificity. The sensitivity needs to be sufficiently high for the test to pick up the persons with cancers. However, the specificity cannot be too low otherwise this can lead to healthy individuals unnecessarily being subjected to invasive investigations and undue stress. Many of the studies identified in the review provide sensitivity and specificity metrics for their reported ‘VOC biomarkers’; however, the majority have relatively small samples sizes, minimal evaluation of exogenous contributions and negligible assessment of confounding factors, all of which are important in the development of new diagnostic tests. There is also a need to develop standardized VOC sampling protocols and address the concerns regarding contamination and degradation of biological samples. Future VOC research should also consider the effects of confounding factors, employ rigorous protocols for recruitment of patients and controls and investigate sampling of multiple biological surrogates (e.g. urine and breath).

2.6.3 Thesis Aims

The primary aim of this thesis is to identify potential VOC biomarkers in oesophago-gastric cancer; the long-term aim of this research being the translational development of VOCs as non-invasive diagnostic tests. For this investigation, Selected-Ion Flow Tube Mass Spectrometry (SIFT-MS) has been employed in the analysis of VOCs from gastric content, urine and exhaled breath of patients with oesophageal and gastric cancer (Figure 3). Method experiments include the optimisation of VOC sampling techniques with evaluation of exogenous contamination and variability measures.

Figure 2.3: Bio-samples analysis strategy in oesophageal and gastric cancer



Chapter 3

Selected Ion Flow Tube Mass Spectrometry

3.1 The Selected Ion Flow Tube Mass Spectrometry Technique

Adams & Smith originally conceptualised and developed the Selected Ion Flow Tube (SIFT) in 1976 to study ion-molecule reactions at thermal energies^{100,101}; it has been fundamental to the advancement of interstellar ion chemistry and the understanding of reactions that occur in the ionized terrestrial atmosphere^{102,103,104}. The SIFT-MS analytical method is borne out of the SIFT and represents a direct mass spectrometric analysis by chemical ionization of the volatile compounds in the air/vapour samples⁵³. Precursor ions are formed in a microwave discharge source and are selected according to their mass-to-charge ratio (m/z) by an upstream quadrupole mass filter; they are injected into flowing helium carrier gas (at a pressure of ~ 100 Pa and a flow rate of typically 400 mL/min at standard temperature & pressure) through a Venturi-type inlet (diameter of 1 mm) where they are convected as a thermalized swarm along a flow tube (typically at a velocity of 100 m/s)⁵³. The selected precursor ions (H_3O^+ , NO^+ and O_2^+) are used to ionize the trace gases in air/vapour samples. These precursor ions are injected into flowing helium carrier gas, and the gas/vapour to be analysed is introduced at a specified flow rate (20 mL/min in the present experiments) via the heated sample inlet port; the capillary within the inlet port is heated to a temperature of 80°C to prevent condensation of water vapour and/or metabolites. The flow speed of the helium carrier gas transporting the ion swarm and the length of the flow tube determines the reaction time of the precursor ions with the analyte molecules. Bimolecular and termolecular reactions occur when a specific precursor ion reacts with the trace compounds in the sample gas/vapour, leading to conversion of a small fraction of the reagent ions to the product ions that are characteristic of the analytes⁵³. The remaining (large fraction) of the reagent ions together with the characteristic product ions are sampled from the flowing swarm *via* a pinhole orifice (~ 0.3 mm diameter) located at the downstream quadrupole mass spectrometer. After mass analysis, absolute quantification of the volatile organic compounds is achieved by detection and counting at the channeltron multiplier/pulse counting system⁵³.

3.2 Ion Chemistry in SIFT-MS

Chemical ionisation (soft ionisation) in combination with mass spectrometry was developed with the specific objective of minimising the fragmentation of complex analyte molecules, resulting in simpler characteristic product ions and thus aiding analyte molecule identification and quantification. However, there are few precursor ions suitable for ambient gas analysis as these ions must be relatively unreactive with the major constituents of air (i.e. N_2 , O_2 , H_2O , CO_2 , Ar) but very reactive with the trace gases being quantified. The study of ion-molecule reactions at thermal energies identified those ions appropriate for this purpose¹⁰⁴. At present, the only suitable reagent ions species for SIFT-MS are H_3O^+ , NO^+ & O_2^+ ; these ions do not undergo bimolecular reactions with the major components of air and react only very slowly via slow termolecular association^{105,106,107,108}. The availability of three reagent ions in SIFT-MS provides improved accuracy for compound identification, which is particularly important when isobaric product ions are formed. The ability to rapidly switch between the H_3O^+ , NO^+ & O_2^+ reagent ions within milliseconds also allows one to obtain analytical data for each reagent ion without delay, which is advantageous in the analysis of transient gaseous samples. Moreover, these precursor ion species are generated efficiently in simple discharge ion sources¹⁰⁹. The potential use of anions as precursor ions for SIFT-MS analyses has previously been explored. The reactivity of OH^- was found to be far less in comparison to H_3O^+ and there was difficulty in obtaining a large enough current of OH^- ions; using negative ions would thus restrict the level of analytical sensitivity of SIFT-MS¹¹⁰.

3.3 The development of the *Profile 3* SIFT-MS instrument

In 1997, the drive to develop a transportable SIFT-MS instrument resulted in the creation of the transportable Mk.1 instrument, which weighed 350kg. The new-found mobility of the SIFT-MS allowed it to be employed for the first time outside the setting of the analytical chemistry laboratory. Online gas analysis was conducted of exhaled breath samples from patients pre- and post-haemodialysis in a hospital¹¹¹. The Mk1 instrument was also employed to study exhaust gas emissions from a diesel engine in a health and safety laboratory¹¹². In 2001, the Mk.2 instrument became the first commercial version of SIFT-MS and consisted of the latest components (pumps and quadrupole mass spectrometers) and improved specifications¹¹³. This allowed for a relaxation in critical parameters such as orifice diameters and improved the detection limit for exhaled breath samples to 2 *ppbv* for a 1 sec integration

time¹¹³. The Mk.2 instrument was essential in the study of important abundant exhaled breath metabolites in healthy subjects including acetone, ammonia, methanol, acetaldehyde, propanol and isoprene^{114,115,116,117}. The considerable thinking and continued efforts of Professors Smith and Španěl regarding the flow dynamics, Venturi design, pumping systems, vacuum housings and the underlying ion chemistry of SIFT-MS led to the production of the *Profile 3* instrument (Instrument Science Ltd, Crewe, UK), which was launched in 2006 in the UK. In particular, the replacement of the large pumps that were utilized in the Mk.1 and Mk.2 instruments with rotary backing pumps, led to a significant reduction in weight and size of the *Profile 3* instrument (weight ~120kg), resulting in a more easily transported instrument.

There have been several important principles that have guided the desire to increase the sensitivities of the SIFT-MS instruments¹¹³. As the reagent ions are convected along the flow tube by the carrier gas, their number density (n_1) near to the upstream ion injection orifice is continuously reduced by diffusive loss to the flow tube walls. This results in a lower number density (n_2) near the downstream ion sampling orifice. After being transported through the ion sampling orifice and subsequent detection at the downstream mass spectrometer, the count rate of the precursor ions at the ion detector is designated I_1 . The ion diffusive loss is decreased by increasing the carrier gas pressure; however, the upper limit of this pressure is dictated by the gas flow that can be tolerated through both orifices to ensure appropriately low pressures within the quadrupole chambers. A reduction in the length of the flowtube (l) and increasing the flow velocity of the carrier gas and therefore the ions (v_i) also reduces diffusive loss. Once a gaseous sample is introduced into the instrument, the precursor ions react with the trace gases within the sample (rate coefficient k) producing characteristic product ions. The equation below expresses the fundamental scientific principles that underpin the design of SIFT-MS, in particular the need for a suitably high enough I_1 and reaction time (l/v_i)¹¹³:

$$I_2 = I_1 k [M] \frac{l}{v_i}$$

(N.B. $[M]$ = number density of the trace gas)

There is a need for balance between all the variable parameters including the diameters of the upstream ion injection and ion sampling orifices diameters, carrier gas pressure, length of flow tube and flow velocity of ions to ensure optimal sensitivity of the SIFT-MS instrument. The flowtube within a *Profile 3* instrument is 5 cm long and 1 cm diameter. The operating helium

carrier gas pressure is ~ 1.0 Torr and the reaction time is about 0.5 msec^{113} . The combination of the shorter flow tube and the higher pressure than in the Mk.1 and Mk.2 instruments has led to a decrease in the diffusive loss of precursor ions to the walls of the flow tube, resulting in I_1 values more than $106/\text{sec}^{113}$. A *Profile 3* instrument was employed for all the experiments described in this thesis (see Figure 3.1)

3.4 Modes of acquisition of SIFT-MS data

SIFT-MS can be operated in two different modes. In the full-scan mode (FS), the ambient air/breath or liquid headspace vapours are introduced at known flow rates into the carrier gas, and the downstream analytical mass spectrometer system is scanned over predetermined m/z ranges for a chosen time period⁵³. The count rates of the ions are then calculated from the actual numbers of counts and the total sampling time for each ion. The count rates are stored, and then displayed on a linear or semi-logarithmic scale. The mass spectra are interpreted by relating the product ion peaks to the trace gases present in the sample. The concentrations of VOCs and trace gases are calculated from the count rates of the precursor ions and their characteristic product ions and the appropriate k values (using the standardized kinetics library) and the flow rates of the helium carrier gas and of the air/breath/headspace sample. In the multi-ion monitoring mode (MIM), the downstream analytical mass spectrometer is rapidly switched between selected m/z values to target selected trace gas species⁵³. From the data obtained, together with the rate coefficients of the reactions as included in the kinetics library, specifically targeted volatile compounds are quantified in real-time. This real-time monitoring is possible because of the instantaneous response of SIFT-MS, which is typically less than 20 ms; the fast response time is dictated by the fast flow rates of the carrier gas along the flow tube and the sample gas along the inlet tube. In addition to the count rates of the characteristic product ions, those of the precursor ions (e.g., H_3O^+ and $\text{H}_3\text{O}^+(\text{H}_2\text{O})_n$) are also continuously recorded, because they continually change with the humidity of the sample gas.

3.5 Sensitivity and limit of detection of the *Profile 3* SIFT-MS instrument

The functionality of the ion sources employed in SIFT-MS are essential in determining the sensitivities and detection limits of the instrument. Španěl *et al.* conducted a study on the ion chemistry occurring within the microwave discharge plasma ion sources that are currently used for the production of currents of the three precursor ions¹¹⁸. They demonstrated that the most

suitable ion source gas composition is a mixture comprising maximal water vapour and minimal air at the lowest total pressure at which the discharge is sustained and stable¹¹⁸. It was also reported that rapid switching between the H_3O^+ , NO^+ and O_2^+ precursor ions is only possible because the SIFT-MS source constantly produces these three reagent ions in similar concentration when the partial pressures of air and water vapour in the discharge are optimised. The continual improvements in the performance of the microwave discharge ion sources have led to count rates of several million per second for the reagent ions and, in turn, superior count rates for the characteristic analytical product ions, which have a first order influence over the sensitivity of SIFT-MS. Most importantly, this has the overall effect of delivering better quality mass spectra. The limit of detection for any trace gas can be evaluated in terms of the reagent ion count rate, the sample flow rate, the reaction rate coefficient and the data acquisition time¹¹⁹. The reagent ion count rate of the *Profile-3* SIFT-MS instrument under optimal conditions is in excess of 10^6 s^{-1} , which results in a detection limit of ~ 0.1 ppbv for 1 sec of integration¹¹³.

3.6 Direct analysis of ambient gaseous samples

A major advantage of the SIFT-MS technique is that the sample and its analytes are not modified by any pre-analysis preparation and remain undisturbed throughout the analytical process. Several approaches for ambient analyses using SIFT-MS have previously been described:

- 1) Through presentation of the sample to the sampling capillary of the instrument when the ambient air is also sampled simultaneously;
- 2) By collecting the gaseous sample in either a collapsible bag in which atmospheric pressure is automatically maintained or in a fixed volume vessel;
- 3) Liquid samples have been placed in a fixed volume container closed by a septum, which is subsequently punctured by a needle connected directly to the sample inlet of the instrument.

Real time analysis of a gaseous sample can be performed by directly connecting it to the sampling inlet, whilst the downstream mass spectrometer allows for data acquisition in either the FS or MIM mode. This is made possible as the sample input pressure remains atmospheric, thereby making the analysis relatively simple. Another valuable feature of the SIFT-MS

method is that the humidity of an air and breath sample can be obtained routinely from the distribution of the hydrated hydronium ions. The humidity measurement is based on the idea that the injected H_3O^+ precursor ions react with H_2O molecules in the sample producing the hydrates $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{1,2,3}$, and it is the fraction of the H_3O^+ that is converted to its hydrates that provides a determination of the H_2O number density in the helium carrier gas and hence the sample humidity. This determination is possible, because the three-body rate coefficient (k_3) for the conversion of H_3O^+ to $\text{H}_3\text{O}^+(\text{H}_2\text{O})$ is known. Since the humidity of exhaled breath in humans (6% at a body temperature of 37°C) is known, this serves as an internal validation for the analytical process¹¹⁹. Prior to any sample being analysed, the operator performs a calibration procedure using exhaled breath acetone (H_3O^+). The flow rate settings of the instrument are standardized to ensure the water content of the operator's exhaled breath measures 6 +/- 0.1%; these checks act as an important quality control step prior to performing any SIFT-MS analysis¹²⁰.

Another unique feature of the SIFT-MS method is that one can analyse several compounds simultaneously from the headspace above a liquid, a Nalophan[®]/Tedlar[®] collected exhaled breath sample or even from a single exhalation directly into the instrument. These compounds can range from water vapour and CO_2 present at percentage levels to trace compounds that are present at *ppbv* concentrations. Figure 3.2a illustrates the online capabilities of SIFT-MS; it includes data obtained for the detection and quantification of water vapour, CO_2 and acetone using a *Profile 3* instrument in MIM mode from six sequential single breath exhalations¹²¹. In Figure 3.2b, the simultaneous analysis of ammonia, acetone and hydrogen cyanide within single exhalations is shown - the rapid time response of the SIFT-MS allows instantaneous measurement of trace compounds and abundant compounds simultaneously and spans a concentration range of 10^7 .

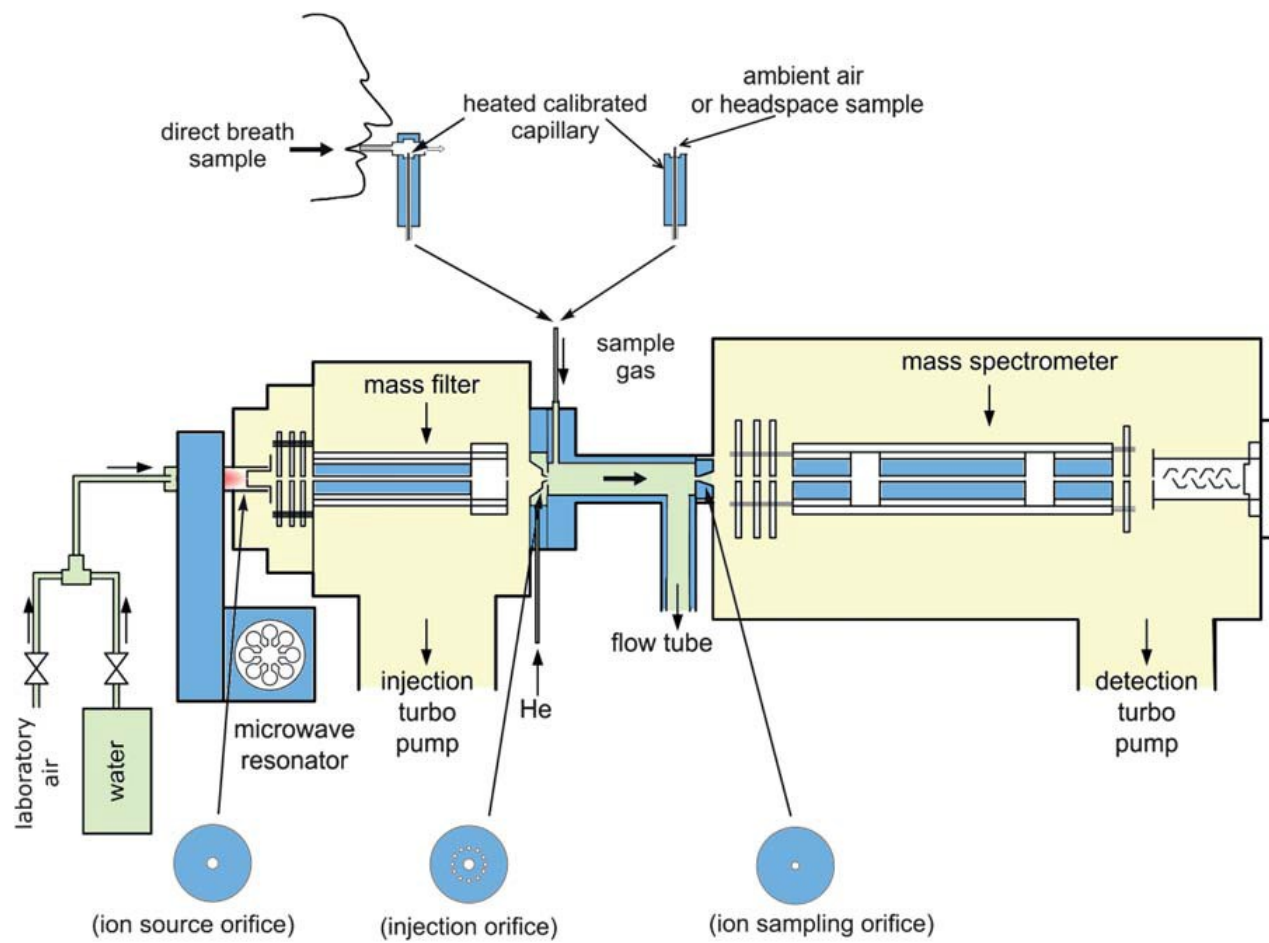


Figure 3.1: A schematic diagram of the *Profile 3* SIFT-MS instrument showing the microwave discharge ion source, injection mass filter and the three metal discs to which ion current can be measured and which support the orifices through which (left to right) ions pass from the ion source into the injection mass filter, mass selected ions enter the flow tube, and ions pass from the carrier gas into the analytical quadrupole mass spectrometer¹²¹.

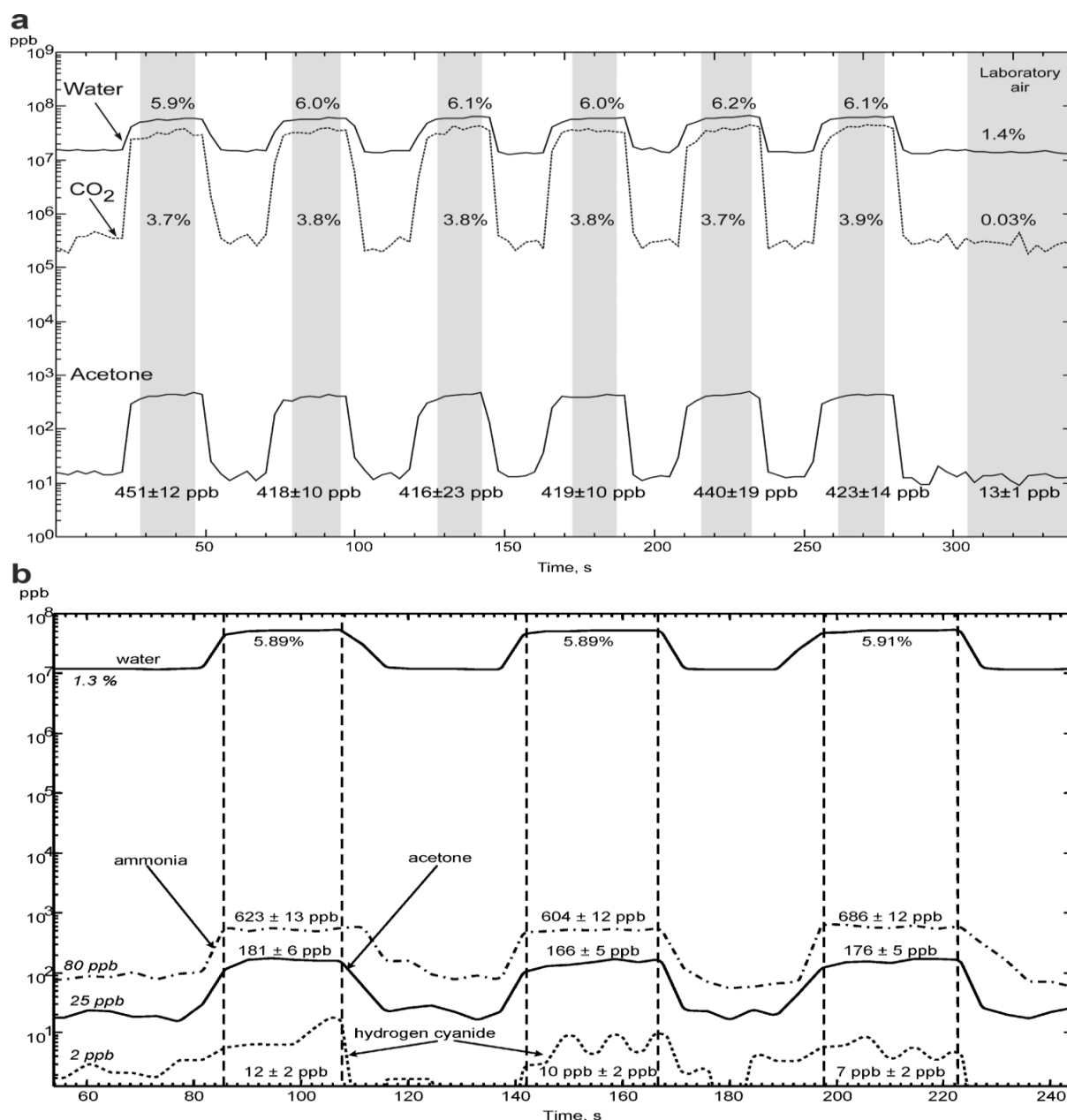


Figure 3.2: Plots of the derived concentrations of water vapour and acetone (in *ppbv*) obtained using the *Profile 3* SIFT-MS instrument in the multi ion monitoring mode for six sequential breath exhalations by one volunteer over the time indicated in seconds. This data illustrates the consistency in the derived levels of these compounds in the alveolar regions of the exhalations (as indicated within the shaded regions) The mean values (with standard deviations): are water vapour, 6.1 (+/-0.1)%, CO₂ 3.8 (+/-0.1)% and acetone 428 (+/-14) *ppbv*. Also indicated are the laboratory air levels of these compounds. b: Simultaneous absolute quantification of water vapour, ammonia, acetone and hydrogen cyanide in three sequential breath exhalations. The water vapour is ~6% of the exhaled breath and that the concentration of the trace compound hydrogen cyanide obtained in the same breath exhalations has a mean value of 10 *ppbv*¹²¹.

3.7 The SIFT-MS Kinetics Library

SIFT-MS quantification of VOC concentrations is calculated using the laws of chemical kinetics and the measured physical quantities. The concentrations are automatically calculated instantaneously using the on-board *Profile 3* SIFT-MS software. There are general equations applying to all compounds which are coded within the algorithms of the SIFT-MS software. Separately, there is a kinetics library of numerical constants which are characteristic and specific to individual compounds. The number densities of analytes diluted by the sample air and by the carrier gas in the flow tube, $[M]$, are calculated using the known rate coefficients for their reactions with the reagent ions, k_1 , the reaction time, t_r , and the ratio of the sum of the count rates of all the product ions to the sum of the count rates of the precursor ions and their hydrates¹²². These count rates are divided by factors that account for differential diffusion (De) and can be optionally modified by factors (f) when specific additional kinetic processes must be accounted for.

To allow for immediate calculation of VOCs, each compound is assigned a corresponding set of reagent ions and characteristic product (analyte) ions together with the rate coefficients and factors (f). This data are stored within the SIFT-MS kinetics library. Examples of the kinetics library entries used for the analysis of acetone by H_3O^+ , NO^+ and O_2^+ are given in Table 3.1:

Compound(ion)	Acetone (H ₃ O ⁺)	Acetone (NO ⁺)	Acetone (O ₂ ⁺)
Number of precursors (<i>m/z k fi</i>)	4 precursors 19 3.9e-9 1.0 37 3.3e-9 1.0 55 2.5e-9 1.0 73 2.4e-9 1.0	1 precursor 30 1.8e-9 1.0	1 precursor 32 3.1e-9 1.0
Number of products (<i>m/z k fp</i>)	3 products 59 1.0 77 1.0 95 1.0	1 product 88 1.0	2 products 43 1.0 58 1.0

Table 3.1: Sample Kinetics Library entries for acetone using all three precursor ions.

The format of each entry within the kinetics library is as follows - after the name of the compound and the symbol of the injected ion, the number of reagent ions is given followed by their *m/z* values with the rate coefficients in units of cm³s⁻¹. The *fi* are simply used to multiply the acquired raw count rate of the ion, with its *m/z* given at the beginning of the row. Normally, *fi* values of 1.0 are used, as indicated in example of acetone. However, such simple calculations do not always provide valid results in the cases when additional ion chemistry occurs (e.g. the removal of the product ions by reactions with H₂O molecules in reverse proton transfer or switching reactions) or when product ions overlap with other ions present¹²². Values other than 1.0 can therefore be incorporated to account for the influence of humidity or to optimise calculations that do not sum all product ions¹²².

3.8 Headspace of Biofluids and Exhaled Breath Sampling

The sampling methodologies and individual experimental protocols employed for SIFT-MS analysis of gastric content, urine and exhaled breath are described in subsequent chapters. The reproducibility of these methods is also evaluated.

Chapter 4

VOCs in the headspace of gastric content

4.1 VOCs within gastric content in OG Cancer

4.1.1 Objective

To evaluate whether the VOCs within gastric content retrieved from patients with oesophago-gastric cancer differs from patients with benign upper gastro-intestinal conditions and those with healthy upper gastro-intestinal tracts.

4.1.2 Introduction

Gastric content is a complex biofluid within the human stomach, which has an important role in digestive processes and is a combination of gastric acid, saliva and bile reflux. Gastric acid forms a major component of gastric content, being secreted by the gastric parietal cells and is a combination of HCl, NaCl and KCl; this results in gastric content most commonly being at the strongly acidic end of the pH range. Gastric content hypoacidity has been implicated in the development of upper gastrointestinal (GI) diseases. Lu *et al.* investigated the pH of fasting gastric content samples from patients with upper GI diseases and healthy controls and observed gastric cancer and gastric ulcer patients have a significantly higher pH than healthy controls¹²³. Gastric content has also been proposed as a potential source of upper GI disease biomarkers. Inada *et al.* found that cobalamin-binding protein was present in the gastric content of all patients with early gastric cancer¹²⁴. Hsu *et al.* reported that gastric content α 1-antitrypsin concentrations in gastric cancer patients are markedly increased compared to healthy subjects and patients with benign gastrointestinal diseases¹²⁵. Deng *et al.* have investigated aromatic amino acids in gastric content samples from patients with early-stage gastric cancer, advanced gastric cancer and non-neoplastic gastric diseases using high performance liquid chromatography mass spectrometry (HPLC-MS)¹²⁶. They reported significantly higher concentrations of tyrosine, phenylalanine, tryptophan and total protein in the gastric content of patients with gastric malignancies¹²⁶.

Biofluids obtained from organ-specific cavities have much potential in assisting in the identification of disease biomarkers. VOCs have previously been investigated in urine, serum, plasma, saliva, sweat and bronchoalveolar lavage^{59,70, 127,128}. Gastric content has been proposed as a contributory factor in the pathogenesis of upper gastro-intestinal diseases and represents the only biofluid to comes into direct contact with oesophageal and gastric cancer cells. Hence,

it represents a promising medium to investigate VOCs of interest in oesophago-gastric cancer. In the present study, headspace analysis of gastric content in upper gastro-intestinal diseases was investigated for the first time.

4.1.3 Methods

4.1.3.1 Patient selection

Patients for this study were recruited through St. Mary's Hospital, Paddington (Imperial College NHS Trust, London) and via the North West Thames regional oesophago-gastric cancer network (UK). Patients referred with upper gastro-intestinal symptoms are commonly sent for an Oesophago-Gastro-Duodenoscopy (OGD) by their clinician. An OGD is a routinely performed camera test in which the upper gastro-intestinal tract is visualized (a description of the procedure and sample retrieval process is provided in the section below). An OGD referral form contains information explaining the indication for the procedure and the patient's known medical history. These referral forms were screened to identify suitable patients for recruitment. The exclusion criteria for this study included patients with:

- Barrett's oesophagus
- Liver disease
- Small bowel or colonic pathology
- A history or presence of another type of cancer
- Presence of acute infection

Local ethics committee approval was obtained for this study. Patients eligible for the study were given information sheets pertaining to the study and informed consent was obtained prior to enrolment in the study. All patients provided a comprehensive medical history, including past medical history, medications as well as smoking status and alcohol intake. Demographic and clinical data for each patient were archived in a linked, anonymized database. Only patients who had requested sedation for their OGD procedure and agreed not to have Xylocaine® (lidocaine hydrochloride) throat spray were included in the study. Patients with a biopsy-confirmed invasive gastric or oesophageal cancer were included in the cancer cohort. Patients with any non-cancer conditions of the upper gastrointestinal tract (e.g. oesophagitis, gastritis and peptic ulcer disease) were included in the positive control cohort. Patients with a normal OGD examination and confirmed *H. pylori* negative were included in a normal Upper GI cohort

group. The comparative analysis of gastric content vapour metabolites was conducted with an oesophago-gastric cancer cohort of 25 patients, a positive control group of 20 patients and a 'Normal Upper GI' group of 18 patients.

4.1.3.2 Gastric content sampling

All patients were fasted for a minimum of 6 hours, which is standard protocol for an OGD procedure. All gastric content samples were taken from patients undergoing an OGD. A gastroscope (Olympus, Essex, UK) is introduced through the mouth, into the oesophagus and then into the stomach. During this procedure, biopsies can be taken for histological examination and fluid can also be retrieved through a suction channel within the gastroscope. The gastric content was retrieved during the first pass of the gastroscope and captured in a disposable suction jar (see Figure 4.1). This suction jar was then taken to the SIFT-MS instrument located in the hospital and all samples were analysed within 1 hour. The pH of gastric content samples was measured immediately after SIFT-MS analysis using a Hanna Piccolo[®] pH meter (Sigma-Aldrich, Missouri, USA).



Figure 4.1: Suction Trap for the collection of gastric content

4.1.3.3 SIFT-MS analysis

In the first phase of this study, the FS mode (please refer to section 3.4) of SIFT-MS was initially operated to reveal which groups of compounds were present in the headspace vapour of the gastric content. However, the MIM mode (please refer to section 3.4) was mainly employed for this study as it is more sensitive and reproducible than the FS mode and allows targeted compound analysis. Prior to MIM headspace analysis of a sample, the calibration procedure using the water content of the operator's exhaled breath ($6 \pm 0.1\%$) was carried out to check the sample flow rate into the instrument. For each measurement, a total of 10mL of gastric content from the original sample jar was aliquoted into a standard 60 mL specimen jar, which was then sealed with parafilm and placed in an incubator at 37°C for five minutes prior to analysis. A 21g Luer-lock sterile hypodermic needle was connected to a 200cm Vygon extension tube, which was directly connected to the SIFT-MS sampling port. The sterile needle end punctured the parafilm layer of the specimen jar to sample the headspace above the gastric content liquid. The headspace vapour automatically flowed into the helium carrier gas at a fixed rate of 20mLs/min through the sampling line which was held at a temperature of 80°C and the analysis was conducted. Throughout the analysis, the specimen jar, needle and extension tubing were all kept in the incubator at 37°C . All experiments in the study were conducted using only the 200cm version of the tubing, after assessment of 50cm versus 200cm Vygon extension tubes for headspace sampling of gastric content was conducted. The 50cm extension tubing was suitable for headspace analysis for a short period of analysis; however, after ~20minutes a droplet of condensation would reach the sample inlet and if this entered the instrument, the analysis would be ruined. For the FS mode, two serials of 9×60 seconds repeated scans were performed using H_3O^+ and NO^+ precursor ions. The O_2^+ precursor ion was not used for the FS analysis as it is more energetic than the others precursor ions resulting in busy spectra that are more difficult to interpret¹²⁹. For the MIM mode, the selected VOCs from the vapour samples were analysed for a total of 60 seconds and the measured concentrations were averaged over the analysis time for each VOC.

4.1.3.4 Statistical analysis

Statistical analysis was performed using IBM SPSS statistics 19 (SPSS Inc., Chicago, IL). Mann-Whitney U test was used to compare the measured concentrations of VOCs between oesophago-gastric cancer patients and healthy controls. A $p\text{-value} \leq 0.05$ was taken as the level to indicate statistical significance. In medicine, a Receiver Operating Characteristic curve (ROC) is a fundamental test used to assess the accuracy of a diagnostic test to discriminate between a diseased group and healthy group. ROC curves were originally developed in the field of statistical decision theory and used by radar operators in World War II to distinguish between enemy targets or ships belonging to the Allied Powers¹³⁰. In a ROC curve for clinical assessment, the sensitivity is plotted against (1-specificity) using the variation to differentiate between the two groups. The cut-off points on the ROC curve indicate when a discrimination threshold is reached with a paired sensitivity and specificity. The area under the curve (AUC) measures the discrimination power of the test, which is the ability of the test to correctly classify subjects into the correct group¹³¹. The higher the AUC value, the better the ability of the test to positively identify those persons with the disease and correctly rule out those persons without the disease. A diagnostic test is considered highly accurate when the AUC is > 0.9 ⁶⁴. The statistically significant VOCs (identified through Mann Whitney U analysis) that showed a difference between the cancer and the healthy cohort were used as included variables to construct the ROC curve. The disease conditions were used as the dependent variable and the statistically significant VOCs were used as the independent variable. A binary logistic regression method was conducted to test different combinations of selected VOCs as the variables to find the model with the highest R^2 value and AUC value for the ROC curve.

4.1.4 Results

4.1.4.1 Analysis of full scan spectra

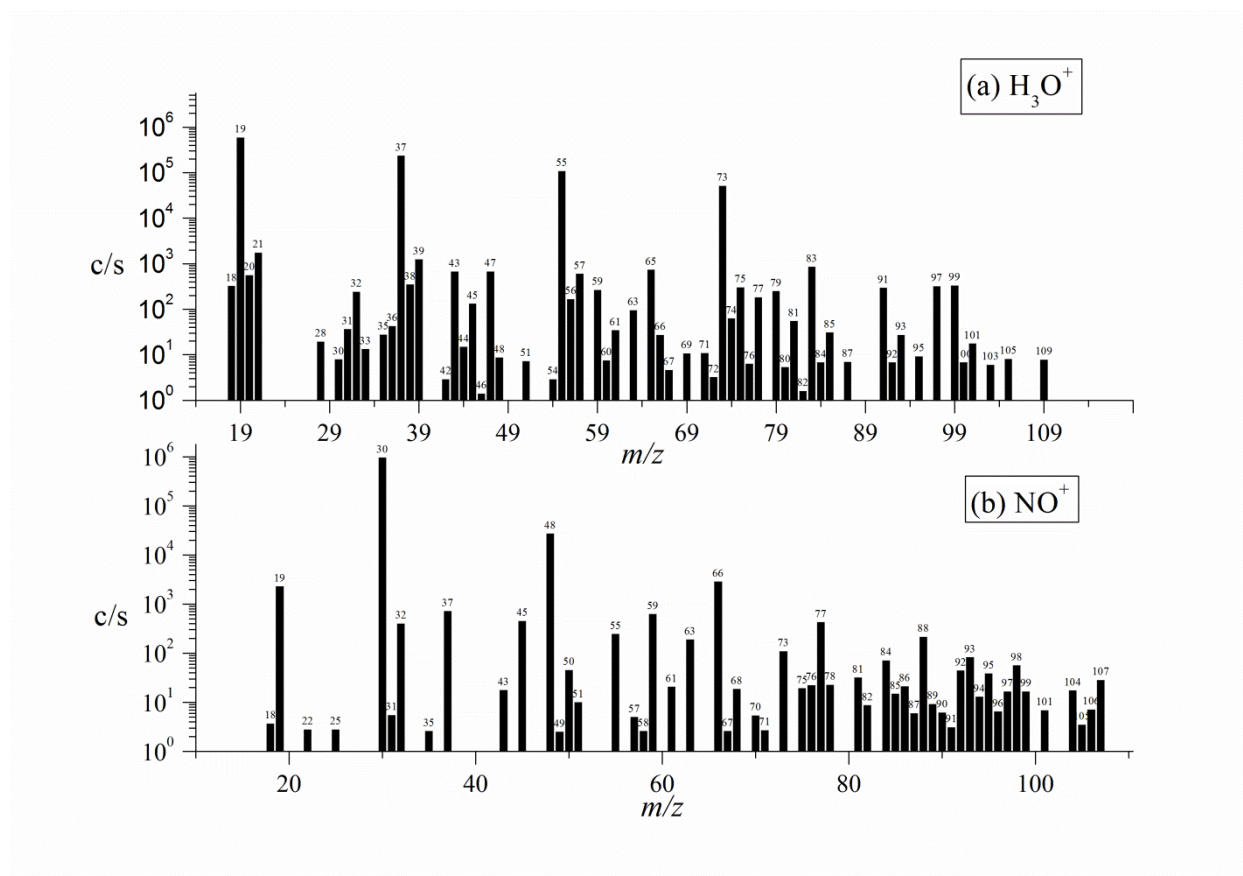


Figure 4.2: Full scan spectra of SIFT-MS analysis of the headspace of gastric content using (a) H_3O^+ precursor ion; (b) NO^+ precursor ion. The individual characteristic product ions and their hydrates were assigned to the analysed molecules. In (a), m/z 19, 37, 55, 73 are $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{0,1,2,3}$; m/z 59, 77, acetone; m/z 33, 51, 69, methanol; m/z 47, 65, 83, ethanol; m/z 45, 81, 83, acetaldehyde; m/z 31, formaldehyde; m/z 35, hydrogen sulphide; m/z 28, hydrogen cyanide; m/z 95, phenol; m/z 18, 36, 54, ammonia. In (b), m/z 19, 37, 55, 73 are $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{0,1,2,3}$; m/z 30, 48, 66, 84 are $\text{NO}^+(\text{H}_2\text{O})_{0,1,2,3}$; m/z 90 is acetic acid.

Full scan (FS) spectra were obtained from analysis of the headspace vapour of gastric content. Representative spectra using H_3O^+ and NO^+ precursor ions are shown in Figure 4.2. A variety of VOCs was detected and several compounds were identified from their characteristic product ions. From full scan spectra, those VOCs that led to characteristic product ions of nominally the same m/z values (overlapping ions) were not selected for MIM mode analysis. For instance,

in the carboxylic acid group, propanoic acid and butanoic acid were not selected for further MIM investigation as their product ions shared overlapping m/z values and these would be difficult to quantify accurately. GC-MS is often not capable of detecting low molecular mass trace volatile organic compounds such as hydrogen sulphide and hydrogen cyanide, but both H_2S and HCN were unambiguously identified from their product ions using SIFT-MS.

4.1.4.2 Reproducibility and validity of the sampling method

The reproducibility of headspace analysis of gastric content was assessed. A total of ten gastric content samples were analysed and with each sample three MIM measurements for selected VOCs were taken. The coefficient of variation (CV) were calculated for the spread of measured concentrations of the VOCs in the gastric content headspace. These CV values in percent are: acetaldehyde, 6.7%; formaldehyde, 11%; acetone, 6.8%; hexanoic acid, 10.7%; methanol, 15.9%; ethanol, 6.2%; phenol, 22.6%; methyl phenol, 13.4%; hydrogen sulphide, 44.7%; hydrogen cyanide, 22.7%; ammonia, 6.8%; nitric oxide, 14.8%. These coefficients of variation demonstrate the reproducibility of the MIM measurement. In contrast to the major metabolites (e.g. acetone, acetaldehyde) the higher CV for hydrogen sulphide and hydrogen cyanide is expected given their extremely low concentrations in gastric content and hence the lower count rates of their characteristic product ions¹¹⁹.

4.1.4.3 MIM analysis of selected VOCs in the headspace of gastric content

Compounds	Molecular formula	Precursor Ions	m/z	Characteristic product ions	REF
Methanol	CH ₄ O	H ₃ O ⁺	33, 51, 69	CH ₃ O ⁺ , CH ₃ O ⁺ (H ₂ O), CH ₃ O ⁺ (H ₂ O) ₂	132
Ethanol	C ₂ H ₆ O	H ₃ O ⁺	47, 65, 83	C ₂ H ₇ O ⁺ , C ₂ H ₇ O ⁺ (H ₂ O), C ₂ H ₇ O ⁺ (H ₂ O) ₂	132
Acetic acid	C ₂ H ₄ O ₂	NO ⁺	90, 108	NO ⁺ ·CH ₃ COOH, NO ⁺ ·H ₂ OCH ₃ COOH	133
Hexanoic acid	C ₆ H ₁₂ O ₂	H ₃ O ⁺	117, 135	C ₆ H ₁₂ O ₂ H ⁺ , C ₆ H ₁₂ O ₂ H ⁺ (H ₂ O)	132
Acetaldehyde	C ₂ H ₄ O	H ₃ O ⁺	45, 81	CH ₃ CHOH ⁺ , CH ₃ CHOH ⁺ (H ₂ O) ₂	134
Formaldehyde	CH ₂ O	H ₃ O ⁺	31	CH ₂ OH ⁺	134
Acetone	C ₃ H ₆ O	H ₃ O ⁺	59, 77	C ₃ H ₆ OH ⁺ , C ₃ H ₆ OH ⁺ (H ₂ O)	134
Acetone	C ₃ H ₆ O	O ₂ ⁺	43, 58	C ₃ H ₆ O ⁺ , C ₂ H ₃ O ⁺	134
Hydrogen sulphide	H ₂ S	H ₃ O ⁺	35, 53	H ₃ S ⁺ , H ₃ S ⁺ (H ₂ O)	129
Phenol	C ₆ H ₆ O	H ₃ O ⁺	95, 113	C ₆ H ₆ OH ⁺ , C ₆ H ₆ OH ⁺ (H ₂ O)	135
Methyl phenol	C ₇ H ₈ O	O ₂ ⁺	108, 126	C ₇ H ₈ O ⁺ , C ₇ H ₈ O ⁺ (H ₂ O)	135
Ammonia	NH ₃	O ₂ ⁺	17	NH ₃ ⁺	83
Hydrogen cyanide	HCN	H ₃ O ⁺	28	H ₂ CN ⁺	136
Nitric oxide	NO	O ₂ ⁺	30	NO ⁺	83

Table 4.1: A summary of analytical information of VOCs analysed by the MIM mode of SIFT-MS using variously H₃O⁺, NO⁺ and O₂⁺ precursor ions and the *m/z* of characteristic product ions.

A total of 13 volatile compounds released from the headspace of gastric content were analysed using the MIM mode of the SIFT-MS. The analytical information, including chemical formula, precursor ions used, and the characteristic product ions and their m/z ratio are given in Table 4.1. H_3O^+ is the most used precursor ion as it readily protonates most VOCs. However, NO^+ and O_2^+ have also been employed for specific molecules, when they produce unambiguous characteristic product ion and the overlapping ions that sometimes occur when using H_3O^+ precursor ions, e.g. propanol and acetic acid.

		OG Cancer		Normal Upper GI tract		Positive controls ¹	
VOCs	Precursor ion	Median	IQR	Median	IQR	Median	IQR
Acetaldehyde	H_3O^+	77	[52-141]	29	[21-55]	63	[27-136]
Acetone	H_3O^+	94	[28-253]	26	[12-67]	24	[16-75]
Formaldehyde	H_3O^+	45	[36-52]	52	[40-72]	49	[36-77]
Acetic acid	NO^+	43	[32-73]	20	[16-39]	30	[19-100]
Hexanoic acid	H_3O^+	36	[20-269]	13	[11-71]	13	[10-196]
Hydrogen sulphide	H_3O^+	2.3	[0.8-5.5]	0.9	[0.4-1.9]	1.7	[0.8-2.2]
Hydrogen cyanide	H_3O^+	15.1	[7.7-25]	6.1	[4-10.5]	7.2	[3.7-13]
Methanol	H_3O^+	74	[53-160]	48	[39-69]	51	[36-91]
Acetone	O_2^+	296	[154-525]	96	[63-128]	178	[86-229]
Methyl phenol	O_2^+	17	[11-22]	8	[5-15]	12	[8-31]

Table 4.2: A summary of median concentrations and interquartile ranges (IQR) for those VOCs that were significantly different between cancer and healthy groups. (The results of positive control group are included for comparison). ¹(The concentrations are measured in part-per-billion by-volume, *ppbv*. The cancer group included all patients with a biopsy-confirmed oesophago-gastric cancer. The positive control group included all patients diagnosed with non-cancer diseases of the upper gastro-intestinal tract. The Normal Upper GI cohort did not have any diseases of the upper gastro-intestinal tract and were *H.pylori* negative).

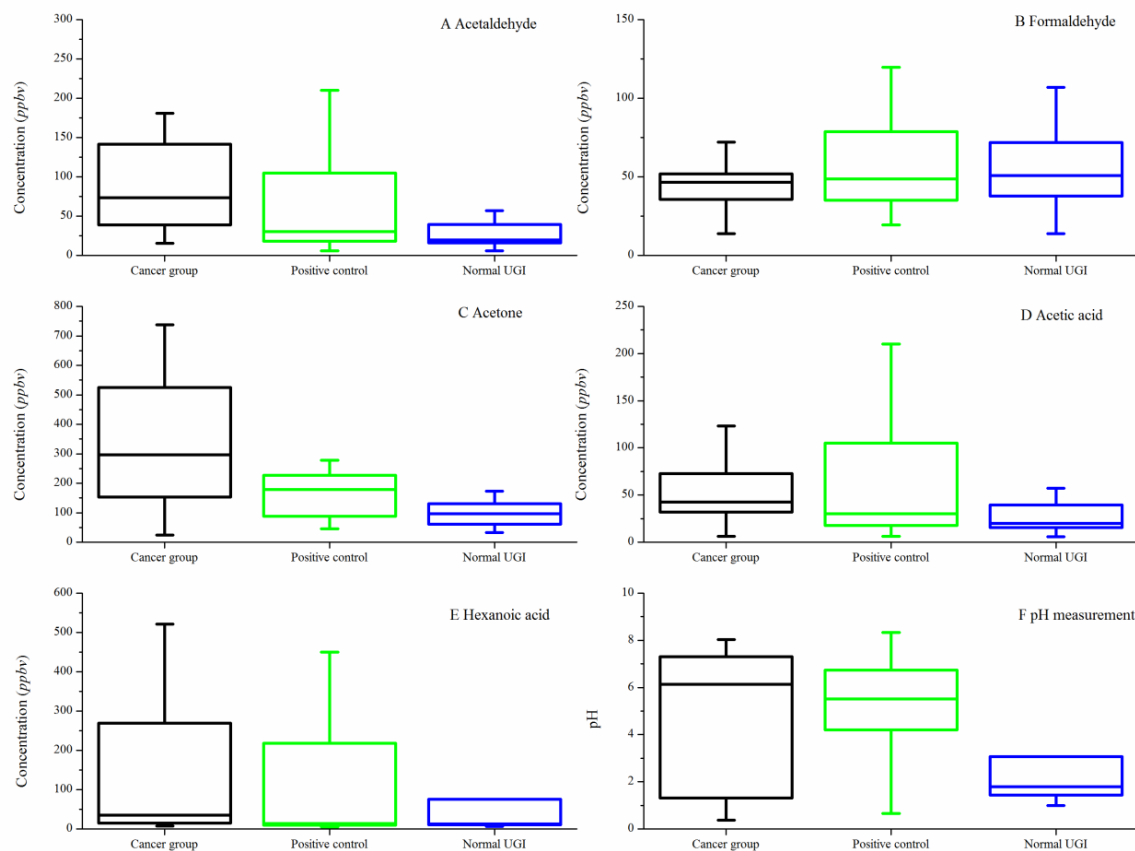


Figure 4.3: Box-Whisker plots of the headspace median concentrations and interquartile ranges (part-per-billion by-volume, *ppbv*) of (A) acetaldehyde, (B) formaldehyde, (C) acetone, (D) acetic acid, (E) hexanoic acid and (F) pH measurements of the gastric content of cancer, positive control and normal upper GI groups. Acetaldehyde and formaldehyde and hexanoic acid were analysed using H_3O^+ precursor ions, acetone were analysed using O_2^+ ions, acetic acid was analysed using NO^+ precursor ions.

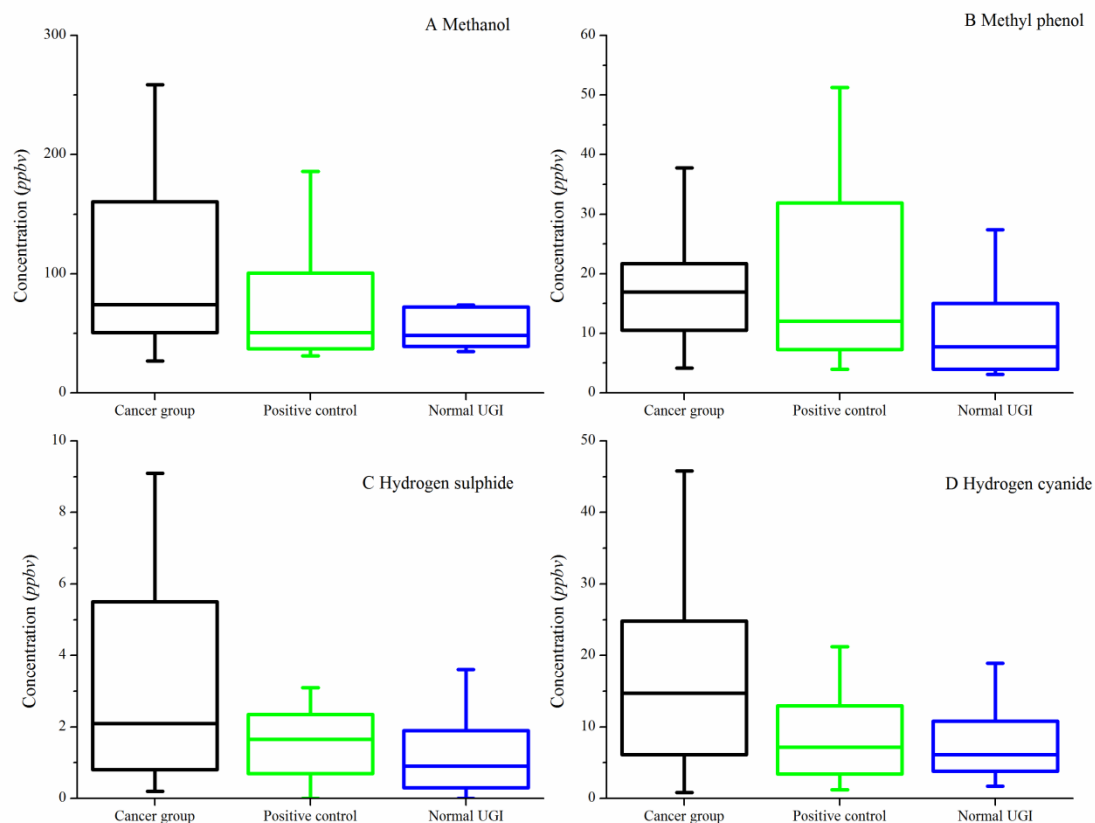


Figure 4.4: Box-Whisker plots of the headspace median concentrations and interquartile ranges (part-per-billion by volume, *ppbv*) of (A) methanol, (B) methyl phenol, (C) hydrogen sulphide and (D) hydrogen cyanide from the gastric content of cancer, positive control and normal upper GI groups. Methanol, hydrogen sulphide and hydrogen cyanide were analysed using H_3O^+ ions and methyl phenol was analysed using O_2^+ precursor ions and in the SIFT-MS.

4.1.4.4 VOCs within gastric content headspace to differentiate cancer from healthy controls

Mann-Whitney U test (p -values ≤ 0.05) was performed for all the volatile organic compounds (listed in Table 4.1) to assess for any differences in median concentrations between oesophago-gastric patients and people with healthy upper gastro-intestinal tracts. Eight metabolites were found to be statistically significantly different across the two groups including:

- acetaldehyde ($p = 0.01$)
- formaldehyde ($p = 0.046$)
- hexanoic acid ($p = 0.036$)
- acetic acid ($p = 0.016$)
- acetone ($p = 0.007$)
- hydrogen sulphide ($p = 0.015$)
- hydrogen cyanide ($p = 0.016$)
- methyl phenol ($p = 0.01$)

Their median concentrations and interquartile ranges from cancer, normal upper GI and positive control groups are given in Table 4.2.

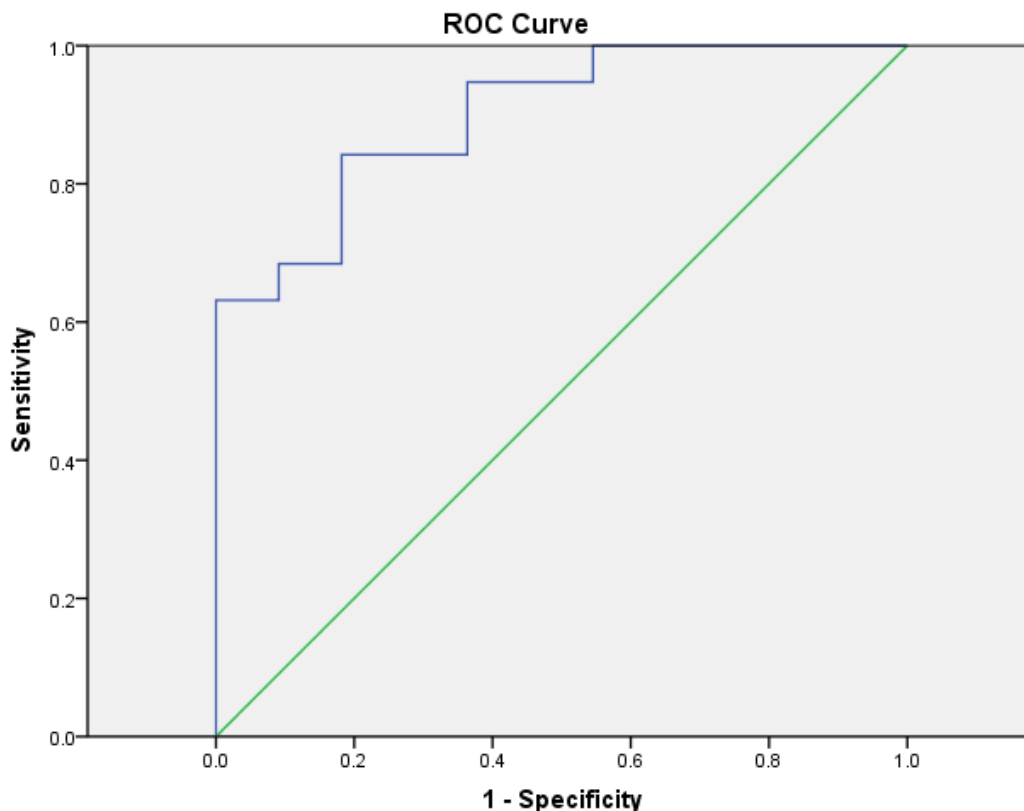


Figure 4.5: Receiver operating characteristics (ROC) curve of discrimination result for the diagnosis of oesophago-gastric cancer from normal upper GI controls using the combined VOCs of acetaldehyde, formaldehyde, hydrogen sulphide and methyl phenol. The integrated area under ROC curve (AUC) is 0.90.

Using the statistically significant VOCs identified in the Mann Whitney U test, binary logistic regression analysis was performed to construct a discrimination model. The combination of acetaldehyde, formaldehyde, hydrogen sulphide and methyl phenol demonstrated the highest area under curve (AUC) for diagnosing cancer with a value of 0.90. Figure 4.5 demonstrates a high degree of discrimination between oesophago-gastric cancer and healthy controls with reasonable sensitivity and specificity using the aforementioned compounds. The result raises the prospect that a VOC profile rather than a single biomarker may be a preferable in the molecular-orientated diagnosis of oesophago-gastric cancer.

4.1.5 Discussion

4.1.5.1 Aldehydes and Ketones

Three short-chain carbonyl compounds, i.e., acetaldehyde, formaldehyde and acetone present in the headspace of gastric content were quantified using SIFT-MS. There is a decreasing trend in the median concentrations of acetaldehyde when comparing cancer, positive control and normal upper GI cohorts (Figure 4.3). The acetaldehyde concentrations of the cancer and positive control patients have similar median values 77 *ppbv* and 63 *ppbv*, as shown in Table 4.2 and Figure 4.3. In comparison, the median value from the normal upper GI cohort (29 *ppbv*) is much lower than the cancer group. Furthermore, the majority of acetaldehyde concentrations in the normal upper GI cohort are less than 55 *ppbv*, while in the cancer and positive control groups, the majority of observed values are above 55 *ppbv*. Acetaldehyde is a metabolite of ethanol oxidation, but in recent years it has been classified as a Class I carcinogen. Several studies have postulated that deficiencies in enzymes involved in aldehyde metabolism may contribute to upper gastro-intestinal carcinogenesis; Oikawa *et al.* demonstrated that those deficient in aldehyde dehydrogenase II (ALDH2) in the presence of gastric hypochlorhydria are at increased risk of oesophageal squamous cell carcinoma¹³⁷. Linderborg *et al.* reported that reducing acetaldehyde exposure in those with an achlorhydric stomach may be an important factor in reducing the risk of gastric cancer¹³⁸.

For formaldehyde, a slight increasing trend in median value from the cancer, positive control and normal upper GI groups is apparent (Figure 4.3). Formaldehyde has been associated with cancer states to varying degrees in previously reported studies. Fuchs *et al.* studied breath gas aldehydes using GC-MS in patients with lung cancer, smokers and healthy volunteers and found that concentrations of exhaled formaldehyde were significantly lower in smokers when compared to healthy volunteers and lung cancer patients⁶³; there was however no difference observed between lung cancer patients and healthy volunteers. Spanel *et al.* studied formaldehyde in the headspace of urine from bladder cancer, prostate cancer and healthy controls⁸³. This study demonstrated that formaldehyde was clearly elevated in the headspace of urine from cancer patients when compared to healthy controls⁸³. The varying findings associated with formaldehyde in different biofluids and exhaled breath studies are difficult to explain and further investigation of formaldehyde in cancer is necessary.

Acetone, the simplest of the ketones, was measured using both H_3O^+ and O_2^+ by SIFT-MS. It is the most abundant ketone body (the other two being beta-hydroxybutyric acid and acetoacetic acid) and is a product of lipolysis¹³⁹. Elevated acetone levels have been observed

in people in starving states and specifically in diabetic ketoacidosis^{140,141,142}. There is a decreasing trend in the median acetone values from the cancer, positive control to the normal upper GI groups (Figure 4.3). Patients in the cancer cohort demonstrated the highest concentrations of acetone. As shown in Figure 4.3, acetone values in the cancer group also have a much wider interquartile range of 154-525 *ppbv* in contrast to the positive control and normal upper GI groups. This observed increase in acetone may be related to the patient's cancer status or may reflect a change in their underlying metabolism in response to disease. Patients with oesophago-gastric cancer often have difficulty in maintaining an adequate oral intake due to dysphagia and given their increased energy requirements, this can induce a state of malabsorption and weight loss. There is subsequently a reliance on ketogenic metabolic processes leading to a reduction in hepatic glycogen stores. The body's natural response to depleted glycogen stores is to cleave triacylglycerol to produce fatty acid chains and glycerol. Through beta-oxidation, these fatty acids are utilized as an alternative source of energy and produce acetyl Co-A. The liver utilizes the excessive acetyl Co-A leading to generation of ketone bodies, acetoacetate, β -hydroxybutyrate and acetone. This mechanism may be partially responsible for the increased concentrations of acetone in these patients.

4.1.5.2 Acids

Two organic acids were measured in the headspace of gastric content using SIFT-MS. Acetic acid is a metabolic intermediate for the generation of the coester acetyl-CoA. Its presence has been shown in the exhaled breath of lung cancer patients using SIFT-MS¹⁴³. As shown in Figure 4.3, acetic acid in the headspace of gastric content from the cancer cohort has a higher median value and wider interquartile range compared to the other two groups. The median value in the cancer group is found to be higher than that in the normal upper GI group, i.e., 43 and 20 *ppbv*. Hexanoic acid has been previously been noted as a potential biomarker for cancer in a study of urine from patients with leukaemia, lymphoma and colorectal cancer⁷⁰. Hexanoic acid is found to be significantly different between cancer and the other two control groups (Figure 4.3). The normal upper GI group, in particular, has a lower median and narrower interquartile range. The median concentration of hexanoic acid in the cancer group is 36 *ppbv*, i.e., approximately 3 times greater than both the positive and normal Upper GI groups. As shown in Figure 4.3, compared to the positive and normal Upper GI controls, a much wider distribution of hexanoic acid values in the cancer group are observed (interquartile range 20-269 *ppbv*).

The observed differences in the headspace acid concentrations between the 3 groups prompted analysis of the pH of the gastric content samples, because it was hypothesised that the higher acid concentrations would correlate with a lower pH of the gastric content. Gastric content can have a variable pH range depending on the respective contribution from gastric acid and bile reflux. This variation in pH, and specifically the hypoacidity, is believed to contribute to the pathogenesis of upper gastro-intestinal diseases. Lu *et al.* investigated gastric content acidity and found that both gastric cancer and gastric ulcer patients had a higher pH (6.6 and 3.4 vs 2.9) than healthy controls¹²³. As shown in Figure 4.3, the gastric content from the normal upper GI group has a median pH value of 1.93. The median pH of gastric content from the positive group is 5.67 and the median value for the cancer group is 6.18. The cancer cohort also has the widest pH range amongst the three groups. There were no differences in proton pump inhibitor or H₂ receptor antagonist use amongst the three groups; many of these patients are started on such medications (i.e. for acid suppression) by their general practitioners prior to being referred for an OGD. However, from an analytical chemistry perspective, it can be postulated that a low pH would facilitate the release of volatile organic acids into the headspace vapour. However, the results from this study do not support this theory. Figure 4.3 shows that both the cancer and positive control groups (with higher median pH values) have higher measured concentrations of acetic acid and hexanoic acid than the normal upper GI controls. The pH of a biofluid represents one variable that may contribute to the measured volatile organic acid concentrations; however, other factors such as bacteria (e.g. *H.pylori*) or cancer cells may have a greater overall influence within gastric content. The potential influence of pH on the release of VOCs from gastric content in cancer and non-cancer is further reported in this chapter.

4.1.5.3 Alcohol and Phenol derivatives

Methanol, ethanol, phenol and methyl phenol were detected in the headspace of the gastric content. Methanol is a low molecular mass highly volatile compound that can be difficult to detect by GC-MS but which is readily detected and quantified by SIFT-MS. The measured median values for methanol demonstrate a decreasing trend from the cancer cohort to the normal upper GI group and positive control cohort (Figure 4.4). The methanol concentration from the gastric content of the cancer cohort shows the widest distribution in contrast to the other two groups. Ethanol levels within the headspace of gastric content were also measured during this study; however these values are difficult to interpret given the potential effects of exogenous ethanol sources (e.g. disinfectant hand gel) within the hospital. Boshier *et al.* have

previously investigated the level of volatile trace gases within the Imperial College NHS Trust environments. Median ethanol levels of 2136 *ppbv* and 733 *ppbv* were observed within ambient air collected from hospital wards in the morning and evening, respectively¹⁴⁴. Therefore, caution must be exercised when interpreting ethanol data from patients in the hospital environment.

As shown in Figure 4.4, the distribution of methyl phenol concentrations is similar among the cancer and positive control groups. Methyl phenol concentration from gastric content in the cancer group is found to have the largest median value (17 *ppbv*), followed by positive control group (12 *ppbv*) and the normal upper GI group (8 *ppbv*).

4.1.5.4 Sulphur and nitrogen-containing compounds

H₂S and HCN are highly volatile compounds with low molecular weights, for which SIFT-MS has an advantage in analysing compared to GC-MS. The measured H₂S and HCN levels showed large variation between cancer, positive control and normal upper GI groups. As shown in Figure 4.4 and Table 4.2, the median concentrations and interquartile ranges of H₂S showed a decreasing trend from the cancer cohort (2.3 *ppbv*) to the normal upper GI group (0.9 *ppbv*). The HCN median value is 15.1 *ppbv* for the cancer group which is significantly higher than the other two groups. The measured median concentrations for positive control and normal upper GI groups were found to be similar, i.e., 6.1 and 7.2 *ppbv*, respectively. The other volatile organic compounds shown in Table 4.1 did not show any significant differences between the groups. Therefore, no further evaluation of these compounds was conducted in this study.

4.1.6 Conclusions

This is the first study of VOC analysis of gastric content from patients with oesophago-gastric cancer, non-cancer diseases of the upper gastro-intestinal tract and those with normal upper gastro-intestinal tracts. A total of 13 VOCs have been analysed in the headspace of gastric content samples. Of these compounds, the headspace concentrations of eight VOCs (acetone, formaldehyde, acetaldehyde, acetic acid, hexanoic acid, hydrogen sulphide, hydrogen cyanide and methyl phenol) were statistically different between the oesophago-gastric cancer group and the healthy cohort. The study highlights that VOC differences exist within the stomach environment between those with cancer and those with normal upper gastro-intestinal tracts. Moreover, gastric content represents a good surrogate biofluid that may improve our understanding of the pathogenesis of upper gastro-intestinal diseases. The clear differences in pH were most interesting given there was no difference in proton pump inhibitor or H₂-receptor antagonist use amongst the 3 groups. The influence of pH on the production of abundant VOCs within urine headspace has previously been investigated by Smith & Španěl¹⁴⁵. Ammonia concentrations significantly increased by making urine alkaline; abnormally high levels of nitric oxide were also observed in the headspace above acidified bacterially infected urine¹⁴⁵. Hence, the influence of pH on the release of VOCs from gastric content is further investigated.

4.2 The effect of pH on VOCs in gastric content headspace

4.2.1 Objective

To evaluate the influence of pH on the VOCs in gastric content headspace, focusing on those VOCs that discriminate patients with oesophago-gastric cancer to those with normal upper GI tracts.

4.2.2 Methods

4.2.2.1 Gastric content sampling and storage

All gastric content samples for this study were taken from patients undergoing an OGD and retrieved using a disposable suction jar (procedure described in section 4.1.3.2). Gastric content samples were stored in a -80°C freezer within 1 hour of collection. Studies have shown there is no decline in the quality of biofluids when stored in this state for several weeks¹⁴⁵.

4.2.2.2 Gastric content sample preparation & SIFT-MS analysis

Gastric content samples were thawed on the day of analysis. A total of 5mL gastric content from the original suction jar was aliquoted into a standard 60 mL specimen jar. The natural pH of a gastric content sample was measured using a Hanna pH meter HI 9124 (Sigma-Aldrich, Missouri, USA). The specimen jar was sealed with parafilm and then placed in an incubator at 37 °C for five minutes prior to SIFT-MS analysis of the headspace. The sterile needle placed within the headspace of the gastric content liquid was connected to the SIFT-MS sampling port via 200cmVygon extension tubing. The headspace vapour was analysed using the same method described in section 1 of this chapter with the specimen jar, needle and extension tubing all being kept within the incubator at 37 °C. All the VOCs in this study - acetaldehyde, formaldehyde, hexanoic acid, acetic acid, acetone, hydrogen sulphide, hydrogen cyanide, methyl phenol - were quantified using the MIM mode and analysed using the same reagent ions, as described in section 4.1.4.3. After completion of MIM analysis of the sample in its natural state, 5mL of the same gastric content sample was taken from its original suction jar for each pH measurement. The sample was aliquoted into a standard 60mL specimen jar and its pH was altered through the addition of appropriate amounts of 0.1M HCl and 0.1M NaOH. Samples were investigated over the expected pH range of gastric content and hence analysis

was conducted at pH 1, 2, 4, 6 & 8; the pH levels were measured to an accuracy of 0.01 using the Hanna pH meter HI 9124 (Sigma-Aldrich, Missouri, USA)

4.2.3 Results

Twenty gastric content samples were analysed at pH 1, 2, 4, 6 & 8. There were 10 samples from patients with oesophago-gastric cancer and 10 samples from non-cancer controls (i.e., those with benign upper GI diseases and normal upper GI tracts). Gastric content samples were retrieved from 8 males & 2 females in OG cancer group and 7 males & 3 females in the non-cancer control group. The median pH in the OG cancer and non-cancer controls was 4.8 (IQR 4.0-7.5) and 1.3 (IQR 1.0-2.0), respectively. Figures 4.7 and 4.8 demonstrate the comparative analysis of the investigated VOCs at pH 1, 2, 4, 6 & 8.

Figure 4.6: Bar graphs of acetaldehyde, formaldehyde, hexanoic acid (HA) and acetic acid (AA) concentrations from the headspace of gastric content from non-cancer controls and patients with oesophago-gastric cancer. Comparative analysis of the groups at pH 1, 2, 4, 6 & 8 is shown.

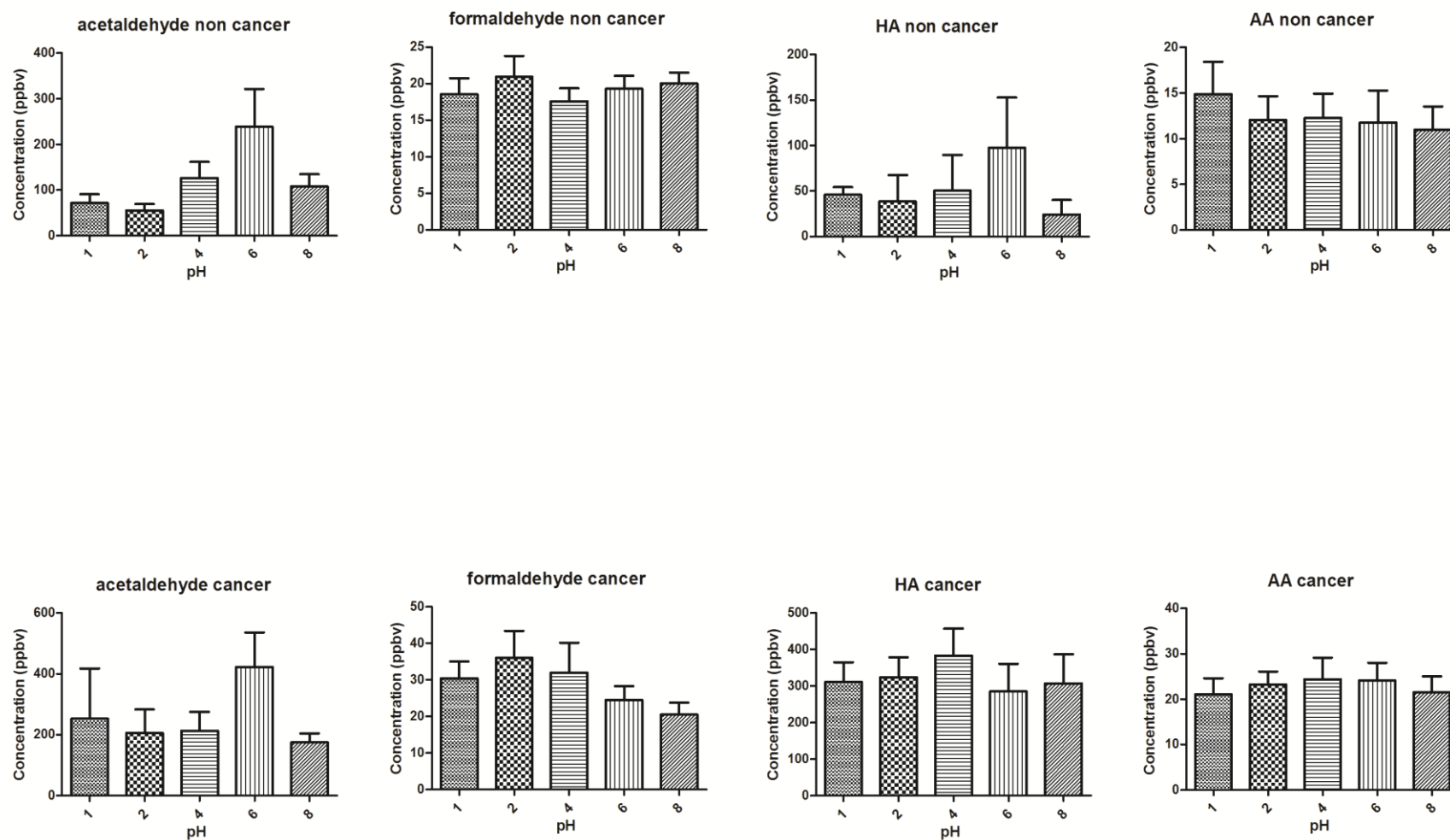
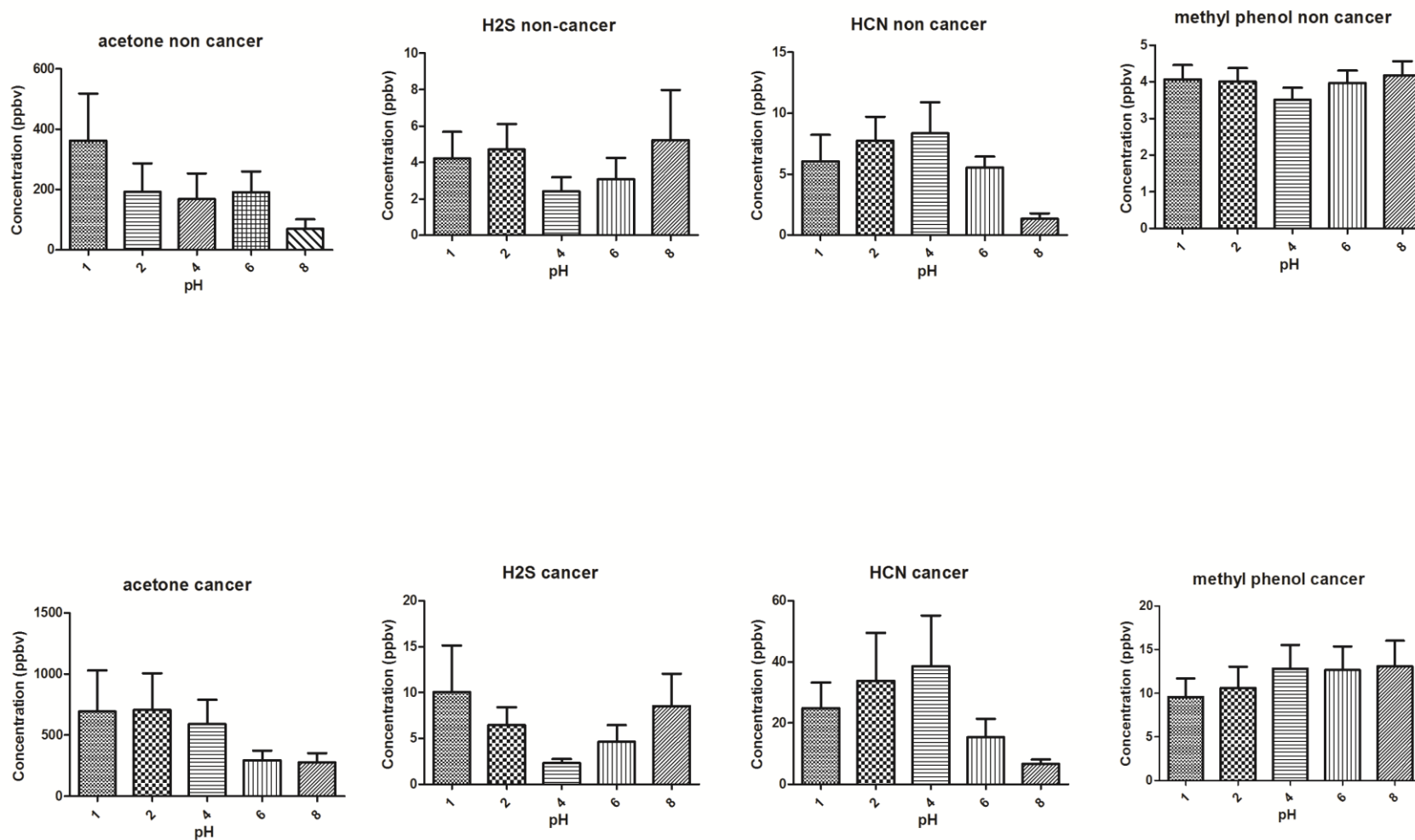


Figure 4.7: Bar graphs of acetone, hydrogen sulphide (H₂S), hydrogen cyanide (HCN) and methyl phenol concentrations from the headspace of gastric content from non-cancer controls and patients with oesophago-gastric cancer. Comparative analysis of the groups at pH 1, 2, 4, 6 & 8 is shown.



4.2.4 Discussion

4.2.4.1 Aldehydes

Acetaldehyde and formaldehyde were significantly increased at all pH levels in the oesophago-gastric cancer group compared to controls (Figure 4.6). The production of acetaldehyde within the headspace vapour of gastric content demonstrated a similar trend in both groups, peaking at pH 6; this may partly explain why increased concentrations of acetaldehyde are observed in the headspace of gastric content of OG cancer and positive controls. The median acetaldehyde concentrations in gastric content headspace were 77 ppbv and 63 ppbv in the OG cancer and positive control groups, respectively. The median pHs in the same study were 6.18 and 5.67 for the aforementioned groups. Jokelainen *et al.* investigated *in vitro* acetaldehyde formation by human colonic bacteria. They reported maximal acetaldehyde production at pH 9.6 but significant concentrations were also observed at pH 7.4 and 6¹⁴⁶. The risk of upper GI malignancy is known to be markedly increased in ALDH2-deficient alcohol users compared to those with the normal enzyme^{93,94}. The reduced ability of ALDH2-deficient persons to process acetaldehyde within the stomach environment may be a contributory factor in the development of these cancers. It has also previously been reported that reducing acetaldehyde exposure may decrease the risk of gastric cancer in patients with chronic gastric acid hyposecretion¹³⁸. Thus the role of acetaldehyde in upper GI disease must be explored further. There was little influence of pH on formaldehyde concentrations within the headspace vapour of gastric content in both OG cancer and non-cancer controls (Figure 4.6).

4.2.4.2 Acids

Hexanoic acid and acetic acid were increased at all pH levels in the oesophago-gastric cancer group compared to non-cancer controls (Figure 4.6). Hexanoic acid concentrations were similar at pHs 1, 2 & 4 in the non-cancer controls with a sudden increase at pH 6; this is followed by a significant decline at pH 8. However, in the cancer cohort, the higher hexanoic acid concentrations within the headspace vapour of gastric content were maintained at pH 8. Acetic acid demonstrated a decreasing trend from pH 1 to 8 in the non-cancer controls (Figure 4.6). In the cancer cohort, acetic acid concentrations were maintained until pH 6 and decreased slightly at pH 8. These phenomena remain difficult to fully explain as one would expect the acid concentrations to decrease at neutral and alkali pH. It is hypothesised that other influential

factors such as *H.pylori* or the cancer cells may be stimulating the production of these acids within gastric content.

4.2.4.3 Acetone & Methyl Phenol

Acetone and methyl phenol concentrations were increased across all pH levels in the OG cancer group compared to non-cancer controls (Figure 4.7). Acetone concentrations were reduced at pH 6 & 8 in both groups. Španěl *et al.* previously reported that pH variation (i.e. from 4 to 8) had no effect on acetone concentrations within the headspace of urine¹⁴⁵. The reduced concentrations of acetone observed at neutral and alkali pH in this study are most likely due to the biochemical differences between gastric content and urine. In healthy individuals, urine is sterile whereas gastric content is a complex biofluid interacting with the gut microbiome. Further studies utilising multiple reagent ions are necessary to verify the influence of pH on acetone within gastric content. There was a negligible effect of pH on the production of methyl phenol in both groups.

4.2.4.4 Sulphur and nitrogen-containing compounds

Hydrogen sulphide and hydrogen cyanide concentrations were increased across all pH levels in the OG cancer group compared to non-cancer controls (Figure 4.7). Hydrogen sulphide demonstrated a varying trend at different pH levels. In both the cancer cohort and non-cancer controls, H₂S demonstrates a decreasing trend from pH 1 towards pH 4, followed by increases at pH 6 & 8 (Figure 4.7). In both the OG cancer group and non-cancer controls, HCN production from the headspace of gastric content peaked at pH 4 followed by significant decreases at pH 6 & 8 (Figure 4.7). The weakly acidic nature of HCN is likely to contribute to this trend.

4.2.5 Conclusions

The results of this study demonstrate a significant influence of pH on the measured headspace levels of VOCs that may be of importance in oesophago-gastric cancer. In particular, the pH changes affecting acetaldehyde production are noteworthy. Acetaldehyde derived from alcohol is a human carcinogen and the observed increased production at pH 6 in OG cancer patients warrants further investigation. Within the stomach environment, there are several variables that may interplay to affect VOC production including gastric content, presence/absence of *H.pylori*, cancer cell metabolism and intra-gastric pH. The higher pH observed in gastric content with patients with benign diseases of the upper GI tract suggests higher intra-gastric pH is likely to be a causative factor preceding cancer development. Future studies should focus on the effects of varying pH on the growth of oesophageal and gastric cancer cell lines; this could provide us with new insights in the oncological proliferation of oesophago-gastric cancer.

Chapter 5

VOCs in the headspace of urine

5.1 Objective

To evaluate whether VOCs within the headspace of urine from patients with oesophago-gastric cancer are present at different concentrations to those with benign upper gastro-intestinal conditions and healthy persons.

5.2 Introduction

In the previous chapter, eight VOCs (viz. acetone, formaldehyde, acetaldehyde, acetic acid, hexanoic acid, hydrogen sulphide, hydrogen cyanide and methyl phenol) within gastric content were identified that were significantly different between cancer patients and a normal upper GI group. The results highlighted potential VOCs that may be important in the molecular orientated diagnosis of oesophago-gastric cancer. However, the retrieval of gastric content as part of a diagnostic test is not straightforward. Gastric content can only be retrieved either during an OGD or via a nasogastric (NG) tube, both of which are invasive procedures. If one is performing an OGD, then a tissue biopsy would be the ‘gold standard’ method for diagnosis confirmation. Insertion of a NG tube also causes patient discomfort and still has a small but non-negligible risk of morbidity¹⁴⁷. The analysis of gastric content remains important to develop our insights into the pathogenesis of upper gastrointestinal malignancy. However, in order to develop a VOC-based diagnostic test for oesophago-gastric cancer, it would be preferable to utilize bio-samples that can be obtained non-invasively.

Urine is a biofluid that can provide good information regarding the physiological or pathophysiological state of a person. Hippocrates (~400 BC), the father of Western Medicine, visually assessed urine to interpret human body functioning and wrote “no other organ system or organ of the human body provides so much information by its excretion as does the urinary system”¹⁴⁸. In ~ 700 AD, Theophilus developed uroscopy (the visual examination of urine) into a diagnostic tool and established systematic methods for its analysis¹⁴⁹. One of the most well-known historical examples of uroscopy relate to King George III (1738-1820), whose attending physicians noted unusual colours in his urine. It is believed that the ‘bilious’ and ‘bluish’ colours observed in the King’s urine and his concurrent episodes of insanity were due to porphyria¹⁵⁰. Since that time, urinalysis has become a quick, cost-effective, routinely employed diagnostic test in clinical medicine. Urinalysis for the assessment of infection, protein, blood and ketone content is well established and urine-based pregnancy testing kits can be purchased over the counter. However, urine vapour is chemically complex with over 280 VOCs present⁴⁶.

Urine, in particular, is considered an ideal biofluid for VOC analysis, because it is obtained by non-invasive sampling and relatively large volumes can be retrieved.

The contemporaneous technological advances in gas phase analytical techniques have allowed scientists to measure VOCs developed in the headspace of urine at concentrations down to parts-per-billion by volume (*ppbv*). Several mass spectrometric techniques have been employed to study metabolites present in urine from patients with specific cancer states. Hu *et al.* developed an experimental model using mice implanted with human gastric cancer SGC-7901 cells to predict gastric cancer metastasis through urinary metabolomic investigations¹⁵¹. They employed GC-MS and demonstrated that 10 metabolites differentiated cancer from controls. Smith & Španěl performed the first study on the analysis of formaldehyde in the headspace of urine samples from patients with bladder cancer and prostate cancer using SIFT-MS⁸³. In this study, they observed elevated levels of formaldehyde in the headspace of urine samples from patients with bladder cancer (mean concentration 85 *ppbv*) and prostate cancer (mean concentration 25 *ppbv*) compared to healthy controls (mean concentration 11 *ppbv*)⁸³. Issaq *et al.* employed liquid chromatography-mass spectrometry (LC-MS) to discriminate between healthy urine specimens and bladder cancer urine samples¹⁵². By employing principal component analysis, PCA, they were able to correctly predict 46 of 48 healthy and 40 of 41 bladder cancer urine samples¹⁵². Cheng *et al.* analysed urine samples from patients with colorectal cancer and healthy controls with time-of-flight mass spectrometry coupled with GC and LC (GC-TOF-MS and LC-TOF-MS)¹⁵³. They observed a distinct urinary metabolic footprint for the colorectal cancer cohort - a panel of metabolite markers including citrate, hippurate, cresol, 2-aminobutyrate, myristate, putrescine, and kynurenate were able to discriminate colorectal cancer patients from controls¹⁵³. Guadagni *et al.* developed a Headspace-SPME/GC-MS method for the analysis of hexanal and heptanal in urine⁶⁹. In this study, higher urinary concentrations of hexanal were found in the lung cancer cohort compared with controls⁶⁹. Hanai *et al.* investigated urinary VOCs in patients with lung cancer using GC-TOF-MS, identifying nine VOCs - tetrahydrofuran, 2-chloroethanol, 2-pentanone, 2-methylpyrazine, cyclohexanone, 2-ethyl-1-hexanol, 2-phenyl-2-propanol, isohorone and 2,6-diisopropylphenol - to be increased in the lung cancer cohort compared with controls¹⁵⁴. They also reported that 2-pentanone may have utility in differentiating between adenocarcinoma and squamous cell carcinomas of the lung¹⁵⁴. The above animal and human studies have all demonstrated promising early results through urine profiling in cancer disease states.

Both GC-MS and LC-MS are offline mass spectrometric techniques requiring appropriate column selection^{155,156}. It is also necessary to optimize storage of urine as well as

undertake pre-analysis preparation methods. Therefore, analytical processing times are lengthy and VOCs with low molecular weight can be difficult to accurately quantify with GC-MS or LC-MS¹⁵⁷. In comparison to these techniques, SIFT-MS has the advantage of online, real-time VOC quantification without the need for any pre-concentration steps. The analysis of VOCs emitted from the headspace of urine using SIFT-MS has been reported in several studies^{83,158,159}.

As mentioned at the beginning of this Introduction, several VOCs that were significantly different between cancer and healthy controls were identified in the headspace vapour of gastric content. For this study, urine was selected as the biofluid of choice for investigation. It was hypothesized that the VOCs of the urine retrieved from patients with oesophago-gastric cancer would differ from non-cancer controls. It was also postulated that several of the VOCs identified in the headspace of the gastric content of oesophago-gastric cancer patients may be present in the headspace of the urine of cancer patients, reflecting a systemic signature of the disease. This study also constitutes the first investigation of VOCs using the SIFT-MS platform for the headspace analysis of urine of oesophago-gastric cancer patients.

5.3 Methods

5.3.1 Patient selection

Patients for this study were recruited through St. Mary's Hospital, Paddington (Imperial College NHS Trust, London) and via the North West Thames regional oesophago-gastric cancer network (UK). As before, all patients provided a comprehensive medical history, including past medical history, medications as well as smoking status and alcohol intake. Demographic and clinical data for each patient were archived in a linked, anonymised database. The same exclusion criteria described in section 4.1.3.1 was exercised for this study. Patients with a current or past medical history of significant kidney/bladder disease were also excluded. The cancer and positive control groups were recruited following the same criteria in section 4.1.3.1. The 'healthy' cohort in this study had no upper gastro-intestinal symptoms and no past medical history. Local ethics committee approval was obtained for this study. Patients eligible for the study were given information sheets pertaining to the study and written informed consent was obtained prior to enrolment in the study. The comparative analysis of urine vapour metabolites was conducted with an oesophago-gastric cancer cohort of 17 patients, a positive control group of 14 patients and a 'healthy' group of 13 patients.

5.3.2 Urine sampling procedures

All patients and healthy participants had fasted for a minimum of 6 hours with no oral intake prior to urine sampling. Urine samples were all retrieved from patients prior to undergoing their scheduled gastroscopies. Study participants were requested to pass urine into a standardized 60 mL urine specimen vial, which was immediately sealed. All samples were taken to the Surgical Sciences laboratory located within the hospital and analysed within 1 hour of collection. The pH of each urine sample was measured immediately after SIFT-MS analysis using a Hanna Piccolo[®] pH meter (Sigma-Aldrich, Missouri, USA). All urine samples also underwent urine dipstick analysis (Siemens Multistix 8G) to assess for any infection. Any samples with dipstick-proven evidence of infection were excluded from final analysis.

5.3.3 SIFT-MS analysis of urine headspace

In this study, the full-scan mode (FS) of the SIFT-MS was initially employed to identify VOCs of interest in the headspace vapour of the urine samples. Using H_3O^+ and NO^+ precursor ions individually, the downstream analytical mass spectrometer system was scanned from m/z 10 to 200 for 9 x 60 seconds period. H_3O^+ and NO^+ full scan spectra were produced that revealed the characteristic product ions, which allowed the identification and quantification of specific VOCs using the standardized kinetics library⁵³. The O_2^+ precursor ion is suitable for the detection and quantification of specific molecules (e.g. NH_3). However, it was not used for FS analysis, as the majority of VOCs investigated in this study were best analysed by H_3O^+ and NO^+ . The O_2^+ reagent ion was only used in MIM mode for the detection and quantification of one VOC - methyl phenol.

For the accurate quantification of selected VOCs of interest, multi-ion monitoring mode (MIM) was primarily employed in this study. Prior to MIM headspace analysis of a urine sample, the calibration procedure using the water content of the operator's exhaled breath (6 +/- 0.1%) was carried out to check the sample flow rate settings of the instrument. The absolute concentrations of the volatile compounds present in the headspace of the urine samples were obtained over a chosen period of analysis using the in-built SIFT-MS software that utilizes the known rate coefficients of the reactions between reagent ions and selected volatile organic compounds contained in the in-built kinetics library⁵³. In the MIM mode, selected VOCs from the urine headspace were analysed for a total of 60 seconds and the measured concentrations were averaged over the analysis time for each VOC. For each measurement, a total of 10 mL urine was aliquoted into a standard 60 mL specimen vial. The specimen vial was sealed with parafilm and then placed in an incubator at 37°C for five minutes prior to analysis. As before, a 21g Luer-lock sterile hypodermic needle was connected to a 200 cm Vygon extension tube that was directly connected to the inlet port of the SIFT-MS instrument. The sterile needle punctured the parafilm layer of the specimen vial to sample the headspace above the liquid. The headspace vapour automatically flowed into the helium carrier gas at a fixed rate through the sampling line that was held at a temperature of 80°C. For the duration of the analysis, the specimen vial, needle and extension tubing were all kept within the incubator at 37°C. Measurements of VOCs within the background laboratory room air were also recorded.

5.3.4 Statistical analysis

Statistical analysis was performed using IBM SPSS statistics 20 (SPSS Inc., Chicago, IL). Kruskal-Wallis one-way analysis of variance was used to compare the measured concentrations of VOCs between cancer patients, positive controls and healthy controls. The Mann-Whitney U test was used to compare the measured concentrations of acetone between the cancer cohort and positive controls. Receiver operating characteristic (ROC) curves were used to determine the accuracy of a diagnostic test in classifying subjects into those with and without disease. The ROC curves were constructed using the same method described in section 4.1.3.4. In this study, the statistically significant VOCs identified by the Kruskal-Wallis test, which showed a difference between cancer and the non-cancer controls (including both positive and healthy controls), were included as variables for the ROC curve. To construct the ROC curve, the disease conditions were used as the dependent variable, and the sum concentration of selected VOCs was used as the independent variable. To select the VOCs, different combinations of the statistically significant ones were tested using a binary logistic regression analysis to obtain the discrimination model with the highest R^2 and the AUC value for the ROC curve. A p -value ≤ 0.05 was taken as the level to indicate statistical significance, in keeping with most clinical studies.

5.4 Results

5.4.1 Analysis of full scan spectra using H_3O^+ and NO^+

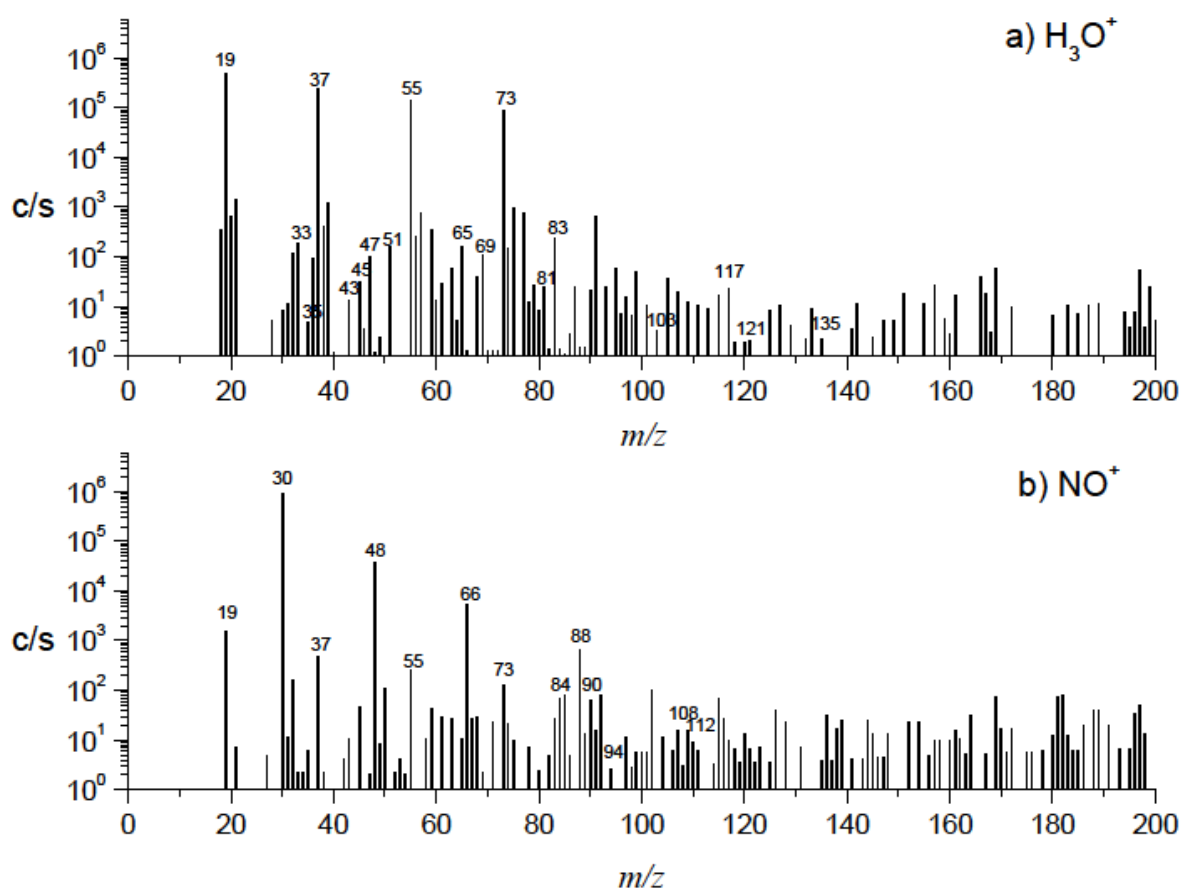


Figure 5.1: Full scan spectra of the headspace of urine using (a) H_3O^+ precursor ion (b) NO^+ precursor ion. The individual characteristic product ions and their hydrates were assigned to the analyzed molecules. In (a), m/z 19, 37, 55, 73 are $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{0,1,2,3}$; m/z 33, 51, 69, methanol; m/z 47, 65, 83, ethanol; m/z 43, propanol; m/z 45, 81, acetaldehyde; m/z 35, hydrogen sulphide; m/z 117, 135, hexanoic acid. In (b), m/z 19, 37, 55, 73 are $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{0,1,2,3}$; m/z 30, 48, 66, 84 are $\text{NO}^+(\text{H}_2\text{O})_{0,1,2,3}$; m/z 88 is acetone; m/z 90, 108 is acetic acid; m/z 94, 112 is phenol.

Full scan (FS) spectra were obtained when the headspace vapour of urine was introduced into the helium carrier gas. Several VOCs of interest were clearly identified. As shown in Figure 5.1(a), the representative spectrum using H_3O^+ clearly shows the presence of the precursor ion H_3O^+ and its water hydrates (m/z 19, 37, 55 and 73), due to the humidity of the urine sample (typically 6% by volume of water vapour). Important VOCs, such as methanol, ethanol, propanol, acetaldehyde and hydrogen sulphide were unambiguously identified by their characteristic product ions. Similarly, Figure 5.1(b) shows the presence of NO^+ precursor ions and its associated hydrates (m/z 30, 48, 66 and 84); acetone, acetic acid and phenol, which are detected using NO^+ ion, are also clearly shown. Furthermore, full scan spectra were also used to identify those VOCs showing overlapped product ions and hence these were excluded from further MIM mode analysis.

5.4.2 Reproducibility and validity of the sampling method

Additional experiments were conducted to assess the inter-sample reproducibility of headspace analysis using the exact methods described above. Ten healthy volunteers (with no medical history) were recruited and requested to provide 2 urine samples each. A total of twenty urine samples were analysed using the MIM mode for selected VOCs. The inter-sample coefficient of variation (CV) was calculated for the measured concentration of several VOCs from the headspace vapour of the urine. These CV values obtained in percent are: acetaldehyde, 8.7%; acetone, 22.6%; acetic acid, 15.8%; hexanoic acid, 27.8%; methanol, 44.5%; ethanol, 9.9%; phenol, 19.3%; methyl phenol, 41.2%; ethyl phenol, 26.0%; hydrogen sulphide, 38.3%. These coefficients of variation demonstrate the reproducibility of the MIM measurement.

5.4.3 MIM analysis of selected VOCs in the headspace of urine

Compounds	Molecular formula	Precursor ions	<i>m/z</i>	Characteristic product ions	REF
Methanol	CH ₄ O	H ₃ O ⁺	33, 51, 69	CH ₃ O ⁺ , CH ₃ O ⁺ (H ₂ O), CH ₃ O ⁺ (H ₂ O) ₂	129
Ethanol	C ₂ H ₆ O	H ₃ O ⁺	47, 65, 83	C ₂ H ₇ O ⁺ , C ₂ H ₇ O ⁺ (H ₂ O), C ₂ H ₇ O ⁺ (H ₂ O) ₂	129
Propanol	C ₃ H ₈ O	H ₃ O ⁺	43	C ₃ H ₇ ⁺	129
Hydrogen sulphide	H ₂ S	H ₃ O ⁺	35, 53	H ₃ S ⁺ , H ₃ S ⁺ (H ₂ O)	129
Hexanoic acid	C ₆ H ₁₂ O ₂	H ₃ O ⁺	117, 135	C ₆ H ₁₂ O ₂ H ⁺ , C ₆ H ₁₂ O ₂ H ⁺ (H ₂ O)	129
Acetaldehyde	C ₂ H ₄ O	H ₃ O ⁺	45, 81	CH ₃ CHOH ⁺ , CH ₃ CHOH ⁺ (H ₂ O) ₂	134
Acetone	C ₃ H ₆ O	NO ⁺	88	NO ⁺ ·C ₃ H ₆ O	134
Acetic acid	C ₂ H ₄ O ₂	NO ⁺	90, 108	NO ⁺ ·CH ₃ COOH, NO ⁺ ·H ₂ OCH ₃ COOH	129
Phenol	C ₆ H ₆ O	NO ⁺	94, 112	C ₆ H ₆ O ⁺ , C ₆ H ₆ O ⁺ (H ₂ O)	135
Ethyl phenol	C ₈ H ₁₀ O	NO ⁺	122, 140	C ₈ H ₁₀ O ⁺ , C ₈ H ₁₀ O ⁺ (H ₂ O)	135
Methyl phenol	C ₇ H ₈ O	O ₂ ⁺	108, 126	C ₇ H ₈ O ⁺ , C ₇ H ₈ O ⁺ (H ₂ O)	135

Table 5.1: A summary of analytical information of 11 volatile organic compounds detected and quantified by the MIM mode of SIFT-MS using H₃O⁺, NO⁺ and O₂⁺ precursor ions.

As shown in Table 5.1, a total of 11 volatile organic compounds released from the headspace of urine were analysed using the MIM mode of the SIFT-MS. The analytical information, including chemical formula, precursor ions, *m/z* values and characteristic product ions are given. H₃O⁺ is the most frequently used precursor ion as it reacts with more compounds than the other two precursor ions. However, NO⁺ has also been employed for specific molecules when it has an advantage in unambiguously detecting and quantifying these compounds over H₃O⁺ (i.e., when there are known overlaps with other product ions).

	Cancer cohort		Positive controls		Healthy controls		Room air		Kruskal-Wallis test
VOCs	Median	IQR	Median	IQR	Median	IQR	Median	IQR	p-value
Acetaldehyde	140	[68-210]	92	[69-140]	58	[49-59]	7	[3-11]	p < 0.001
Acetone	1326	[646-3281]	882	[552-1584]	265	[206-357]	6	[3-10]	p < 0.001
Methanol	216	[172-277]	210	[159-276]	70	[56-102]	38	[30-50]	p < 0.001
Phenol	12	[9-29]	14	[11-35]	34	[28-50]	2	[1-3]	p = 0.019
Acetic acid	40	[27-90]	59	[46-103]	25	[20-36]	5	[3-5]	p = 0.004
Hexanoic acid	66	[27-263]	20	[11-41]	18	[11-32]	1	[0-3]	p = 0.006
Hydrogen sulphide	11	[6-26]	8	[4-13]	3	[1-4]	0	0	p < 0.001

Table 5.2: A summary of median concentrations and interquartile ranges (in *ppbv*) for the 7 VOCs that were significantly different between oesophago-gastric cancer and non-cancer controls. Results of Kruskal-Wallis one-way analysis of variance for the VOCs that are different between cancer, positive control and healthy samples are included. There are 17 patients in the cancer cohort, 14 patients in the positive control group and 13 patients in the ‘healthy’ control group. Laboratory room air concentrations for each of the VOCs are also included.

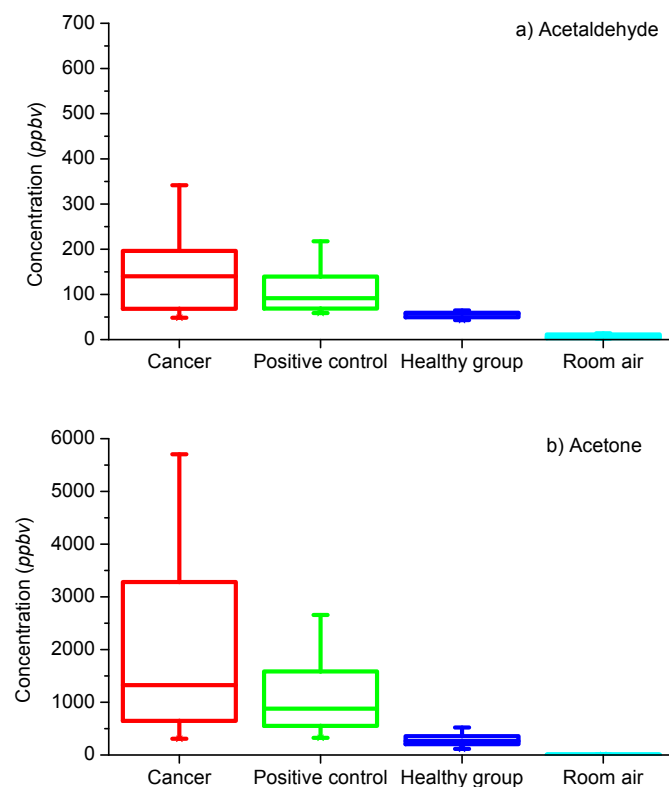


Figure 5.2: Box-Whisker plots of the median concentrations and interquartile ranges (part-per-billion by-volume, *ppbv*) of (a) acetaldehyde, (b) acetone measurement of the headspace vapour of urine of oesophago-gastric cancer patients, positive controls and healthy groups. Laboratory room air concentrations for each of the volatile organic compounds are also included for comparison.

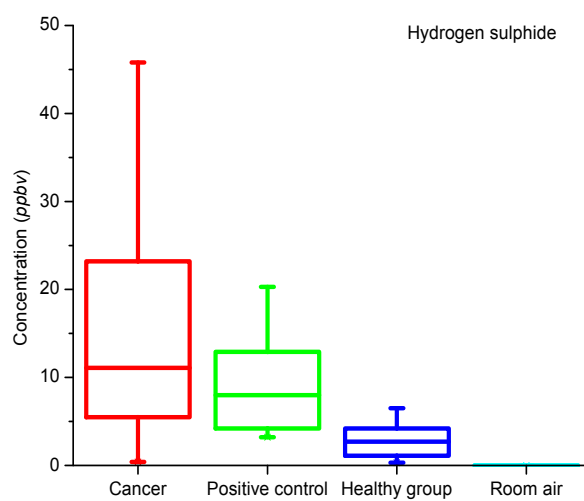


Figure 5.3: Box-Whisker plots of the median concentrations and interquartile ranges (part-per-billion by-volume, *ppbv*) of hydrogen sulphide measurement of the headspace vapour of urine of oesophago-gastric cancer patients, positive controls and healthy groups.

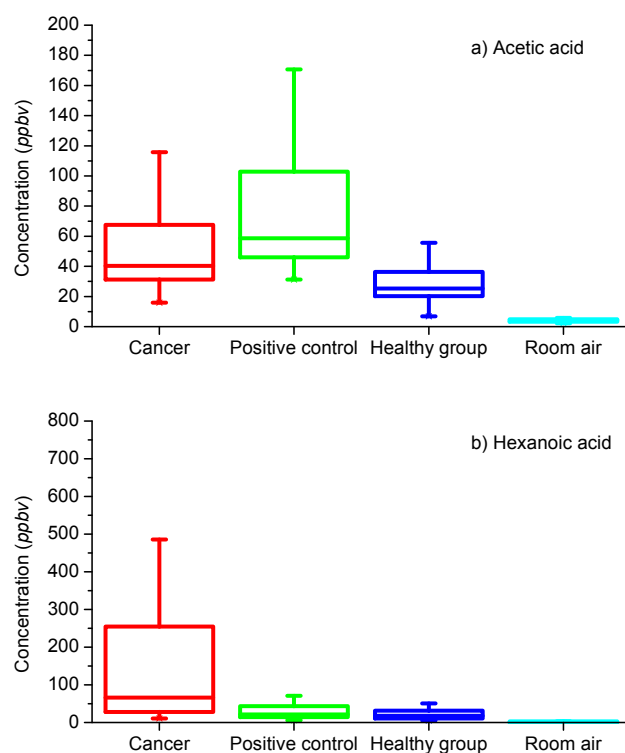


Figure 5.4: Box-Whisker plots of the median concentrations and interquartile ranges (part-per-billion by-volume, *ppbv*) of (a) acetic acid (b) hexanoic acid measurement of the headspace vapour of urine of OG cancer patients, positive controls and healthy groups. Laboratory room air values for each of the volatile organic compounds are also included for comparison.

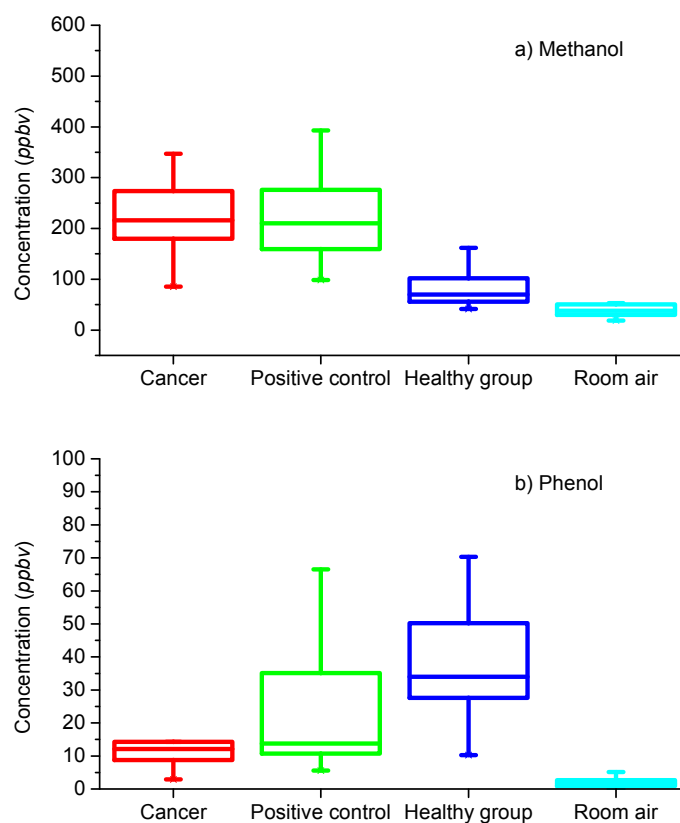


Figure 5.5: Box-Whisker plots of the median concentrations and interquartile ranges (part-per-billion by-volume, *ppbv*) of (a) methanol, (b) phenol, measurement from the headspace vapour of urine of OG cancer, positive control and healthy groups. Laboratory room air concentrations for each of the VOCs are also included for comparison.

5.4.4 pH measurements of urine

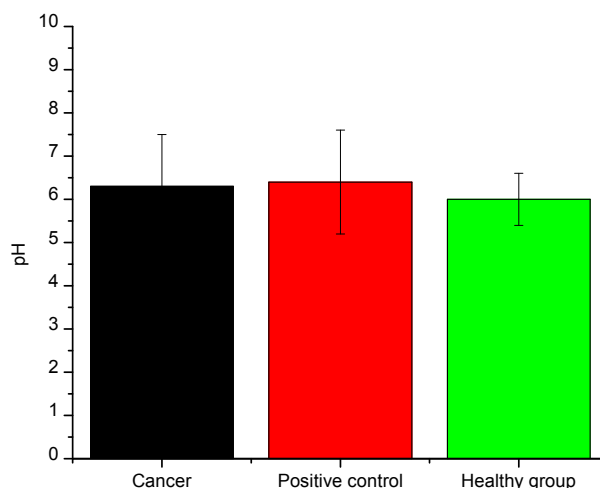


Figure 5.6: pH measurement of urine of oesophago-gastric cancer patients, positive controls and the healthy group.

The normal pH of the urine is approximately 6.7. Španěl & Smith have reported that pH variations (from 4 to 8) do not influence the measured concentrations of several VOCs (including acetone, ethanol and methanol), which are released from urine to the headspace vapour¹⁴⁵. As shown in Figure 5.6, the urine from the healthy group has a pH value of 6.0 ± 0.6 . The mean pH of urine from the positive group is 6.4 ± 1.2 and the mean value for the cancer group is 6.3 ± 1.2 . The cancer cohort demonstrated the widest pH range amongst the three groups. However, there was no significant difference in pH across the three groups and therefore no further investigation of the influence of pH on VOC concentrations within urine was undertaken.

5.4.5 VOCs within urine headspace to differentiate OG cancer from non-cancer controls

Kruskal-Wallis one-way analysis of variance (p -value ≤ 0.05) was performed for all the volatile organic compounds listed in Table 5.1, to assess any differences in median concentrations between the three groups. Seven metabolites were found to be statistically significant different across the three groups, including acetaldehyde ($p < 0.001$), acetone ($p < 0.001$), acetic acid ($p = 0.004$), hexanoic acid ($p = 0.006$), hydrogen sulphide ($p < 0.001$), methanol ($p < 0.001$) and phenol ($p = 0.019$), five of these compounds also being significant in the headspace analysis of gastric content. Their median concentrations and interquartile ranges from cancer, healthy and positive control groups are given in Table 5.2.

Using the six statistically significant VOCs, which demonstrated an increased concentration in the cancer cohort, a binary logistical regression analysis was performed to construct a ROC model. The sum concentrations of acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulphide and methanol demonstrated the highest area under the curve, discriminating cancer from both control groups with a value of 0.904. Figure 5.7 below demonstrates a high degree of discrimination between oesophago-gastric cancer and non-cancer controls (both positive and healthy control groups) with a good sensitivity and specificity using the aforementioned compounds.

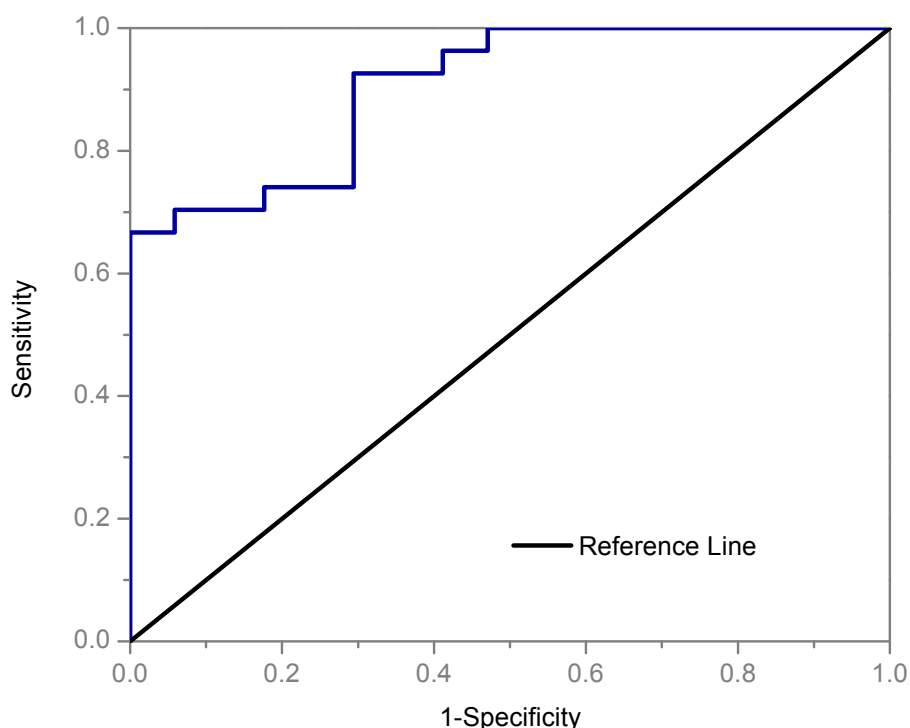


Figure 5.7: Receiver operating characteristics (ROC) curve of discrimination result for the diagnosis of oesophago-gastric cancer from non-cancer controls using the sum concentrations of acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulphide and methanol exhibited, respectively. The AUC is 0.904.

Of note, is that both acetaldehyde and hydrogen sulphide were previously identified as VOCs in the ROC curve constructed in the study on headspace analysis of gastric content. Given these VOCs have been influential in separate ROC curves from 2 different biofluids in oesophago-gastric cancer, further studies are necessary to establish whether certain VOCs are systemically raised in cancer states. Moreover, this result provides further evidence that a VOC profile may be most appropriate for molecular orientated diagnosis in OG Cancer; it could also have a potential therapeutic monitoring role in the surveillance of disease recurrence.

5.5 Discussion

5.5.1 Aldehydes and Ketones

Two short-chain carbonyl compounds, i.e., acetaldehyde, and acetone, were detected in the headspace of urine. There is a significant decreasing trend in the median concentrations of acetaldehyde when comparing cancer, positive control and healthy cohorts, which is the same trend that was observed in the analysis of the headspace of gastric content (Figure 5.2). The median acetaldehyde concentration of the cancer is 140 *ppbv*, and positive control patients have a median value of 92 *ppbv*. In comparison, the median value from the healthy group is 58 *ppbv*, which is much lower than that of the cancer group. Furthermore, the majority of acetaldehyde values in the healthy cohort are less than 60 *ppbv*, while in the cancer and positive control groups, the majority of observed values are above 60 *ppbv*. As shown in Table 5.2, in sampling our laboratory air, we observed a median acetaldehyde concentration of 7 *ppbv*; this value highlights the negligible contribution from the surrounding environment. There is significant evidence linking acetaldehyde and carcinogenesis. Fanelli *et al.* examined the effects of acetaldehyde accumulation in rat mammary tissue and concluded that this may be a critical event in the development of breast cancer⁹⁶. Väkeväinen *et al.* reported that microbial ethanol metabolism in achlorhydric atrophic gastritis after alcohol ingestion leads to high intra-gastric acetaldehyde levels¹⁶⁰. The group postulates that the higher levels of intra-gastric acetaldehyde may be one of the factors contributing to the increased risk of gastric cancer in atrophic gastritis patients. The evidence from both headspace studies for acetaldehyde within biofluids in patients with oesophago-gastric cancer as well as significant animal model and epidemiological evidence indicate that the potential role of acetaldehyde should be further evaluated.

Acetone is generally observed as the most abundant ketone body in urine (the other two being beta-hydroxybutyric acid and acetoacetic acid)¹⁶¹. As shown in Figure 5.2, the acetone values in the cancer group also have a much higher interquartile range of 646 to 3281 *ppbv* in comparison to the healthy cohort, 206 to 357 *ppbv*. There is a clear decreasing trend of acetone concentration in the headspace of urine in the cancer, positive control and healthy groups which is the same trend observed in the previous study on gastric content. In contrast to the cancer group, although the positive control group has a smaller median concentration of acetone (882 *ppbv* vs 1326 *ppbv*), the interquartile range of acetone from positive control (552 to 1584 *ppbv*) falls within the range of the oesophago-gastric cancer group (646 to 3281 *ppbv*). A Mann-Whitney U test was performed to compare the acetone concentrations between the cancer and

positive group and the difference was not found to be significant ($p = 0.250$). Several studies have also reported increased acetone concentrations in urine, blood and breath in various cancer states. Shariff *et al.* characterized urinary biomarkers in hepatocellular carcinoma (HCC) in patients of Nigerian descent using Nuclear Magnetic Resonance (NMR) spectroscopy¹⁶². In this study, acetone was identified as one of the most influential metabolites in multivariate models that distinguished HCC from patients with cirrhosis and healthy individuals¹⁶². Hasim *et al.* reported markedly increased blood plasma acetone concentrations in patients with poorly differentiated oesophageal cancer compared with patients with well differentiated oesophageal cancer¹⁶³. Zhang *et al.* employed NMR spectroscopy to distinguish pancreatic cancer from chronic pancreatitis and healthy individuals¹⁶⁴. Their group reported elevated levels of acetone in blood plasma of patients with pancreatic cancer compared with controls¹⁶⁴. Ulanowska *et al.* observed an increased concentration of acetone in the exhaled breath of patients with lung cancer compared with controls (including active smoker, passive smoker and non-smokers)¹⁶⁵. The observed increase in acetone in this study may be related to the patient's underlying cancer status or may reflect a changed physiological state of the patient in response to disease. Furthermore, patients with oesophago-gastric cancer can often be malnourished due to dysphagia or the systemic effects of cancer, resulting in a dependence on ketogenic metabolic processes.

5.5.2 Sulphur compounds

Hydrogen sulphide (H_2S) is produced by anaerobic bacteria in the colon or can be synthesized endogenously from L-cysteine^{166,167}. Ramasamy *et al.* reported that increased levels of hydrogen sulphide in the human colon may be a contributory factor in the development of colorectal cancer¹⁶⁸. Chávez-Piña *et al.* studied the effects of endogenous hydrogen sulphide in a rat model¹⁶⁹. They demonstrated that inhibition of endogenous hydrogen sulphide protects against ethanol-induced gastric damage¹⁶⁹. Yamagishi *et al.* reported the concentrations of sulphur-containing compounds (including hydrogen sulphide) in the flatus of colorectal cancer patients were significantly higher than in healthy individuals¹⁷⁰. Their study also demonstrated that gaseous sulphur-containing compounds are generated in tumour tissues¹⁷⁰. H_2S is a highly volatile compound and has a low molecular weight; GC-MS is often not capable of detecting H_2S . SIFT-MS has an advantage in analysing these types of compounds in comparison to GC-MS techniques. The measured H_2S concentrations of the cancer and positive control groups showed similar values, 11 and 8 *ppbv* respectively. However, only trace quantities of H_2S were

detected in the headspace of the urine from healthy control group, i.e., 3 *ppbv*. There was a large variation in the measured H₂S levels between oesophago-gastric cancer, positive control and healthy groups, with those from the cancer cohort having the widest interquartile range (Figure 5.3). The decreasing trend of H₂S in the headspace of urine from the cancer cohort to the healthy control group was also observed in the study on the headspace of gastric content. It is postulated that the increased concentrations of hydrogen sulphide observed in both biofluids of the cancer patients may reflect H₂S production by cancer cells (as observed by Yamagishi *et al.*¹⁷⁰) or may be due to alterations in gut microbiota in response to cancer.

5.5.3 Acids

Acetic acid and hexanoic acid were detected in the headspace of the urine samples. Acetic acid can be quantified using both H₃O⁺ and NO⁺ using SIFT-MS. However, the true concentration of acetic acid can be lost among the interferences from the possible isobaric compounds such as methyl formate and propanol which also have a molecular weight of 60. This problem is increased when humid exhaled breath is analysed as these compounds can form hydrates of characteristic product analytical ions¹¹³. Pysanenko *et al.* studied the ion chemistry of acetic acid using SIFT-MS; they reported that the most reliable identification and quantification of acetic acid from methyl formate within a vapour is achieved using NO⁺ precursor ion¹⁵⁸. Hence, the NO⁺ precursor ion was selected to analyse acetic acid in this study. As shown in Figure 5.4, acetic acid in the headspace of urine from the cancer cohort and positive controls has a higher median value and wider interquartile range, in contrast to the healthy group. The median concentration in the cancer group was ~ 40% higher than that in the healthy group, i.e., 40 and 25 *ppbv*, a virtually identical trend to that observed in the headspace analysis of gastric content. Hasim *et al.* also reported elevated levels of acetate in the NMR profile of urine in patients with oesophageal cancer compared with healthy controls¹⁶³. Urinary 5-hydroxyindole acetic acid (5-HIAA) is a metabolite that is routinely measured in the diagnosis and monitoring of carcinoid tumours¹⁷¹. However, it is evident that the increased acetic acid concentrations within urine are not oesophago-gastric cancer-specific, given the similar median values observed in both the cancer and positive control groups.

Hexanoic acid is found to be significantly different between cancer and the other two control groups (Figure 5.4). Both positive and healthy control groups have similar medians and interquartile ranges. The median concentration of hexanoic acid in the cancer group is 66 *ppbv*, whilst the positive and healthy controls have median concentrations of 20 and 18 *ppbv*,

respectively. As shown in Figure 5.4, compared to the positive and healthy controls, a much wider distribution of hexanoic acid values is observed in the cancer cohort (interquartile range 27-263 *ppbv*). This result again displays a similar trend to the hexanoic acid concentrations reported in the headspace of gastric content from patients with oesophago-gastric cancer. Silva *et al.* have reported lower levels of hexanoic acid in the urine of patients with colorectal cancer, leukaemia and lymphoma compared to normal controls⁷⁰; however these samples were analysed with GC-MS and there were no patients with upper gastrointestinal malignancy included in this study⁷⁰. Hexanoic acid is increased within the headspace of gastric content and urine of the cancer patients in these studies and hence this VOC may be specific to oesophago-gastric cancer.

5.5.4 Alcohols and Phenol derivatives

Methanol, ethanol, propanol, phenol, methyl phenol and ethyl phenol were investigated in the headspace of the urine samples. Of these compounds, only methanol and phenol demonstrated a statistically significant difference across the three groups. The methanol concentration in the headspace of urine samples of the cancer cohort and positive control groups demonstrated a similar median concentration (216 *ppbv* and 210 *ppbv* respectively). The measured median values for methanol demonstrate a significant decreasing trend from the cancer cohort to the healthy group, as shown in Figure 5.5. A similar decreasing trend for methanol was also observed in the headspace analysis of gastric content in oesophago-gastric cancer patients, although this result was not statistically significant. In comparison to the other significant VOCs, a higher background laboratory room air value for methanol was observed. However, as shown in Table 5.2, this value is much lower than the median concentrations measured in both cancer and positive control groups. Garner *et al.* studied VOCs emitted from stool samples collected from asymptomatic donors and patients with infective diarrhoea and ulcerative colitis¹⁷². They investigated the headspace of the faeces samples using GC-MS and observed that methanol was present more frequently in patients with *Campylobacter jejuni* infection and ulcerative colitis compared to the asymptomatic donors¹⁷². Caldwell studied the effects of methanol on the growth of representative, anaerobic gut bacteria¹⁷³. He demonstrated that the growth of many organisms was completely inhibited by higher concentrations of methanol. Many of these bacteria have a protective role in the human gut and the higher concentrations of methanol may have a role in the pathogenesis of gastro-intestinal diseases.

In this study, phenol was measured with the NO^+ precursor ion to avoid $\text{acetoneH}^+(\text{H}_2\text{O})_2$ at m/z 95 that occurs with the H_3O^+ precursor ion. The measured phenol

concentrations in cancer and positive controls were similar (12 and 14 *ppbv*), while the healthy group has a higher phenol concentration in comparison (34 *ppbv*), as shown in Figure 5.5. In the previous study on the headspace analysis of gastric content, there was a significant difference in the concentration of methyl phenol between the cancer cohort and controls. Phenol is also the most frequently observed VOC in the gastric content of patients with oesophago-gastric cancer (Appendix B). Ling *et al.* have reported that urinary phenols can be up to 10-times higher than the concentrations of phenol found in the serum of the same individuals¹⁷⁴. McDonald *et al.* hypothesized that the higher concentrations of phenol observed in the urine reflect efficient removal of phenol from the blood and subsequent excretion¹⁷⁵. It is postulated that the lower concentrations and tight interquartile range of urinary phenols observed in the cancer group in this study may be due to impaired phenol excretion. Dysfunction of phenol or its derivatives may be due to alterations in tyrosine metabolism by gut bacteria or impairment in the mechanisms that remove these compounds from the blood.

The other volatile organic compounds shown in Table 5.1 did not show any significant differences between the groups. Hence, no further consideration was given to these compounds in this study. Ten measurements of laboratory room air concentrations of acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulphide, methanol and phenol were also recorded. As shown in the figures above, the room air concentrations of all the VOCs are trace in comparison to the measured concentrations from the headspace of patients' urine samples. Thus, the potential exogenous contribution of background laboratory room air to the measured concentrations of the VOCs in the headspace of the urine samples is deemed negligible.

5.6 Conclusions

This is the first study of headspace analysis of urine from patients with oesophago-gastric cancer and non-cancer controls using the SIFT-MS platform. Real-time analysis of urine headspace without the need for pre-concentration treatment was conducted using this platform; this is particularly advantageous in a clinical environment setting whereby samples can be readily retrieved and processed within an hour. A total of 11 VOCs were analysed in the headspace of urine obtained from the above cohorts. Of these compounds, the concentrations of 7 VOCs (acetaldehyde, acetone, methanol, hexanoic acid, acetic acid, hydrogen sulphide and phenol) were statistically different between the oesophago-gastric cancer group and non-cancer controls. Several VOCs which differentiated oesophago-gastric cancer from healthy controls from the headspace analysis of gastric content have also been identified. The increased concentrations of specific VOCs observed in two separate biofluids in patients with oesophago-gastric cancer are noteworthy and may reflect the systemic effects of the disease. OG Cancer is known to have poor 5-year cancer survival outcomes^{18,19}. One of the preferred approaches for improving oncological survival lies in the earlier diagnosis of these cancers through targeted screening. Urine is an ideal biofluid that would be suitable for development as a non-invasive diagnostic test in oesophago-gastric cancer. It can be retrieved without causing any patient discomfort and relatively large volumes are readily available. The rationale for developing a VOC-based urine screening test would be to identify patients at highest risk of OG cancer, who could subsequently be sent for further investigations.

Chapter 6

VOCs and Exhaled Breath

6.1 Exhaled Breath VOCs in OG Cancer

The analysis of exhaled breath through the identification and accurate quantification of volatile organic compounds (VOCs) is an evolving scientific field. It has great potential for both disease diagnosis and measuring physiological response to treatment. The major appeal of exhaled breath analysis for both the clinicians and patients is that the technique is both non-invasive and painless. The easy availability of exhaled breath as a bio-sample is also a plus point; it means that sampling and testing could eventually be conducted in a physician's consultation room. The development of clinical breath tests for certain diseases could also potentially reduce the number of invasive procedures undertaken and decrease hospital visits. As well as the potential financial savings in a nationalised healthcare system, these indirect benefits are particularly advantageous in certain patients groups, including young children and the elderly. However, there remain limited medical applications of human breath analysis at present. Quantification of blood alcohol levels is the most widely employed breath test. The ^{13}C urea breath testing for *H. pylori* bacterium, hydrogen breath test for small bowel bacterial overgrowth and the measurement of exhaled nitric oxide for airways inflammation in asthma represent the only mainstream medical breath tests^{48,49,50}. However, there remains much scientific interest and research effort in developing exhaled breath analysis of VOCs for cancer.

Human breath has been demonstrated to be a complex gaseous mixture containing hundreds of volatile organic compounds⁴⁶. The systematic review (Section 2.4) highlighted the potential of exhaled breath analysis of VOCs in several cancer states, including lung, prostate, breast and colon cancers^{68,72}. These studies have employed various chemical analytical techniques and have demonstrated some very promising VOC-disease associations. To date, GC-MS has been the most commonly employed analytical instrument for the detection of VOCs from exhaled breath. However, it remains an analytical technique which requires trace gas adsorption devices, column calibration for desired analytes as well as lengthy sample processing times. GC-MS is also less preferable to quantify important trace compounds that are present in exhaled breath. Bajtarevic *et al.* investigated VOCs within the exhaled breath of lung cancer patients and reported SPME GC-MS to be a relatively insensitive method for analysis of breath metabolites¹⁷⁶. In recent years, mass spectrometry techniques such as proton transfer reaction mass spectrometry (PTR-MS) and selected ion flow tube mass spectrometry (SIFT-MS) have garnered more interest as they offer real time VOC quantifications of exhaled breath^{53,177}. PTR-MS most commonly utilises H_3O^+ as the reagent ion for chemical ionisation and represents semi-quantitative analysis, but the most recent PTR-MS instruments can also utilize NO^+ and

O_2^+ reagent ions¹⁷⁷. However, since its inception, SIFT-MS has employed three reagent ions, H_3O^+ , NO^+ and O_2^+ and offers absolute quantification of VOCs in exhaled breath. It allows the simultaneous quantification of several VOCs in air or exhaled breath without the need for sample pre-treatment⁵³. The abundant volatile organic compounds within exhaled breath including acetone, methanol, ethanol, acetaldehyde, propanol and isoprene have been extensively investigated in healthy individuals using SIFT-MS^{114,115,116,117}. It has also been employed for breath analysis in disease diagnosis and therapeutic monitoring. Hryniuk *et al.* measured the exhaled breath metabolites of patients with coeliac disease using SIFT-MS¹⁷⁸. Endre *et al.* measured breath ammonia using SIFT-MS in patients with chronic kidney failure undergoing haemodialysis¹⁷⁹. Boshier *et al.* have also employed a SIFT-MS instrument intra-operatively in which they measured trace gases in the ventilated exhalations of anaesthetised patients during laparoscopic surgery¹⁸⁰. These studies all highlight the capabilities and advantages of SIFT-MS as a chemical analytical tool in the field of breath analysis.

6.1.1 Objective

To investigate whether VOCs within exhaled breath from patients with oesophago-gastric cancer differ from patients with benign upper gastro-intestinal conditions and those with healthy upper gastro-intestinal tracts using the SIFT-MS analytical method.

6.1.2 Introduction

The VOCs in the headspace vapour of biofluids retrieved from patients with oesophago-gastric cancer, positive controls (i.e., non-cancer diseases of the upper gastro-intestinal tract) and a 'healthy' cohort (i.e., those with no diseases present within the upper gastro-intestinal tract) were previously investigated. In the study on the headspace of gastric content (section 4.1.4), the concentrations of eight VOCs - acetone, formaldehyde, acetaldehyde, hexanoic acid, acetic acid, hydrogen sulphide, hydrogen cyanide and methyl phenol were identified that were significantly different between the cancer patients and healthy control group. In the study on the headspace of urine (section 5.4), the concentrations of seven VOCs - acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulphide, methanol and phenol were found to be significantly different between cancer, positive control and healthy groups. The results of these studies on biofluids identified potential compounds that may be important in VOC profiling of oesophago-gastric cancer and thus warrant further investigation. In this study, exhaled breath has been selected as the bio-sample for investigation. A breath sampling methodology has been developed and validated including assessment of Nalophan[®] bag contribution, water level degradation and presence of trace VOCs. It is postulated that the VOCs within exhaled breath obtained from patients with oesophago-gastric cancer will differ from controls. It is also hypothesised that several of the VOCs detected at increased concentrations in the headspace of biofluids from oesophago-gastric cancer patients may be present in the exhaled breath of these patients. The present study reports the first investigation of VOCs using the SIFT-MS platform of exhaled breath from patients with oesophageal and gastric cancer. VOCs were selected for analysis based on their presence in biofluids in the previous studies on oesophago-gastric cancer or if they have been detected in exhaled breath in other cancer states.

6.1.3 Methods

6.1.3.1 Patient selection

Patients were recruited through the endoscopy suite at St. Mary's Hospital, Paddington (Imperial College NHS Trust, London) and via the North West Thames regional oesophago-gastric cancer network (UK). OGD referral forms were screened to identify suitable patients for recruitment into this study. Patients were age-matched across the 3 groups and requested to provide a comprehensive medical history, smoking status and alcohol intake history. Patients with the following conditions were excluded from recruitment:

- Barrett's oesophagus
- Liver disease/oesophageal varices/portal hypertension
- Small bowel or colonic pathology
- A history or presence of another type of cancer
- Presence of acute infection

All patients included in this study underwent their OGD on the day of breath collection. Patients with a biopsy-confirmed invasive gastric or oesophageal cancer were included in the cancer cohort. Patients with biopsy confirmed non-cancer conditions of the upper gastrointestinal tract (e.g., esophagitis, gastritis and peptic ulcer disease) were included in the positive control cohort. Patients who had a normal OGD test (as interpreted by the endoscopist) and a negative rapid urease test were included in the 'healthy upper gastrointestinal tract' cohort. Local ethics committee approval was granted for this study and written informed consent was obtained from all patients prior to enrolment in the study. Demographic and clinical data were archived in a linked, anonymized database. The comparative analysis of exhaled breath VOCs was conducted with an oesophago-gastric cancer cohort of 18 patients, a positive control group of 18 patients and a 'healthy upper gastrointestinal tract' cohort group of 17 patients.

6.1.3.2 Breath Sample Collection

Exhaled breath samples were collected using a standardised method in the same room in the admissions area of the endoscopy suite. All study participants were fasted for a minimum of 6 hours in keeping with the standard protocol for an OGD with no oral intake (including water) prior to their breath sample collection. All subjects had been rested in the same area for at least 20 minutes prior to exhaled breath collection with all samples retrieved prior to OGD procedures. Samples of mixed alveolar breath were collected into secure double thickness (2x 25µm) Nalophan® (Kalle UK Ltd., Witham, UK) bags via a 1mL Luer lok syringe (Terumo Europe, Leuven, Belgium). The syringe is made up of 2 parts: the first part is the syringe barrel, which is fixed to the inside of the Nalophan® bag with a pull tight security seal. The second part is the syringe plunger that when removed, allows one to breathe into the secure bag via the aperture of the syringe barrel. The sampling bags used in these experiments were constructed in the Surgical Sciences Laboratory and subject to strict quality control protocols; every breath bag had a fixed volume of 2 litres and had been washed with dry synthetic air (BOC Ltd, Guildford, UK) prior to its use (Figure 6.1A). All subjects were seated throughout the breath sampling process. All participants were requested to perform a single deep nasal inhalation (as close to total lung capacity as possible) followed by complete exhalation via their mouth. This exhaled breath was not collected for analysis and served the purpose of demonstrating the breath manoeuvre required for sampling to the subject. The syringe plunger was only then removed from the syringe barrel and subjects were requested to repeat this procedure using the same breath manoeuvre directly into the Nalophan® bag. Following complete exhalation, the syringe plunger was immediately re-inserted into the barrel and breath samples were taken directly to the Surgical Sciences laboratory located within the hospital (Figure 6.1B). All breath samples were analysed immediately after collection.

A) Pre-sampling



B) Post-sampling



Figure 6.1: Nalophan[®] bags employed for the collection of mixed alveolar breath samples. A) An example of a Nalophan[®] bag pre-sampling B) The same Nalophan[®] bag containing an exhaled breath sample

The studies identified in the systematic review in chapter 2 predominantly employed off-line bag sampling methodologies (i.e, Nalophan[®] or Tedlar[®] bags) for the collection of exhaled breath. Although direct exhaled breath sampling (which is feasible with the SIFT-MS platform) would be preferable, there were several reasons that an off-line breath sampling method was employed:

- Standardisation of the direct exhaled breath analytical method is difficult. Many patients can often find it difficult to maintain the continuous respiratory effort or requested breath manoeuvre needed to obtain accurate on-line measurements.
- Direct breath sampling only provides a small window of opportunity to capture exhaled breath VOC data. Although multiple exhalations are possible, patients easily become tired with this approach. A far larger number of VOCs can be measured with off-line bag sampling rather than with direct exhalations.
- Exhaled breath sample using off-line bag samples can be acquired from patient within one minute. This allowed a larger number of patients to be recruited. The minimal time required to collect an exhaled breath sample also gave the sampling method a good acceptability rating amongst patients; hence there was an extremely low drop-out rate for participation in the studies.

- The SIFT-MS instrument is located in the Surgical Sciences Laboratory within the hospital. Hence, Nalophan[®] bag samples were collected in the endoscopy suite and analysed within an hour, minimising potential VOC degradation. Gilchrist *et al.* have previously investigated the suitability of double thickness (2x 25µm) Nalophan[®] bags for the collection and storage of breath samples¹⁸¹. They reported that the correlation of on-line and off-line measurements at 37°C for trace compounds such as HCN was good up to 6 hours and for abundant compounds such as acetone was good up to 24 hours¹⁸¹.

6.1.3.3 SIFT-MS analysis of exhaled breath samples

A *Profile-3* SIFT-MS instrument (Instrument Science, Crewe, UK) located in the Surgical Sciences laboratory at St. Mary's Hospital, Paddington (Imperial College Healthcare NHS Trust, London) was employed to analyse all exhaled breath samples in this study. Identification of exhaled breath VOCs including acetone, methanol and isoprene have been extensively reported in the literature^{53,121}. These studies employed the more accurate multiple ion monitoring (MIM) mode of SIFT-MS for targeted analysis of selected VOCs of interest in exhaled breath, and the MIM mode was also adopted in the current study. A full description of the MIM mode of analysis is given in a section 3.4. The majority of VOCs in this study were analysed using the H_3O^+ and NO^+ precursor ions. The O_2^+ precursor ion was only used for the detection of methyl phenol and ammonia in the MIM mode, as this reagent ion provides more accurate quantification for these trace compounds than does the H_3O^+ and NO^+ ions⁵³. For each measurement, the syringe plunger was removed from the 1 mL Luer lok syringe and the Nalophan[®] bag was directly connected via the syringe barrel to the sample inlet arm of the SIFT-MS instrument. The breath sample automatically flowed into the helium carrier gas at a fixed rate (0.2 TorrL/s) through the sampling line that was held at a temperature of 80 °C. For the duration of the analysis, breath sample bags were enclosed within an incubator held at 37 °C. For the MIM mode, selected VOCs from the breath samples were analysed for a total of 60 seconds and the measured concentrations were averaged over this analysis time for each VOC.

6.1.3.4 Statistical analysis

Statistical analysis was performed using IBM SPSS statistics 20 (SPSS Inc., Chicago, IL). A p -value ≤ 0.05 was taken as the level to indicate statistical significance. The Mann-Whitney U test was used to compare the measured concentrations of VOCs between the oesophago-gastric cancer cohort and control groups. The Mann-Whitney U test was also applied to compare the measured concentrations between positive control and healthy upper gastrointestinal tract control groups. The ROC curves were constructed using the same method described in section 4.1.3.4. In the current study, the statistical significance of differences in VOC concentrations between the cancer and positive controls was determined and then only those with $p \leq 0.05$ were included as variables for the proposed diagnostic test. The aim of comparing these groups was to assess the feasibility of VOC profiles within exhaled breath to differentiate between two distinct (biopsy confirmed) disease groups of the upper gastrointestinal tract.

6.1.4 Results

6.1.4.1 Water vapour level of breath samples

Ten healthy volunteers (with no co-morbidities) were recruited to provide four breath samples for the purposes of assessing the water level degradation employing Nalophan[®] bag (2x 25µm) sampling. All four breath samples were collected consecutively with a standardized 1-minute interval between each sample and immediately stored in a sealed 9-litre Really Useful Box (395 x 255 x 155mm) at room temperature. Breath samples were analysed at 30, 60, 120 and 240 minutes from the time of collection. Each sample bag was only removed from the storage container at their specific time point. All samples were analysed in the order that they were collected in (i.e., at the 30, 60, 120 and 240 minutes time points). Table 6.1 demonstrates the median water level of the exhaled breath samples at each of the designated time points. With appropriate storage, the water level within double thickness (2 x 25µm) Nalophan[®] bags is relatively stable with a median water level of 5.3% at 30 minutes and 4.6% at 240 minutes. This is an important consideration in terms of calibration of the SIFT-MS instrument and analysis of Nalophan[®] collected breath samples. Since the absolute humidity of exhaled breath is close to 6% (the saturation vapour pressure of water at body temperature) and so the measurement of the breath humidity acts as an internal check on the proper set up of the instrument⁵³. Furthermore, the observed water levels within double thickness Nalophan[®] bags in this study are similar to those recorded in on-line studies. In a recent study by Gilchrist *et al.*, median (IQR) on-line water vapour concentration was 5 (4.9-5.2)%¹⁸¹. The results in this study demonstrate that double thickness Nalophan[®] bags, that are appropriately stored, maintain a suitable water level up to 4 hours after collection.

Analysis time (minutes)	Median (%)	Interquartile range (%)
30	5.3	5.1-5.5
60	5.1	4.2-5.7
120	4.9	4.0-5.7
240	4.6	3.8-5.0

Table 6.1: Median water vapour level and interquartile range within Nalophan[®] bags at 30, 60 120 and 240 minutes from the time of breath collection taken from 10 healthy volunteers.

6.1.4.2 Reproducibility and validity of the sampling method

Experiments were also conducted to assess the inter-sample reliability and reproducibility of the aforementioned breath sampling methodology. Ten healthy volunteers (with no co-morbidities) were recruited and requested to provide 2 breath samples obtained within a minute of each other. A total of 20 breath samples were analyzed using the MIM mode for the selected VOCs. The inter-sample coefficient of variation (CV) was calculated for the measured concentration of several VOCs within the exhaled breath samples. These CV values obtained in percentage are: acetone, 10.7%; isoprene, 11.5%; methanol, 7.2%; hexanoic acid, 15.7%; phenol 14.9%; ethyl phenol, 17.2%; methyl phenol, 14.6%. These coefficients of variation demonstrate the reproducibility of the MIM measurement for these VOCs within exhaled breath.

6.1.4.3 Surgical Sciences Laboratory room air and Nalophan bag measurements

Ten measurements of the Surgical Sciences laboratory room air concentrations for acetone, isoprene, methanol, hexanoic acid, phenol, methyl phenol and ethyl phenol were also recorded. As shown in Table 6.3 and the figures later in the ‘Discussion’, the room air concentrations of all the VOCs are at trace levels in comparison to the measured concentrations within exhaled breath. It was noted that methanol has a higher background laboratory room air value, but this is much lower compared to the median concentrations observed within exhaled breath. In a recent study by Španěl *et al.*, it was reported that there is minimal retention of exogenous methanol in exhaled breath¹⁸². Hence, the potential exogenous contribution of background laboratory room air to the measured concentrations of the VOCs in exhaled breath samples in these experiments is deemed negligible.

The potential contribution of the Nalophan[®] bag surface to exhaled breath samples has also been investigated. Ten Nalophan[®] bags were filled with dry synthetic air (BOC Ltd, Guildford, UK) and analysed using the MIM mode of the SIFT-MS. All the VOCs within the Nalophan[®] bags were present in trace quantities in comparison to the median concentrations within exhaled breath. Both these experiments highlight that the important trace VOCs identified in this study are not significantly influenced by exogenous sources or experimental methodology.

6.1.4.4 Analysis of selected VOCs in exhaled breath

As shown in Table 6.2, a total of 17 VOCs present within exhaled breath were analysed using the MIM mode of the SIFT-MS. The analytical information, including chemical formula, precursor ions, m/z values and characteristic product ions are given. H_3O^+ is the most frequently used precursor ion as it reacts with more compounds than the other 2 precursor ions. However, NO^+ and O_2^+ have also been employed for specific molecules when they have an advantage in unambiguously detecting and quantifying these compounds over H_3O^+ (i.e., when there are known overlaps with other product ions).

Compounds	Molecular formula	Precursor ions	<i>m/z</i>	Characteristic product ions	REF
Methanol	CH ₄ O	H ₃ O ⁺	33, 51	CH ₅ O ⁺ , CH ₅ O ⁺ (H ₂ O)	53
Propanol	C ₃ H ₈ O	H ₃ O ⁺	43	C ₃ H ₇ ⁺	53
Butanol	C ₄ H ₁₀ O	H ₃ O ⁺	57	C ₄ H ₉ ⁺	53
Pentanol	C ₅ H ₁₂ O	H ₃ O ⁺	71	C ₅ H ₁₁ ⁺	53
Pentanoic acid	C ₅ H ₁₀ O ₂	H ₃ O ⁺	103	C ₅ H ₁₀ O ₂ H ⁺	53
Hexanoic acid	C ₆ H ₁₂ O ₂	H ₃ O ⁺	117, 135	C ₆ H ₁₂ O ₂ H ⁺ , C ₆ H ₁₂ O ₂ H ⁺ (H ₂ O)	53
Hydrogen sulphide	H ₂ S	H ₃ O ⁺	35	H ₃ S ⁺	53
Hydrogen cyanide	HCN	H ₃ O ⁺	28	H ₂ CN ⁺	136
Acetaldehyde	C ₂ H ₄ O	H ₃ O ⁺	45, 81	CH ₃ CHOH ⁺ , CH ₃ CHOH ⁺ (H ₂ O) ₂	53
Formaldehyde	H ₂ CO	H ₃ O ⁺	31	H ₂ COH ⁺	53
Acetone	C ₃ H ₆ O	NO ⁺	88	NO ⁺ ·C ₃ H ₆ O	53
Acetic acid	C ₂ H ₄ O ₂	NO ⁺	90, 108	NO ⁺ ·CH ₃ COOH, NO ⁺ ·(H ₂ O)CH ₃ COOH	53
Isoprene	C ₅ H ₈	NO ⁺	68	C ₅ H ₈ ⁺	53
Phenol	C ₆ H ₆ O	NO ⁺	94, 112	C ₆ H ₆ O ⁺ , C ₆ H ₆ O ⁺ (H ₂ O)	183
Methyl phenol	C ₇ H ₈ O	O ₂ ⁺	108, 126	C ₇ H ₈ O ⁺ , C ₇ H ₈ O ⁺ (H ₂ O)	183
Ethyl phenol	C ₈ H ₁₀ O	NO ⁺	122, 140	C ₈ H ₁₀ O ⁺ , C ₈ H ₁₀ O ⁺ (H ₂ O)	183
Ammonia	NH ₃	O ₂ ⁺	17, 35	NH ₃ ⁺ , NH ₃ ⁺ (H ₂ O)	53

Table 6.2: A summary of the characteristic ions of the 17 VOCs analysed by the MIM mode of SIFT-MS using H₃O⁺, NO⁺ and O₂⁺ precursor ions.

	Cancer cohort		Positive controls		Healthy controls		Room air		Nalophan [®] bag	
VOCs	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Acetone	307	[116-1205]	449	[261-698]	372	[164-958]	6	[3-10]	1	[0-2]
Isoprene	71	[46-175]	72	[46-117]	52	[45-80]	3	[1-5]	2	[1-3]
Methanol	235	[170-328]	210	[170-304]	159	[125-195]	38	[30-50]	29	[22-37]
Hexanoic acid*	19	[12-37]	7	[4-13]	10	[8-18]	1	[0-3]	1	[1-2]
Phenol*	17	[8-26]	7	[5-9]	6	[3-9]	2	[1-3]	1	[0.95-1.35]
Methyl phenol*	12	[9-19]	3	[2-6]	5	[3-10]	1	[0-2]	1	[1-3]
Ethyl phenol*	9	[5-23]	5	[2-11]	6	[2-15]	0	[0-2]	2	[1-2]

Table 6.3: Summary of median concentration and interquartile range (IQR) measured in parts-per-billion by volume (*ppbv*) for the most abundant VOCs present in exhaled breath and the 4 VOCs that were significantly different between the oesophago-gastric cancer cohort and controls as denoted by the asterix*. Room air and Nalophan[®] bags (inflated with dry synthetic air) measurements of each VOC are also included. (N.B. Throughout the ‘Results and Discussion’ Section, the Healthy Upper GI tract cohort is abbreviated to the ‘Healthy’ cohort)

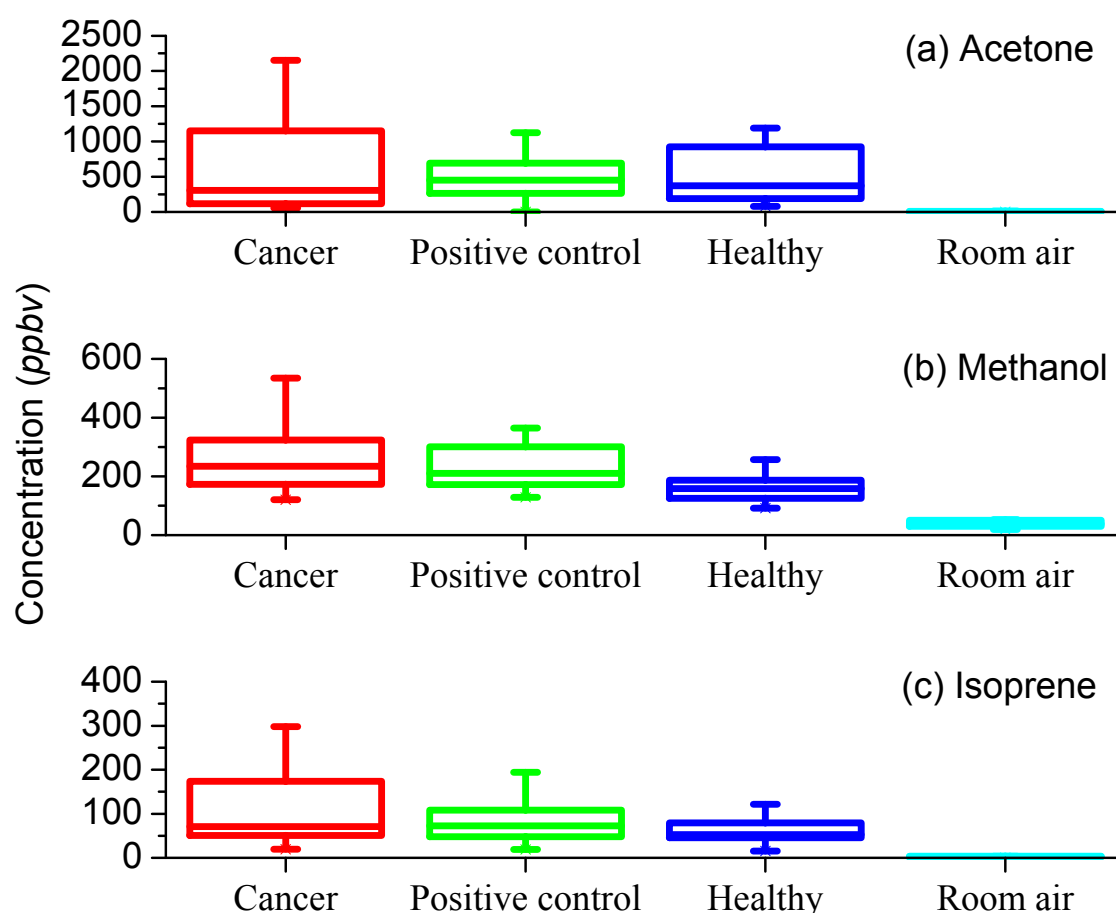


Figure 6.2: Box-Whisker plots of the median concentrations and interquartile range (in *ppbv*) of (a) acetone, (b) methanol and (c) isoprene from the exhaled breath of oesophago-gastric cancer patients, positive controls and the healthy upper GI tract cohort. Room air concentrations from the Surgical Sciences laboratory for each VOC are also included.

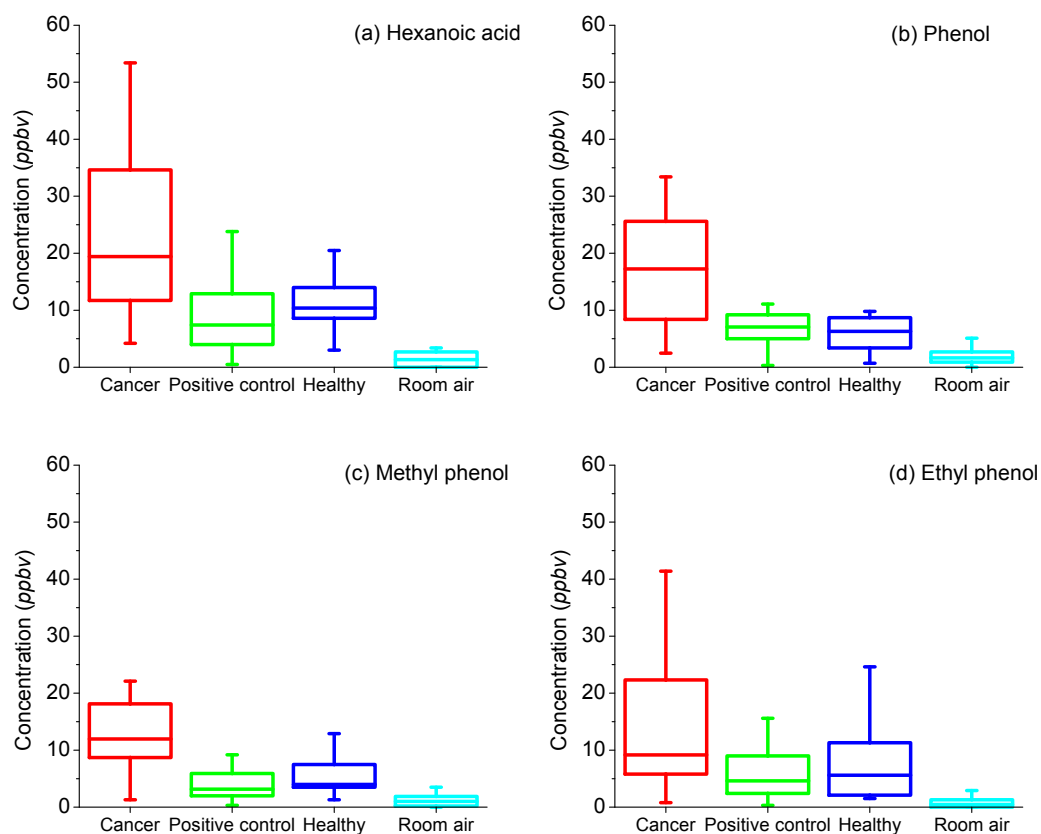


Figure 6.3: Box-Whisker plots of the median concentrations and interquartile range (in *ppbv*) of (a) hexanoic acid, (b) phenol, (c) methyl phenol and (d) ethyl phenol from the exhaled breath of oesophago-gastric cancer patients, positive controls and healthy controls. Room air concentrations from the Surgical Sciences laboratory for each VOC are also included.

6.1.4.5 Exhaled Breath VOCs to discriminate cancer from non-cancer controls

Mann-Whitney U test ($p\text{-value} \leq 0.05$) was performed for all the volatile organic compounds listed in Table 6.2, to assess for any differences in median concentrations between the oesophago-gastric cancer cohort and positive control group. Additional tests were also performed between the cancer cohort and healthy control group as well as the positive control and healthy control groups. Between the cancer and positive control groups, four VOCs were found to be statistically significantly different, i.e. hexanoic acid ($p=0.001$), phenol ($p=0.001$), methyl phenol ($p<0.001$) and ethyl phenol ($p=0.044$). Similar results were obtained between the cancer and healthy control groups including methanol ($p=0.007$), hexanoic acid ($p=0.006$), phenol ($p=0.006$) methyl phenol ($p=0.001$). However, methanol ($p=0.007$) was the only VOC to be statistically significantly different in the exhaled breath of the positive control and healthy control groups. Using the four statistically significant VOCs identified in the Mann-Whitney U test, the binary logistical regression analysis resulted in a linear combination of concentrations of hexanoic acid (coefficient 0.134), phenol (0.176), methyl phenol (0.228) and ethyl phenol (-0.187). The AUC of the ROC using this model equation is 0.910, demonstrating a high discriminating power between the oesophago-gastric cancer and positive control groups with good sensitivity and specificity (Figure 6.4). This result is encouraging and further supports the theory that non-invasive VOC profiling may have a diagnostic role for risk stratification in oesophago-gastric cancer.

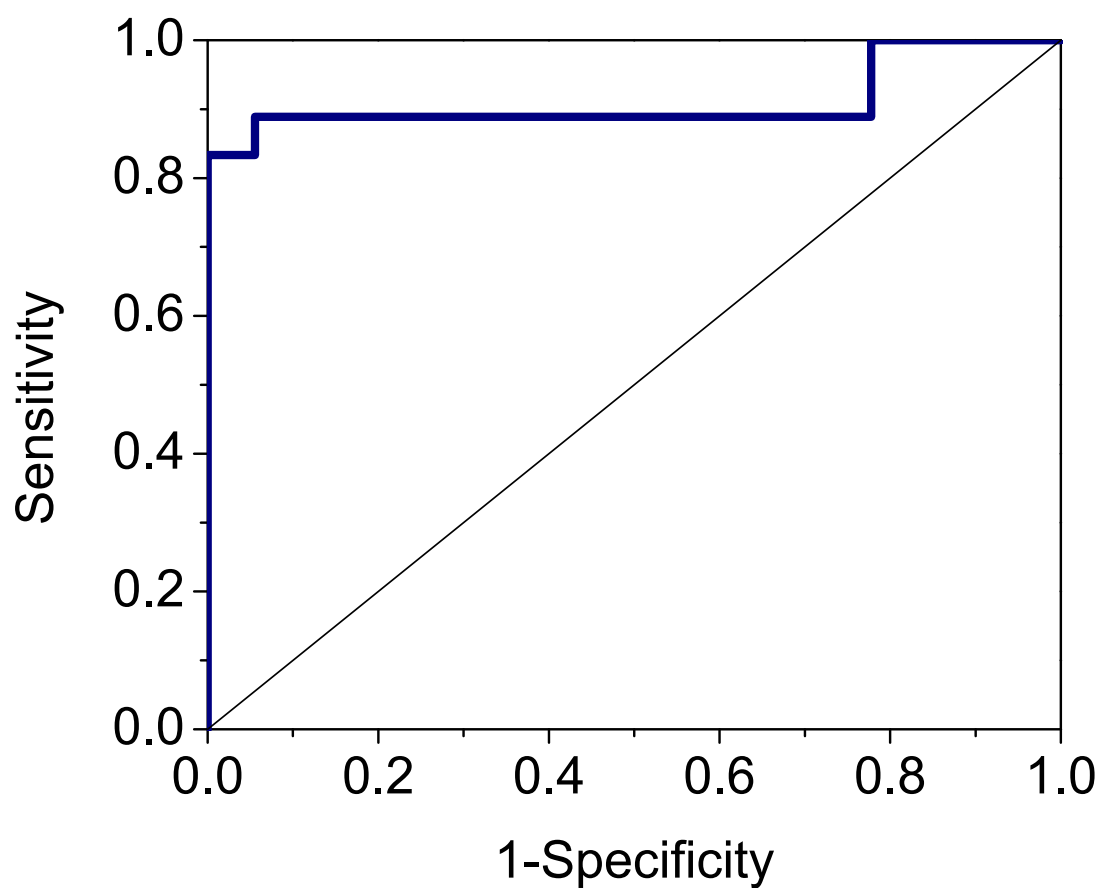


Figure 6.4: ROC curve of discrimination results for the diagnosis of oesophago-gastric cancer from positive controls using a combination of the following exhaled breath VOCs - hexanoic acid, phenol, methyl phenol and ethyl phenol exhibited, respectively. The AUC is 0.910.

6.1.5 Discussion

6.1.5.1 Abundant breath VOCs

Abundant breath VOCs are those compounds present at concentrations of several hundred *ppbv* in the exhaled breath of the majority of the general population. Acetone, isoprene and methanol are considered abundant metabolites within exhaled breath, all of systemic origin, that have been formerly investigated using SIFT-MS^{183,184}. Acetone (2-propanone, β -ketopropane) is a product of mitochondrial β -oxidation and the most abundant ketone present in exhaled breath¹³⁹. Endogenous production of acetone in humans occurs by two processes - the decarboxylation of acetoacetate and dehydrogenation of isopropanol¹⁸⁵. In this study, the quantification of acetone was achieved using the NO^+ precursor ion, which produces a single characteristic product ion (m/z of 88) and no further cluster ions. The accurate quantification of acetone in exhaled breath using the NO^+ precursor ion of SIFT-MS has been previously reported¹⁸⁶. In the previous studies on the headspace of gastric content and urine, higher concentrations of acetone were observed in both biofluids in the cancer cohort compared to non-cancer controls. Pabst *et al.* reported increased concentrations of exhaled breath acetone during cardiac surgery and proposed this may be related to the simultaneous increase in oxidative stress¹⁸⁷. Marcondes-Braga *et al.* reported that exhaled breath acetone levels correlate with heart failure severity¹⁸⁸. As shown in Figure 6.2, the measured exhaled acetone concentrations of the cancer cohort has a larger interquartile range in comparison to the non-cancer controls. However, no significant differences in median concentrations were observed between the oesophago-gastric cancer cohort (307 *ppbv*), positive control (449 *ppbv*) and the healthy (372 *ppbv*) cohort. Turner *et al.* have previously studied acetone in single breath exhalations by healthy volunteers using SIFT-MS over a 6-month period¹¹⁴. In this on-line study, they reported a median acetone concentration of 462 *ppbv* with a range of 148-2744 *ppbv* in 30 healthy volunteers¹¹⁴. The results are comparable with the exhaled breath acetone concentrations observed in this study and further validate the aforementioned off-line bag sampling method.

The exhaled breath concentrations of methanol demonstrate a clear decreasing trend from the oesophago-gastric cancer group to the healthy cohort, a trend that was observed in both the studies on the headspace of gastric content and urine. The oesophago-gastric cancer cohort and positive control group have very close median concentrations of methanol, i.e., 235 *ppbv* and 210 *ppbv*, respectively (Figure 6.2). The healthy control group has a lower median methanol

concentration of 159 *ppbv*, which falls within quoted ranges for online measurements of exhaled breath methanol¹¹⁷. There is no statistically significant difference in methanol concentration between the cancer cohort and positive control group by Mann-Whitney U test. However, methanol was found to be statistically significantly different between the cancer cohort and healthy control group ($p=0.006$). Of note, it is also the only VOC that demonstrated a statistically significant difference between the positive control and healthy control groups ($p=0.007$). Methanol has been observed at a consistently higher median concentration in the exhaled breath, urine and gastric content of subjects with upper gastrointestinal diseases compared to those with healthy upper gastrointestinal tracts. The majority of endogenous methanol is produced within the gastro-intestinal tract. Caldwell demonstrated higher concentrations of methanol completely inhibit predominant anaerobic gut bacteria; many of these bacteria are present in a protective role in the human gut and their inhibition can lead to gut dysfunction¹⁷³. Garner *et al.* also observed that methanol was present more frequently in faeces samples of patients with *Campylobacter jejuni* infection and ulcerative colitis compared to the asymptomatic donors¹⁷². Although exhaled breath methanol is not a specific VOC to oesophago-gastric cancer, it is increased in both upper gastrointestinal disease groups. It therefore may represent a marker of gut dysfunction and thus warrants further investigation in this regard.

Isoprene (2-methyl-1,3-butadiene) is an abundant hydrocarbon within human exhaled breath and known metabolite of cholesterol synthesis¹⁸⁹. Spänzel & Smith have investigated the accurate quantification of isoprene in exhaled breath using SIFT-MS¹⁹⁰. They reported that all three precursor ions can react with isoprene. For H_3O^+ precursor ions, the characteristic product ion $C_5H_9^+$ is formed by a protonation reaction with both H_3O^+ and its monohydrate ion. $C_5H_9^+$ has a m/z value of 69, which is overlapped with the hydrate of protonated methanol, $CH_3OH.H^+(H_2O)_2$, which is often observed in humid breath samples. In contrast to the reaction of O_2^+ with isoprene, in which two dominant product ions are produced, NO^+ precursor ion only ionises isoprene into $C_5H_8^+$ at m/z 68 by a simple non-dissociative charge transfer reaction and no other common breath metabolites will produce m/z 68 products using NO^+ . Therefore, exhaled breath isoprene was analysed using the NO^+ precursor ion in this study. As shown in Figure 6.2 and table 6.3, the median concentration of isoprene within exhaled breath is 71, 72 and 52 *ppbv* in the oesophago-gastric cancer cohort, positive control and healthy control groups, respectively. Turner *et al.* have previously studied breath isoprene through online SIFT-MS analysis in healthy volunteers¹¹⁵. In this study, they reported a mean breath isoprene

level of 118 ± 68 ppbv with significant variation amongst apparently healthy individuals¹¹⁵. Boshier *et al.* reported significantly increased concentrations of exhaled isoprene in patients undergoing laparoscopic surgery and postulated that this may represent a stress response¹⁸⁰. Xu *et al.* recently analysed exhaled breath samples from patients with gastric complaints using a nanomaterial-based gas sensor¹⁹¹. Complementary chemical analysis was performed on these breath samples using GC-MS and they observed significantly elevated levels of isoprene in patients with gastric cancer and/or peptic ulcer compared to those with less severe gastric conditions¹⁹¹. Higher median concentrations of isoprene were observed in both the cancer cohort and positive control group compared to healthy controls, however this result did not reach statistical significance ($p=0.08$)¹⁹¹. Given the abundance of acetone, methanol and isoprene within exhaled breath and the levels observed in this study, it is evident these VOCs are unlikely to be specific breath biomarkers for oesophago-gastric cancer.

6.1.5.2 Hexanoic Acid

Hexanoic acid was analyzed using the H_3O^+ precursor ion with 2 characteristic product ions (m/z 117 and 135 at a typical ratio of 27/29 counts per second). The median concentration of hexanoic acid in the exhaled breath of the oesophago-gastric cancer cohort (19 ppbv) is significantly increased in comparison to the positive control (7 ppbv) and healthy control groups (10 ppbv). As shown in Figure 6.3, the cancer cohort also have a much wider interquartile range (12-37 ppbv) range in comparison to both control groups. As shown in Table 6.3, there is also a negligible hexanoic acid contribution from both laboratory room air and the surface of Nalophan[®] sampling bags. In the previous studies on the headspace of gastric content and urine, the same trend of significantly increased concentrations of hexanoic acid in the cancer cohort in comparison to non-cancer controls was observed. Saturated fatty acid vapours (up to C14) have previously been observed in human breath using electrospray ionization and atmospheric pressure ionization-mass spectrometry¹⁹². However, there remains limited data published regarding concentrations of these compounds (including hexanoic acid) within exhaled breath. Silva *et al.* have previously analysed urine samples from patients with leukaemia, lymphoma and colorectal cancer using SPME GC-MS⁷⁰. They observed hexanoic acid less frequently in urine samples from these cancer groups compared with the healthy controls⁷⁰. Hexanoic acid has been consistently elevated in biological samples from patients with oesophago-gastric cancer when analysed with SIFT-MS using the H_3O^+ precursor ion with 2 characteristic product ions (m/z 117 and 135); this may reflect this VOC being systemically

increased in oesophago-gastric cancer. To further investigate this, experiments using standards of hexanoic acid isomers were conducted at the SIFT-MS laboratory in Keele University. Hexanoic acid standards gave the following % signals at the characteristic product ions and the associated dihydrate using SIFT-MS – m/z 117: 5-23%, m/z 135: 27-63% and at m/z 153: 31-69% - (D.Smith, priv. comm.). Analysis of the signals within exhaled breath samples from OG cancer patients with increased concentrations of hexanoic acid gave the following median % signals – m/z 117: 16%, m/z 135: 45% and at m/z 153: 37% (Table 6.4). Though the count rates were lower in exhaled breath, which is to be expected, the median % signals all fall within the appropriate ranges for the hexanoic acid standards. Further research utilising multiple analytical platforms (e.g., GC-MS) is necessary to verify the compound responsible for these increased concentrations is only hexanoic acid and whether a specific isomer of this compound is also consistently increased in OG cancer.

Exhaled breath sample	Water level (%)	m/z 117	m/z 135	m/z 153
OG 1	4	18%	17%	66%
OG 2	4.6	17%	39%	45%
OG 3	3.79	7%	45%	48%
OG 4	4.62	23%	35%	42%
OG 5	5.75	25%	44%	31%
OG 6	4.33	7%	56%	37%
OG 7	4.49	18%	46%	36%
OG8	4.48	6%	29%	26%
OG 9	4.81	4%	59%	37%
OG 10	4.56	15%	65%	20%
Median	4.53	16%	45%	37%

Table 6.4: The water level and individual % signals for m/z 117, m/z 135 & m/z 153 for the measurement of hexanoic acid in exhaled breath samples from patients with OG cancer

6.1.5.3 Phenol and phenol derivatives

The phenol, methyl phenol and ethyl phenol measurements are statistically significantly different in the exhaled breath of the oesophago-gastric cancer cohort compared to non-cancer controls. The median concentration of phenol in the exhaled breath of the cancer cohort is 17 *ppbv* compared to 7 *ppbv* in the positive control group and 6 *ppbv* in the healthy control group. As shown in Figure 6.3, the cancer cohort also demonstrate a much wider interquartile range (8-26 *ppbv*) in comparison to both control groups which have tighter interquartile ranges. Of note is the upper quartile value of 9 *ppbv* for phenol in the positive and healthy control groups. Turner *et al.* have previously demonstrated the presence of phenol in the exhaled breath of healthy volunteers using thermal desorption GC-MS¹⁹³. In the study on gastric content headspace using SIFT-MS, phenol was increased in patients with OG cancer compared to controls but this did not reach statistical significance. In an analysis of gastric content headspace using SPME/GC-MS (Appendix B), phenol was the most frequently observed VOC in gastric content from patients with oesophago-gastric cancer. In the headspace of urine, significantly lower concentrations of phenol were observed in the cancer cohort compared with controls. McDonald *et al.* have previously hypothesised that higher concentrations of phenol observed in urine may reflect its efficient removal from the blood and subsequent excretion¹⁷⁵. It is postulated that the variations observed in phenol concentrations in the OG cancer cohort when compared to controls may be due to dysfunctional phenol excretion secondary to alterations in tyrosine metabolism by gut bacteria or impairment in the removal mechanisms in blood. Phenol is a known breakdown product of tyrosine through the process of proteolytic fermentation¹⁹⁴. Deng *et al.* studied aromatic amino acids in gastric juice samples from patients with early-stage gastric cancer, advanced gastric cancer and non-neoplastic gastric diseases using high performance liquid chromatography mass spectrometry (HPLC-MS)¹²⁶. They found that high levels of aromatic amino acids including tyrosine and phenylalanine were present in gastric juice samples from patients with gastric cancer¹²⁶. It is hypothesised that the significantly increased concentrations of phenol in the exhaled breath of the oesophago-gastric cancer cohort may be associated with increased protein catabolism, changes in gut microbiota, up-regulation of tyrosine metabolism or a combination of all three.

As shown in Figure 6.3, the median concentrations and trends of methyl phenol and ethyl phenol across the three groups are very similar. Exhaled breath methyl phenol and ethyl phenol from the cancer cohort are at higher median concentrations and wider interquartile ranges in comparison to non-cancer controls. For instance, the measured methyl phenol concentration

from the cancer cohort at 12 *ppbv* is significantly increased compared to the positive control and healthy control groups which have similar median concentrations of 3 *ppbv* and 5 *ppbv*, respectively. Ethyl phenol has a median concentration of 9 *ppbv* in the cancer cohort whilst the positive control and healthy control groups have median concentrations of 5 *ppbv* and 6 *ppbv*, respectively. In the study on the headspace of urine, no significant differences were observed in the measured concentrations of methyl phenol or ethyl phenol between the groups. In the study on the headspace of gastric content, a significantly increased concentration of methyl phenol was observed in the oesophago-gastric cancer cohort compared with healthy controls (17 *ppbv* vs. 8 *ppbv*). Silva *et al.* reported increased levels of methyl phenol in urine samples from patients with cancer compared to controls¹⁹⁵. Phenol, methyl phenol and ethyl phenol within exhaled breath did not demonstrate any significant differences between the positive control and healthy control groups. As shown in Table 6.3, there were negligible phenol, methyl phenol or ethyl phenol contributions from both laboratory room air and Nalophan[®] sampling bags.

The other VOCs that were investigated in Table 6.2 did not show any significant differences between the groups in this study and hence no further analysis was conducted at this time.

6.1.6 Conclusions

This is the first study of VOCs within exhaled breath of patients with oesophago-gastric cancer. Comparative analysis has been made with control groups, all of whom had their upper gastrointestinal tracts visualised on the day of breath sampling. A total of 17 VOCs were analysed in exhaled breath samples from patients with OG cancer, positive controls and a healthy control group. Of these compounds, the median concentrations of 4 VOCs (hexanoic acid, phenol, methyl phenol and ethyl phenol) were statistically different between the oesophago-gastric cancer patients and the positive control group. Several of the same VOCs that differentiated oesophago-gastric cancer from control groups in the previous work of headspace analysis of gastric content and urine have also been identified in this study (Chapters 4 and 5). Of note, hexanoic acid has been observed at increased concentrations in all biological samples from patients with oesophago-gastric cancer. In comparison to the abundant breath metabolites, it is evident from this study that trace compounds are more likely to be influential in VOC profiling in exhaled breath in OG cancer. The SIFT-MS technique allows real-time detection and quantification of trace volatile compounds in human biological samples. In the current study, the VOCs within exhaled breath have been readily analysed without any sample preparation and therefore minimal delay. This is particularly advantageous within the clinical environment where samples are retrieved and real-time VOC measurements can be made with minimal concern for potential sample degradation. Multiple chemical analytical platforms are necessary to identify other VOCs that may be important in differentiating oesophago-gastric cancer from other diseases. Most importantly, chemical analytical research could lead to the development of a non-invasive VOC-based breath (and/or urine) screening test.

The differences in the concentrations of particular VOCs observed in the exhaled breath of the oesophago-gastric cancer cohort when compared with control groups in this pilot study are promising. The results of VOC levels within the environment and from the surfaces of Nalophan[®] bags also gave me confidence in analysing the trace compounds within exhaled breath. Acetaldehyde and formaldehyde did not demonstrate any significant statistically significant differences between the cancer cohort and controls in this study. Smith *et al.* have investigated acetaldehyde in patients with non-small cell lung cancer. They observed increased concentrations of acetaldehyde in these patients but quickly realised that the major fraction for this VOC was being emitted from the bag surface¹⁴³. However, aldehydes remain important VOCs to investigate in OG cancer

6.2 Exhaled breath Aldehydes in a Healthy population

6.2.1 Objective

To investigate the concentrations of C3-C10 aldehydes in the exhaled breath of healthy persons and elucidate ‘normal values’ or reference ranges for these compounds.

6.2.2 Introduction

The results of the systematic review from Chapter 2 identified aldehydes as VOCs that were linked to several types of cancer. Aldehydes produced within the human body are most commonly oxidised to carboxylic acids by aldehyde dehydrogenase enzymes. The C3-C10 aldehydes have previously been reported to be associated with lung and colorectal cancer. Fuchs *et al.* reported elevated pentanal, hexanal, octanal and nonanal concentrations in the exhaled breath of lung cancer patients compared to healthy volunteers; they postulated that the increases in these exhaled breath aldehydes reflected oxidative stress and tumour-specific tissue metabolism⁶³. Using GC-Time Of Flight MS, Buszewski *et al.* reported a higher concentration of butanal in the exhaled breath of lung cancer patients compared with controls¹⁹⁶. Altomare *et al.* employed an off-line breath collection method (employing Tedlar[®] bags) to investigate VOC patterns within the exhaled breath of patients with colorectal cancer. Using thermal desorption GC-MS and a probabilistic neural network (PNN) analysis, they identified nonanal and decanal as two of the discriminating variables to distinguish colorectal cancer patients from controls¹⁹⁷. The aforementioned studies all employed GC-MS for the analysis of exhaled breath aldehydes. Apart from the studies described in Chapters 4 and 5, SIFT-MS has previously been employed to study aldehydes in bladder, prostate and lung cancer. Španěl *et al.* demonstrated elevated levels of formaldehyde in the headspace of urine samples from patients with bladder and prostate cancer⁸³. Smith *et al.* also identified higher levels of acetaldehyde in the lung cancer cell line SK-MES and CALU-1 cell lines⁸². SIFT-MS has great potential in the targeted, real-time analysis of short chain aldehydes from exhaled breath. In order for exhaled breath analysis to play a more prominent role in clinical diagnostics, standardization of breath collection methodologies and establishment of ‘normal values’ for VOCs of interest in healthy persons are necessary. Given there was no published data for C3-C10 aldehydes within exhaled breath using the SIFT-MS platform, it was considered important to investigate levels of these compounds in healthy subjects. In this study,

the detection and quantification of C3-C10 saturated aldehydes (propanal, butanal, pentanal, heptanal, hexanal, heptanal, octanal, nonanal and decanal) in exhaled breath has been investigated using the SIFT-MS platform.

6.2.3 Methods

6.2.3.1 Subjects

Twenty six healthy volunteers were enrolled in this study, including 11 male and 15 female participants. Their mean age was 45.8 ± 16.7 years. All subjects had no significant past medical history. There were 4-smokers and 22 non-smokers amongst the study participants.

6.2.3.2 Breath sample collection

All subjects were fasted for a minimum of 8 hours with no oral intake (including water) prior to sample collection. They had been requested to refrain from alcohol 24 hours prior to exhaled breath sampling. Off-line mixed alveolar breath samples were collected from all participants in same location (i.e., the Surgical Sciences laboratory) between 8am and 10am. All subjects provided samples into secure double thickness (2 x 25 μ m) Nalophan[®] (Kalle UK Ltd., Witham, UK) bags via the same method described earlier in this chapter.

Ten of the healthy volunteers were selected at random to provide four breath samples for the purposes of assessing the repeatability and reproducibility of the Nalophan[®] bag sampling methodology for saturated aldehydes. All four breath samples were collected consecutively with a standardized 1-minute interval between each sample and immediately stored in a sealed 9-litre Really Useful Box (395 x 255 x 155mm) at room temperature. Breath samples were analysed at 30, 60, 120 and 240 minutes from the time of collection.

6.2.3.3 SIFT-MS analysis

In this study, the FS mode was initially employed to demonstrate that SIFT-MS could detect the C3-C10 aldehydes within off-line exhaled breath samples. However, the MIM mode was the main *modus operandi* for quantification of the selected aldehydes. Calibration check procedures of the instruments and SIFT-MS analysis of exhaled breath samples were conducted as previously described in Section 6.1.3. For the FS mode, 9 x 60 seconds repeated scans were performed using the NO^+ precursor ion only. For the MIM mode, the selected aldehydes from the breath samples were analysed for a total of 60 seconds and the measured concentrations were averaged over the analysis time for each compound. In addition, exhaled breath VOCs, including acetone, acetic acid and isoprene were simultaneously analysed using NO^+ precursor ions.

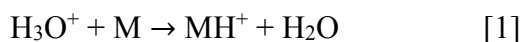
6.2.3.4 Ion chemistry

C3 to C10 aldehydes, including propanal, butanal, pentanal, hexanal, heptanal, octanal, nonanal and decanal from exhaled breath were analysed using the NO^+ reagent ion only. Acetone, acetic acid and isoprene were also measured using NO^+ reagent ions (Table 6.5).

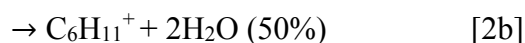
Aldehyde	Formulae	Precursor ions	Characteristic ions	<i>m/z</i>
Acetone	$\text{C}_3\text{H}_6\text{O}$	NO^+	$\text{NO}^+ \cdot \text{C}_3\text{H}_6\text{O}$	88
Acetic acid	$\text{C}_2\text{H}_4\text{O}_2$	NO^+	$\text{NO}^+ \cdot \text{CH}_3\text{COOH}$, $\text{NO}^+ \cdot (\text{H}_2\text{O})\text{CH}_3\text{COOH}$	90, 108
Isoprene	C_5H_8	NO^+	C_5H_8^+	68
Propanal	$\text{C}_2\text{H}_5\text{CHO}$	NO^+	$\text{C}_3\text{H}_5\text{O}^+$	57
Butanal	$\text{C}_3\text{H}_7\text{CHO}$	NO^+	$\text{C}_4\text{H}_7\text{O}^+$	71
Pentanal	$\text{C}_4\text{H}_9\text{CHO}$	NO^+	$\text{C}_5\text{H}_9\text{O}^+$	85
Hexanal	$\text{C}_5\text{H}_{11}\text{CHO}$	NO^+	$\text{C}_6\text{H}_{11}\text{O}^+$	99
Heptanal	$\text{C}_6\text{H}_{13}\text{CHO}$	NO^+	$\text{C}_7\text{H}_{13}\text{O}^+$	113
Octanal	$\text{C}_7\text{H}_{15}\text{CHO}$	NO^+	$\text{C}_8\text{H}_{15}\text{O}^+$	127
Nonanal	$\text{C}_8\text{H}_{17}\text{CHO}$	NO^+	$\text{C}_9\text{H}_{17}\text{O}^+$	141
Decanal	$\text{C}_9\text{H}_{19}\text{CHO}$	NO^+	$\text{C}_{10}\text{H}_{19}\text{O}^+$	155

Table 6.5: Characteristic analyte and *m/z* of acetic acid, acetone, isoprene and the C3-C10 aldehydes analysed using SIFT-MS.

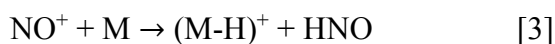
H₃O⁺ reagent ion can also be used to ionise C3-C10 aldehydes via proton transfer reaction¹¹⁶ [Equation 1].



However, the parent product ions (MH⁺)⁺ often dissociate into fragments (M-H₂O)⁺ with loss of a water molecule. For instance, the ultimate product ions from hexanal and H₃O⁺ will be a mixture of C₅H₁₁CHOH⁺ and C₆H₁₁⁺ [Equation 2]:



Hence, NO⁺ precursor ion is used for their quantification as only a single product ion is produced (M-H)⁺ via hydride transfer reaction¹³⁴ [Equation 3]. It is also found that the hydride transfer reaction is the only process that occurs between a saturated aldehyde and NO⁺.¹⁸⁰. Španěl *et al.* also reported that the produced (M-H)⁺ shows very little degree of association with water and hence the water molecules from the sample have no influence on the quantification of these saturated aldehydes¹⁹⁸. Therefore, NO⁺ remains the most suitable precursor ion to analyse the C3-C10 aldehydes.



6.2.4 Results

6.2.4.1 Aldehydes quantification and variations within healthy subjects

For the accurate quantification of exhaled breath aldehydes using SIFT-MS, specific precursor ions and characteristic product ions must be carefully selected for analysis¹²¹. As mentioned above, all eight aldehydes (C3 to C10) in this study were analysed using NO^+ reagent ions to give a single product ion, which does not overlap with any other ions (Table 6.5). Propanal was the most abundant and has the smallest molecular weight amongst these eight compounds. Conventional GC-MS techniques are not suitably sensitive for the detection and quantification of aldehydes with less than four carbon atoms. However, SIFT-MS has been demonstrated to have a good sensitivity for these compounds in headspace experiments of foodstuffs^{129,199}. Both H_3O^+ and NO^+ reagent ions can react with propanal to produce a single parental product ion. However, when H_3O^+ is used, the product ion of propanal is isobaric with acetone, which will have the same m/z value¹¹². Therefore, the NO^+ ion is used for the quantification of propanal as acetone reacts with NO^+ to produce the m/z 88 product ion, whilst the product ion of propanal is 57^{129,134}. Figure 6.5 below shows that propanal can be identified from m/z 57 when NO^+ reagent ion is used in the FS mode of SIFT-MS.

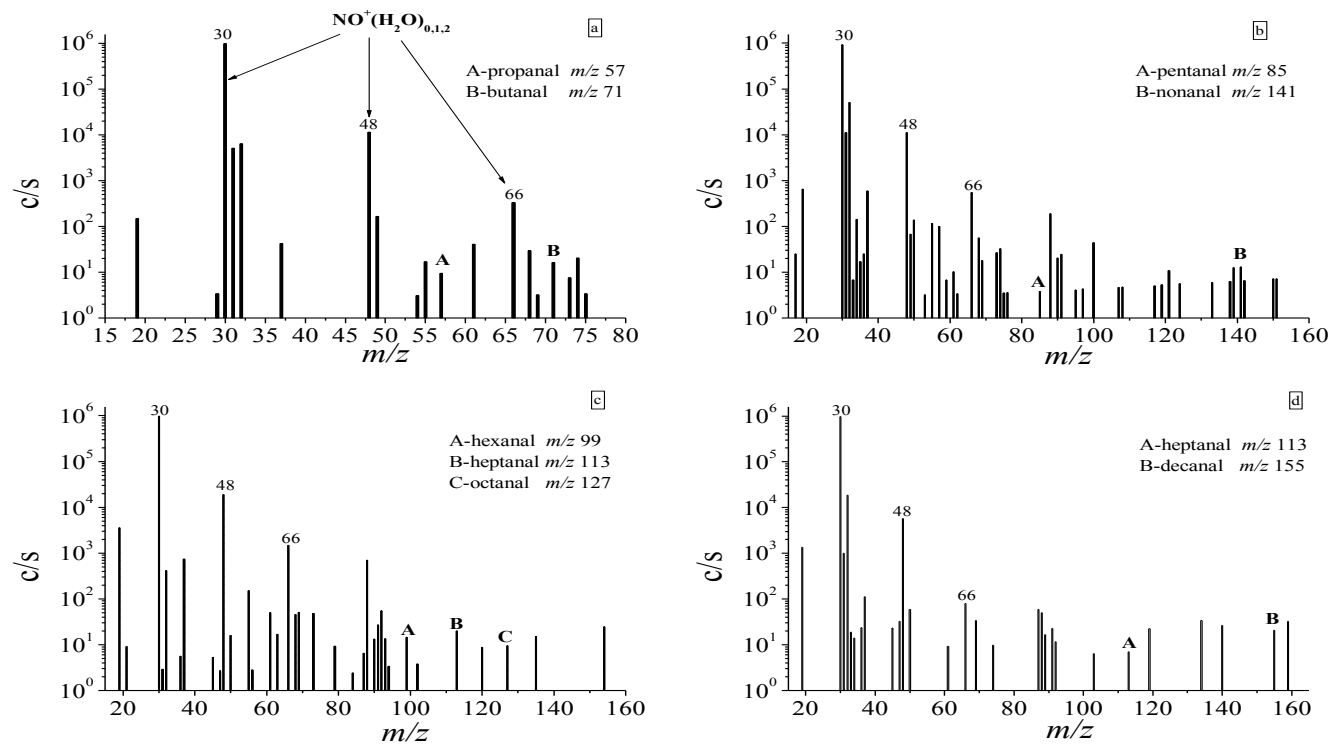


Figure 6.5: Full scan spectrum of exhaled breath obtained using the NO^+ reagent ion. m/z 30, 48 and 66 are the NO^+ reagent ions and its associated hydrate ion. m/z 57 is propanal, m/z 71 is butanal, m/z 85 is pentanal, m/z 99 is hexanal, m/z 113 is heptanal, m/z 127 is octanal, m/z 141 is nonanal and m/z 155 is decanal.

Similarly, figure 6.5 also demonstrates the presence of butanal at m/z 71 ($C_4H_7O^+$), when the exhaled breath was analysed using NO^+ in FS mode. All the other aldehydes were detected from their single product ion using the NO^+ . However, the ion count rate is relatively low (as expected) in comparison to the abundant VOCs (i.e., acetone, methanol, isoprene) within exhaled breath. Hence, the MIM mode remained the preferable method for the accurate quantification of C3-C10 aldehydes.

6.2.4.2 C3 to C10 aldehyde concentrations in exhaled breath of healthy subjects

The repeatability of the sampling method was assessed, through calculating the coefficient of variation (CV%) of the measured concentrations of C3-C10 aldehydes from the four exhaled breath samples collected from ten healthy volunteers. All four breath samples were collected consecutively and immediately stored in a sealed 9-litre Really Useful Box (395 x 255 x 155mm) at room temperature. All breath samples were analysed in the order they had been collected in. Each sample was only removed from the storage container at the time of analysis. The eight aldehydes were observed at trace quantities in the exhaled breath of the 10 healthy persons, and hence their associated CV% values are, as expected, relatively larger in comparison to abundant breath VOCs. As shown in Table 6.6, propanal has the lowest CV of 28.4%. Butanal, pentanal, hexanal, heptanal, octanal and decanal all have CV values in a range of 40% to 60%. Trace exhaled breath VOCs often have a higher coefficient of variation as the exhaled concentrations are also close to the instrument detection limit. For instance, the CV% of exhaled breath HCN has been reported to be around 50% with a maximum of 85%; HCN has been proposed as a potential biomarker for *Pseudomonas aeruginosa* bacterial infection in patients with cystic fibrosis¹⁸¹. However, if a trace VOC is to be considered a suitable volatile breath biomarker for a specific disease, its concentration would need to be sufficiently increased or decreased in comparison to background levels. At the same time, it must also overcome any variation resulting from the sampling process or analytical method employed.

Aldehyde	Median (IQR)	CV%
Propanal	10.8(6.9-18.1)	28
Butanal	2.5(1-3.9)	43
Pentanal	1.9(0.8-2.5)	55
Hexanal	2.3(1.5-4.4)	51
Heptanal	1.5(0.6-3.6)	58
Octanal	1.5(0-2.4)	63
Nonanal	0.9(0.2-2.2)	34
Decanal	0.9(0.3-2.7)	41

Table 6.6: Median concentrations and interquartile range (IQR) (ppbv) of C3 to C10 aldehydes from 26 healthy volunteers. The intra-individual CV% values have been calculated from four exhaled breath samples taken from 10 of the healthy volunteers.

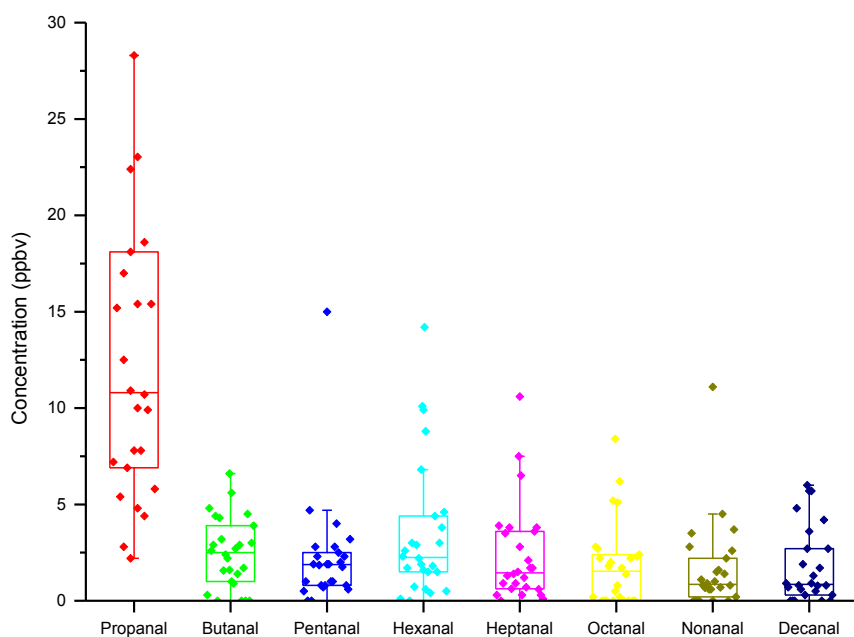


Figure 6.6: Box-Whisker plots of the median concentrations and IQR (ppbv) of C3 to C10 aldehydes within exhaled breath samples from the 26 healthy volunteers.

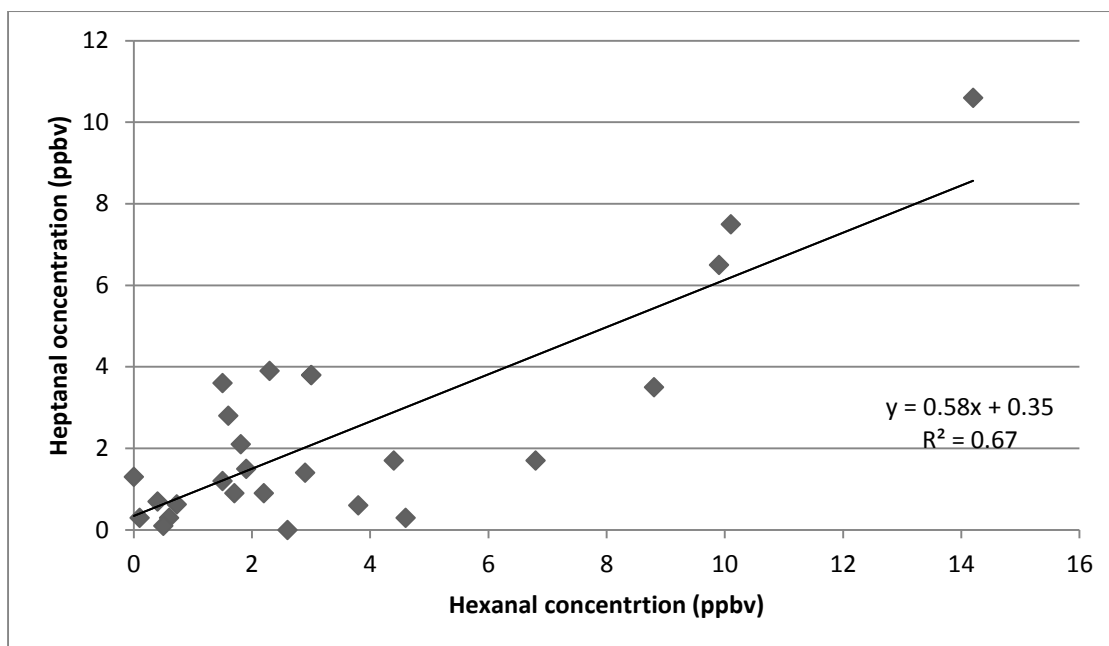


Figure 6.7: Correlation of the measured concentrations of hexanal with heptanal, from the exhaled breath of 26 healthy volunteers.

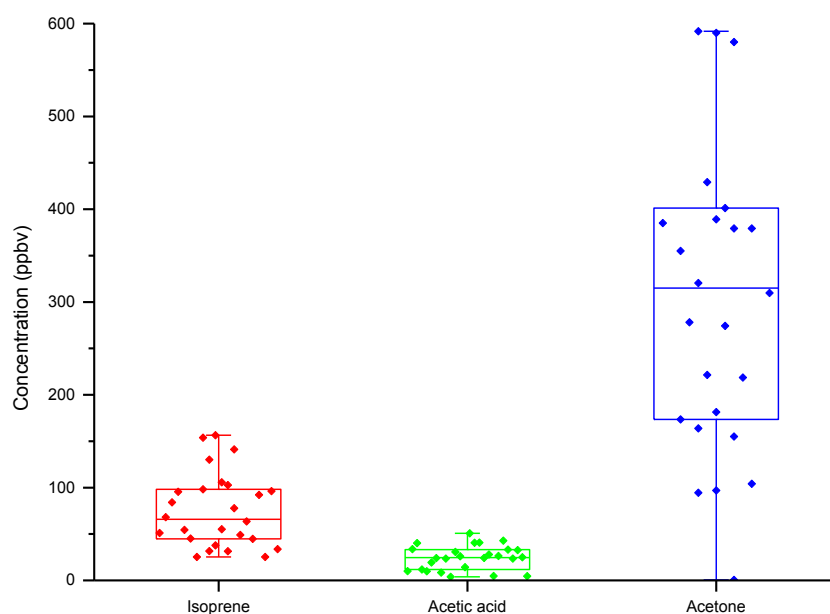


Figure 6.8: Box-Whisker plots of isoprene, acetic acid and acetone from the exhaled breath of 26 healthy subjects.

6.2.5 Discussion

In this study, it is observed that C4-C10 aldehydes are present at trace concentrations in the exhaled breath of healthy subjects using SIFT-MS. The concentrations of exhaled propanal to decanal from 26 healthy volunteers were measured using the MIM mode of SIFT-MS. Propanal was found to be the most abundant amongst the investigated aldehydes, with a median concentration of 10.8 *ppbv* (IQR 6.9-18.1). The kinetics library entry for propanal has accounted for the isotopologue of $(\text{H}_3\text{O}^+)_2$ at m/z 55. In contrast to propanal, saturated aldehydes from butanal to decanal were present at trace concentrations in the exhaled breath of healthy persons. Each of these aldehydes has a median concentration of less than 3 *ppbv*, with a narrow interquartile range, in the exhaled breath of all healthy volunteers. Several aldehydes had concentrations of < 1 *ppbv* recorded within exhaled breath; at these levels it is likely that the limit of detection for the instrument is being reached for these compounds. In contrast to abundant VOCs such as acetone, these aldehydes did not demonstrate a normal or log-normal distribution within a healthy cohort. The median concentration and interquartile range (IQR) for these aldehydes is shown in Table 6.6 & Figure 6.6. Amongst the investigated VOCs, the measured concentrations of hexanal and heptanal were the only aldehydes that demonstrated a reasonable correlation (Figure 6.7). The square correlation coefficient R^2 is 0.67. There was no correlation of age or gender with any of the C3-C10 aldehydes, a finding that previously been reported^{63,200}. There were 4 smokers in this study out of 26 healthy subjects. Using Mann-Whitney U test, no aldehyde demonstrated a significant difference between smokers and non-smokers, although it is noted that there were relatively few smokers in this study cohort. Jareno-Esteban *et al.* have previously reported that nonanal is influenced by smoking status²⁰¹. Fuchs *et al.* reported that the exhaled of propanal, butanal, heptanal and decanal concentrations were not significantly different between smokers and healthy volunteers using GC-MS⁶³.

The Box-Whisker plot of the measured abundant VOCs from the exhaled breath of 26 healthy volunteers is shown in Figure 6.8. The median and IQR concentrations of isoprene, acetic acid and acetone were 66 *ppbv* (IQR 45-97), 25 *ppbv* (IQR 12-33) and 315 *ppbv* (IQR 175-398), respectively. Pysanenko *et al.* have previously reported a concentration range of 30 to 60 *ppbv* for acetic acid, from the exhaled breath of five healthy volunteers analysed using SIFT-MS¹⁵⁸. In a study on exhaled breath samples from 30 healthy volunteers, Turner *et al.* reported a mean concentration of 462 *ppbv* with a range of 148-2744 *ppbv* and 118 ± 68 *ppbv* for acetone and isoprene, respectively^{114,115}. The measured concentrations of acetic acid, acetone and isoprene in this study are in agreement with the previous exhaled breath studies on

healthy volunteers using SIFT-MS, further underlining the reliability of the Nalophan[®] sampling method employed.

Ligor *et al.* have previously analysed the exhaled breath of healthy volunteers using SPME/GC-MS²⁰². Their study demonstrated the presence of propanal, hexanal and heptanal in the exhaled breath of healthy volunteers²⁰². Fuchs *et al.* have demonstrated that C1–C10 aldehydes could be detected in the exhaled breath of all their study subjects which included healthy volunteers, smokers and lung cancer patients using SPME/GC-MS⁶³. Using this method, they reported aldehyde concentrations ranging from 7 pmol/l (161 pptv) for butanal to 71 nmol/l (1,582 ppbv) for formaldehyde. Aldehydes are known to have both mutagenic and carcinogenic potential and hence these compounds warrant further investigation in the exhaled breath of persons with underlying disease to assess their potential as volatile breath biomarkers. Several studies have already demonstrated higher concentrations of aldehydes in respiratory diseases. In the study by Fuchs *et al.*, they found significantly higher concentrations of pentanal, hexanal, octanal and nonanal in the exhaled breath of lung cancer patients compared with smokers and healthy controls⁶³. Corradi *et al.* investigated aldehydes in exhaled breath condensate from patients with chronic obstructive pulmonary disease using liquid chromatography tandem mass spectrometry²⁰³. They found increased levels of malondialdehyde, hexanal, and heptanal in patients as compared with non-smoking control subjects. Bartoli *et al.* also demonstrated significantly higher levels of malondialdehyde in the exhaled breath condensate of those with chronic airways diseases vs. control subjects²⁰⁴. Further studies with large patient numbers are necessary to assess whether the promising aldehyde associations observed in these studies could be applied in mainstream clinical practice.

6.2.6 Conclusions

In this study, C3-C10 aldehydes have been quantified within the exhaled breath of healthy subjects using the SIFT-MS platform. Propanal is the most abundant aldehyde present within the exhaled breath of healthy volunteers. The C4-C10 aldehydes are present at trace quantities in the exhaled breath of healthy subjects at median concentrations of less than 3 *ppbv*. The concentrations of the C3-C10 aldehydes observed in this study provide some guidance with respect to ‘normal values’ or reference ranges in absolutely healthy subjects. Knowledge of such ‘baseline’ concentrations in healthy persons is important, particularly in the context of exhaled breath aldehydes being investigated in those with underlying disease.

6.3 Exhaled breath VOC risk prediction model in oesophageal and gastric adenocarcinoma

6.3.1 Objective

The aims of this study were 1) to validate the exhaled breath VOCs identified in the pilot study in Section 6.1 in a larger patient cohort 2) to identify and quantify C3-C10 aldehydes in patients with oesophageal and gastric adenocarcinoma 3) to develop a VOC-based risk prediction model to distinguish oesophageal and gastric adenocarcinoma from non-cancer controls.

6.3.2 Introduction

The lack of signs and symptoms in the early stages of upper gastrointestinal cancer lead to delays in diagnosis. The results of the National OG Cancer Audit demonstrated that 65% of patients with oesophago-gastric cancer were not eligible for curative treatment due to spread of their disease²⁰. However, many of these patients have soft symptoms in the earliest stages of their cancer. It is very challenging for general practitioners to differentiate patients with early cancer from those with benign causes for their symptoms, without referring all patients for endoscopy. This would not represent a cost-effective strategy in the Western world and hence there remains a significant clinical need to develop novel techniques to identify patients with upper gastro-intestinal cancer.

The results of the pilot study (in Section 6.1) on exhaled breath VOCs identified compounds of interest in OG cancer, following on from the studies on biofluids. A total of 29 VOCs have been investigated in this study including the C3-C10 aldehydes in a large patient cohort. In the cancer group, patients with both oesophageal and gastric adenocarcinoma were recruited for investigation. All patients with gastro-oesophageal junctional tumours (i.e., Siewert Type I-III) were classified within the oesophageal cancer group²⁰⁵. Patients with Barrett's metaplasia were an additional control group that were also recruited. Demographic and potential clinical confounding factors that may affect exhaled breath VOC profiles have also been evaluated.

6.3.3 Methods

6.3.3.1 Study Population

Eligible patients referred for upper gastrointestinal endoscopy were recruited through Imperial College NHS Trust between November 2011 and August 2013. Exhaled breath samples were collected when patients were attending for one of the following investigations - staging laparoscopy and Oesophago-Gastro-Duodenoscopy (OGD), Upper Gastro-intestinal Endoscopic Ultrasound (EUS) or OGD only. The OGD component of the investigations and histological confirmation formed the reference standard for classification of patients within the study groups. Patients with histologically confirmed invasive oesophageal or gastric adenocarcinoma (prior to the commencement of any treatment) and considered for an 'intention to cure' treatment pathway were included in the cancer groups. The non-cancer controls were classified according to endoscopy findings and biopsy results into Barrett's metaplasia, benign disease and a normal upper gastro-intestinal (UGI) tract with negative urease test. Exclusion criteria for this study included patients with squamous cell carcinoma (SCC) of the upper gastro-intestinal (UGI) tract, previously diagnosed liver disease, small bowel/colonic pathology, any non-UGI cancer; and those with any signs/symptoms of acute infection. SCCs of the UGI tract were excluded from this study due to the relatively few number of patients treated with this pathology in our centre. Approval for the study was obtained from the institutional ethics review committee. Fully informed, written consent was obtained from all participants prior to enrolment in the study. Demographic and clinical data were archived in a linked, anonymised database.

6.3.3.2 Procedures

Breath sampling and analysis procedures were conducted using the method described in Section 6.1.3.2. All patients were fasted for a minimum of 6 hours in keeping with the standard protocol for an OGD with no oral intake (including water) prior to their breath sample collection. Mixed alveolar breath samples were collected in secure double thickness (2x25µm) Nalophan® (Kalle UK Ltd., Witham, UK) bags prior to the patients' scheduled clinical investigations. Collected breath samples were taken directly to the Surgical Sciences laboratory and analysed within an hour of collection.

6.3.3.3 SIFT-MS Analysis

The Multiple Ion Monitoring Mode (MIM) of the *Profile-3* SIFT-MS instrument was employed to analyse all exhaled breath samples. The calibration check procedure using the water content of the operator's exhaled breath (6 +/- 0.1%) was conducted prior to each analysis. In this study, an additional check procedure using a Red-y Compact Mass Flow Meter (GCM-A4SA-BN00, Icenta Controls Ltd, Salisbury, UK) was performed. This external flow meter was employed to check the flow rate at the sample inlet of the SIFT-MS instrument prior to each analysis. The Nalophan[®] bag containing the breath sample was directly connected to the sample inlet arm of the SIFT-MS instrument and enclosed within an incubator held at 37 °C for the duration of the analysis. The 29 selected VOCs from exhaled breath samples were analysed for a total of 60 seconds and the measured concentrations were averaged over this analysis time for each VOC. The running order for the VOCs and the reagent ions chosen for analysis of the VOCs is given in Table 6.7. The running order was strictly adhered to for each breath sample analysis.

Sample Run	Ion	Volatile organic compounds
1	H_3O^+	Methanol, Propanol
2	H_3O^+	Butanol, Pentanol, Diethyl ether
3	H_3O^+	Propanoic acid, Buytric acid
4	H_3O^+	Pentanoic acid, Hexanoic acid, Dimethyl disulphide
5	H_3O^+	Dimethyl sulphide, Hydrogen sulphide, Hydrogen cyanide
6	NO^+	Isoprene, Phenol, Acetic acid
7	NO^+	Acetone, Ethyl phenol
9	$\text{O}_2^{+ \cdot}$	Carbon disulphide, Ammonia, Methyl phenol
10	NO^+	Propanal, Butanal, Pentanal, Hexanal
11	NO^+	Heptanal, Octanal, Nonanal, Decanal

Table 6.7: The running order for the analysis of the VOCs and the characteristic ions employed.

6.3.3.4 Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics 20 (Chicago, IL, USA). A P value < 0.05 was considered significant and all statistical tests were two-sided. Mann-Whitney U tests were used to compare the concentrations of the measured VOCs. The comparisons made were: gastric adenocarcinoma versus normal UGI tract, oesophageal adenocarcinoma versus normal UGI tract, gastric adenocarcinoma versus non-cancer controls (benign controls & normal UGI tract), oesophageal adenocarcinoma versus non-cancer controls (Barrett's metaplasia & benign controls & normal UGI tract), gastric adenocarcinoma versus oesophageal adenocarcinoma and finally benign controls versus normal UGI tract. The significantly different VOCs between the aforementioned groups were included in a binary logistic regression model; the accuracies of the models were assessed using the area under the curve (AUC) of the Receiver Operating Characteristic (ROC) curves. The VOCs that demonstrated a statistically significant difference (set at a level of $P < 0.05$) between each cancer group and non-cancer controls were included as variables ROC analysis. The aim of comparing these groups was to assess the feasibility of VOC profiles within exhaled breath to differentiate:

- 1) Gastric adenocarcinoma from those with a normal upper gastro-intestinal tract
- 2) Oesophageal adenocarcinoma from normal upper gastro-intestinal tract
- 3) Gastric adenocarcinoma from non-cancer controls (benign controls & normal upper gastro-intestinal tract)
- 4) Oesophageal adenocarcinoma from non-cancer (Barrett's metaplasia, benign controls & normal upper gastro-intestinal tract)

The linear combination of their concentrations was then compared with a variable threshold in order to construct the ROC curve. Optimum weights of the concentrations of individual VOC providing the maximal area under the curve (AUC) value for the ROC curve were obtained by a binary logistical regression analysis in which the disease conditions were used as the dependent variable and the VOCs were used as the independent variable. Using the linear combination of significant VOCs, the predicted probability (p) obtained from the binary logistical regression analysis is:

$$\ln(p/1-p) = \text{coefficient}_n \times \text{VOC}_n + \text{constant}.$$

Potential confounding factors across the study groups were also evaluated by employing the Kruskal-Wallis test for age and χ^2 test to assess for differences in gender, ischaemic heart disease, diabetes mellitus, respiratory disorders, renal disease, medications, *H. pylori* status, smoking status and alcohol intake. Linear regression models to assess any influence of patient demographic factors, upper gastrointestinal factors and medications on VOC concentrations were also created. Each individual VOC was used as a dependent variable and the possible confounding factors and cancer disease status as the independent variables. Age was continuous, whilst the remaining independent variables were categorical; medications, *H. pylori* status and cancer disease status were binary, either present or absent, smoking status and alcohol intake were ternary (smoker, ex-smoker, non-smoker and no alcohol, within guidance, excess of guidance).

6.3.3.5 Risk Prediction Model

A diagnostic prediction model was created for OG adenocarcinoma versus non-cancer controls using binary logistic regression analysis based on subset of VOCs statistically relevant for discrimination. The cancer group included both oesophageal and gastric adenocarcinoma; this combined group was tested as these cancers are considered comparable subtypes which have frequently been grouped together in neo-adjuvant chemotherapy trials²⁰⁶. The non-cancer group included the Barrett's metaplasia, benign disease and the normal UGI tract controls.

Step 1: a 10-fold cross-validation procedure for all VOCs was performed. Using this analysis, 10% of randomly selected patients are left out at a time and the model is constructed based on the remaining 90% of the data. As part of this model building process, we employed ANOVA analysis in order to select the VOCs relevant for discrimination between the 2 groups based on the *P* value threshold of 0.05. The same procedure is iteratively applied until the 'cancer' versus 'non-cancer' status of each patient has been predicted. Using a 10-fold cross-validation analysis results in 10 separate models being constructed, and a frequency histogram demonstrating the number of times a given VOC has been selected for discrimination using ANOVA is produced.

Step 2: the data was then split through a random number generation process; 2/3 of data were used to create the prediction model (model subset) and 1/3 to test the model (validation subset). The statistically significant VOCs were entered into a logistic regression analysis in a stepwise fashion. The most significant predictors were used to create the diagnostic prediction model. The accuracy of the resultant models was assessed using the area under ROC curve, sensitivity and specificity measures. In order to estimate variability, we performed a Monte Carlo simulation by repeating the aforementioned procedure 10 times with the resultant AUC presented as a mean +/- standard deviation.

6.3.4 Results

6.3.4.1 Patient Characteristics

210 consecutive patients (81 cancer cases and 129 controls) were enrolled into the study during the recruitment period. There were 48 patients with oesophageal adenocarcinoma and 33 patients in the gastric adenocarcinoma group. The control groups had 16 patients with Barrett's metaplasia and, 62 patients classified as benign controls and 51 patients were included in the normal UGI tract group. Patients within the benign group had the following diagnoses: oesophagitis (N=15), oesophageal stricture (N=5), oesophageal candidiasis (N=2), gastritis (N=20), gastric ulcer (N=6), gastro-jejunostomy-induced gastritis (N=2), duodenitis (N=9), duodenal ulcer (N=3). Table 6.8 outlines the demographic and clinical data for the patient groups. Details of the clinical staging of the cancer groups are provided in Table 6.9. The groups were well balanced in terms of baseline characteristics, except for gender. There is a male preponderance in the cancer groups (though not statically significant), which is to be expected given the disease is more common in men. There were no statistically significant differences in potential confounding factors across the groups except for *H.pylori* status; which, as expected, was higher in gastric cancer and benign groups. With regards to medications, aspirin was the only drug that demonstrated a statistical difference across the groups (Table 6.10). All medications which demonstrated a *P* value < 0.2 between the study groups (i.e., Proton Pump Inhibitors (PPI), aspirin, folic acid, salbutamol, salmeterol, Selective Serotonin Reuptake Inhibitors (SSRIs), iron supplements & senna) were further assessed by linear regression analysis.

	Normal upper Gastro-intestinal tract (N = 51)	Benign Controls (N = 62)	Oesophageal Adenocarcinoma (N = 48)	Gastric Adenocarcinoma (N = 33)	Barrett's metaplasia (N = 16)	P value†
Age (IQR)*	61 (50-73)	64.5 (50-72)	63.5 (55.3-72.8)	58 (50.5-70.5)	67 (58.8-72.8)	0.47‡
Gender (M:F)	30:21	38:24	40:8	24:9	11:5	0.07
Smoking						0.31
Non Smoker	33 (64.7)	32 (51.6)	21 (43.8)	14 (42.4)	8 (50.0)	
Ex-Smoker	11 (21.6)	17 (27.4)	17 (35.4)	15 (45.5)	4 (25.0)	
Smoker	7 (13.7)	13 (21.0)	10 (20.8)	4 (12.1)	4 (25.0)	
Alcohol Intake						0.76
Within Guidance	48 (94.1)	58 (93.6)	43 (89.6)	32 (97.0)	15 (93.8)	
Excess	3 (5.9)	4 (6.5)	5 (10.4)	1 (3.0)	1 (6.3)	
PPI/H ₂ Antagonist	26 (51.0)	35 (56.5)	20 (41.7)	18 (54.5)	12 (75.0)	0.20
Diabetes	13 (25.5)	8 (12.9)	4 (8.3)	6 (18.2)	2 (12.5)	0.18
Hypertension/IHD	23 (45.1)	22 (35.5)	23 (47.9)	13 (39.4)	7 (43.8)	0.72
Respiratory disorder	5 (9.8)	8 (12.9)	5 (10.4)	4 (12.1)	2 (12.5)	0.99
Renal Disease	0 (0.0)	2 (3.2)	2 (4.2)	2 (6.1)	1 (6.3)	0.55
<i>H. Pylori</i> Status	0 (0.0)	4 (6.5)	0 (0.0)	4 (12.1)	0 (0.0)	0.07

Table 6.8: Patient characteristics across study groups. *Values are median (IQR), otherwise n (%). † χ^2 test, except ‡ Kruskal-Wallis Test. Percentages are given within the brackets.

Characteristics	Oesophageal Adenocarcinoma (N = 48)	Gastric Adenocarcinoma (N = 33)
Site of tumour:		
Lower Oesophagus	23 (47.9%)	N/A
Gastro-oesophageal junction	25 (52.1%)	N/A
Proximal Stomach	N/A	17 (51.5%)
Distal Stomach	N/A	16 (48.5%)
T Stage:		
T1	11 (22.9%)	4 (12.1%)
T2	7 (14.6%)	11 (33.3%)
T3	29 (60.4%)	12 (36.4%)
T4	1 (2.1%)	6 (18.2%)
Lymph Node Status		
N0	19 (40.0%)	10 (30.3%)
N1	27 (56.3%)	11 (33.3%)
N2	2 (4.17%)	12 (36.4%)
N3	0 (0.0)	0 (0.0)

Table 6.9: Clinical staging of the oesophageal and gastric adenocarcinoma groups.

Medications	Normal upper GI Tract	Benign Controls	Oesophageal cancer	Gastric cancer	Barrett's metaplasia	P values
Proton Pump Inhibitors (Omeprazole/Lansoprazole/Pantoprazole)	26	33	21	18	13	0.14
H ₂ Receptor Antagonist (Ranitidine/Famotidine)	1	4	1	1	1	0.67
Gaviscon	1	1	0	0	0	0.78
Sulfonylurea (Gliclazide/Glibenclamide/Glimperide)	3	4	3	1	1	0.97
Metformin	9	5	4	5	1	0.41
Repaglinide	1	0	0	0	0	0.54
Sitagliptin	1	0	1	0	0	0.69
Insulin	2	1	1	0	0	0.72
Statin(Simvastatin/Atorvastatin/Pravastatin & Rosuvastatin)	19	16	13	5	5	0.27
Ezetimide	1	1	0	0	0	0.78
ACE Inhibitors (Lisinopril, Perindopril, Ramipril, Enalapril)	12	9	6	4	4	0.43
Angiotension II Inhibitors (Irbesartan, Losartan, Candesartan)	4	5	5	1	3	0.45
β-Blocker (Atenolol, Bisoprolol, Metoprolol)	6	10	6	5	3	0.93
Furosemide	3	3	0	0	0	0.25
Dihydropyridine Calcium Channel Blocker (Amlodipine, Nifedipine, Felodipine)	11	11	8	5	3	0.953
Diltiazem	0	2	1	0	0	0.56
Thiazide Diuretics (Bendroflumethiazide, Indapamide)	6	8	2	1	1	0.32
Isosorbide Mononitrate	1	0	0	0	0	0.54

Digoxin	0	1	1	0	0	0.75
Amiodarone	1	0	0	0	0	0.54
Spironolactone	1	1	0	0	1	0.42
Aspirin	12	9	4	2	6	0.01
Clopidogrel	2	1	2	2	2	0.41
Dipyridamole	2	1	0	0	0	0.45
Warfarin	1	1	2	1	0	0.86
B12 Injections	0	2	1	1	0	0.72
Thiamine	2	2	0	0	0	0.47
Folic Acid	2	0	0	0	1	0.14
Salbutamol	0	6	2	2	0	0.15
Salmeterol	0	3	1	3	0	0.17
Tiotropium	1	2	0	1	0	0.73
Prednisolone	2	4	0	0	0	0.20
Levothyroxine	2	2	3	0	1	0.66
Carbimazole	0	1	0	0	0	0.66
Bisphosphonates (Alendronic Acid, Risedronate)	4	3	0	0	1	0.21
ADCAL D ₃	2	2	1	0	1	0.74
Cholecalciferol	1	4	0	1	2	0.37
Calchichew	3	1	0	0	1	0.21
Allopurinol	1	0	0	0	0	0.54
Latanoprost	2	1	0	0	0	0.45
Tramadol	1	1	0	2	0	0.36
Cocodamol/Co-dydramol	3	0	1	0	1	0.20
Paracetamol	1	0	0	1	0	0.51
NSAIDS (Diclofenac/Naproxen)	1	1	0	0	0	0.78
Selective Serotonin Reuptake Inhibitors (SSRIs) (Citalopram/Fluoxetine/Paroxetine/Sertraline)	3	3	3	0	3	0.13
Tricyclic Antidepressants (Amitriptyline, Nortriptyline)	3	1	2	0	0	0.44
Gabapentin	1	2	0	0	0	0.57

Benzodiazepines (Lorazepam/Diazepam/Temazepam)	2	1	2	1	0	0.86
Iron Supplements (Ferrous Sulphate, Ferrous Fumurate)	4	3	0	5	1	0.09
Doxazosin	3	1	2	1	2	0.39
Tamsulosin	0	2	0	0	0	0.31
Mycophenolate	1	0	0	0	0	0.54
Methotrexate	1	0	0	0	0	0.54
Domperidone	1	4	2	0	0	0.41
Metoclopramide	0	0	0	1	0	0.25
Senna	1	0	0	0	1	0.15
Baclofen	1	0	0	0	0	0.54
Dilsufuriam	0	0	1	0	0	0.50
Mebeverine	0	0	0	1	0	0.25

Table 6.10: Assessment of medication use across the study groups

6.3.4.2 Volatile Organic Compound Analysis

A total of 29 VOCs from 8 distinct chemical groups present within exhaled breath (including alcohols, phenols, ketones, fatty acids, aldehydes, sulphur-containing compounds, nitrogen compounds and ether) were detected and quantified. The VOCs and the characteristic ions used for their analysis in MIM mode by SIFT-MS are listed in Table 6.11. The heat map demonstrated higher concentrations in the region of the fatty acids, phenols and aldehydes within both cancer groups (Figure 6.9). Univariate Mann-Whitney U analysis revealed 12 VOCs at significantly higher concentrations in both the gastric and oesophageal adenocarcinoma groups when compared to non-cancer controls. These were pentanoic acid, hexanoic acid, phenol, methyl phenol, ethyl phenol, butanal, pentanal, hexanal, heptanal, octanal, nonanal and decanal (Figures 6.10, 6.11 and 6.12). In addition to these VOCs, butyric acid was significantly higher in the oesophageal adenocarcinoma group when compared to the non-cancer controls. There was no significant difference in the abundant breath VOCs (methanol, acetone, isoprene) and ammonia across the study groups (Figure 6.13 & Table 6.12). Mann-Whitney U tests of all VOCs of oesophageal adenocarcinoma compared with gastric adenocarcinoma did not demonstrate any significant differences. Between the benign controls and the normal UGI tract group, there was a significant difference in nonanal ($P < 0.001$) and decanal ($P = 0.013$). Details of the concentrations of the significant VOCs from each specific group and abundant breath metabolites are given in Table 6.12.

	Molecular formula	Precursor ion	<i>m/z</i>	Characteristic product ions	Reference
Acetic acid	C ₂ H ₄ O ₂	NO ⁺	90, 108	NO ⁺ ·CH ₃ COOH, NO ⁺ ·(H ₂ O)CH ₃ COOH	53
Propanoic acid	C ₃ H ₆ O ₂	H ₃ O ⁺	75, 93	C ₃ H ₆ O ₂ H ⁺ , C ₃ H ₆ O ₂ H ⁺ (H ₂ O)	53
Butyric acid	C ₄ H ₈ O ₂	H ₃ O ⁺	89, 107	C ₄ H ₈ O ₂ H ⁺ , C ₄ H ₈ O ₂ H ⁺ (H ₂ O)	53
Pentanoic acid	C ₅ H ₁₀ O ₂	H ₃ O ⁺	103	C ₅ H ₁₀ O ₂ H ⁺	53
Hexanoic acid	C ₆ H ₁₂ O ₂	H ₃ O ⁺	117, 135	C ₆ H ₁₂ O ₂ H ⁺ , C ₆ H ₁₂ O ₂ H ⁺ (H ₂ O)	53
Propanal	C ₃ H ₆ O	NO ⁺	57	C ₃ H ₅ O ⁺	53
Butanal	C ₄ H ₈ O	NO ⁺	71	C ₄ H ₇ O ⁺	53
Pentanal	C ₅ H ₁₀ O	NO ⁺	85	C ₅ H ₉ O ⁺	53
Hexanal	C ₆ H ₁₂ O	NO ⁺	99	C ₆ H ₁₁ O ⁺	53
Heptanal	C ₇ H ₁₄ O	NO ⁺	113	C ₇ H ₁₃ O ⁺	53
Octanal	C ₈ H ₁₆ O	NO ⁺	127	C ₈ H ₁₅ O ⁺	53
Nonanal	C ₉ H ₁₈ O	NO ⁺	141	C ₉ H ₁₇ O ⁺	53
Decanal	C ₁₀ H ₂₀ O	NO ⁺	155	C ₁₀ H ₁₉ O ⁺	53
Methanol	CH ₄ O	H ₃ O ⁺	33, 51	CH ₃ O ⁺ , CH ₃ O ⁺ (H ₂ O)	53
Propanol	C ₃ H ₈ O	H ₃ O ⁺	43	C ₃ H ₇ ⁺	53
Butanol	C ₄ H ₁₀ O	H ₃ O ⁺	57	C ₄ H ₉ ⁺	53
Pentanol	C ₅ H ₁₂ O	H ₃ O ⁺	71	C ₅ H ₁₁ ⁺	53
Phenol	C ₆ H ₆ O	NO ⁺	94, 112	C ₆ H ₆ O ⁺ , C ₆ H ₆ O ⁺ (H ₂ O)	135
Methyl phenol	C ₇ H ₈ O	O ₂ ⁺	108, 126	C ₇ H ₈ O ⁺ , C ₇ H ₈ O ⁺ (H ₂ O)	135
Ethyl phenol	C ₈ H ₁₀ O	NO ⁺	122, 140	C ₈ H ₁₀ O ⁺ , C ₈ H ₁₀ O ⁺ (H ₂ O)	135
Acetone	C ₃ H ₆ O	NO ⁺	88	NO ⁺ ·C ₃ H ₆ O	53
Dimethyl sulphide	C ₂ H ₆ S	H ₃ O ⁺	63	C ₂ H ₆ SH ⁺	53
Dimethyl disulphide	C ₂ H ₆ S ₂	H ₃ O ⁺	95	C ₂ H ₆ S ₂ H ⁺	53
Hydrogen sulphide	H ₂ S	H ₃ O ⁺	35	H ₃ S ⁺	53
Carbon disulphide	CS ₂	O ₂ ⁺	76	CS ₂ ⁺	53
Ammonia	NH ₃	O ₂ ⁺	17, 35	NH ₃ ⁺ , NH ₃ ⁺ (H ₂ O)	53
Hydrogen cyanide	HCN	H ₃ O ⁺	28	H ₂ CN ⁺	53
Isoprene	C ₅ H ₈	NO ⁺	68	C ₅ H ₈ ⁺	53
Diethyl ether	C ₄ H ₁₀ O	H ₃ O ⁺	75	C ₄ H ₁₀ OH ⁺	53

Table 6.11: The analytical chemistry information of the 29 VOCs, including their molecular formula, mass to charge ratio, the precursor ions, and characteristic product ions used for their analysis by SIFT-MS.



Figure 6.9: Heat map of the discriminatory VOCs from all patients across the study groups. Each column represents one patient (N = 210). The study groups include the normal upper gastro-intestinal tract, the benign disease controls, Barrett's metaplasia, oesophageal cancer and gastric cancer groups. Each value in the unit is the measured concentration (*ppbv*) of the specific VOC. Within the colour scheme the 25th lower quartile, median and the 75th upper quartile of the measured concentrations for specific VOC are defined as pure green, yellow and red, respectively.

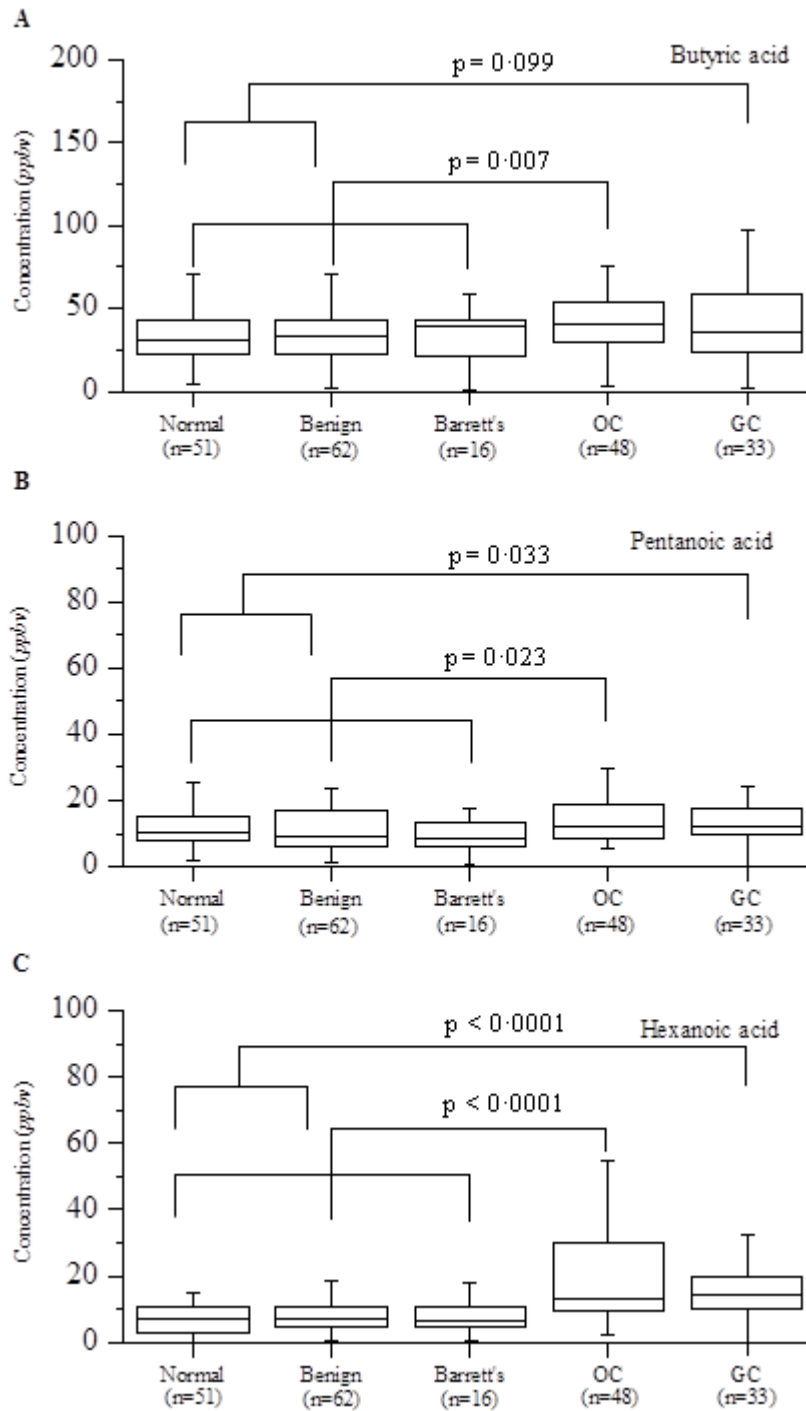


Figure 6.10: Box-Whisker plots of the measured concentrations (*ppbv*) of (A) butyric, (B) pentanoic and (C) hexanoic acids. Normal represents the normal UGI tract group; benign represents the benign disease group; Barrett's is the Barrett's metaplasia group; OC is the oesophageal cancer group and GC is the gastric cancer group.

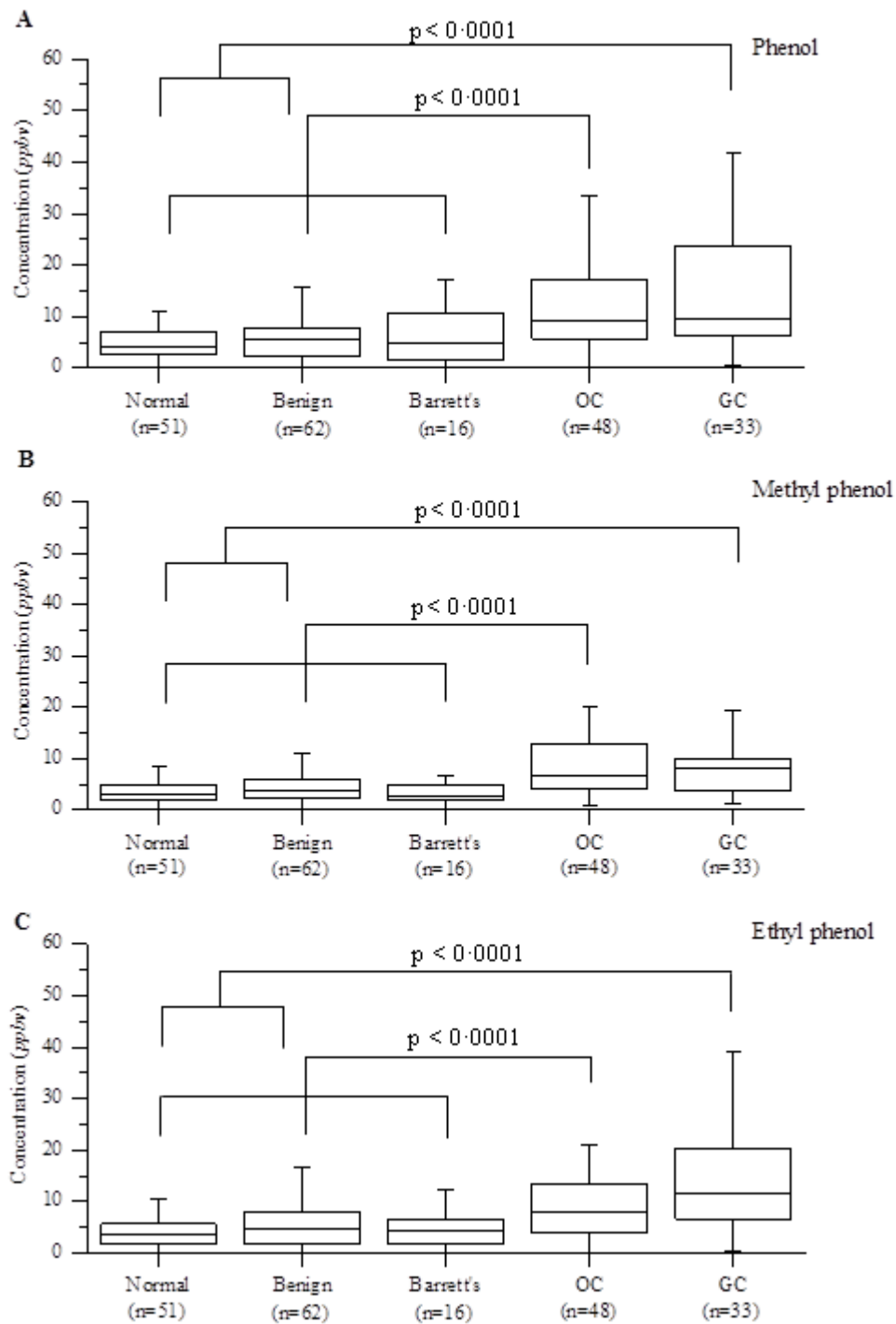


Figure 6.11: Box-Whisker plots of the measured concentrations (ppbv) of (A) phenol, (B) methyl phenol and (C) ethyl phenol. Normal represents the normal UGI tract group; benign represents the benign disease group; Barrett's is the Barrett's metaplasia group; OC is the oesophageal cancer group and GC is the gastric cancer group.

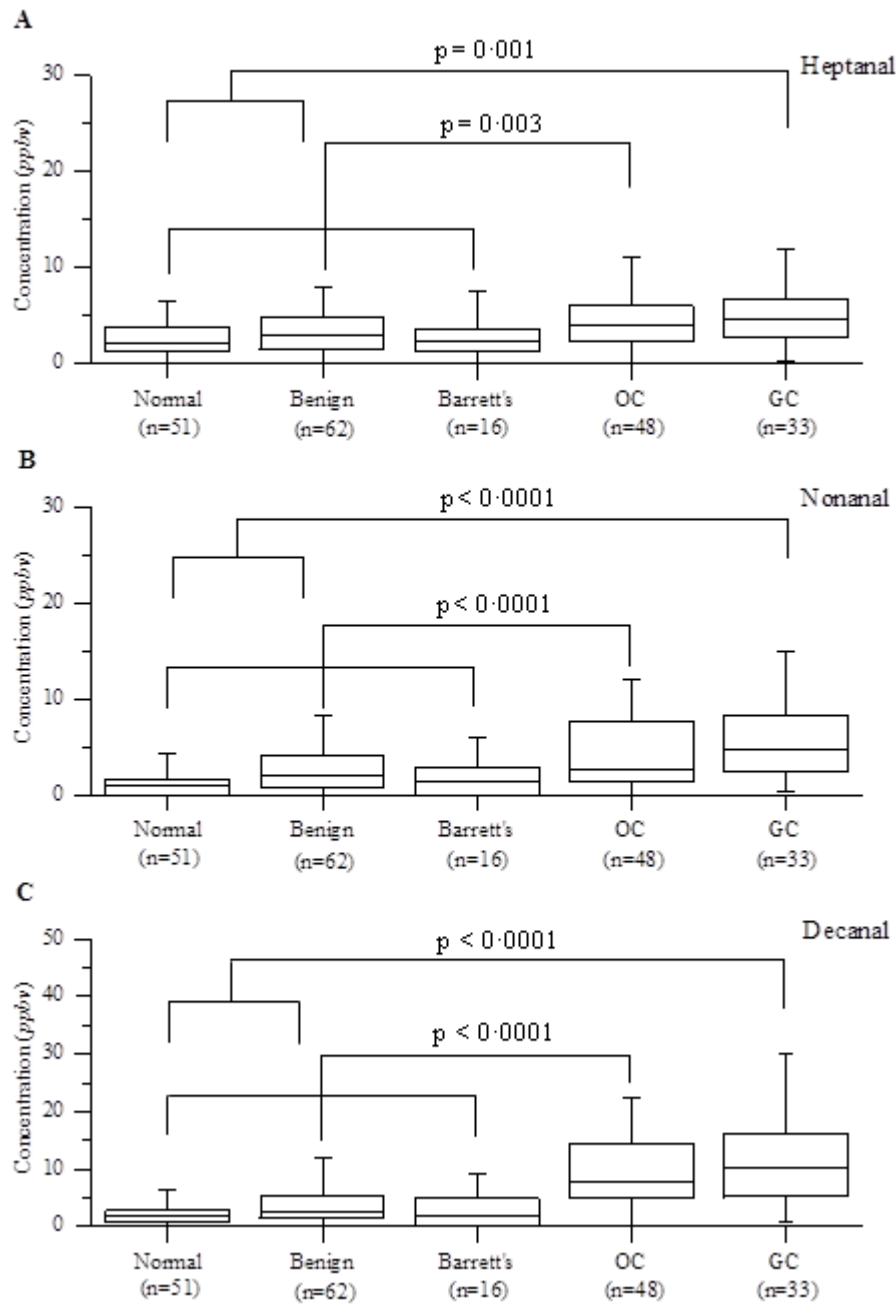


Figure 6.12: Box-Whisker plots of the measured concentrations (ppbv) of (A) heptanal, (B) nonanal and (C) decanal. Normal represents the normal UGI tract group; benign represents the benign disease group; Barrett's is the Barrett's metaplasia group; OC is the oesophageal cancer group and GC is the gastric cancer group.

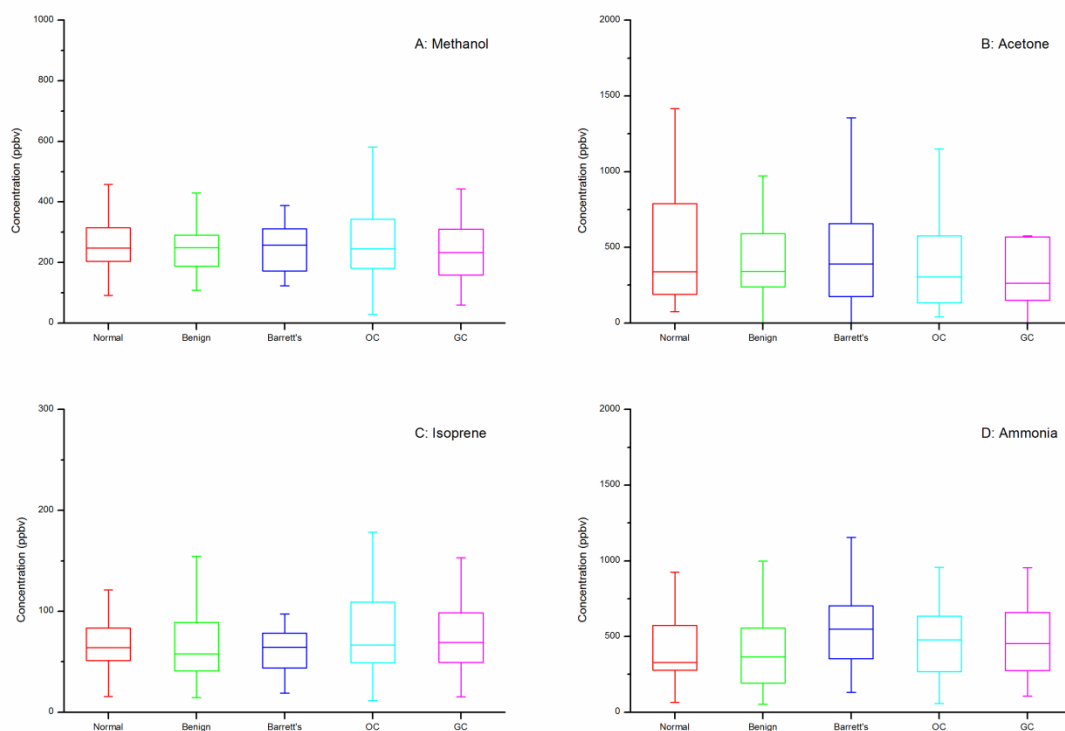


Figure 6.13: Box-Whisker plots of the measured concentrations (*ppbv*) of the abundant breath metabolites - (A) methanol, (B) acetone, (C) isoprene and (D) ammonia. Normal represents the normal UGI tract group; benign represents the benign disease group; Barrett's is the Barrett's metaplasia group; OC is the oesophageal cancer group and GC is the gastric cancer group.

	Gastric cancer		Oesophageal cancer		Barrett's metaplasia		Benign Controls		Normal upper gastro-intestinal tract		P - values (Mann-Whitney U test)		
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Oesophageal cancer vs. non-cancer ¹	Gastric cancer vs. non-cancer ²	Adenocarcinoma ³ versus non-cancer ¹
Butyric acid	36	23-61	40	29-55	40	24-43	33	23-43	31	22-43	0.007	0.10	0.005
Pentanoic acid	12	9-18	12	8-19	9	6-14	9	6-17	10	8-15	0.02	0.03	0.004
Hexanoic acid	14	10-22	13	9-30	6	5-11	7	5-10	7	3-10	<0.001	<0.001	<0.001
Phenol	10	6-25	9	6-18	5	1-11	5	2-8	4	3-7	<0.001	<0.001	<0.001
Methyl phenol	8	3-10	7	4-13	3	2-5	4	2-6	3	2-5	<0.001	<0.001	<0.001
Ethyl phenol	12	7-21	8	4-14	4	2-7	5	2-8	4	2-6	<0.001	<0.001	<0.001
Butanal	4	2-7	4	2-7	3	2-4	3	1-6	2	1-4	0.008	0.04	0.002
Pentanal	2	2-7	3	2-7	2	1-4	2	0-3	2	0-3	<0.001	<0.001	<0.001
Hexanal	5	3-10	6	3-11	5	1-6	4	2-7	3	2-5	<0.001	0.009	<0.001
Heptanal	5	2-7	4	2-6	2	1-4	3	1-5	2	1-4	0.003	0.001	<0.001
Octanal	5	3-9	3	1-7	2	1-5	2	1-3	1	1-2	0.001	<0.001	<0.001
Nonanal	5	2-10	3	1-8	1	0-3	2	1-4	1	0-2	<0.001	<0.001	<0.001
Decanal	10	5-17	8	5-15	2	0-5	3	2-5	2	1-3	<0.001	<0.001	<0.001
Methanol	233	157-333	245	180-344	258	171-317	248	187-291	248	204-315	0.87	0.65	0.89
Acetone	262	146-573	305	133-588	389	157-674	341	233-596	338	190-789	0.17	0.17	0.09
Isoprene	69	49-100	67	49-109	64	41-79	58	41-89	64	51-83	0.18	0.38	0.14
Ammonia	452	275-687	478	266-635	549	346-705	367	190-570	330	277-575	0.21	0.11	0.11

Table 6.12: The statistically significant VOCs and abundant breath metabolites across the study groups. Median and Interquartile range (IQR) are measured by parts-per-billion by volume (ppbv). 1: the non-cancer controls comprise of the normal upper gastro-intestinal tract, benign and Barrett's metaplasia groups. 2: the non-cancer controls comprises of normal upper gastro-intestinal tract and benign groups. 3: the carcinoma group comprises gastric adenocarcinoma and oesophageal adenocarcinoma.

ROC analysis applying the significant VOCs identified in univariate analysis gave an area under the curve (AUC) for discriminating gastric cancer and oesophageal cancer from the normal UGI tract group of 0.98 and 0.97 respectively (Figure 6.14). The AUC for the ROC curves to discriminate gastric cancer and oesophageal cancer from non-cancer controls were 0.92 and 0.90 respectively (Figure 6.15). A summary of the ROC analysis, sensitivity and specificity of discriminating cancer groups from controls is given in Table 6.13. Detailed information of the coefficients and constants derived from the binary logistic regression equation for the ROC analysis is given in Appendix C.

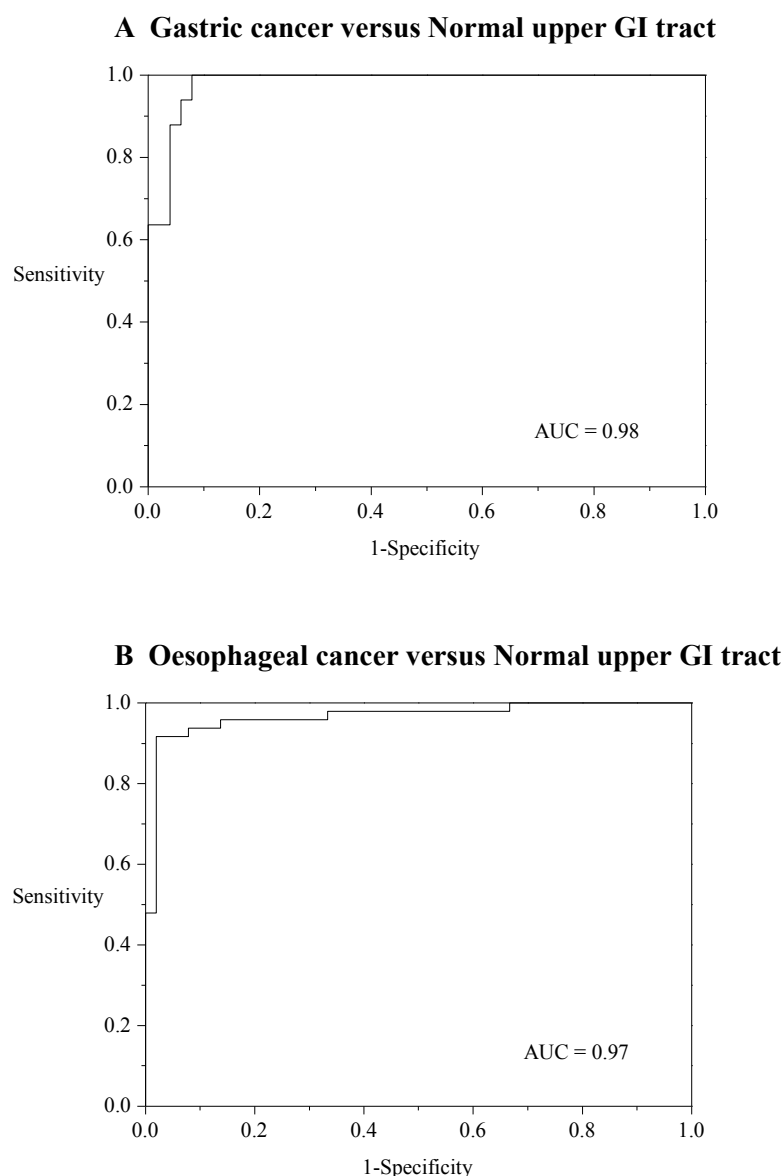


Figure 6.14: ROC curves for (A) gastric and (B) oesophageal cancer versus normal upper GI tract

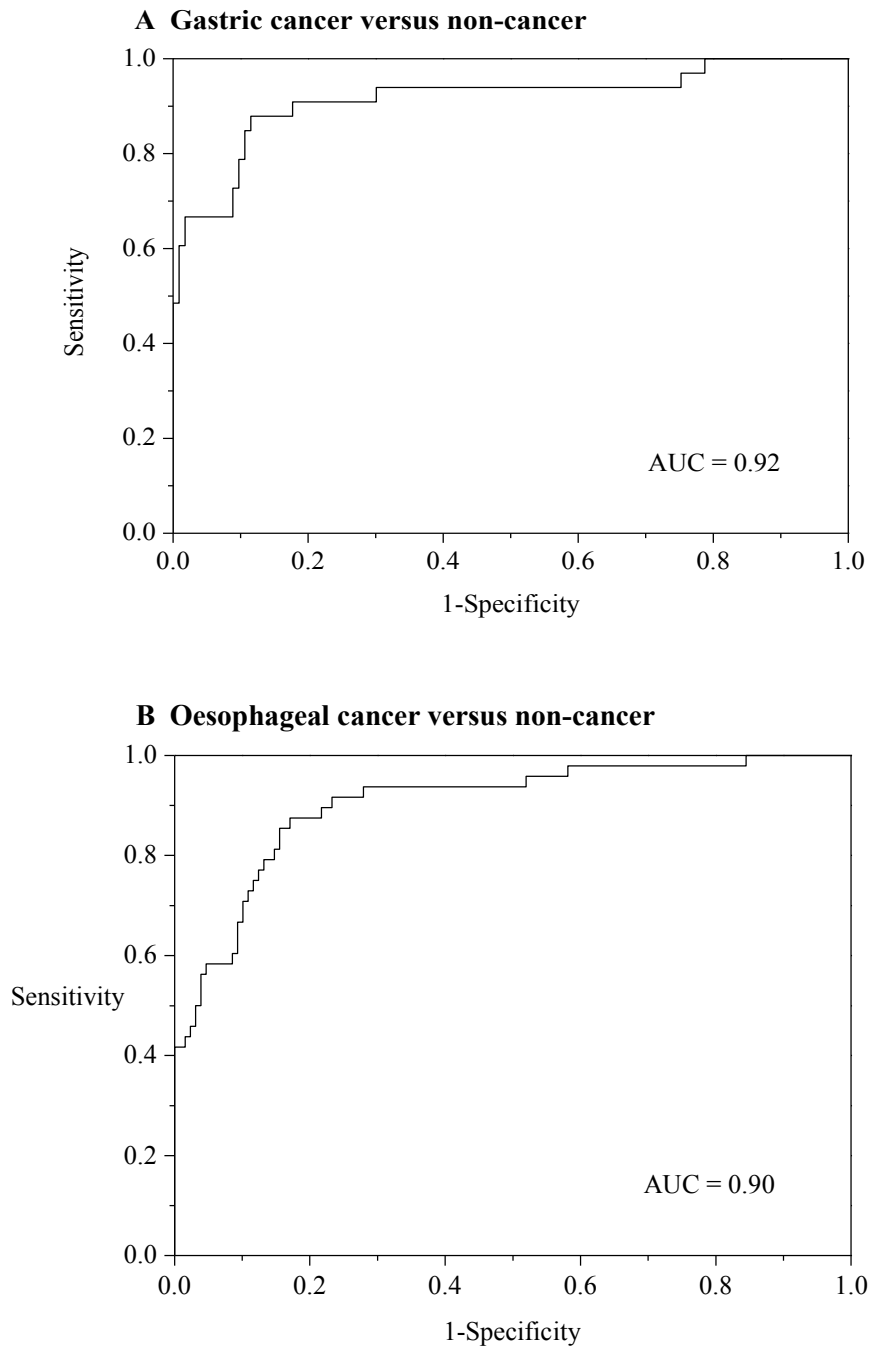


Figure 6.15: ROC curves for (A) gastric and (B) oesophageal cancer versus all non-cancer controls.

Table 6.13: Summary of ROC analysis, sensitivity and specificity of discriminating cancer groups from control groups.

	Gastric Adenocarcinoma versus normal upper gastro- intestinal tract	Oesophageal Adenocarcinoma versus normal upper gastro- intestinal tract	Gastric Adenocarcinoma versus non-cancer control¹	Oesophageal Adenocarcinoma versus non-cancer control²	Diagnostic prediction model Model subset	Diagnostic prediction model Validation subset
AUC	0.98	0.97	0.92	0.90	0.92 ± 0.01	0.87 ± 0.03
95% CI	0.96-1.00	0.93-1.00	0.85-0.98	0.85-0.96	N/A	N/A
Optimum probability	0.35	0.55	0.20	0.21	N/A	N/A
Sensitivity (%)	100%	98%	87.9%	87.5%	89.3%	86.7%
Specificity (%)	92.2%	91.7%	88.5%	82.9%	83.7%	81.2%

^{1.} Benign control and normal upper gastro-intestinal tract.

^{2.} Benign control, Barrett's metaplasia and normal upper gastro-intestinal tract.

^{3.} Gastric adenocarcinoma and oesophageal adenocarcinoma

The linear regression model demonstrated cancer disease status to be the strongest predictor of high concentrations for all the significant VOCs; disease status did not correlate with known abundant exhaled breath VOCs except isoprene. None of the patient demographic factors, *H. pylori* status, alcohol intake, smoking status or medications were independent predictors for any significant VOC. *H. pylori* status and PPI use were independent predictors for breath ammonia concentrations, which further emphasises the reliability of the sampling methodology and sensitivity of the analytical technique employed in this study. Detailed information of the binary logistic regression analysis for the potential confounding factors is given in Appendix D. The potential exogenous VOC contamination within the experimental methodology has also been investigated; all potential VOC biomarkers were present at trace concentrations or at the detection limit of the instrument within laboratory room air and the synthetic air-filled Nalophan[®] bags in comparison to the concentrations observed within exhaled breath. These experiments confirmed that the ambient air trace compounds and the sampling bag surface emissions have negligible overall contributions to the measured VOCs (Table 6.14).

VOCs	Room air		Nalophan® Bag	
	Median	IQR	Median	IQR
Butyric acid	1.3	1-4.3	3.2	1.7-4.1
Hexanoic acid	1	0-3	1	1-2
Phenol	2	1-3	1	1-1.4
Methyl phenol	1	0-2	1	1-3
Ethyl phenol	0	0-2	2	1-2
Butanal	1	0-1.3	0.5	0.2-0.6
Pentanal	0.8	0-1.3	0.5	0.2-1.0
Hexanal	0.8	0-1.3	0.9	0.6-1.4
Heptanal	0	0-0.3	1.1	0.8-1.6
Octanal	0.9	0-1.8	0.6	0.4-1.0
Nonanal	0.7	0-1	0.5	0.2-0.9
Decanal	0.5	0-1.3	0.7	0.7-1.0
Methanol	38	30-50	29	22-37
Acetone	6	3-10	1	0-2
Ammonia	75	30-200	50.4	21.2-84.3
Isoprene	3	1-5	2	1-3

Table 6.14: Room air and Nalophan® bags (inflated with dry synthetic air) measurements of each VOC. Median and Interquartile range (IQR) concentrations in parts-per-billion by volume (*ppbv*) have been calculated from ten measurements of laboratory room air taken on ten separate days; the measurement for the Nalophan® bag contribution was also calculated by applying the same method.

6.3.4.3 Risk Prediction Model

A diagnostic prediction model to discriminate OG adenocarcinoma (N = 81) from non-cancer controls (including Barrett's metaplasia, benign group & normal UGI tract N = 129) was constructed. The frequency histogram of the VOCs identified as discriminators between cancer and non-cancer is provided (Figure 6.16). Noteworthy, is that the significant VOCs identified by both univariate and multivariate analysis all came from the same chemical groups - fatty acids, phenols and aldehydes. The eight most significant predictors for adenocarcinoma identified from stepwise logistic regression were: butyric acid, hexanoic acid, phenol, ethyl phenol, methyl phenol, heptanal, nonanal and decanal. After being subjected to Monte Carlo simulation, the AUC curve for the model subset was 0.92 ± 0.01 with a sensitivity of 89.3% (95% CI, 77.0 – 95.7) and specificity of 83.7% (95% CI, 74.5 – 90.9). The AUC for the validation subset was 0.87 ± 0.03 , with a sensitivity of 86.7% (95% CI, 71.6 – 97.6) and specificity of 81.2% (95% CI, 64.0 – 90.0) (Table 6.13 and Figure 6.17). The model subset, on average, correctly classified 120 out of 139 patients (diagnostic accuracy $86.0\% \pm 0.02$) with 6 false negatives. The validation subset, on average, correctly classified 59 out of 71 patients (diagnostic accuracy $83.3\% \pm 0.04$), with 3 false negatives.

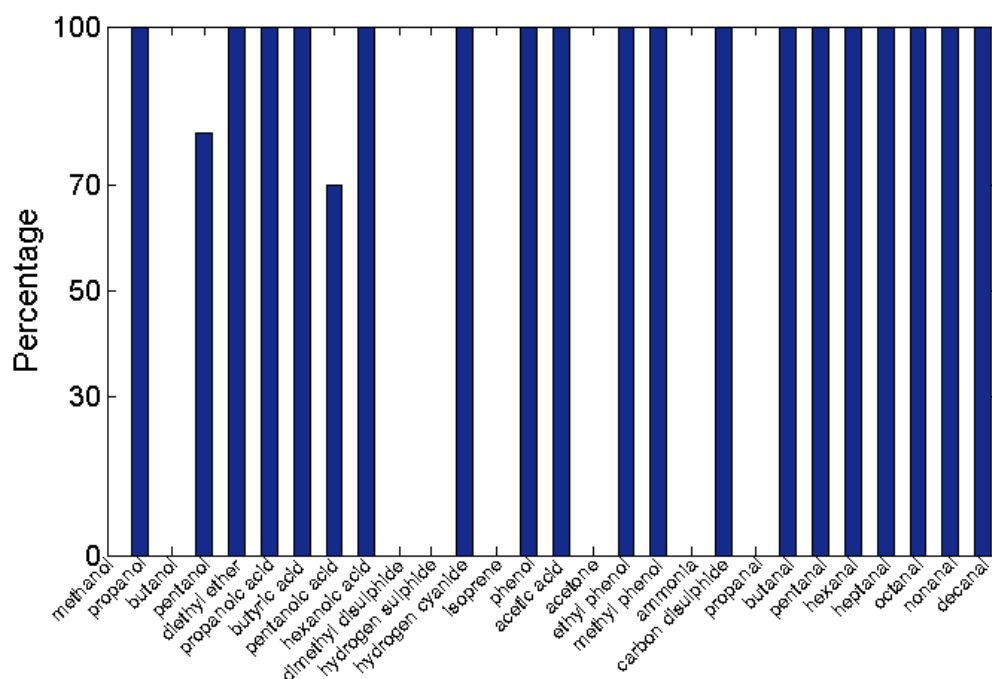


Figure 6.16: Frequency histogram of VOCs used in the diagnostic model to discriminate OG adenocarcinoma from non-cancer controls.

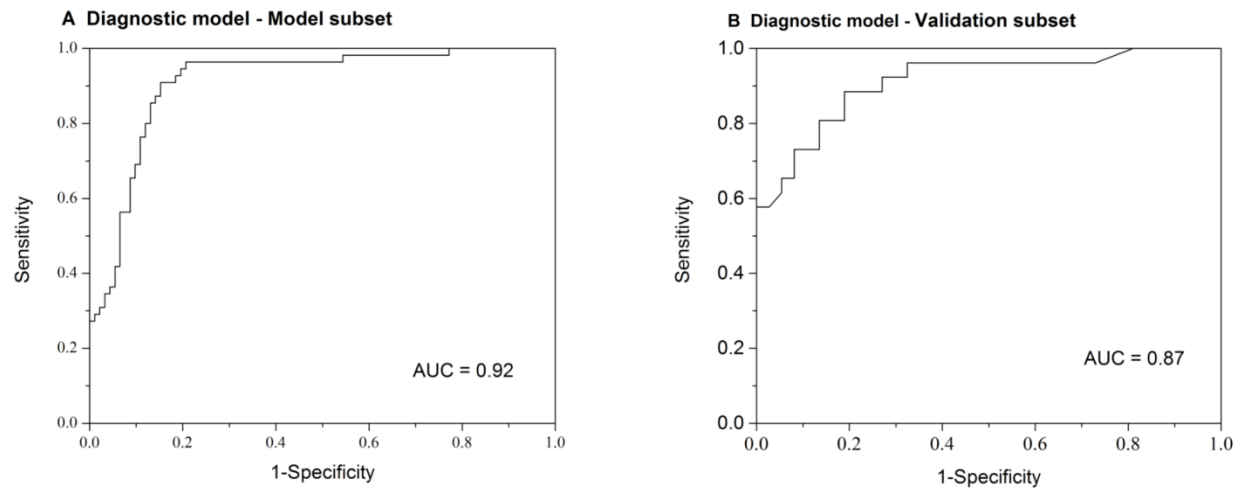


Figure 6.17: ROC curves for (A) Model subset and (B) Validation subset of the diagnostic prediction model to discriminate upper gastro-intestinal adenocarcinoma from non-cancer controls.

6.3.5 Discussion

In this study, increased concentrations of fatty acid, phenol and aldehyde VOCs have been observed in patients with oesophageal and gastric adenocarcinoma. The diagnostic prediction model based on butyric acid, hexanoic acid, phenol, ethyl phenol, methyl phenol, heptanal, nonanal and decanal differentiated cancer from controls with good diagnostic accuracy and with relatively few false negatives. The significant VOCs (i.e., phenol, methyl phenol, ethyl, hexanoic acid) previously identified (Section 6.1) have all replicated the same trend in this larger study on patients with oesophageal and gastric adenocarcinoma. The control groups recruited are also representative of a typical upper GI patient population, many of whom present with non-specific upper gastro-intestinal symptoms; in this group, it is very difficult to reliably identify those at highest risk of an upper gastrointestinal malignancy without referring the majority of such patients for endoscopy. The results of this study demonstrate that quantification of exhaled breath VOCs has the potential for development as a diagnostic test for oesophageal and gastric adenocarcinoma.

The most significant VOC predictors for oesophageal and gastric adenocarcinoma all came from the fatty acid, phenol and aldehyde groups. These compounds were all observed at varying concentrations in the earlier studies on the headspace of gastric content and urine of patients with oesophago-gastric cancer (Chapters 4 and 5). The upregulation of fatty acids has been reported in both oesophageal cancer tissue and in the serum of patients with colorectal cancer^{207,208}. Proteolytic fermentation of tyrosine results in the production of phenols with dysfunction in these pathways being linked to carcinogenesis¹⁹⁴; high levels of tyrosine have previously been reported in the gastric juice of patients with gastric cancer and in oesophageal cancer tissue^{126, 207}. Increased concentrations of aldehydes have been reported in patients with lung and liver cancer within exhaled breath and blood, respectively^{63,74}. Similarly, in this study, we have observed increased concentrations for aldehydes within exhaled breath in oesophageal and gastric cancer. Elevated aldehyde dehydrogenase activity has been demonstrated as a function of oesophageal cancer tumorspheres and aldehydes have also been shown to be metabolised by human hepatocellular carcinoma cells^{209,210}. The majority of VOCs are thought to be of systemic origin; the concentrations of these VOCs in exhaled breath are most likely related to levels observed in blood¹⁸³. The higher concentrations of specific VOCs in the cancer cohort may be a result of the underlying malignancy or the host response to the disease; future studies will need to explore the biochemical mechanisms responsible for this.

The confounding factors, co-existing illnesses and medications that may affect exhaled breath VOC profiles in this cancer patient population have been extensively evaluated. The exclusion criteria were adopted to eliminate any potential influence from other intra-abdominal conditions on the investigated exhaled breath VOCs. Linear regression analysis of upper gastrointestinal factors, smoking status, alcohol intake and medications did not demonstrate any influence on the significant volatile breath biomarkers. *H. pylori* status predicted for increased concentrations of breath ammonia, irrespective of disease status; this underlines the reliability of the breath sampling technique and sensitivity of the VOC analytical method employed. Ulanowska *et al.* have previously analysed exhaled breath samples from patients with *H. pylori* and healthy controls using SPME/GC-MS²¹¹. They observed isobutane, 2-butanone and ethyl acetate in the exhaled breath of *H-pylori*-infected patients and from cultures of *H.pylori* strains; they did not observe the presence of any fatty acids, phenols or aldehydes, indicating that their increased concentrations in this study are a primarily cancer-related effect²¹¹. The concentrations of abundant breath metabolites (i.e., acetone, methanol, isoprene and ammonia) did not demonstrate any significant differences between the cancer patients and non-cancer controls. On regression analysis, isoprene though did correlate with cancer disease status. Isoprene is a known product of lipid peroxidation and hence a surrogate for oxidative stress, but it is observed in the exhaled breath of all healthy people except neonates^{187,212}. Exhaled breath isoprene has previously been reported at increased concentrations in patients with gastric cancer and in head & neck squamous cell carcinomas^{191,213}. Isoprene levels have also been observed at increased concentrations intra-operatively and to subsequently reduce significantly after surgical resection of lung cancer²¹⁴. The elevated levels of isoprene represent are likely to be a systemic oxidative stress response and are not specific to oesophageal and gastric adenocarcinoma. Comparison of early (Stage I) and late stages (Stage II & III) within the cancer groups did not demonstrate any significant VOC differences. However, all the recruited cancer patients had localized disease with no evidence of distant or peritoneal spread; this may account for the minimal differences observed between cancer stages. Phillips *et al.* have previously reported no differences in VOC concentrations between cancer stages in their study investigating exhaled breath in lung cancer⁶⁰.

The key advantage of the breath collection method that has been employed for these studies is that samples are easily acquired using a non-invasive approach; this gives it a high acceptability rating amongst patients. Sample processing (without any need for pre-analysis preparation) and acquisition of results takes ~ 15 minutes, which is advantageous in the clinical setting. The

SIFT-MS technique has allowed accurate quantification of VOCs within human biological samples down to 1 part-per-billion by volume. In this study, the clinical classification of the patients is also rigorous given everyone underwent an OGD on the day of sampling with subsequent histological confirmation. All cancer patients recruited into this study were on an ‘intention to cure’ pathway (prior to starting chemotherapy treatment) and hence the results are applicable to a patient population that would benefit from curative surgery. An important limitation of the study is that further studies are required to evaluate the potential influence of other gastro-intestinal diseases as well as non-GI cancers on the volatile breath biomarkers. Investigation of exhaled breath VOCs in squamous cell carcinoma will also be necessary to give global relevance for any potential diagnostic test.

6.3.6 Conclusions

In conclusion, fatty acid, phenol and aldehyde VOCs have been identified as potential breath biomarkers in oesophageal and gastric adenocarcinoma. These VOCs represent suitable chemical targets that could be used in the development of a point-of-care sensor-based device following the same concept as breathalysers for blood alcohol quantification. Following this principle, a VOC-based breath test could be employed in the primary care setting to risk-stratify patients with non-specific upper gastro-intestinal symptoms, identifying those at highest risk of oesophageal and gastric cancer and hence in need of urgent endoscopy. OGD will remain the 'gold standard' procedure for diagnosis of upper gastro-intestinal malignancy given it allows both direct visualization and tissue biopsy. Any VOC-based test would be utilised as a diagnostic adjunct to endoscopic investigation, being employed in a screening capacity. However, development of VOCs for this purpose would also need to take account of the World Health Organisation's guidelines on screening²¹⁵:

1. The condition should be an important health problem.
2. There should be a treatment for the condition.
3. Facilities for diagnosis and treatment should be available.
4. There should be a latent stage of the disease.
5. There should be a test or examination for the condition.
6. The test should be acceptable to the population.
7. The natural history of the disease should be adequately understood.
8. There should be an agreed policy on whom to treat.
9. The total cost of finding a case should be economically balanced in relation to medical expenditure as a whole.
10. Case-finding should be a continuous process, not just a "once and for all" project

If these criteria could be satisfied, this translational approach has the potential to improve the pick-up rate for oesophageal and gastric cancer. Moreover, if a higher proportion of 'treatable' cancer cases are identified, this would lead to an increased number of patients being managed on an intention to cure basis and potentially improve oncological survival rates in the long term. A secondary benefit could also be a resultant decrease in the negative OGD rate and a reduction in the number of patients undergoing unnecessary invasive procedures. Thus, more studies with large patient cohorts will be necessary to evaluate exhaled breath VOCs for their applicability as potential cancer diagnostic tools.

6.4 The effect of acute infection on cancer-related VOCs

6.4.1 Objective

To investigate whether the significant exhaled breath VOCs in oesophageal and gastric cancer are elevated in the presence of acute infection.

6.4.2 Introduction

Community acquired pneumonia and urinary tract infections (UTI) are medical conditions that frequently necessitate hospital admission. The incidence of both infections continues to rise in line with increases in our aging population. The most common organisms responsible for community acquired bacteraemias include *E.coli*, *Staphylococcus aureus* and *Streptococcus Pneumoniae*²¹⁶. Kunze *et al.* investigated *E.coli* cultures using multi-capillary column-ion mobility spectrometry (MCC-IMS); they reported that decan-1-ol, ethanol, indole and octan-1-ol signals were important in differentiating *E.coli* from the Lysogeny Broth growth medium in which the bacterium was grown²¹⁷. Filipiak *et al.* investigated VOC release from the headspace of *Staphylococcus aureus* and *Pseudomonas aeruginosa* cultures²¹⁸. They observed increased concentrations of acetic acid, 3-methylbutanal, 2-methylpropanal, acetaldehyde and methyl-2-butanal in the headspace of *Staph. aureus* cultures²¹⁸. It was also reported that propanal, butanal, octanal and nonanal were taken up by *Pseudomonas aeruginosa* cultures²¹⁸. Julák *et al.* analysed bronchoalveolar lavage samples from patients with bacterial pneumonia using SIFT-MS and GC-MS²¹⁹. They reported acetone and methanol levels of 1800-7200 ppbv and 170-1000 ppbv, respectively²¹⁹. The above studies identified potential VOCs that may be of importance in the non-invasive diagnosis of infectious pathogens. The main aim of this study was to assess the concentrations of fatty acids, phenols and aldehydes within the exhaled breath of persons admitted to hospital with community acquired infections and UTIs.

6.4.3 Methods

Patients were recruited for the study through Imperial College Healthcare NHS Trust, London after a diagnosis of community acquired pneumonia or urinary tract infection (UTI) was reported by an independent physician. Fully informed, written consent was obtained from all participants prior to enrolment in the study. Exhaled breath samples were collected from every patient within the first 24 hours of their admission to hospital. Samples of mixed alveolar breath were collected into Nalophan[®] bags employing the exact method described in section 6.1.3.2. All breath samples were analysed within 60 minutes of collection. SIFT-MS analysis of the exhaled breath samples was conducted using the MIM mode. The twelve VOCs present at significantly higher concentrations in patients with gastric and oesophageal adenocarcinoma (i.e., pentanoic acid, hexanoic acid, phenol, methyl phenol, ethyl phenol, butanal, pentanal, hexanal, heptanal, octanal, nonanal and decanal) were analysed in this study. In addition to these VOCs, concentrations of the abundant breath metabolites (i.e., methanol, acetone, isoprene and ammonia) were also recorded. Comparative analysis is made with the OG cancer cohorts and the upper gastro-intestinal control groups from section 6.3.3.1.

6.4.4 Results

Ten patients with community acquired pneumonia (7 male: 3 female) and 10 patients with UTI (2 male: 8 female) were recruited for this pilot study. The mean ages were 63.8 ± 22.4 years and 52.1 ± 19.6 years in the pneumonia and UTI groups, respectively. The median admission day C-reactive protein (CRP) levels were 217 (IQR 103-273) and 157 (IQR 119-274) in the pneumonia and UTI groups, respectively. *Streptococcus pneumoniae* growth was confirmed on blood culture in 3 of the pneumonia patients. *E.coli* was the most frequently cultured organism in the UTI group (3 out of 10 patients). Sputum, urine and blood cultures from patients admitted with pneumonia or UTI do not always confirm growth of a specific organism. Huijskens *et al.* reported that sampling strategy may influence diagnostic yield of bacterial pathogens in community acquired pneumonia²²⁰. They reported detection of a pathogen in 49.6% cases of pneumonia when employing conventional culture techniques²²⁰.

Off-line bag sampling was observed as an important methodology in patients with pneumonia, as many of them would have been unable to achieve the necessary Forced Expiratory Volume (FEV1) required to record accurate on-line measurements. The median concentrations and interquartile ranges of the same compounds in the community acquired pneumonia and UTI groups are shown in Table 6.15. Box and Whisker plots of the median concentrations and interquartile range (in *ppbv*) of the 12 VOCs are shown in Figures 6.18, 6.19, 6.20 and 6.21. Comparison of the pneumonia & UTI cohorts is made with the groups from the previous study. In the Box and Whisker plots, the x-axis labels refer to the following - Normal represents the normal UGI tract group; benign represents the benign Upper GI disease group; Barrett's is the Barrett's metaplasia group; OC is the oesophageal cancer group and GC is the gastric cancer group.

	Gastric cancer		Oesophageal cancer		Community Acquired Pneumonia		Urinary Tract Infection	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Pentanoic acid	12	9-18	12	8-19	9	6-13	10	7-12
Hexanoic acid	14	10-22	13	9-30	8	6-11	8	7-9
Phenol	10	6-25	9	6-18	6	3-9	5	2-8
Methyl phenol	8	3-10	7	4-13	6	4-7	3	1-7
Ethyl phenol	12	7-21	8	4-14	6	4-8	3	2-8
Butanal	4	2-7	4	2-7	5	3-6	2	2-5
Pentanal	2	2-7	3	2-7	3	2-4	2	1-4
Hexanal	5	3-10	6	3-11	4	2-6	3	1-5
Heptanal	5	2-7	4	2-6	3	1-5	3	2-3
Octanal	5	3-9	3	1-7	2	1-4	2	1-3
Nonanal	5	2-10	3	1-8	2	1-3	3	2-5
Decanal	10	5-17	8	5-15	4	3-5	2	1-2
Methanol	233	157-333	245	180-344	176	154-289	296	226-397
Acetone	262	146-573	305	133-588	373	190-570	534	401-1246
Ammonia	452	275-687	478	266-635	437	308-477	277	236-402
Isoprene	69	49-100	67	49-109	84	70-110	129	92-163

Table 6.15: Concentrations of fatty acids, phenols, aldehydes and abundant breath VOCs comparing oesophageal and gastric cancer with levels observed in the community acquired pneumonia and urinary tract infection groups.

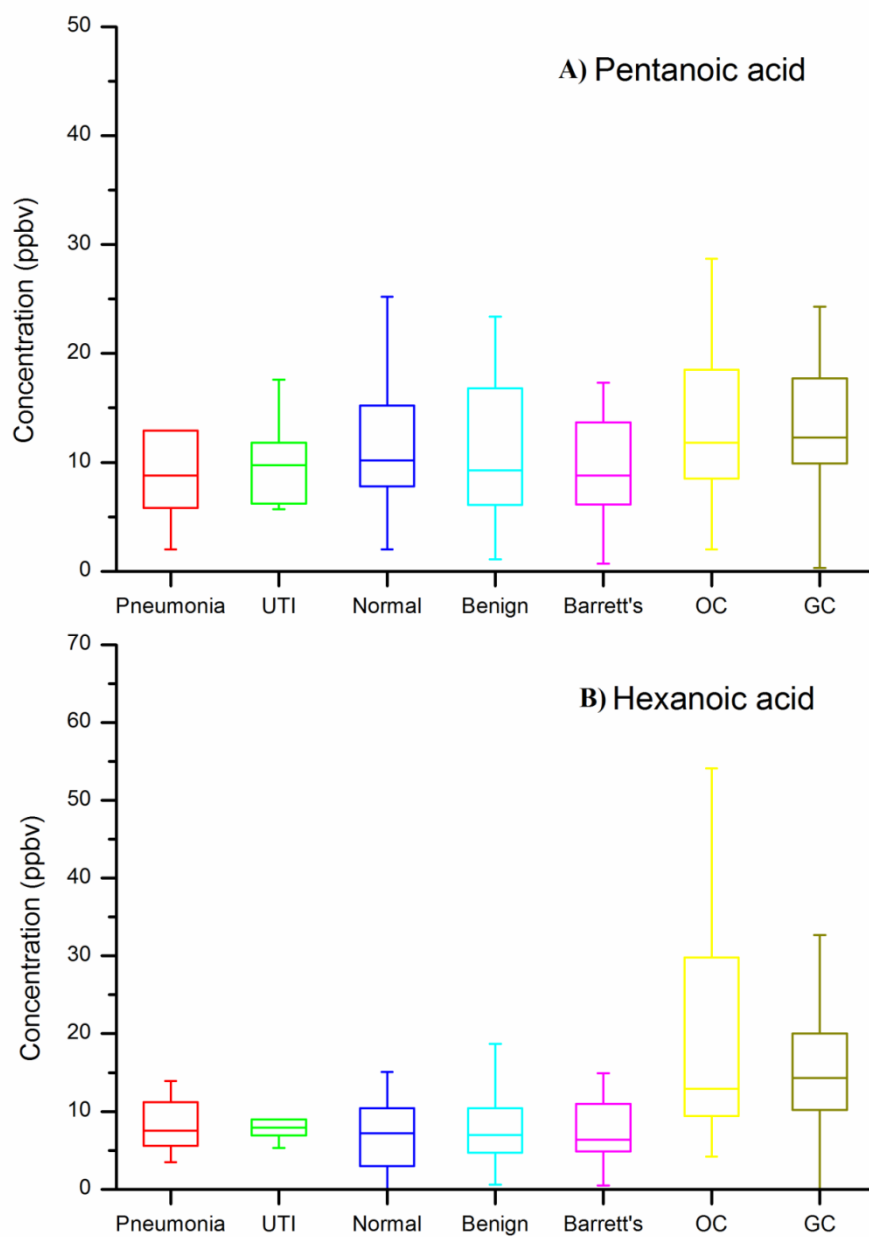


Figure 6.18: Box-Whisker plots of the measured concentrations (*ppbv*) of the fatty acids

A) Pentanoic acid and B) Hexanoic acid.

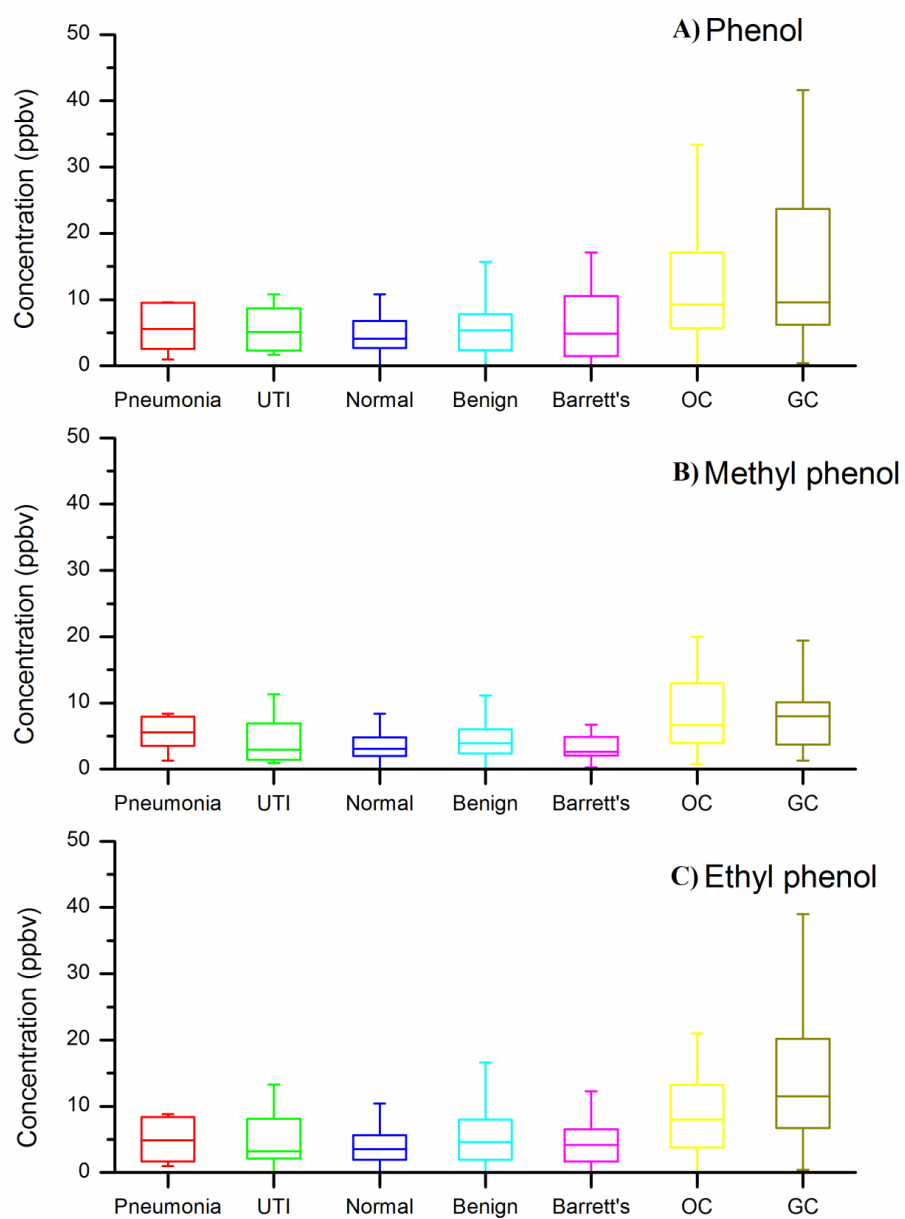


Figure 6.19: Box-Whisker plots of the measured concentrations (*ppbv*) of A) Phenol B) Methyl Phenol and C) Ethyl phenol.

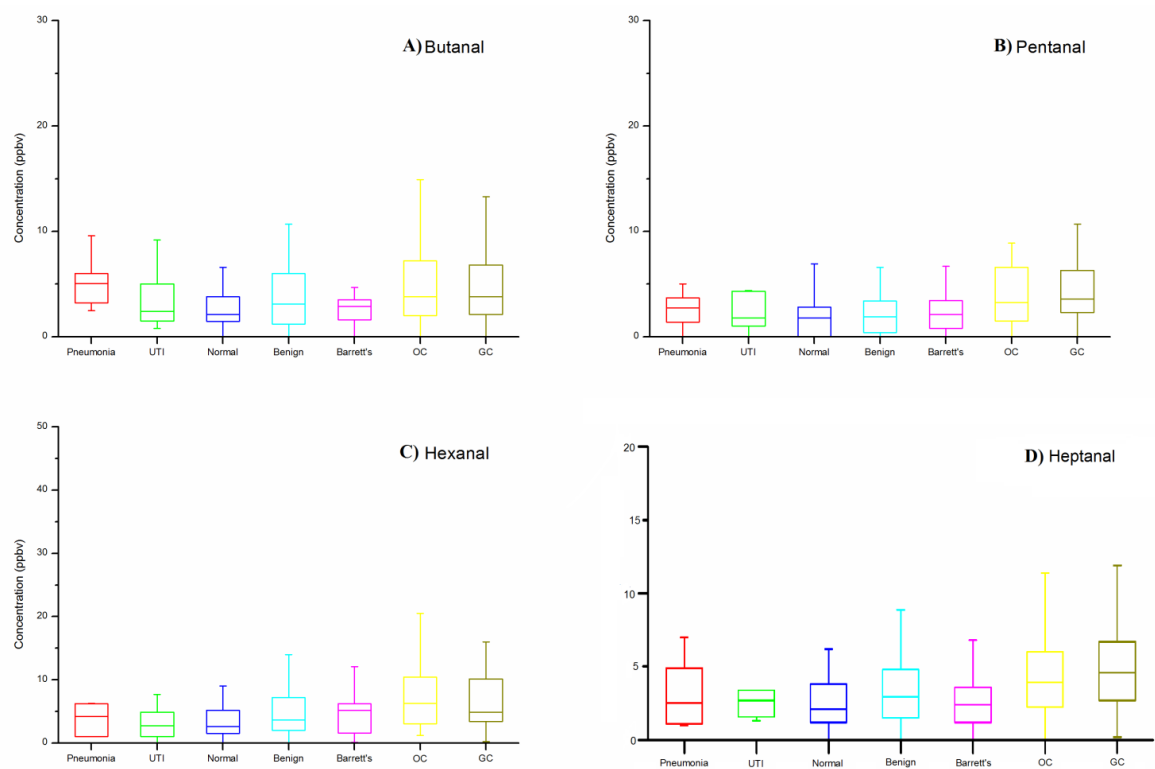


Figure 6.20: Box-Whisker plots of the measured concentrations (*ppbv*) of the aldehydes

A) Butanal, B) Pentanal, C) Hexanal and D) Heptanal.

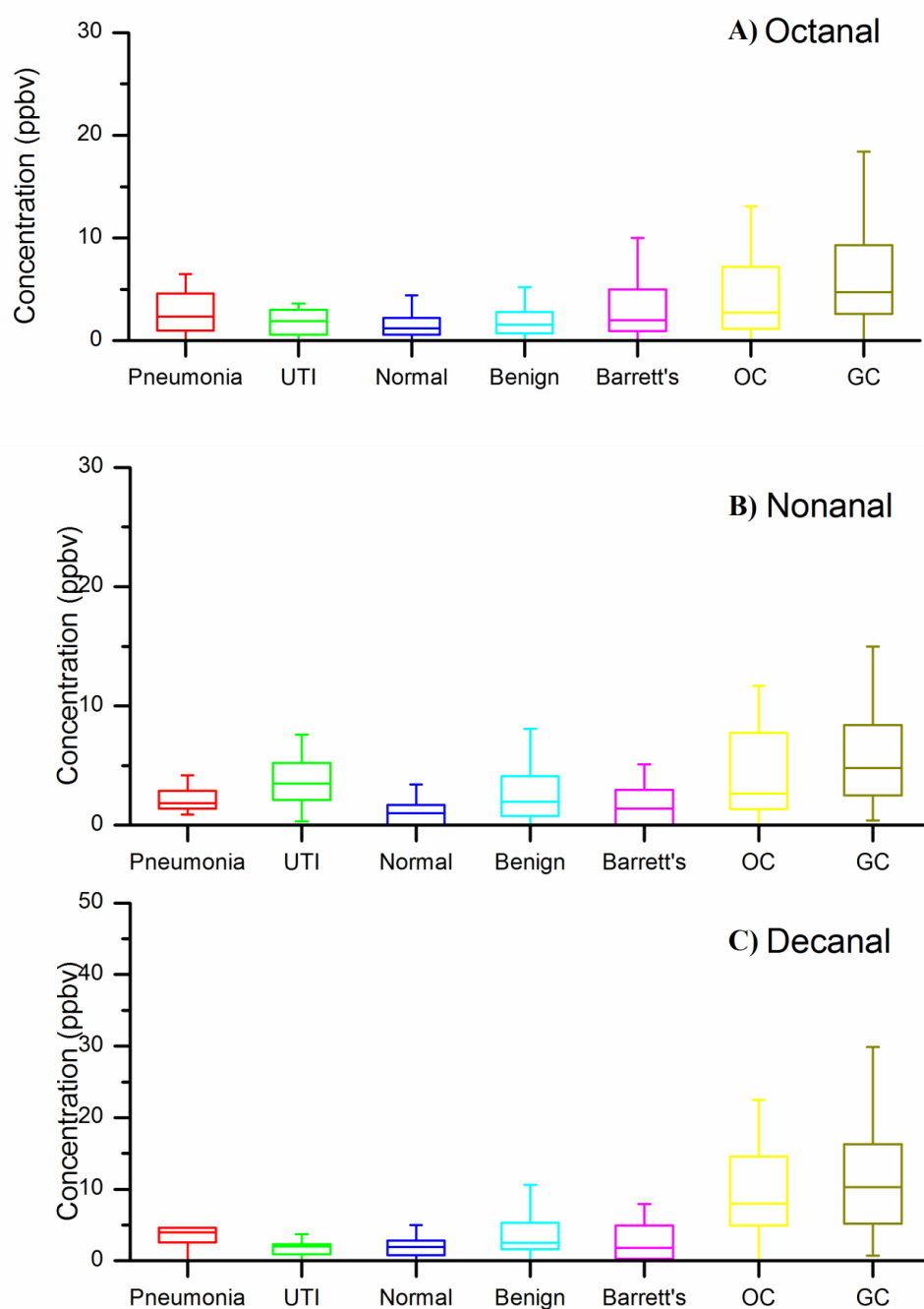


Figure 6.21: Box-Whisker plots of the measured concentrations (*ppbv*) of the aldehydes

A) Octanal, B) Nonanal and C) Decanal.

6.4.5 Discussion

The majority of the fatty acid, phenol and aldehyde compounds in pneumonia and urinary tract infection cohorts demonstrated median concentrations in the same range as the upper gastrointestinal control groups. All the VOCs of interest, except for butanal, were significantly different ($p < 0.05$) between the cancer groups and both infection cohorts. Butanal demonstrated an increased median concentration in the pneumonia group of 5 *ppbv* compared with 4 *ppbv* in the cancer cohorts and 2 *ppbv* in the UTI group. This isolated increase of butanal in the pneumonia group is notable given that none of the other investigated aldehydes demonstrated higher concentrations than the cancer cohorts. Filipiak *et al.* investigated *Streptococcus pneumoniae* and *Haemophilus influenza* cultures using GC-MS²²¹. Both bacteria demonstrated significantly increased concentrations of butanal in their headspace. The concentration of butanal was 18 *ppbv* in the medium (blood agar plates and in tryptic soy broth) used to grow *Streptococcus pneumoniae*; after 7.5 hrs of growth, the butanal concentration within the headspace was 33 *ppbv*²²¹. *Staphylococcus aureus* cultures have been reported to produce 3-methylbutanal²¹⁸. Buszewski *et al.* also reported a higher concentration of butanal in the exhaled breath of lung cancer patients compared with controls¹⁹⁶. Given the bacteria would have a local effect on VOC concentrations within the lungs in pneumonia, this may account for the increased concentrations of butanal within exhaled breath in the community acquired pneumonia group. Alternatively, given the increased concentrations of butanal have also been observed in lung cancer, this may represent a physiological response of diseased lung parenchyma. With regards to the abundant breath metabolites, lower concentrations of ammonia (277 *ppbv*) and higher concentrations of acetone (534 *ppbv*) and isoprene (129 *ppbv*) were observed in the UTI group when compared to the other study groups. However, a limitation of this study is that there were relatively fewer patients in the pneumonia and UTI cohorts in comparison to the other study groups.

6.4.6 Conclusions

The results of this pilot study demonstrate minimal influence of acute infection on the significant VOCs in oesophageal and gastric adenocarcinoma, except for butanal. The potential effects of individual pathogens on VOCs must be further evaluated as these could be confounding factors in the application of these metabolites as potential cancer screening tools. The identification of VOCs that are produced by specific bacteria also raises the prospect that such compounds could have a non-invasive diagnostic or therapeutic monitoring role in community acquired infections. This could have significant clinical impact particularly if these VOCs are able to discriminate bacterial infection from viral aetiologies.

Chapter 7

Summary, Conclusions and Future work

The present mass spectrometric interrogation of multiple bio-samples from patients with oesophageal and gastric cancer has identified several trace VOCs including fatty acids, phenols and aldehydes. The analysis of gastric content, urine and exhaled breath have clearly demonstrated that such trace compounds are more influential than abundant metabolites (i.e., acetone, isoprene, methanol) in upper gastro-intestinal malignancy. The results of this wide ranging study also highlighted that VOC profiles rather than individual compounds are likely to be more accurate for the molecular-orientated diagnosis of oesophageal and gastric cancer. The investigation of VOCs within gastric content has been described for the first time; this has included development of the headspace sampling and analysis, assessment of reproducibility and validation of measured VOC concentrations. Gastric content represents an important bio-fluid to investigate so that we can improve our understanding of the pathogenesis of upper gastro-intestinal disease. Significant differences in VOCs within the stomach environment have been observed between patients with upper gastro-intestinal malignancy and those with normal upper GI tracts. Analysis of gastric content samples has also indicated hypoacidity to be a feature of upper gastro-intestinal disease. The influence of varying pH of this bio-fluid demonstrated a significantly increased production of (known carcinogen) acetaldehyde at pH 6 within gastric content. Further investigation of the potentiating effects of *H.pylori* and cancer cells on intra-gastric pH is necessary to establish which of these factors may be predominantly responsible for the production of specific VOCs in oesophago-gastric cancer.

In the headspace analysis of urine, increased concentrations of acetaldehyde, acetone, hexanoic acid and hydrogen sulphide in OG cancer were observed. These VOCs were also observed at higher concentrations in gastric content from patients with OG cancer and may reflect the systemic effects of the cancer state. The markedly lower concentrations of phenol within the urine of OG cancer patients were also noteworthy. Future studies on the headspace of urine in OG cancer should also investigate short chain aldehydes as potential VOCs of interest. With regards to exhaled breath analysis, a reproducible and reliable off-line bag sampling method for trace VOCs was validated and tested in a clinical setting. The C4-C10 aldehydes were all shown to be present at < 3 ppbv within the exhaled breath of healthy volunteers; this study was particularly important given the minimal published literature available on quantifying aldehydes within exhaled breath. The study on exhaled breath in 210 patients considered a representative upper gastro-intestinal population, all of whom were classified according to endoscopy and histology. An extensive analysis of confounding factors including co-morbidities and medications revealed that the increased concentrations of specific fatty acids, phenols and aldehydes are related primarily to cancer status. Future studies will

need to explore whether these increased concentrations arise from the cancer cells themselves, the surrounding micro-environment, the patient's physiological response to malignancy or a combination of these. There will also need to be investigation of patients with metastatic or peritoneal disease secondary to oesophageal or gastric cancer and how cancer spread can affect these VOCs. The preliminary study on the effect of acute infection on fatty acids, phenols and aldehydes demonstrated that butanal is increased in community acquired pneumonia and hence is not a cancer-specific VOC. This research has highlighted the significant capabilities of on-line mass spectrometry in the clinical environment, however there remain certain questions to be answered before VOCs can fulfil their translational potential.

In all the studies, bio-samples were analysed within 1 hour of collection in the same hospital. This protocol was adopted to allow for standardization in experimental methodology and prevent any significant VOC degradation. In the next phase of this research, external validation studies in different hospitals are necessary to verify the promising findings. In the first instance, this will require transport of samples to the location of the SIFT-MS instrument laboratory. Hence, temporal assessment of VOC decay/stability of fatty acids, phenols and aldehydes will be necessary; this will be particularly important for exhaled breath samples. As part of the same experiments, an investigation of the suitability of different sampling bag materials (i.e., (2 x 70µm) Nalophan[®] and/or Tedlar[®]) could also be undertaken to check their retention efficiency to the different VOCs. Gilchrist *et al.* have previously investigated HCN concentrations within different bag materials; they reported Tedlar[®]-based bags provide optimal storage for exhaled breath HCN at 37°C at 24hrs¹⁸¹. Similar experiments should be conducted for the fatty acids, phenols and aldehydes to evaluate optimal storage conditions for these compounds.

In the present research, patients with non-upper GI disease were excluded from recruitment at the outset. This exclusion criteria was adopted to allow investigation of VOCs in only one clinically relevant patient population. However, in order to develop non-invasive VOC profiling as potential cancer-screening tools, there will need to be systematic assessment of other diseases that may influence the same VOCs. Several of the detected short chain aldehydes have previously been observed at higher concentrations in the exhaled breath of patients with lung cancer and there remains the possibility the identified VOCs may be increased in different cancer states⁶³. The next organ system to be investigated should be the lower gastro-intestinal tract to allow direct comparison with the results of the present research. Previous studies have investigated exhaled breath VOC patterns in colorectal cancer and exhaled breath pentane in inflammatory bowel disease^{197,222}. Both studies were entirely novel

in their approach but the findings have not been investigated in a large patient cohort or in heterogenous clinical groups. In patients with diseases of the colon, there will also be methodological issues to overcome with regards to sample retrieval. For example, VOC analysis should not be conducted on the day of a colonoscopy in the manner of this research. Patients undergoing a complete colonic camera evaluation are required to take bowel preparation 24 hours prior to the procedure to clear the colon of faeces. This creates an artificial environment within the colon and affects the natural flora, which may impact on VOC concentrations within exhaled breath and urine. However, sampling could be conducted in patients undergoing a flexible sigmoidoscopy (prior to phosphate enema) and hence diseases of the left colon may represent an appropriate starting point for comparative analysis. Liver diseases will also warrant investigation given gastro-intestinal cancers metastasize to the liver and its significant role in metabolic activities within the human body. Twelve significant VOCs having previously been reported in the exhaled breath of patients with cirrhosis - four ketones (2-butanone, 2- or 3-pentanone, C8-ketone, C9-ketone), two terpenes (monoterpene, monoterpene), four organosulphur & organonitrogen compounds and two alcohols (heptadienol, methanol)²²³. Three significant VOCs have also been reported in blood from patients with liver cancer - hexanal, 1-octen-3-ol and octane⁷⁴. In addition to these two organ systems, there will also need to be separate studies of other potential confounding factors for VOC profiles in oesophageal and gastric cancer including chronic diseases (i.e., hypertension, diabetes mellitus), nutrition/diet and exercise.

Potential VOC biomarkers have been identified in this research for upper GI cancer and across several gastro-intestinal conditions²²⁴. It is hypothesized that there are likely to be organ-generated VOCs, disease-specific VOCs and certain compounds that are systemically elevated. In order to answer these questions, tracing the source of the VOCs and identifying potential mechanisms responsible for their production will also be necessary. The efforts of developing non-invasive VOCs as diagnostic tools in cancer should ideally be for exhaled breath and urine given the significant translational potential and numerous clinical applications. However, correlation of VOCs within blood and biofluids from the target organs should also be evaluated to improve our understanding of the origins of these compounds. Moreover, there is also a need to explore the underlying pathways that may be contributing to increased concentrations of specific VOCs. The lower concentrations of phenol in urine coupled with the significantly increased concentrations of phenol, methyl phenol and ethyl phenol in exhaled breath led to a consideration of the following mechanisms that may account for phenol dysfunction. Phenol production is linked to aromatic amino acid metabolism and several studies have reported

increased phenylalanine and tyrosine in oesophageal and gastric cancer states^{194,225}. Phenylalanine is converted to tyrosine by phenylalanine hydroxylase and tyrosine is subsequently converted to phenol through the action of the bacterial enzyme tyrosine phenol-lyase. In a study of gastric content using NMR and LC-MS, median concentrations of phenylalanine and tyrosine were significantly higher in the gastric cancer group compared to controls¹²⁶. Investigation of matched cancer and healthy mucosa tissue samples from patients with oesophageal cancer demonstrated significantly increased levels of tyrosine in cancer tissue compared to normal mucosa²⁰⁷. Significantly higher levels of phenylalanine and tyrosine have also been reported in gastric cancer tissue when compared to controls²²⁶. This possible relationship between increased phenol production and aromatic amino acid dysfunction in oesophageal and gastric cancer warrants further investigation. Assessment of these pathways will require employing multiple analytical platforms and investigation of the enzymatic intermediates.

Further consideration must also be given to the roles of fatty acids and aldehydes in oesophageal and gastric cancer. Lipid peroxidation pathways activated by reactive oxygen species dysfunction in the Barrett's metaplastic-dysplastic cascade have previously been reported, indicating that numerous fatty acids and aldehydes are produced^{227,228}. Targeting of these compounds at a cellular level and understanding their potential contributions to VOC production should be considered. Evaluation of the key enzymes involved in their biological pathways is also necessary. Fatty acid synthase (FASN) generates medium chain fatty acid compounds from acetyl-CoA subunits under hypoxic conditions in the early stages of cancer growth²²⁹. FASN is upregulated by hypoxic-inducible factor-1 and over-expression of cytosolic FASN has also been reported in oesophageal adenocarcinoma^{230,231}. Similarly, the aldehyde dehydrogenase superfamily of enzymes should also be investigated. Upregulation of aldehyde dehydrogenases have previously been reported in oesophageal cancer²⁰⁹; further studies should focus on ALDH1, ALDH2 & ALDH5 subtypes in oesophageal and gastric cancer. Through understanding the mechanisms that lead to VOC production, it will be possible to improve our insights of tumour biology as well as developing more robust diagnostic tools.

There also remain technological issues to address to improve VOC analysis for clinical diagnostics. The sub-*ppbv* concentrations of specific VOCs observed in healthy study participants represent the current limit of detection for the SIFT-MS technology, but there are prospects of lowering this towards 0.1 *ppbv*. Lower sensitivities can be achieved using sample pre-concentration techniques such as SPME/GC-MS methods, but longer analysis times and

compromised accuracy of quantification are inevitable. However, SIFT-MS allows on-line, real-time detection and quantification of VOCs without the need for bio-sample modification in both headspace and exhaled breath analysis, thus providing immediate analyses. Though GC-MS and LC-MS will remain important techniques for VOC analysis, they both require the use of solvents and extensive sample preparation. This limits their potential applications within a clinical setting and prevents a biological sample being analysed in its natural state. At the present time, the limitations for the SIFT-MS technique are the relatively fewer number of VOCs that can be targeted and measured in real time; isomeric differentiation is also not feasible with the relatively low resolution analytical quadrupole mass spectrometers currently used in SIFT-MS instruments.

The online analytical capabilities of the SIFT-MS for breath analysis have also not been exploited in this research given the number of VOCs being investigated and the inability to standardize direct breath manoeuvres in patients. If cross-platform validation of the highly significant VOCs in OG cancer occurs, then a direct exhaled breath method could be developed targeting only a few VOCs. At the same time, improvements in SIFT-MS technology will result in lighter, more readily portable instruments potentially with internal calibration mechanisms. Much smaller SIFT-MS instruments could be moved from the laboratory to the clinical consultation room and potentially be utilised by non-specialists. However, a second strategy for the clinical setting could be to develop electrochemical sensor-based devices targeting specific VOCs. Following a similar strategic approach to this research, mass spectrometry could be employed initially to detect and quantify VOCs in a particular cancer type. After validation on another analytical platform, these VOCs could then be targeted in a point-of-care hand-held type sensor device. ‘E-noses’ have previously demonstrated significant promise in the analysis of exhaled breath and in the diagnosis of bladder cancer from urine samples^{232,233}. A significant advantage of sensor devices is that they can be designed to provide simple read-outs of the outcome for the clinician/health care professional obviating the need for interpreting complex mass spectral data.

Development of a non-invasive VOC-based screening tool to identify or at least risk-stratify patients with suspected oesophageal and gastric cancer could have significant clinical impact. The potential benefits of undertaking such a strategy include:

- Improving early diagnosis of oesophageal and gastric cancer.

- Reducing the number of patients referred for OGD and increasing the pick-up rate of oesophageal and gastric cancer. This could also result in cost-savings for the NHS.
- Improving quality of life and patient experience for those diagnosed with oesophageal or gastric cancer.

However, there is also a need to consider Barrett's oesophagus and squamous cell carcinoma of the oesophagus for VOC analysis in upper gastro-intestinal malignancy. The frequency of assessment of Barrett's oesophagus with endoscopy remains a much discussed subject. In a case-control study of 8272 patients with Barrett's oesophagus, surveillance (within 3 years) of the Barrett's oesophagus was not associated with a significant decrease in death from oesophageal adenocarcinoma²³⁴. In the study of exhaled breath, patients with Barrett's metaplasia demonstrated VOC profiles more in the range of the control groups rather than cancer. However, investigation of Barrett's dysplasia is also essential as a significant proportion of these patients will progress and develop oesophageal adenocarcinoma. Identifying VOC profiles specifically for early oesophageal cancer could provide another tool in the surveillance of patients with Barrett's dysplasia and in terms of monitoring cancer recurrence. There are also potential roles for non-invasive VOC analysis beyond diagnostics. With the establishment of 'normal values' for increasing numbers of compounds, there is the possibility that VOCs could be employed as therapeutic monitoring tools to measure response to treatment.

In conclusion, this research has identified VOCs that are elevated in concentration in biogenic fluids (i.e., fatty acids, phenol and aldehydes) in oesophageal and gastric cancer. There are several avenues of research discussed above that warrant investigation and there remains much promise for VOCs as non-invasive markers of disease. By employing a multi-platform approach to VOC analysis, biomarker and technological break-throughs are more likely to happen resulting in clinically applicable tests that could help save the lives of patients with cancer.

REFERENCES

1. The WHO Global Status report on Non-communicable Diseases (2010).
2. The World Health Organization's fight against cancer: strategies that prevent, cure and care (2007).
3. Office for National Statistics, 2010 Mortality Statistics: Deaths registered in England & Wales (2008).
4. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *Am J Epidemiol.* 107(1), 71-76 (1978).
5. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* 127, 2893–2917 (2010).
6. Day NE. Some aspects of the epidemiology of esophageal cancer. *Cancer Res.* 35, 3304–3307 (1975).
7. Pohl H, Sirovich B, Welch HG. Esophageal adenocarcinoma incidence: are we reaching the peak? *Cancer Epidemiol Biomarkers Prev.* 19, 1468–1470 (2010).
8. Lepage C, Rachet B, Jooste V, Faivre J, Coleman MP. Continuing rapid increase in esophageal adenocarcinoma in England and Wales. *Am. J. Gastroenterol.* 103, 2694–2699 (2008).
9. Steevens J, Botterweck AA, Dirx MJ, van den Brandt PA, Schouten LJ. Trends in incidence of oesophageal and stomach cancer subtypes in Europe. *Eur. J. Gastroenterol. Hepatol.* 22, 669–678 (2010).
10. Bosetti C, Levi F, Ferlay J, Garavello W, Lucchini F, Bertuccio P, Negri E, La Vecchia C. Trends in oesophageal cancer incidence and mortality in Europe. *Int. J. Cancer.* 122, 1118–1129 (2008).
11. Devesa SS, Blot WJ, Fraumeni JF Jr. Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. *Cancer.* 83, 2049–2053 (1998).
12. Lagergren J. Oesophageal cancer and gastro-oesophageal reflux: what is the relationship? *Gut.* 53, 1064–1065 (2004).
13. Houghton J, Wang TC. Helicobacter pylori and gastric cancer: a new paradigm for inflammation-associated epithelial cancers. *Gastroenterology.* 128, 1567-1578 (2005).
14. Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of eighteen major cancers in 1985. *Int J Cancer.* 54, 594-606 (1993).
15. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global Cancer statistics. *CA Cancer J Clin.* 61, 69-90 (2011).
16. Howe HL, Wu X, Ries LA, Cokkinides V, Ahmed F, Jemal A, Miller B, Williams M, Ward E, Wingo PA, Ramirez A, Edwards BK. Annual report to the nation on the status of cancer, 1975-2003, featuring cancer among U.S. Hispanic/Latino populations. *Cancer.* 107, 1711-1742 (2006).
17. Goldstein, DB, Hirschhorn JN. In genetic control of disease, does 'race' matter? *Nat Genet.* 36, 1243-1244 (2004).
18. Office for National Statistics (ONS). Cancer survival in England: Patients diagnosed 2005-2009 and followed up to 2010. London: ONS (2011).
19. Howlader N, Noone AM, Krapcho M, Neyman N, Aminou R, Altekruse SF, Kosary CL, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Chen HS, Feuer EJ, Cronin K. *SEER Cancer Statistics Review, 1975-2009*; National Cancer Institute, Bethesda, MD (2011).
20. National Institute of Clinical Excellence Referral Guidelines for suspected cancer <http://www.nice.org.uk/nicemedia/live/10968/29814/29814.pdf>
21. SCOTTISH REFERRAL GUIDELINES FOR SUSPECTED CANCER http://www.sehd.scot.nhs.uk/mels/HDL2007_09.pdf
22. The NHS Information Centre, 2013. National Oesophago-gastric Cancer Audit 2010. (at <http://www.hqip.org.uk/assets/NCAPOP-Library/NCAPOP-2013-14/NOGCA-2013-INTERACTIVE-PDF27062013.pdf>.)
23. Prasad GA, Wu TT, Wagle DA, Buttar NS, Wongkeesong LM, Dunagan KT, Lutzke LS, Borkenhagen LS, Wang KK. Endoscopic and surgical treatment of mucosal (T1a) esophageal adenocarcinoma in Barrett's esophagus. *Gastroenterology.* 137, 815-2 (2009).
24. Hamashima C, Ogoshi K, Okamoto M, Shabana M, Kishimoto T, Fukao A. A community-based, case-control study evaluating mortality reduction in gastric cancer from endoscopic screening in Japan. *PLoS One.* 8 (11), (2013).
25. Kim BJ, Heo C, Kim BK, Kim JY, Kim JG. Effectiveness of gastric cancer screening programs in South Korea: organized vs opportunistic model. *World J Gastroenterol.* 19, 736-41(2013).
26. Vakil N, Talley N, van Zanten SV, Flook N, Persson T, Björck E, Lind T, Bolling-Sternevald E. Cost of detecting malignant lesions by endoscopy in 2741 primary care dyspeptic patients without alarm symptoms. *Clin Gastroenterol Hepatol.* 7, 756–61 (2009).

27. Vakil N, Moayyedi P, Fennerty MB, Talley NJ. Limited value of alarm features in the diagnosis of upper gastrointestinal malignancy: systematic review and meta-analysis. *Gastroenterology* 131, 390-401 (2006).
28. Shawihi M, Thompson E, Kapoor N, Powell G, Sturgess RP, Stern N, Roughton M, Pearson MG, Bodger K. Variation in gastroscopy rate in English general practice and outcome for oesophagogastric cancer: retrospective analysis of Hospital Episode Statistics. *Gut*. 63(2), 250-261 (2013).
29. Rolff HC, Simonsen LR, Rosenberg J. Clinical findings confirm national guidelines regarding primary gastroscopy for upper gastrointestinal symptoms. *Dan Med Bull*. 58, A4363 (2011).
30. Banki F, Mason RJ, Oh D, Hagen JA, DeMeester SR, Lipham JC, Tanaka K, Danenberg KD, Yacoub WN, Danenberg PV, DeMeester TR. Plasma DNA as a molecular marker for completeness of resection and recurrent disease in patients with esophageal cancer. *Arch. Surg*. 142, 533-538 (2007).
31. Dulak AM, Stojanov P, Peng S, Lawrence MS, Fox C, Stewart C, Bandla S, Imamura Y, Schumacher SE, Shefler E, McKenna A, Carter SL, Cibulskis K, Sivachenko A, Saksena G, Voet D, Ramos AH, Auclair D, Thompson K, Sougnez C, Onofrio RC, Guiducci C, Beroukhi R, Zhou Z, Lin L, Lin J, Reddy R, Chang A, Landrenau R, Pennathur A, Ogino S, Luketich JD, Golub TR, Gabriel SB, Lander ES, Beer DG, Godfrey TE, Getz G, Bass AJ. Exome and whole-genome sequencing of esophageal adenocarcinoma identifies recurrent driver events and mutational complexity. *Nat. Genet*. 45, 478-486 (2013).
32. Streppel MM, Lata S, DelaBastide M, Montgomery EA, Wang JS, Canto MI, Macgregor-Das AM, Pai S, Morsink FH, Offerhaus GJ, Antoniou E, Maitra A, McCombie WR. Next-generation sequencing of endoscopic biopsies identifies ARID1A as a tumor-suppressor gene in Barrett's esophagus. *Oncogene*. 33, 347-357 (2014).
33. Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, Davies H, Jones D, Lin ML, Teague J, Bignell G, Butler A, Cho J, Dalglish GL, Galappaththige D, Greenman C, Hardy C, Jia M, Latimer C, Lau KW, Marshall J, McLaren S, Menzies A, Mudie L, Stebbings L, Largaespada DA, Wessels LF, Richard S, Kahnoski RJ, Anema J, Tuveson DA, Perez-Mancera PA, Mustonen V, Fischer A, Adams DJ, Rust A, Chan-on W, Subimerb C, Dykema K, Furge K, Campbell PJ, Teh BT, Stratton MR, Futreal PA. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene *PBRM1* in renal carcinoma. *Nature*. 469, 539-542 (2011).
34. Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, Aoki M, Hosono N, Kubo M, Miya F, Arai Y, Takahashi H, Shirakihara T, Nagasaki M, Shibuya T, Nakano K, Watanabe-Makino K, Tanaka H, Nakamura H, Kusuda J, Ojima H, Shimada K, Okusaka T, Ueno M, Shigekawa Y, Kawakami Y, Arihiro K, Ohdan H, Gotoh K, Ishikawa O, Ariizumi S, Yamamoto M, Yamada T, Chayama K, Kosuge T, Yamaue H, Kamatani N, Miyano S, Nakagama H, Nakamura Y, Tsunoda T, Shibata T, Nakagawa H. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet*. 44, 760-764 (2012).
35. Zang ZJ, Cutcutache I, Poon SL, Zhang SL, McPherson JR, Tao J, Rajasegaran V, Heng HL, Deng N, Gan A, Lim KH, Ong CK, Huang D, Chin SY, Tan IB, Ng CC, Yu W, Wu Y, Lee M, Wu J, Poh D, Wan WK, Rha SY, So J, Salto-Tellez M, Yeoh KG, Wong WK, Zhu YJ, Futreal PA, Pang B, Ruan Y, Hillmer AM, Bertrand D, Nagarajan N, Rozen S, Teh BT, Tan P. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. *Nat. Genet*. 44, 570-574 (2012).
36. Jin Z, Hamilton JP, Yang J, Mori Y, Olaru A, Sato F, Ito T, Kan T, Cheng Y, Paun B, David S, Beer DG, Agarwal R, Abraham JM, Meltzer SJ. Hypermethylation of the AKAP12 promoter is a biomarker of Barrett's-associated esophageal neoplastic progression. *Cancer Epidemiol Biomarkers Prev*. 17, 111-117 (2008).
37. Jin Z, Olaru A, Yang J, Sato F, Cheng Y, Kan T, Mori Y, Mantzur C, Paun B, Hamilton JP, Ito T, Wang S, David S, Agarwal R, Beer DG, Abraham JM, Meltzer SJ. Hypermethylation of tachykinin-1 is a potential biomarker in human esophageal cancer. *Clin Cancer Res* 13, 6293-6300 (2007).
38. Hamilton JP, Sato F, Jin Z, Greenwald BD, Ito T, Mori Y, Paun BC, Kan T, Cheng Y, Wang S, Yang J, Abraham JM, Meltzer SJ. Reprimo methylation is a potential biomarker of Barrett's-associated esophageal neoplastic progression. *Clin Cancer Res*. 12, 6637-6642 (2006).
39. Ahn HS, Shin YS, Park PJ, Kang KN, Kim Y, Lee HJ, Yang HK, Kim CW. Serum biomarker panels for the diagnosis of gastric adenocarcinoma. *Br J Cancer*. 106, 733-9 (2012).
40. Kim TY, Lee HJ, Hwang KS, Lee M, Kim JW, Bang YJ, Kang GH. Methylation of RUNX3 in various types of human cancers and premalignant stages of gastric carcinoma. *Lab Invest*. 84, 479-484 (2004).
41. Sakakura C, Hamada T, Miyagawa K, Nishio M, Miyashita A, Nagata H, Ida H, Yazumi S, Otsuji E, Chiba T, Ito K, Ito Y. Quantitative analysis of tumor-derived methylated RUNX3 sequences in the serum of gastric cancer patients. *Anticancer Res*. 29, 2619-2625 (2009).

42. Abbaszadegan MR, Moaven O, Sima HR, Ghafarzadegan K, A'rabi A, Forghani MN, Raziee HR, Mashhadinejad A, Jafarzadeh M, Esmaili-Shandiz E, Dadkhah E. P16 promoter hypermethylation: a useful serum marker for early detection of gastric cancer. *World J Gastroenterol.* 14, 2055–2060 (2008).
43. Wang YC, Yu ZH, Liu C, Xu LZ, Yu W, Lu J, Zhu RM, Li GL, Xia XY, Wei XW, Ji HZ, Lu H, Gao Y, Gao WM, Chen LB. Detection of RASSF1A promoter hypermethylation in serum from gastric and colorectal adenocarcinoma patients. *World J Gastroenterol.* 14, 3074–3080 (2008).
44. Qin DX, Wang GQ, Zuo JH, Zhang XH, Yuan FL, Li MS, Wu CR, Ju CL. Screening of esophageal and gastric cancer by occult blood bead detector. *Cancer.* 71, 216–218 (1993).
45. Roth MJ, Liu SF, Dawsey SM, Zhou B, Copeland C, Wang GQ, Solomon D, Baker SG, Giffen CA, Taylor PR. Cytologic detection of esophageal squamous cell carcinoma and precursor lesions using balloon and sponge samplers in asymptomatic adults in Linxian, China. *Cancer.* 80, 2047–2059 (1997).
46. Pauling L, Robinson AB, Teranishi R, Cary P. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proc Natl Acad Sci U S A.* 68, 2374-6 (1971).
47. Manolis A. The diagnostic potential of breath analysis. *Clin Chem.* 29, 5-15 (1983).
48. Kawai T, Kawakami K, Kataoka M, Itoi T, Takei K, Moriyasu F, Takagi Y, Aoki T, Serizawa H, Rimbar E, Noguchi N, Sasatsu M. A study of the relationship between *Helicobacter pylori* microbial susceptibility, ¹³C-urea breath test values. *Hepatogastroenterology.* 55, 786-90 (2008).
49. Kerlin P, Wong L. Breath hydrogen testing in bacterial overgrowth of the small intestine *Gastroenterology.* 95, 982-8 (1988).
50. Kharitonov SA, Yates D, Robbins RA, Logan-Sinclair R, Shinebourne EA, Barnes PJ. Increased nitric oxide in exhaled air of asthmatic patients. *Lancet.* 343, 133-5 (1994).
51. Sonoda H, Kohnoe S, Yamazato T, Satoh Y, Morizono G, Shikata K, Morita M, Watanabe A, Morita M, Kakeji Y, Inoue F, Maehara Y. Colorectal cancer screening with odour material by canine scent detection. *Gut.* 60, 814-9 (2011).
52. Ehmann R, Boedeker E, Friedrich U, Sagert J, Dippon J, Friedel G, Walles T. Canine scent detection in the diagnosis of lung cancer: revisiting a puzzling phenomenon. *Eur Respir J.* 39, 669-76 (2012).
53. Smith D, Španěl P. Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass Spectrom Rev.* 24, 661-700 (2005).
54. Miekisch W, Schubert JK, Noeldge-Schomburg, GF. Diagnostic potential of breath analysis--focus on volatile organic compounds. *Clin Chim Acta.* 347, 25-39 (2004).
55. Alderson P, Green S, Higgins JPT. Cochrane reviewers handbook 4.2.3. Chichester UK: John Wiley & Sons Ltd, (2005).
56. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 155(8), 529-36 (2011).
57. Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of Variation and Bias in Studies of Diagnostic Accuracy. *Ann Intern Med.* 140(3), 189-202 (2004).
58. Pawliszyn J. Theory of solid-phase microextraction. *J Chromatogr Sci.* 38, 270-278 (2000).
59. Deng C, Zhang X, Li N. Investigation of volatile biomarkers in lung cancer blood using solid-phase microextraction and capillary gas chromatography-mass spectrometry. *J Chromatogr B.* 808, 269-77 (2004).
60. Phillips M, Gleeson K, Hughes JM, Greenberg J, Cataneo RN, Baker L, McVay WP. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. *Lancet.* 353, 1930-3 (1999).
61. Poli D, Carbognani P, Corradi M, Goldoni M, Acampa O, Balbi B, Bianchi L, Rusca M, Mutti A. Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross sectional and nested short-term follow-up study. *Respir Res.* 6, 71 (2005).
62. Poli D, Goldoni M, Corradi M, Acampa O, Carbognani P, Internullo E, Casalini A, Mutti A. Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. *J Chromatogr B.* 878, 2643-51 (2010).
63. Fuchs P, Loeseken C, Schubert JK, Miekisch W. Breath gas aldehydes as biomarkers of lung cancer. *Int J Cancer.* 126, 2663-70 (2010).
64. Song G, Qin T, Liu H, Xu GB, Pan YY, Xiong FX, Gu KS, Sun GP, Chen ZD. Quantitative breath analysis of volatile organic compounds of lung cancer patients. *Lung Cancer.* 67 (2), 227-231 (2010).
65. de Gennaro G, Dragonieri S, Longobardi F, Musti M, Stallone G, Trizio L, Tutino M. Chemical characterization of exhaled breath to differentiate between patients with malignant pleural mesothelioma from subjects with similar professional asbestos exposure. *Anal Bioanal Chem.* 398, 3043-50 (2010).

66. Phillips M, Cataneo RN, Ditzhoff BA, Fisher P, Greenberg J, Gunawardena R, Kwon CS, Tietje O, Wong C. Prediction of breast cancer using volatile biomarkers in the breath. *Breast Cancer Res Treat.* 99, 19-21 (2006).
67. Phillips M, Altorki N, Austin JH, Cameron RB, Cataneo RN, Greenberg J, Kloss R, Maxfield RA, Munawar MI, Pass HI, Rashid A, Rom WN, Schmitt P. Prediction of lung cancer using volatile biomarkers in breath. *Cancer Biomark.* 3, 95-109 (2007).
68. Qin T, Liu H, Song Q, Song G, Wang HZ, Pan YY, Xiong FX, Gu KS, Sun GP, Chen ZD. The screening of volatile markers for hepatocellular carcinoma. *Cancer Epidemiol Biomarkers Prev.* 19, 2247-2253 (2010).
69. Guadagni R, Miraglia N, Simonelli A, Silvestre A, Lamberti M, Feola D, Acampora A, Sannolo N. Solid-phase microextraction-gas chromatography-mass spectrometry method validation for the determination of endogenous substances: Urinary hexanal and heptanal as lung tumor biomarkers. *Anal Chim Acta.* 701, 29-36 (2011).
70. Silva CL, Passos M, Câmara JS. Investigation of urinary volatile organic metabolites as potential cancer biomarkers by solid-phase microextraction in combination with gas chromatography-mass spectrometry. *Br J Cancer.* 105, 1894-904 (2011).
71. Dragonieri S, Annema JT, Schot R, van der Schee MP, Spanevello A, Carratù P, Resta O, Rabe KF, Sterk PJ. An electronic nose in the discrimination of patients with non-small cell lung cancer and COPD. *Lung Cancer.* 64, 166-70 (2009).
72. Peng G, Hakim M, Broza YY, Billan S, Abdah-Bortnyak R, Kuten A, Tisch U, Haick H. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. *Br J Cancer.* 103, 542-51 (2010).
73. Mazzone PJ, Hammel J, Dweik R, Na J, Czich C, Laskowski D, Mekhail T. Diagnosis of lung cancer by the analysis of exhaled breath with a colorimetric sensor array. *Thorax.* 62, 565-8 (2007).
74. Xue R, Dong L, Zhang S, Deng C, Liu T, Wang J, Shen X. Investigation of volatile biomarkers in liver cancer blood using solid-phase microextraction and gas chromatography/mass spectrometry. *Rapid Commun Mass Spectrom* 22, 1181-1186 (2008).
75. Filipiak W, Sponring A, Mikoviny T, Ager C, Schubert J, Miekisch W, Amann A, Troppmair J. Release of volatile organic compounds (VOCs) from the lung cancer cell line CALU-1 in vitro. *Cancer Cell Int.* 8, 17 (2008).
76. Sponring A, Filipiak W, Mikoviny T, Ager C, Schubert J, Miekisch W, Amann A, Troppmair J. Release of volatile organic compounds from the lung cancer cell line NCI-H2087 in vitro. *Anticancer Res.* 29, 419-26 (2009).
77. Filipiak W, Sponring A, Filipiak A, Ager C, Schubert J, Miekisch W, Amann A, Troppmair J. TD-GC-MS analysis of volatile metabolites of human lung cancer and normal cells in vitro. *Cancer Epidemiol Biomarkers Prev.* 19(1), 182-95 (2010).
78. Sponring A, Filipiak W, Ager C, Schubert J, Miekisch W, Amann A, Troppmair J. Analysis of volatile organic compounds (VOCs) in the headspace of NCI-H1666 lung cancer cells. *Cancer Biomark.* 7, 153-61 (2010).
79. Taylor AJ, Linforth RST, Harvey BA, Blake A. Atmospheric pressure chemical ionisation mass spectrometry for in vivo analysis of volatile flavour release. *Food Chemistry.* 71, 327 (2000).
80. Lindinger W, Hansel A, Jordan A. On-line monitoring of volatile organic compounds at pptv levels by means of proton-transfer-reaction mass spectrometry (PTR-MS) Medical applications, food control and environmental research. *Int J Mass Spectrom Ion Process.* 173, 191-241 (1998).
81. Španěl P, Smith D. Selected ion flow tube mass spectrometry for on-line trace gas analysis in biology and medicine. *Eur J Mass Spectrom.* 13, 77-82 (2007).
82. Smith D, Wang T, Sulé-Suso J, Španěl P, El Haj A. Quantification of acetaldehyde released by lung cancer cells in vitro using selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom.* 17, 845-50 (2003).
83. Španěl P, Smith D, Holland TA, Al Singary W, Elder JB. Analysis of formaldehyde in the headspace of urine from bladder and prostate cancer patients using selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom.* 13, 1354-1359 (1999).
84. Wehinger A, Schmid A, Mechtcheriakov S, Ledochowski M, Grabmer C, Gastl G, Amann A. Lung cancer detection by proton transfer reaction mass-spectrometric analysis of human breath gas. *International Journal of Mass Spectrometry.* 265, 49-59 (2007).
85. Brunner C, Szymczak W, Höllriegl V, Mörtl S, Oelmez H, Bergner A, Huber RM, Hoeschen C, Oeh U. Discrimination of cancerous and non-cancerous cell lines by headspace-analysis with PTR-MS. *Anal Bioanal Chem.* 397, 2315-24 (2010).

86. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y. Phases of biomarker development for early detection of cancer. *J Nat Cancer Inst.* 93, 1054-61 (2001).
87. Ligor M, Ligor T, Bajtarevic A, Ager C, Pienz M, Klieber M, Denz H, Fiegl M, Hilbe W, Weiss W, Lukas P, Jamnig H, Hackl M, Buszewski B, Miekisch W, Schubert J, Amann A. Determination of volatile organic compounds in exhaled breath of patients with lung cancer using solid phase microextraction and gas chromatography mass spectrometry. *Clin Chem Lab Med.* 47, 550-60 (2009).
88. Zhang L, Steinmaus C, Eastmond DA, Xin XK, Smith MT. Formaldehyde exposure and leukemia: a new meta-analysis and potential mechanisms. *Mutat Res.* 681, 150-68 (2009).
89. Teng S, Beard K, Pourahmad J, Moridani M, Easson E, Poon R, O'Brien PJ. Formaldehyde metabolic detoxification enzyme systems and molecular cytotoxic mechanism in isolated rat hepatocytes. *Chem Biol Interact.* 130, 285-96 (2001).
90. Wang M, Cheng G, Villalta PW, Hecht SS. Development of liquid chromatography electrospray ionization tandem mass spectrometry methods for analysis of DNA adducts of formaldehyde and their application to rats treated with N-nitrosodimethylamine or 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Chem Res Toxicol.* 20, 1141-8 (2007).
91. Rager JE, Smeester L, Jaspers I, Sexton KG, Fry RC. Epigenetic changes induced by air toxics: formaldehyde exposure alters miRNA expression profiles in human lung cells. *Environ Health Perspect.* 119, 494-500 (2011).
92. Seitz HK, Stickel F. Acetaldehyde as an underestimated risk factor for cancer development: role of genetics in ethanol metabolism. *Genes Nutr.* 5, 121-128 (2009).
93. Salaspuro M. Acetaldehyde and gastric cancer. *J Dig Dis.* 12, 51-9 (2011).
94. Yang SJ, Wang HY, Li XQ, Du HZ, Zheng CJ, Chen HG, Mu XY, Yang CX. Genetic polymorphisms of ADH2 and ALDH2 association with esophageal cancer risk in southwest China. *World J Gastroenterol.* 13, 5760-4 (2007).
95. Lee CH, Lee JM, Wu DC, Goan YG, Chou SH, Wu IC, Kao EL, Chan TF, Huang MC, Chen PS, Lee CY, Huang CT, Huang HL, Hu CY, Hung YH, Wu MT. Carcinogenetic impact of ADH1B and ALDH2 genes on squamous cell carcinoma risk of the esophagus with regard to the consumption of alcohol, tobacco and betel quid. *Int J Cancer.* 122, 1347-56 (2008).
96. Fanelli SL, Maciel ME, Díaz Gómez MI, Delgado de Layño AM, Bietto FM, Castro JA, Castro GD. Further studies on the potential contribution of acetaldehyde accumulation and oxidative stress in rat mammary tissue in the alcohol drinking promotion of breast cancer. *J Appl Toxicol.* 31, 11-9 (2011).
97. Secretan B, Straif K, Baan R, Grosse Y, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Coglianò V. A review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol.* 10, 1033-1034 (2009).
98. Voulgaridou GP, Anestopoulos I, Franco R, Panayiotidis MI, Pappa A. DNA damage induced by endogenous aldehydes: current state of knowledge. *Mutat Res.* 711, 13-27 (2011).
99. Office for National Statistics, June 2012 (<http://www.ons.gov.uk/ons/rel/vsob1/cancer-statistics-registrations--england--series-mb1-/index.html>).
100. Adams NG, Smith D. The Selected Ion Flow Tube (SIFT): A technique for studying thermal energy ion-neutral reactions. *Int J Mass Spectrom Ion Physics.* 21, 349-359 (1976).
101. Smith D, Adams NG. The selected ion flow tube (SIFT): Studies of ion-neutral reactions. *Adv Atom Mol Phys.* 24, 1-49 (1987).
102. Smith D, Adams NG. Elementary plasma reactions of environmental interest. *Top Curr Chem.* 89, 1-43 (1980).
103. Smith D. The ion chemistry of interstellar clouds. *Chem Rev.* 92, 1473-1485 (1992).
104. Smith D, Španěl P. Ions in the terrestrial atmosphere and in interstellar clouds. *Mass Spectrom Rev.* 14, 255-278 (1996).
105. Španěl P, Smith D. Reactions of hydrated hydronium ions and hydrated hydroxide ions with some hydrocarbons and oxygen-bearing organic molecules. *J Phys Chem.* 99, 15551-15556 (1995).
106. Španěl P, Smith D. A selected ion flow tube study of the reactions of NO⁺ and O₂⁺ ions with some organic molecules: The potential for trace gas analysis of air. *J Chem Phys.* 104, 1893-1899 (1996).
107. Smith D, Španěl P. Application of ion chemistry and the SIFT technique to the quantitative analysis of trace gases in air and on breath. *Int Reviews Phys Chem.* 15, 231-271 (1996).
108. Španěl P, Smith D. Selected ion flowtube: A technique for quantitative trace gas analysis of air and breath. *Med Biol Eng Comput.* 34, 409-419 (1996).
109. Španěl P, Hall EFH, Workman CT, Smith D. A direct coupled monolithic rectangular resonator forming a robust microwave plasma ion source for SIFT-MS. *Plasma Sources Sci. Technol.* 13, 282-284 (2004).

110. Španěl P, Pavlík M, Smith D. Reactions of H_3O^+ and OH^- ions with some organic molecules; applications to trace gas analysis in air. *Int J Mass Spectrom Ion Process.* 145, 177–186 (1995).
111. Davies SJ, Španěl P, Smith D. A new 'online' method to measure increased exhaled isoprene in end-stage renal failure. *Nephrol Dial Transpl.* 16, 836–839 (2001).
112. Smith D, Španěl P, Dabill D, Cocker J, Rajan B. On-line analysis of diesel engine exhaust gases by selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom.* 18, 2830–2838 (2004).
113. Španěl P, Smith D. Progress in SIFT-MS: breath analysis and other applications. *Mass Spectrom Rev.* 30, 236–67 (2011).
114. Turner C, Španěl P, Smith D. A longitudinal study of ammonia, acetone and propanol in the exhaled breath of 30 subjects using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol Meas.* 27, 321–337 (2006).
115. Turner C, Španěl P, Smith D. A longitudinal study of breath isoprene in healthy volunteers using selected ion flow tube mass spectrometry (SIFT-MS). *Physiol Meas.* 27, 13–22 (2006).
116. Turner C, Španěl P, Smith D. A longitudinal study of ethanol and acetaldehyde in the exhaled breath of healthy volunteers using selected ion flow-tube mass spectrometry. *Rapid Commun Mass Spectrom.* 20, 61–68 (2006).
117. Turner C, Španěl P, Smith D. A longitudinal study of methanol in the exhaled breath of 30 healthy volunteers using selected ion flow tube mass spectrometry, SIFT-MS. *Physiol Meas.* 27, 637–648 (2006).
118. Španěl P, Dryahina K, Smith D. Microwave plasma ion sources for selected ion flow tube mass spectrometry: Optimizing their performance and detection limits for trace gas analysis. *Int J Mass Spectrom.* 267, 117–124 (2007).
119. Španěl P, Dryahina K, Smith D. A general method for the calculation of absolute trace gas concentrations in air and breath from selected ion flow tube mass spectrometry data. *Int J Mass Spectrom.* 249/250, 230–239 (2006).
120. Španěl P, Smith D. On-line measurement of the absolute humidity of air, breath and liquid headspace samples by selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.* 15, 563–569 (2001).
121. Smith D, Španěl P. Ambient analysis of trace compounds in gaseous media by SIFT-MS. *Analyst.* 136, 2009–2032 (2011).
122. Španěl P, Smith D. Selected Ion Flow Tube Mass Spectrometry, SIFT-MS *Current Analytical Chemistry* 4, 525–539 (2013).
123. Lu PJ, Hsu P, Chen CH, Hsiao M, Chang WC, Tseng, HH, Lin, KH, Chuah SK, Chen HC. Gastric juice acidity in upper gastrointestinal disease. *World J Gastroenterol.* 16, 5496 (2010).
124. Inada M, Miyake T, Kudo H, Wakatsuki Y, Murakami M, Seki M, Mizuno M, Tsukada H, Okae S. Cobalamin-binding protein in gastric juice as a new tumor marker in gastric cancer. *Cancer Detect Prev.* 10, 285–91 (1987).
125. Hsu PI, Chen CH, Hsiao M, Wu DC, Lin CY, Lai KH, Lu PJ. Diagnosis of gastric malignancy using gastric juice alpha1-antitrypsin. *Cancer Epidemiol Biomarkers Prev.* 19, 405–11 (2010).
126. Deng K, Lin S, Zhou L, Li Y, Chen M, Wang Y, Li Y. High level of Aromatic Amino Acid in Gastric Juice during the Early Stages of Gastric Cancer Progression. *PLoS One.* 7, 49434 (2012).
127. Shirasu M, Touhara K. The scent of disease: volatile organic compounds of the human body related to disease and disorder. *J Biochem.* 150, 257–66 (2011).
128. Julák J, Stránská E, Rosová V, Geppert H, Španěl P, Smith D. Bronchoalveolar Lavage examined by solid phase microextraction, gas chromatography-mass spectrometry and selected ion flow tube mass spectrometry. *J. Microbiol. Methods.* 65, 76–86 (2006).
129. Olivares A, Dryahina K, Navarro JL, Flores M, Smith D, Španěl P. Selected ion flow tube-mass spectrometry for absolute quantification of aroma compounds in the headspace of dry fermented sausages. *Anal Chem.* 82, 5819–5829 (2010).
130. Collison P. Of bombers, radiologists, and cardiologists: time to ROC. *Heart.* 80, 215–217 (1998).
131. Zweig MH, Campbell GC. *Clin Chem.* 39, 561 (1993).
132. Španěl P, Smith D. SIFT studies of reactions of H_3O^+ , NO^+ , and O_2^+ with a series of alcohols. *Int J Mass Spectrom Ion Process.* 167–168, 375–388 (1997).
133. Smith D., Wang T, Španěl P. Analysis of ketones by selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom.* 17, 2655 (2003).
134. Španěl P, Ji Y, Smith D. SIFT studies of the reactions of H_3O^+ , NO^+ , and O_2^+ with a series of aldehydes and ketones. *Int J Mass Spectrom Ion Process* 165–166, 25–37 (1997).

135. Wang T, Španěl P, Smith D. A selected ion flow tube study of the reactions of H_3O^+ , NO^+ , and O_2^+ with some phenols, phenyl alcohols and cyclic carbonyl compounds in support of SIFT-MS and PTR-MS. *Int J of Mass Spectrom.* 239, 139-146 (2004).
136. Španěl P, Wang T, Smith D. Quantification of hydrogen cyanide in humid air by selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom.* 18, 1869-1873 (2004).
137. Oikawa T, Iijima K, Koike T, Uno K, Horii T, Iwai W, Abe Y, Asano N, Imatani A, Shimosegawa T. Deficient aldehyde dehydrogenase 2 is associated with increased risk for esophageal squamous cell carcinoma in the presence of gastric hypochlorhydria. *Scand J of Gastroenterol.* 45, 1338-1344 (2010).
138. Linderborg K, Marvola T, Marvola M, Salaspuro M, Färkkilä M, Väkeväinen S. Reducing carcinogenic acetaldehyde exposure in the achlorhydric stomach with cysteine. *Alcohol Clin and Exp Res.* 35, 516-522 (2011).
139. Laffel L. Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. *Diabetes/Metab Res Rev.* 15, 412-426 (1999).
140. Španěl P, Dryahina K, Rejšková A, Chippendale TW, Smith D. Breath acetone concentration: biological variability and the influence of diet. *Physiol Meas.* 32, N23-N31 (2011).
141. Smith D, Španěl P, Davies S. Trace gases in breath of healthy volunteers when fasting and after a protein-calorie meal: a preliminary study. *J Appl Physiol.* 87, 1584-1588 (1999).
142. Crofford OB, Mallard RE, Winton RE, Rogers NL, Jackson JC, Keller U. Acetone in breath and blood. *Trans Am Clin Climatol Assoc.* 88, 128-139 (1977).
143. Smith D, Španěl P, Sulé-Suso J. Advantages of Breath Testing for Early Diagnosis of Lung Cancer. *Expert Re Mol Diagn.* 10, 255-257 (2010).
144. Boshier PR, Cushnir JR, Priest OH, Marczin N, Hanna GB. Variation in the levels of volatile trace gases within three hospital environments: implications for clinical breath testing. *J Breath Res.* 4, 031001 (2010).
145. Smith D, Španěl P, Holland TA, al Singari W, Elder JB. Selected ion flow tube mass spectrometry of urine headspace. *Rapid Commun Mass Spectrom.* 13, 724-729 (1999).
146. Jokelainen K, Roine RP, Väänänen H, Färkkilä M, Salaspuro M. In vitro acetaldehyde formation by human colonic bacteria. *Gut.* 35, 1271-1274 (1994).
147. Anderson O, Buckle P, Hanna GB. Ergonomic risk assessment of nasogastric tube placement and implications for design and training. *Work.* 41 Suppl 1, 4689-91 (2012).
148. Dimopoulos C, Gialas A, Likourinas M, Androutsos G, Kostakopoulos A. Hippocrates: founder and pioneer of urology. *Br J Urol.* 52, 73-74 (1980).
149. Kouba EF, Wallen EM, Pruthi RS. Uroscopy by Hippocrates and Theophilus: prognosis versus diagnosis. *J Urol.* 177, 50-52 (2007).
150. Hindmarsh JT. King George III and acute porphyria. *Lancet.* 349, 364 (1997).
151. Hu JD, Tang HQ, Zhang Q, Fan J, Hong J, Gu JZ, Chen JL. Prediction of gastric cancer metastasis through urinary metabolomic investigation using GC/MS. *World J Gastroenterol.* 17, 727-34 (2011).
152. Issaq HJ, Nativ O, Waybright T, Luke B, Veenstra TD, Issaq EJ, Kravstov A, Muller M. Detection of bladder cancer in human urine by metabolomic profiling using high performance liquid chromatography/mass spectrometry. *J Urol.* 179, 2422-2426 (2008).
153. Cheng Y, Xie G, Chen T, Qiu Y, Zou X, Zheng M, Tan B, Feng B, Dong T, He P, Zhao L, Zhao A, Xu LX, Zhang Y, Jia W. Distinct urinary metabolic profile of human colorectal cancer. *J Proteome Res.* 11, 1354-1363 (2012).
154. Hanai Y, Shimono K, Matsumura K, Vachani A, Albelda S, Yamazaki K, Beauchamp GK, Oka H. Urinary volatile compounds as biomarkers for lung cancer. *Biosci Biotechnol Biochem.* 76, 679-684 (2012).
155. Roux A, Xu Y, Heilier JF, Olivier MF, Ezan E, Tabet JC, Junot C. Annotation of the human adult urinary metabolome and metabolite identification using ultra high performance liquid chromatography coupled to a linear quadrupole ion trap-Orbitrap mass spectrometer. *Anal. Chem.* 84, 6429-6437 (2012).
156. Spagou K, Wilson ID, Masson P, Theodoridis G, Raikos N, Coen M, Holmes E, Lindon JC, Plumb RS, Nicholson JK, Want EJ. HILIC-UPLC-MS for exploratory urinary metabolic profiling in toxicological studies. *Anal. Chem.* 83, 382-390 (2010).
157. Banday KM, Pasikanti KK, Chan EC, Singla R, Rao KV, Chauhan VS, Nanda RK. Use of Urine Volatile Organic Compounds to Discriminate Tuberculosis Patients from Healthy Subjects. *Anal. Chem.* 83, 5526-5534 (2011).
158. Pysanen A., Španěl P, Smith D. Analysis of the isobaric compounds propanol, acetic acid and methyl formate in humid air and breath by selected ion flow tube mass spectrometry, SIFT-MS. *Int J of Mass Spectrom.* 285, 42-48 (2009).
159. Storer MK, Hibbard-Melles K, Davis B, Scotter J. Detection of volatile compounds produced by microbial growth in urine by selected ion flow tube mass spectrometry. *J. Microbiol. Methods.* 87, 111-113 (2011).

160. Väkeväinen S, Mentula S, Nuutinen H, Salmela KS, Jousimies-Somer H, Färkkilä M, Salaspuro M. Ethanol-driven Microbial Production of Carcinogenic Acetaldehyde in Achlorhydric Atrophic Gastritis. *Scan J. Gastroenterol.* 37, 648-655 (2002).
161. Wang T, Španěl P, Smith D. Selected ion flow tube mass spectrometry of 3-hydroxybutyric acid, acetone and other ketones in the headspace of aqueous solution and urine. *Int. J. Mass Spectrom.* 272, 78–85 (2008).
162. Shariff MI, Ladep NG, Cox IJ, Williams HR, Okeke E, Malu A, Thillainayagam AV, Crossey MM, Khan SA, Thomas HC, Taylor-Robinson SD. Characterization of urinary biomarkers of hepatocellular carcinoma using magnetic resonance spectroscopy in a Nigerian population. *J Proteome Res.* 9, 1096-1103 (2010).
163. Hasim A, Ma H, Mamtimin B, Abudula A, Niyaz M, Zhang LW, Anwer J, Sheyhidin I. Revealing the metabonomic variation of EC using ¹H-NMR spectroscopy and its association with clinicopathological characteristics. *Mol. Biol. Rep.* 39, 8955-8964 (2012).
164. Zhang L, Jin H, Guo X, Yang Z, Zhao L, Tang S, Mo P, Wu K, Nie Y, Pan Y, Fan D. Distinguishing pancreatic cancer from chronic pancreatitis and healthy individuals by ¹H nuclear magnetic resonance-based metabonomic profiles. *Clin Biochem.* 45, 1064-1069 (2012).
165. Ulanowska A, Kowalkowski T, Trawińska E, Buszewski B. The application of statistical methods using VOCs to identify patients with lung cancer. *J. Breath Res.* 5, 10 (2011).
166. Suarez F, Furne J, Springfield J, Levitt M. Insights into human colonic physiology obtained from the study of flatus composition. *Am J Physiol.* 272, 1028-33 (1997).
167. Picton R, Eggo MC, Langman MJ, Singh S. Impaired detoxication of hydrogen sulfide in ulcerative colitis? *Dig Dis Sci.* 52, 373-378 (2007).
168. Ramasamy S, Singh S, Taniere P, Langman MJ, Eggo MC. Sulfide-detoxifying enzymes in the human colon are decreased in cancer and upregulated in differentiation. *Am J Physiol Gastrointest Liver Physiol.* 291, 23 (2006).
169. Chávez-Piña, AE, Tapia-Alvarez GR, Navarrete A. Inhibition of endogenous hydrogen sulfide synthesis by PAG protects against ethanol-induced gastric damage in the rat. *Eur. J. Pharmacol.* 630, 131-136 (2010).
170. Yamagishi K, Onuma K, Chiba Y, Yagi S, Aoki S, Sato T, Sugawara Y, Hosoya N, Saeki Y, Takahashi M, Fuji M, Ohsaka T, Okajima T, Akita K, Suzuki T, Senawongse P, Urushiyama A, Kawai K, Shoun H, Ishii Y, Ishikawa H, Sugiyama S, Nakajima M, Tsuboi M, Yamanaka T. Generation of gaseous sulfur-containing compounds in tumour tissue and suppression of gas diffusion as an antitumour treatment. *Gut.* 61, 554-561 (2012).
171. Perry H, Keevil B. Online extraction of 5-hydroxyindole acetic acid from urine for analysis by liquid chromatography-tandem mass spectrometry. *Ann Clin Biochem.* 45, 149-52 (2008).
172. Garner CE, Smith S, de Lacy Costello B, White P, Spencer R, Probert CS, Ratcliffe NM. Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease. *FASEB J.* 21, 1675-1688 (2007).
173. Caldwell DR. Effect of methanol on the growth of gastrointestinal anaerobes. *Can. J. Microbiol.* 35, 313-317 (1989).
174. Ling WH, Hanninen O. Shifting from a conventional diet to an uncooked vegan diet reversibly alters fecal hydrolytic activities in humans. *J Nutr.* 122, 924 (1992).
175. McDonald TA, Holland NT, Skibola C, Duramad P, Smith MT. Hypothesis: phenol and hydroquinone derived mainly from diet and gastrointestinal flora activity are casual factors in leukemia. *Leukemia* 15, 10 (2001).
176. Bajtarevic A, Ager C, Pienz M, Klieber M, Schwarz K, Ligor M, Ligor T, Filipiak W, Denz H, Fiegl M, Hilbe W, Weiss W, Lukas P, Jamnig H, Hackl M, Haidenberger A, Buszewski B, Miekisch W, Schubert J, Amann A. Noninvasive detection of lung cancer by analysis of exhaled breath. *BMC Cancer.* 9, 348 (2009).
177. Zhan X, Duan J, Duan Y. Recent development of proton-transfer reaction mass spectrometry (PTR-MS) and its applications in medical research. *Mass Spectrom Rev.* 32, 143-65 (2013).
178. Hryniuk A, Ross BM. A preliminary investigation of exhaled breath from patients with celiac disease using selected ion flow tube mass spectrometry. *J Gastrointest Liver Dis.* 19, 15-20 (2010).
179. Endre ZH, Pickering JW, Storer MK, Hu WP, Moorhead KT, Allardyce R, McGregor DO, Scotter JM. Breath ammonia and trimethylamine allow real-time monitoring of haemodialysis efficacy. *Physiol Meas.* 32, 115-30 (2011).
180. Boshier PR, Cushnir JR, Mistry V, Knaggs A, Španěl P, Smith D, Hanna GB. On-line, real time monitoring of exhaled trace gases by SIFT-MS in the perioperative setting: a feasibility study. *Analyst.* 136, 3233-3237 (2011).

181. Gilchrist FJ, Razavi C, Webb AK, Jones AM, Španěl P, Smith D, Lenney W. An investigation of suitable bag materials for the collection and storage of breath samples containing hydrogen cyanide. *J Breath Res.* 6, 036004 (2012).
182. Španěl P, Dryahina K, Smith D. A quantitative study of the influence of inhaled compounds on their concentrations in exhaled breath. *J Breath Res.* 7, 017106 (2013).
183. Wang T, Pysanenko A, Dryahina K, Španěl P, Smith D. Analysis of breath, exhaled via the mouth and nose, and the air in the oral cavity. *J Breath Res.* 2, 037013 (2008).
184. Smith D, Turner C, Španěl P. Volatile metabolites in the exhaled breath of healthy volunteers: their levels and distributions. *J Breath Res.* 1, 014004 (2007).
185. Kalapos MP. On the mammalian acetone metabolism: from chemistry to clinical implications. *Biochim Biophys Acta.* 1621, 122-139 (2003).
186. Pysanenko A, Wang T, Španěl P, Smith D. Acetone, butanone, pentanone, hexanone and heptanone in the headspace of aqueous solution and urine studied by selected ion flow tube mass spectrometry. *Rapid Commun. Mass Spectrom.* 23, 1097-1104 (2009).
187. Pabst F, Miekisch W, Fuchs P, Kischkel S, Schubert JK. Monitoring of oxidative and metabolic stress during cardiac surgery by means of breath biomarkers: an observational study. *J Cardiothorac Surg.* 2, 37 (2007).
188. Marcondes-Braga FG, Gutz IG, Batista GL, Saldiva PH, Ayub-Ferreira SM, Issa VS, Mangini S, Bocchi EA, Bacal F. Exhaled acetone as a New Biomarker of Heart Failure Severity. *Chest.* 142, 457-466 (2012).
189. Hyspler R, Crhová S, Gasparic J, Zadák Z, Cízková M, Balasová V. Determination of isoprene in human expired breath using solid-phase microextraction and gas chromatography-mass spectrometry. *J Chromatogr B* 739, 183-190 (2000).
190. Španěl P, Davies S, Smith D. Quantification of breath isoprene using the selected ion flow tube mass spectrometric analytical method. *Rapid Commun Mass Spectrom.* 13, 1733-1738 (1999).
191. Xu ZQ, Broza YY, Ionsecu R, Tisch U, Ding L, Liu H, Song Q, Pan YY, Xiong FX, Gu KS, Sun GP, Chen ZD, Leja M, Haick H. A nanomaterial-based breath test for distinguishing gastric cancer from benign gastric conditions. *Br J Cancer.* 108, 941-950 (2013).
192. Martínez-Lozano P, Fernández de la Mora J. Direct Analysis of Fatty Acid Vapors in Breath by Electrospray Ionization and Atmospheric Pressure Ionization-Mass Spectrometry. *Anal Chem.* 80, 8210-8215 (2008).
193. Turner MA, Guallar-Hoyas C, Kent AL, Wilson ID, Thomas CL. Comparison of metabolomics profiles obtained using chemical ionization and electron ionization MS in exhaled breath. *Bioanalysis.* 24, 2731-2738 (2011).
194. Geypens B, Claus D, Evenepoel P, Hiele M, Maes B, Peeters M, Rutgeerts P, Ghooys Y. Influence of dietary protein supplements on the formation of bacterial metabolites in the colon. *Gut.* 41, 70-76 (1997).
195. Silva CL, Passos M, Câmara JS. Solid phase microextraction, mass spectrometry and metabolomic approaches for detection of potential urinary cancer biomarkers-a powerful strategy for breast cancer diagnosis. *Talanta.* 89, 360-8 (2012).
196. Buszewski B, Ligor T, Jezierski T, Wenda-Piesik A, Walczak M, Rudnicka J. Identification of volatile lung cancer markers by gas chromatography-mass spectrometry: comparison with discrimination by canines. *Anal Bioanal Chem.* 404, 141-146 (2012).
197. Altomare DF, Di Lena M, Porcelli F, Trizio L, Travaglio E, Tutino M, Dragonieri S, Memeo V, de Gennaro G. Exhaled volatile organic compounds identify patients with colorectal cancer. *British Journal of Surgery.* 100, 144-150 (2013).
198. Španěl P, Doren JMV, Smith D. A selected ion flow tube study of the reactions of H_3O^+ , NO^+ , and O_2^+ with saturated and unsaturated aldehydes and subsequent hydration of the product ions. *Int. J. Mass Spectrom.* 213, 163-176 (2002).
199. Davis BM, Senthilmohan ST, Wilson PF, McEwan MJ. Major volatile compounds in head-space above olive oil analysis by selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom.* 19, 2272-2278 (2005).
200. Phillips M, Cataneo RN, Greenberg J, Gunawardena R, Naidu A, Rahbari-Oskoui F. Effect of age on the breath methylated alkane contour, a display of apparent new markers of oxidative stress. *J Lab Clin Med* 136, 243-249 (2000).
201. Jareño-Esteban JJ, Muñoz-Lucas MÁ, Carrillo-Aranda B, Maldonado-Sanz JÁ, de Granda-Orive I, Aguilar-Ros A, Civera-Tejuca C, Gutiérrez-Ortega C, Callol-Sánchez LM. Volatile organic compounds in exhaled breath in a healthy population: effect of tobacco smoking. *Arch Bronconeumol.* 49, 457-461 (2013).

202. Ligor T, Ligor M, Amann A, Ager C, Bachler M, Dzien A, Buszewski B. The analysis of healthy volunteers' exhaled breath by the use of solid-phase microextraction and GC-MS. *J Breath Res.* 2, 15 (2008).
203. Corradi M, Rubinstein I, Andreoli R, Manini P, Caglieri A, Poli D, Alinovi R, Mutti A. Aldehydes in Exhaled Breath Condensate of Patients with Chronic Obstructive Pulmonary Disease. *Am J of Resp and Crit Care Med.* 167, 1380-1386 (2003).
204. Bartoli ML, Novelli F, Costa F, Malagrinò L, Melosini L, Bacci E, Cianchetti S, Dente FL, Di Franco A, Vagaggini B, Paggiaro PL. Malondialdehyde in exhaled breath condensate as a marker of oxidative stress in different pulmonary diseases. *Mediators Inflamm.* 2011, 891752 (2011).
205. Rüdiger Siewert J, Feith M, Werner M, Stein HJ. Adenocarcinoma of the Esophagogastric Junction: Results of Surgical Therapy Based on Anatomical/Topographic Classification in 1,002 Consecutive Patients. *Ann Surg.* 232, 353–361 (2000).
206. Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJ, Nicolson M, Scarffe JH, Loft FJ, Falk SJ, Iveson TJ, Smith DB, Langley RE, Verma M, Weeden S, Chua YJ. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med* 355, 11-20 (2006).
207. Wang L, Chen J, Chen L, Deng P, Bu Q, Xiang P, Li M, Lu W, Xu Y, Lin H, Wu T, Wang H, Hu J, Shao X, Cen X, Zhao YL. ¹H-NMR based metabolomic profiling of human esophageal cancer tissue. *Mol Cancer.* 12, 25 (2013).
208. Tan B, Qiu Y, Zou X, Chen T, Xie G, Cheng Y, Dong T, Zhao L, Feng B, Hu X, Xu LX, Zhao A, Zhang M, Cai G, Cai S, Zhou Z, Zheng M, Zhang Y, Jia W. Metabonomics identifies serum metabolite markers of colorectal cancer. *J Proteome Res.* 12, 3000-3009 (2013).
209. Zhang G, Ma L, Xie YK, Miao XB, Jin C. Esophageal cancer tumorspheres involve cancer stemlike populations with elevated aldehyde dehydrogenase enzymatic activity. *Mol Med Rep* 6, 519-524 (2012).
210. Mochalski P, Sponring A, King J, Unterkofler K, Troppmair J, Amann A. Release and uptake of volatile organic compounds by hepatocellular carcinoma cells (HepG2) in vitro. *Cancer Cell Int.* 13, 72 (2013).
211. Ulanowska A, Kowalkowski T, Hryniewicz K, Jackowski M, Buszewski B. Determination of volatile organic compounds in human breath for *Helicobacter pylori* detection by SPME-GC/MS. *Biomed Chromatogr.* 25, 391-7 (2011).
212. Nelson N, Lagesson V, Nosratabadi AR, Ludvigsson J, Tagesson C. Exhaled isoprene and acetone in newborn infants and in children with diabetes mellitus. *Pediatr Res.* 44, 363-367 (1998).
213. Schmutzhard J, Rieder J, Deibl M, Schwentner IM, Schmid S, Lirk P, Abraham I, Gunkel AR. Pilot study: volatile organic compounds as a diagnostic marker for head and neck tumors. *Head Neck.* 30, 743-749 (2008).
214. Poli D, Goldoni M, Caglieri A, Ceresa G, Acampa O, Carbognani P, Rusca M, Corradi M. Breath analysis in non- small cell lung cancer patients after surgical tumour resection. *Acta Biomed.* 79, Suppl 1, 64-72 (2008).
215. Wilson JMG, Jungner G. Principles and Practice of Screening for Disease. *WHO Chronicle* Geneva: World Health Organization. 22, 473. Public Health Papers, 34, (1968).
216. Hounsom L, Grayson K, Melzer M. Mortality and associated risk factors in consecutive patients admitted to a UK NHS trust with community acquired bacteraemia. *Postgrad Med J.* 87, 757-762 (2011).
217. Kunze N, Göpel J, Kuhns M, Jünger M, Quintel M, Perl T. Detection and validation of volatile metabolic patterns over different strains of two human pathogenic bacteria during their growth in a complex medium using multi-capillary column-ion mobility spectrometry (MCC-IMS). *Appl Microbiol Biotechnol.* 97, 3665-3676 (2013).
218. Filipiak W, Sponring A, Baur MM, Filipiak A, Ager C, Wiesenhofer H, Nagl M, Troppmair J, Amann A. Molecular analysis of volatile metabolites released specifically by *Staphylococcus aureus* and *Pseudomonas aeruginosa*. *BMC Microbiol.* 12, 113 (2012).
219. Julák J, Stránská E, Rosová V, Geppert H, Spanel P, Smith D. Bronchoalveolar lavage examined by solid phase microextraction, gas chromatography--mass spectrometry and selected ion flow tube mass spectrometry. *J. Microbiol. Methods,* 65, 76-86 (2006).
220. Huijskens EG, Rossen JW, Kluytmans JA, van der Zanden AG, Koopmans M. Evaluation of yield of currently available diagnostics by sample type to optimize detection of respiratory pathogens in patients with a community-acquired pneumonia. *Influenza Other Respir Viruses.* 8, 243-249 (2014).
221. Filipiak W, Sponring A, Baur MM, Ager C, Filipiak A, Wiesenhofer H, Nagl M, Troppmair J, Amann A. Characterization of volatile metabolites taken up by or released from *Streptococcus pneumoniae* and *Haemophilus influenzae* by using GC-MS. *Microbiology.* 158, 3044-3053 (2012).

222. Dryahina K, Španěl P, Pospíšilová V, Sovová K, Hrdlička L, Machková N, Lukáš M, Smith D. Quantification of pentane in exhaled breath, a potential biomarker of bowel disease, using selected ion flow tube mass spectrometry. *Rapid Commun Mass Spectrom*. 27, 1983-1992 (2013).
223. Morisco F, Aprea E, Lembo V, Fogliano V, Vitaglione P, Mazzone G, Cappellin L, Gasperi F, Masone S, De Palma GD, Marmo R, Caporaso N, Biasioli F. Rapid "breath-print" of liver cirrhosis by proton transfer reaction time-of-flight mass spectrometry. A pilot study. *PLoS One*. 8, 59658 (2013).
224. Probert CS, Ahmed I, Khalid T, Johnson E, Smith S, Ratcliffe NM. Volatile organic compounds as diagnostic biomarkers in gastrointestinal and liver diseases. *J Gastrointest Liver Dis*. 18, 337-343 (2009).
225. Wu H, Xue R, Lu C, Deng C, Liu T, Zeng H, Wang Q, Shen X. Metabolomic study for diagnostic model of oesophageal cancer using gas chromatography/mass spectrometry. *J Chromatogr B*. 877, 3111-3117 (2009).
226. Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, Kinoshita T, Saito N, Ochiai A, Tomita M, Esumi H, Soga T. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res*. 69, 4918-4925 (2009).
227. Hong J, Resnick M, Behar J, Wang LJ, Wands J, DeLellis RA, Souza RF, Spechler SJ, Cao W. Acid-induced p16 hypermethylation contributes to development of esophageal adenocarcinoma via activation of NADPH oxidase NOX5-S. *Am J Physiol Gastrointest Liver Physiol*. 299, 697-706 (2010).
228. Farhadi A, Fields J, Banan A, Keshavarzian A. Reactive oxygen species: are they involved in the pathogenesis of GERD, Barrett's esophagus, and the latter's progression toward esophageal cancer? *Am J Gastroenterol*. 97, 22-26 (2002).
229. Kuhajda FP. Fatty acid synthase and cancer: new application of an old pathway. *Cancer Res*. 66, 5977-5980 (2006).
230. Menendez JA, Vellon L, Oza BP, Lupu R. Does endogenous fatty acid metabolism allow cancer cells to sense hypoxia and mediate hypoxic vasodilatation? Characterization of a novel molecular connection between fatty acid synthase (FAS) and hypoxia-inducible factor-1alpha (HIF-1alpha)-related expression of vascular endothelial growth factor (VEGF) in cancer cells overexpressing her-2/neu oncogene. *J Cell Biochem*. 94, 857-863 (2005).
231. Zhou Y, Niu C, Li Y, Gao B, Zheng J, Guo X, Ma W. Fatty acid synthase expression and esophageal cancer. *Mol Biol Rep*. 39, 9733-9739 (2012).
232. Costello BP, Ewen RJ, Ratcliffe NM. A sensor system for monitoring the simple gases hydrogen, carbon monoxide, hydrogen sulfide, ammonia and ethanol in exhaled breath. *J Breath Res*. 2, 037011 (2008).
233. Khalid T, White P, De Lacy Costello B, Persad R, Ewen R, Johnson E, Probert CS, Ratcliffe N. A pilot study combining a GC-sensor device with a statistical model for the identification of bladder cancer from urine headspace. *PLoS One*. 8, 69602 (2013).
234. Corley DA, Mehtani K, Quesenberry C, Zhao W, de Boer J, Weiss NS. Impact of endoscopic surveillance on mortality from Barrett's esophagus-associated esophageal adenocarcinomas. *Gastroenterology*. 145, 312-319 (2013).
235. Mazzone PJ, Wang XF, Xu Y, Mekhail T, Beukemann MC, Na J, Kemling JW, Suslick KS, Sasidhar M. Exhaled breath analysis with colorimetric sensor array for the identification and characterization of lung cancer. *J Thorac Oncol*. 7, 137-142 (2012).
236. Darwiche K, Baumbach JI, Sommerwerck U, Teschler H, Freitag L. Bronchoscopically obtained volatile biomarkers in lung cancer. *Lung*. 189, 445-52 (2011).
237. Patterson SG, Bayer CW, Hendry RJ, Sellers N, Lee KS, Vidakovic B, Mizaikoff B, Gabram-Mendola SG. Breath analysis by mass spectrometry: a new tool for breast cancer detection? *Am Surg*. 77, 747-751 (2011).
238. Rudnicka J, Kowalkowski, T, Ligor T, Buszewski B. Determination of volatile organic compounds as biomarkers of lung cancer by SPME-GC-TOF/MS and chemometrics. *J Chromatogr B*. 879, 3360-3366 (2011).
239. Shuster G, Gallimidi Z, Reiss AH, Dovgolevsky E, Billan S, Abdah-Bortnyak R, Kuten A, Engel A, Shiban A, Tisch U, Haick H. Classification of breast cancer precursors through exhaled breath. *Breast Cancer Res Treat*. 126, 791-796 (2011).
240. D'Amico A, Pennazza G, Santonico M, Martinelli E, Roscioni C, Galluccio G, Paolesse R, Di Natale C. An investigation on electronic nose diagnosis of lung cancer. *Lung Cancer*. 68, 170-6 (2010).
241. Phillips M, Cataneo RN, Saunders C, Hope P, Schmitt P, Wai J. Volatile biomarkers in the breath of women with breast cancer. *J Breath Res*. 4, 026003 (2010).

242. Westhoff M, Litterst P, Freitag L, Urfer W, Bader S, Baumbach JI. Ion mobility spectrometry for the detection of volatile organic compounds in exhaled breath of patients with lung cancer: results of a pilot study. *Thorax*. 64, 744-8 (2009).
243. Barash O, Peled N, Hirsch FR, Haick H. Sniffing the unique "odor print" of non-small-cell lung cancer with gold nanoparticles. *Small*. 5, 2618-2624 (2009).
244. Chen X, Xu F, Wang Y, Pan Y, Lu D, Wang P, Ying K, Chen E, Zhang W. A study of the volatile organic compounds exhaled by lung cancer cells in vitro for breath diagnosis. *Cancer*. 110, 835-844 (2007).
245. Yu H, Xu L, Wang P. Solid phase microextraction for analysis of alkanes and aromatic hydrocarbons in human breath. *J Chromatogr B*. 826, 69-74 (2005).
246. Machado RF, Laskowski D, Deffenderfer O, Burch T, Zheng S, Mazzone PJ, Mekhail T, Jennings C, Stoller JK, Pyle J, Duncan J, Dweik RA, Erzurum SC. Detection of lung cancer by sensor array analyses of exhaled breath. *Am J Respir Crit Care Med*. 171, 1286-1291 (2005).
247. Phillips M, Cataneo RN, Cummin AR, Gagliardi AJ, Gleeson K, Greenberg J, Maxfield RA, Rom WN. Detection of lung cancer with volatile markers in the breath. *Chest*. 123, 2115-23 (2003).
248. Rieder J, Lirk P, Ebenbichler C, Gruber G, Prazeller P, Lindinger W, Amann A. Analysis of volatile organic compounds: possible applications in metabolic disorders and cancer screening. *Wien Klin Wochenschr*. 113, 181-185 (2001).
249. Abaffy T, Duncan R, Riemer DD, Tietje O, Elgart G, Milikowski C, DeFazio RA. Differential volatile signatures from skin, naevi and melanoma: novel approach to detect a pathological process. *PLoS One*. 5, 13813 (2010).
250. Buszewski B, Ulanowska A, Ligor T, Jackowski M, Kłodzińska E, Szeliga J. Identification of volatile organic compounds secreted from cancer tissues and bacterial cultures. *J Chromatogr B*. 868, 88-94 (2008).
251. Bartolazzi, A. Santonico M, Pennazza G, Martinelli E, Paolesse R, D'Amico A, Di Natale C. A sensor array and GC study about VOCs and cancer cells. *Sensors and Actuators B*. 146, 4830 (2010).
252. Hu YJ, Qiu YH, Chen EG, Ying KJ, Yu J, Wang P. Determination of volatile organic compounds in lung cancer cell lines and lung cancer tissue. *Zhejiang Da Xue Xue Bao Yi Xue Ban*. 39, 278-284 (2010).

APPENDICES

Appendix A

Exhaled Breath Studies

Study (Author & Year)	Number and Types of Patients	Analytical Technique	Sampling Technique	REF
Mazzone <i>et al.</i> (2012)	N=92 lung cancer N=137 controls	Colometric Sensor Array	All subjects performed tidal breathing, inhaling unfiltered air through their nose and exhaling through their mouth into disposable corrugated tubing for a total of 5 minutes. The exhaled breath was drawn across the sensor array at 200mL/min.	235
Darwiche <i>et al.</i> (2011)	N=19 lung cancer	Ion Mobility Spectrometry	Samples were collected during bronchoscopy with the patient under conscious sedation with midazolam and topical administration of xylocaine. Excess fluid or mucus was removed. Air from the non-tumour side was aspirated for 20 seconds followed by aspiration of air from the tumour side.	236
Patterson <i>et al.</i> (2011)	N= 20 breast cancer N=20 healthy controls	Thermal desorption GC-MS	Exhaled breath was collected in a commercially available Teflon/valved sampler equipped with a rapid passive diffusive sampler five times at 5-minute intervals, trapping alveolar breath VOCs. Patients were fasted for 2 hours prior to collection of a single breath sample.	237
Rudnicka <i>et al.</i> (2011)	N=23 lung cancer N=31 controls	SPME GC-TOF/MS	Breath samples were collected in a 1L Tedlar [®] bag using a breath sampler. After collection, a 10mL sample was transferred into a glass vial. Ambient air samples were taken for blank measurements.	238

Shuster <i>et al.</i> (2011)	N=7 women with negative mammography N=16 benign breast conditions N=13 breast cancers	Nanoscale artificial NOSE	The dead space air was collected into a separate bag prior to the collection of alveolar air. Samples were collected on 2 separate visits to the hospital.	239
D'amico <i>et al.</i> (2010)	N=56 lung cancer N=36 healthy control	Electronic Nose	Patients were requested to breathe into 2 Tedlar[®] bags through a mouthpiece. By controlling a valve, the first volume of air fills the smaller bag. As the bag is filled, the increased air resistance opens the valve and so the larger bag is filled. The VOCs in the second, larger bag were subsequently analysed. No food/drink the night before sampling.	240
Peng <i>et al.</i> (2010)	N=30 lung cancer N=26 colon cancer N=22 breast cancer N=18 prostate cancer N=22 healthy controls	Gold nanoparticles GC-MS	3-5 mins repeatedly inhaling to lung capacity through a mouthpiece fitted with a cartridge that removes >99.99% exogenous VOCs from air. Exhalation was through a separate port against 10-15cm of water to ensure closure of vellum and hence no nasal entrainment was included. Alveolar breath was collected in a 750mL bag. Patients were requested to refrain from coffee and alcohol for at least 12 hours prior to collection. 2 samples were collected from each patient and all samples were analysed within 4 days.	72

Poli <i>et al.</i> (2010)	N= 40 lung cancer (NSCLC) N=38 healthy, asymptomatic controls	SPME GC-MS	Single, slow vital capacity breath collected into a Bio-Vac tube trap. At least 150mL of breath was collected stored at 4°C and analysed within 2 hours.	62
Phillips <i>et al.</i> (2010)	N=54 breast cancer (pre-treatment) N=204 cancer-free controls	GC-MS	Subjects wore noseclips and breathed normally for 2 mins through a mouth piece connected to a portable breath collection apparatus. VOCs in 1L breath and 1L room air collected on separate sorbent traps. All breath VOC samples were collected at the Royal Perth Hospital Breast Clinic and sent by express mail for analysis to the Breath Research Laboratory of Menssana Research Inc., Newark, NJ. (N.B. Pilot studies performed in normal volunteers prior to the clinical study demonstrated that breath VOC samples remained stable with no detectable loss or contamination for up to 4 weeks at room temperature and were not affected by international express shipping.)	241
De Gennaro <i>et al.</i> (2010)	N=13 malignant pleural mesothelioma N=13 long term exposure to asbestos N=13 No asbestos exposure	Thermal desorption GC-MS	Patients were requested not to eat or drink for at least 3 hours prior to breath collection. Patients breathed tidally through a mouthpiece, connected to a three-way non-rebreathing valve and an inspiratory VOC-filter (A2, North Safety, NL) for 5 mins. After a single deep inspiration, the patient exhaled a single vital capacity volume into a Tedlar® bag connected to the expiratory port. Two samples of exhaled breath were collected. The second sample was retrieved 2 mins after the 1 st .	65
Qin <i>et al.</i> (2010)	N=30 hepatocellular carcinoma N=27 cirrhosis	GC-MS	All patients were fasted overnight prior to sampling. There was a 3 step process for breathing into the Tedlar® bags (4L-T2PV/L, Delin Company): take a deep breath; hold it for 4 seconds; then exhale smoothly into the bag. Simultaneously, the samples of ambient air were collected for reference. The samples were taken to the laboratory and analysed immediately after collection.	68

	N=36 healthy volunteers			
Song <i>et al.</i> (2010)	N=43 lung cancer (NSCLC) N=41 controls	SPME GC-MS	All patients were fasted overnight prior to sampling. The patients were requested to deeply breathe in Tedlar® gas bags which trapped 4L breath gas through a disposable mouthpiece. Mixed expiratory samples were collected. All samples were analysed within 6 hours of collection.	64
Fuchs <i>et al.</i> (2010)	N=12 lung cancer N=12 healthy smokers N=12 healthy volunteers	SPME GC-MS	10mL of exhaled air was drawn into a gastight syringe under visual control of expired PCO ₂ and immediately transferred into an evacuated sealed 20mL headspace vial. All measurements were repeated once and breath gas samples were processed within 6 hrs of collection.	63
Bajtarevic <i>et al.</i> (2009)	2 separate cohorts for each technique N=220 lung cancer N=441 healthy volunteers N=65 lung cancer N=31 healthy volunteers	PTR-MS GC-MS	Samples of mixed alveolar breath were collected in Tedlar® bags (SKC Inc, Eighty Four, PA) with parallel collection of ambient air (also in Tedlar® bags). Breath gas samples were obtained after a volunteer has been seated for ~5 minutes. Each subject provided 1 or 2 breath samples by use of a straw, all of which were processed within 3-6 hours. Mixed alveolar breath (instead of only alveolar breath) was collected. Before collection of exhaled breath, all bags were thoroughly cleaned to remove any residual contaminants by flushing with nitrogen gas (purity of 99.9999%), and then finally filled with nitrogen and heated at 85°C for more than 8 hours. The bags were completely evacuated at the end of this process.	176
Dragonieri <i>et al.</i> (2009)	N=10 lung cancer (NSCLC)	Electronic Nose	Patients were requested to refrain from food and drink for 2 hours prior to breath sample collection. Patients breathed tidally through a mouthpiece connected to a 3-way non-breathing valve and inspiratory VOC filter for 5 mins. Patient were asked to deeply inspire and then exhale a single tidal volume into a Tedlar® bag	71

	N=10 COPD N=10 healthy controls		connected to expiratory port, followed by a 1 min sampling of exhaled air through a silica filled drying chamber. 2 samples were collected per patient excluding the COPD group. The 2 nd sample was collected 2-5 mins after the 1 st sample.	
Westhoff <i>et al.</i> (2009)	N=32 lung cancer N=54 controls	Ion Mobility Spectrometry	Subjects exhaled through a mouthpiece connected to a Teflon bulb. The breath was passed through an unheated sampling loop connected to a miniaturised suction pump to ensure homogeneous breath sample flow. 10mL of exhaled breath was analysed.	242
Barash <i>et al.</i> (2009)	N=40 lung cancer N=56 healthy subjects	Gold nanoparticles	All participants were requested to abstain from alcohol for 12 hrs and coffee for 1 hour prior to sampling. A subject was asked to inhale through a mouth piece containing a cleaning cartridge on the inspiration port, which removed more than 99.99% of VOCs from the air during inspiration. Participants were asked to exhale against 10cm of H ₂ O to ensure closure of the vellum and so exclude nasal entrainment of gas. Exhaled gas was collected through a separate exhalation port into 750mL Tedlar [®] gas-sampling bag and a minimum of 5 bags were collected for each participant.	243
Ligor <i>et al.</i> (2009)	N=65 lung cancer N= 31 healthy controls	SPME GC-MS	Mixed expiratory and indoor air samples were collected.	87
Wehinger <i>et al.</i> (2007)	N=12 lung cancer N=170 controls	PTR-MS	The patients were requested to provide at least two breath samples that were collected the same day into 3L Tedlar [®] (polyvinylfluoride) bags. All patients provided their samples prior to the first chemotherapy cycle. Mixed expiratory breath samples were taken, with no restriction on a particular part of breath. While donating breath, all patients and controls were sitting at rest, inflating a bag through a disposable mouthpiece. To obtain reasonably constant fractions of alveolar and dead space breath in one bag, all individuals were advised to breathe normally during sample collection.	84
Mazzone <i>et al.</i> (2007)	N=49 lung cancer (NSCLC) N=18 COPD	Colometric Sensor Array	Each participant performed tidal breathing of unfiltered room air for a total of 12 mins. During this time, they inhaled through their nose and exhaled through their mouth into the breath collection device. The exhaled breath was drawn over the sensor array by a pump at ~ 250 mL/min. The sensor array was held in place on a flat-bed scanner. The system was contained in an incubator set to 37°C. At the end of the 12 mins breath	73

	N=15 idiopathic pulmonary fibrosis N=20 pulmonary arterial hypertension N=20 sarcoidosis N=21 controls		collection, all tubing and the sensor array were changed and a sample of room air was drawn across the system for a further 12 mins.	
Phillips <i>et al.</i> (2007)	N=193 primary lung cancer N=211 controls	Thermal desorption GC-MS	All patients were fasted overnight. Patients were requested to breathe normally through the disposable mouthpiece of a portable breath collection apparatus for 2 mins. VOCs within 1L alveolar breath and 1L room air were captured onto separate sorbent traps.	67
Chen <i>et al.</i> (2007)	N=29 lung cancer N=7 chronic bronchitis N=13 healthy controls	SPME GC-MS	No food intake was allowed for 2 hours prior to breath collection. Patients were requested to take a deep breath, followed by holding their breath for 1s and exhaling directly into the sample bag.	244
Phillips <i>et al.</i> (2006)	N=51 histological evidence of breast cancer N=50 abnormal mammogram N=42 age matched controls	Thermal desorption GC-MS	Patients were requested to breathe normally through the disposable mouthpiece of a portable breath collection apparatus for 2 mins. VOCs within 1L alveolar breath and 1L room air were captured onto separate sorbent traps.	66

Yu <i>et al.</i> (2005)	N = 15 lung cancer N= 15 healthy controls	SPME GC	Patient were requested not to smoke for one month prior to exhaled breath sample donation. 5L Tedlar [®] bags were used to collect the exhaled breath. These were cleaned with acetone, rinsed with dionised water, filled with nitrogen and dried at 80°C for 12 hours prior to use. Subjects wore a nose clip and exhaled deeply into the bag. The environmental air was collected in a separate sampling bag for assessment of background VOCs.	245
Poli <i>et al.</i> (2005)	N=36 lung cancer (NSCLC) N=25 COPD N=35 smokers N=50 non-smokers	SPME GC-MS	After 60 mins rest, the participants were requested to perform a single slow vital capacity breath into a one-way valve connected to a Teflon-bulb, which traps the last portion of exhaled air (150 mL). 20 environmental samples were taken from the rooms in which the subjects performed the test in order to compare breath and ambient air VOC levels. Twenty-six of the NSCLC patients agreed to repeat the breath collection during a follow-up visit 15–30 days after surgery; the other 10 patients were excluded from the follow- up study due to significant deterioration in their clinical conditions.	61
Machado <i>et al.</i> (2005)	2 separate cohorts for each phase N=14 lung cancer N=45 healthy controls for training phase N=14 lung cancer N=62 healthy controls	Electronic nose, containing 32 polymer composite sensors	The subject inhaled through a mouth piece containing a cartridge on the inspiration port which removes more than 99.99% of VOCs from the air during inspiration, clearing the air of any ambient contaminants. Exhaled against 10cm of H ₂ O to ensure closure of the vellum and so exclude nasal entrainment of gas. Exhaled gas was collected through a separate exhalation port into a non-reactive Mylar gas-sampling bags. Each sample was analysed a minimum of 5 times.	246

Deng <i>et al.</i> (2004)	N=10 lung cancer (NSCLC) N=10 healthy controls	SPME GC-MS	Minimal details specified.	59
Phillips <i>et al.</i> (2003)	N=87 lung cancer N=91 non-cancer patients N=41 healthy volunteers	Thermal desorption GC-MS	Patients were requested to breathe normally through the disposable mouthpiece of a portable breath collection apparatus for 2 mins. VOCs within 1L alveolar breath were captured onto a sorbent trap.	247
Reider <i>et al.</i> (2001)	N=16 various cancers N=100 healthy volunteers	PTR-MS	All subject were fasted overnight. Smoking and chewing of gum was prohibited 2 hour before exhaled breath sampling. Subjects were asked to take 2 deep breaths prior to exhaling into the sample bags.	248
Phillips <i>et al.</i> (1999)	N=60 lung cancer N=48 controls	GC-MS	All participants were fasted overnight. Alveolar breath samples were collected from patients 24 hours prior to bronchoscopy. Participants wore a nose clip while breathing in and out of the portable electrical device, via a disposable mouthpiece for 5 mins. A 10L sample was pumped through a sorbent trap, which contained activated carbon and trapped the VOCs for analysis. Ambient room air was also collected.	60

Tissue Studies

Study (Author & Year)	Number of samples	Analytical Technique	Sampling techniques	REF
Abaffy <i>et al.</i> (2010)	N= 5 naevi (skin) N= 3 melanoma (skin)	HS-SPME GC-MS	Specimens large enough to obtain two parallel 3-mm samples were used for both histology (H&E staining) and volatiles collection. Analysis for volatile compounds was completed within 3 hours of excision.	249
Buszewski <i>et al.</i> (2008)	N= 5 gastric cancer (3 male, 2 female) Normal gastric tissue was taken from same patients	SPME GC-MS	The presence of cancerous cells in tissues was confirmed or excluded by histopathological examination. The stomach tissues were collected in 20 mL vials after surgical resection then transported to laboratory and analysed. Vials were tightly crimped. Before analysis the vials were weighed (~2 grams). The chromatographic analysis was done no longer than 25 mins after surgical resection.	250
Poli <i>et al.</i> (2008)	N= 27 lung cancer (NSCLC) patients	SPME GC-MS	During surgical resection, 50-100 mg of cancerous and macroscopically unaffected lung tissue samples were collected and stored until analysis at -20°C in RNA-later™ solution (Ambion, USA). Before analysis, each sample was weighed and placed in a 4mL SPME-vial containing 0.5 g of NaCl and buffer solution.	214
Chen <i>et al.</i> (2007)	N = 10 lung cancer patients (16 samples retrieved from 4 cell types (NSCLC, squamous cell, adenocarcinoma, broncho-alveolar)	SPME GC-MS	Excised tissue was rinsed in phosphate-buffered saline (PBS) several times and cut into ~ 1mm pieces. The PBS contained 8.0g/L of NaCl, 0.20g/L KCl, 1.38g/L Na ₂ HPO ₄ and 0.20g/L KH ₂ PO ₄ . After rinsing with PBS, the cells were digested with 0.25% trypsin for 45 minutes at 37°C. The cells were freed with gentle agitation, then 10% bovine serum was added to neutralise the trypsin. The solution was filtered and the collected cells were washed with PBS, suspended and counted. The cells were then added to 10% foetal calf serum, 100U/mL penicillin and 100 U/mL streptomycin. The cells were then plated and cultivated at 37°C in 5% CO ₂ .	244

Blood Studies

Study (Author & Year)	Number of blood samples	Analytical Technique	Sampling techniques	REF
Xue <i>et al.</i> (2008)	N= 19 liver cancer N= 18 controls	SPME GC-MS	The blood samples were immediately frozen at -80°C after sampling and thawed at room temperature before analysis. They were all collected within 1 week and analysed in one batch within one day. The time between sampling and analysis was less than 2 weeks.	74
Deng <i>et al.</i> (2004)	N= 10 lung cancer N= 10 controls	SPME GC-MS	5mL of blood from lung cancer patients and controls was introduced into different headspace glass bottles (15mL). A CAR-PDMS fibre was used for extraction of the volatile compounds from the blood. Extraction was carried out at 60 °C for 15mins, with a stirring ratio of 1100rpm.	59

Urine Studies

Study (Author & Year)	Number of samples	Analytical Technique	Sampling techniques	REF
Silva <i>et al.</i> (2011)	N=14 leukaemia N=12 colorectal N=7 lymphoma N=12 cancer free	Dynamic solid-phase microextraction in headspace mode (dHS-SPME) GC-MS	Each participant provided a sample of morning urine (after overnight fasting) in a 20 mL sterile PVC container. The samples were frozen at -80°C and retained until needed for the experiments. Before extraction, the pH values of the 25mL urine samples were adjusted to 1–2. Three dHS-SPME experimental parameters that influence extraction efficiency (fibre coating, extraction time and temperature of sampling) were optimised using a univariate optimisation design. The highest extraction efficiency was obtained when sampling was performed at 50°C for 60 min using samples with high ionic strengths and under agitation.	70
Guadagni <i>et al.</i> (2011)	N = 10 lung cancer N= 25 healthy subjects	SPME GC-MS	Spot urine specimens were collected in 100mL sterile containers and immediately stored at -20°C. Samples were analysed within a week of collection. After thawing, 3 mL of urine was transferred into sealed vials previously added with 1g NaCl, then 30 µL of the 1 ng µL ⁻¹ IS solution was added and the sample was analyzed by HS-SPME/GC/MS. Hexanal and heptanal were extracted from aqueous or urinary samples by SPME, as follows: 10mL sealed vials containing 1g NaCl, 3 mL of sample and 30 µL of the 1 ng µL ⁻¹ IS solution were kept in an ultrasonic bath at 60°C. PDMS/DVB fibres were directly exposed on the head space above the sample for 20 min. At the end of the sampling time, the fibres was immediately inserted into the GC injector (at 200°C).	69
Spanel <i>et al.</i> (1999)	N= 14 bladder Cancer	SIFT-MS	Samples were immediately frozen at -70°C prior to analysis. Tests showed that there was no apparent deterioration of these urine samples even when stored for several weeks. In these experiments, 10 mL aliquots of urine were introduced into glass bottles of volume approximately 200 mL closed with a septum and, by the addition of appropriate amounts of 0.1M HCl and 0.1 M NaOH, urine samples at pH values of 4 and 8 (in addition to samples at their unchanged pH) were prepared. The bottles with the samples were then placed in a thermostatically controlled	83

	N= 24 prostate Cancer N= 14 healthy volunteers		oven to raise their temperature to 40°C. The headspace above each urine sample was then sampled by puncturing the septum with a needle connected directly to the inlet port of the SIFT mass spectrometer. The mass spectra were then acquired using H_3O^+ precursor ions over 50s integration periods.	
--	--	--	--	--

Cell line Studies

Study (Author & Year)	Number and type of cell lines	Analytical Technique	Sampling techniques	REF
Brunner <i>et al.</i> (2010)	N = 2 lung cancer cell lines (A549 & EPLC)N= 2 controls	PTR-MS	The 2 controls for this study were human bronchial epithelial cells (BEAS2B) and immortalized retinal pigment epithelium cells (RPE). The human A-549 lung epithelial cell line was propagated in DMEM (high-glucose) medium containing 10% heat-inactivated foetal bovine serum, 2 mmol L ⁻¹ L-glutamine (PAA Laboratories), 40 UmL ⁻¹ penicillin G, and 40 µg mL ⁻¹ streptomycin (Invitrogen). RPE and A-549 were grown in culture flasks with a breeding area of 175cm ² and a volume of 640cm ³ until 70–80% confluence, which corresponds to approximately 2 million cells per flask. Both EPLC & BEAS2B cell lines were seeded on to 75-mL flasks RPMI 1640 culture medium (Seromed, Heidelberg, Germany) supplemented with 10% new foetal bovine serum, 2 mmolL ⁻¹ L-glutamine, 24 UmL ⁻¹ penicillin, and 24 µg mL ⁻¹ streptomycin. The cells were grown until semiconfluent to avoid growth arrest because of cell–cell contact during overnight sampling.	85
Filipiak <i>et al.</i> (2010)	N= 1 lung cancer cell line (A549) N= 2 controls	Thermal desorption GC-MS	The 2 controls for this study were human bronchial epithelial cells (HBEpC) and human fibroblast cells (hFB). Both samples had approximately 50 million cells. Samples methods were as described below by Filipiak et al (2008) (Ref 75). The only difference was that the fermenters were sealed for 21 hrs in this study.	77
Bartolazzi <i>et al.</i> (2010)	N=5 cancer (3 melanoma, 1 sarcoma and 1 thyroid cancer)	GC-MS	Cells were derived from the respective primary human tumours and cultured in standard conditions in petri dishes with RPMI 1640 medium supplemented with 2mM glutamine, 10% foetal calf serum, penicillin and streptomycin at 37°C and 5% CO ₂ atmosphere.	251
Hu <i>et al.</i> (2010)	N=3 lung cancer cell lines (A549,	SPME GC-MS	The cells were grown at 37°C in an atmosphere of 5% CO ₂ for 7 days. The culture medium was changed for a final time 24 hours before the VOCs were analysed.	252

	SK-MEM-1, NCIH446)			
	N=1 control (BEAS-2B)			
Sponring <i>et al.</i> (2010)	N=1 (Non-small cell lung cancer cell line NCI-H1666)	Thermal desorption GC-MS	The cells were derived from bronchoalveolar adenocarcinoma. The cells were cultivated in RPMI-1640 culture medium supplemented with 10% foetal calf serum (FCS), penicillin (100,000units/L), streptomycin (100 mg/L) and L-glutamine (293mg/L). Cells were cultivated under standard conditions in a conventional incubator at 37°C in a humidified atmosphere with 92.5% air/7.5% CO ₂ . Prior to VOC measurements 25, 50 and 75 million trypsinized cells were inoculated in 100mL phenol red-free RPMI 1640 medium with 5% FCD, 100000 units/L penicillin, 100mg/ml streptomycin, 293 mg/L L-glutamine in a cell culture fermenter, which was flushed with clean, synthetic air taken from a gas cylinder containing 5% CO ₂ for 10 mins and then sealed for 21 hrs. At the end of the incubation time, 200mL of air from the headspace was retrieved for analysis.	78
Sponring <i>et al.</i> (2009)	N= 1 (Lung cancer cell line (NCI-H2087))	Thermal desorption GC-MS	The lung tumour cell line NCI-H2087 was grown in RPMI-1640 culture medium supplemented with 5% foetal calf serum (FCS), penicillin (100,000units/L), streptomycin (100 mg/L) and L-glutamine (293 mg/l). Cells were cultivated under standard conditions in a conventional incubator at 37°C in a humidified atmosphere with 92.5% air/7.5% CO ₂ . For VOC measurements, 100×10 ⁶ trypsinized cells were inoculated in 200mL phenol red-free medium in a cell culture fermenter, which was flushed with clean, synthetic air taken from a gas cylinder (50L defined gas mixture) containing 5% CO ₂ for at least 90 min and then sealed for 16 hrs. At the end of the incubation time, 200mL of air from the headspace was used for GC-MS analyses.	76
Barash <i>et al.</i> (2009)	N=7 lung cancer cell lines (CALU3; H1650; H4006; H1435; H820;	SPME GC-MS	Grown in 100mm cell culture dish to 95% confluency (7×10 ⁶ cells) using a medium under standard conditions (RPMI 1640 medium + 10% foetal bovine serum; 5% CO ₂ environment). A medium with the same incubation time and conditions but without cells served as the control (in duplicate).	243

	H1975; A549)			
	N=2 controls			
Filipiak <i>et al.</i> (2008)	N= 1 (CALU-1)	Thermal desorption GC-MS	The cells were grown in DMEM high glucose (4.5 g/L) culture medium containing pyruvate (PAA) and supplemented with 10% FCS, penicillin (100,000 units/L), streptomycin (100 mg/L) and L-glutamine (293mg/L). Cells were cultivated under standard conditions in a conventional incubator at 37°C in a humidified atmosphere with 92.5% air/7.5% CO ₂ . For VOC measurements, 50 million trypsinized cells were inoculated in 100 mL phenol red free medium (DMEM high glucose) in a cell culture fermenter, flushed with clean, synthetic air taken from a gas cylinder (50L defined gas mixture) containing 5% CO ₂ for at least 10 min at a flow rate of 100mL/min and sealed for 4 to 18 hrs. At the end of the incubation time, 200mL of air from the headspace was retrieved for analysis.	75
Smith <i>et al.</i> (2003)	N= 2 (CALU-1 and SK- MES)	SIFT-MS	Cell lines were grown in complete media (DMEM, Dulbecco/Vogt modified Eagle's minimal essential medium, supplemented with 10% foetal calf serum, 1% glutamine and antibiotic/antimycotic). Cell growth was conducted in tissue culture flasks (Sarstedt, UK) at 37°C in a 5% CO ₂ incubator. The culture medium was changed every 3–4 days. The cells were detached from the tissue culture flasks before reaching confluence using trypsin/EDTA. Complete medium was then added and the cells/medium were spun in a centrifuge at 1200rpm for 7 mins. The supernatant was removed and complete medium was again added. The cell numbers were counted, and the total was apportioned into four 150mL glass bottles, each sealed with a septum. The sealed bottles/cell cultures were placed in an incubator at 37°C overnight together with a 5th bottle that contained the same volume of media only (no cells). The headspaces above the medium-only and the cell/medium cultures were allowed to develop overnight, and were then sampled directly into the SIFT-MS system.	82

Appendix B

Objective

This pilot study was conducted to identify VOCs of interest within the gastric content from patients with oesophago-gastric cancer using a secondary analytical platform, SPME/GC-MS.

Methods

The following experiments were conducted in the laboratory of Professor Patrik Španěl at the J. Heyrovsky Institute of Physical Chemistry (Academy of Sciences of the Czech Republic). I am indebted to Dr. Kristýna Sovová and Professor Patrik Španěl for their method development of SPME GC-MS analysis for gastric content.

Gastric content sampling

Ten gastric content samples from patients with oesophago-gastric cancer were retrieved from storage in a -80°C freezer. These samples were transported by a professional courier company on dry ice from the Surgical Sciences Laboratory at Imperial College, London to the Academy of Sciences of the Czech Republic. Prospective approval for this project was given by the Imperial College Tissue Bank Committee (Project R13057).

SP-ME GC-MS Analysis of gastric content

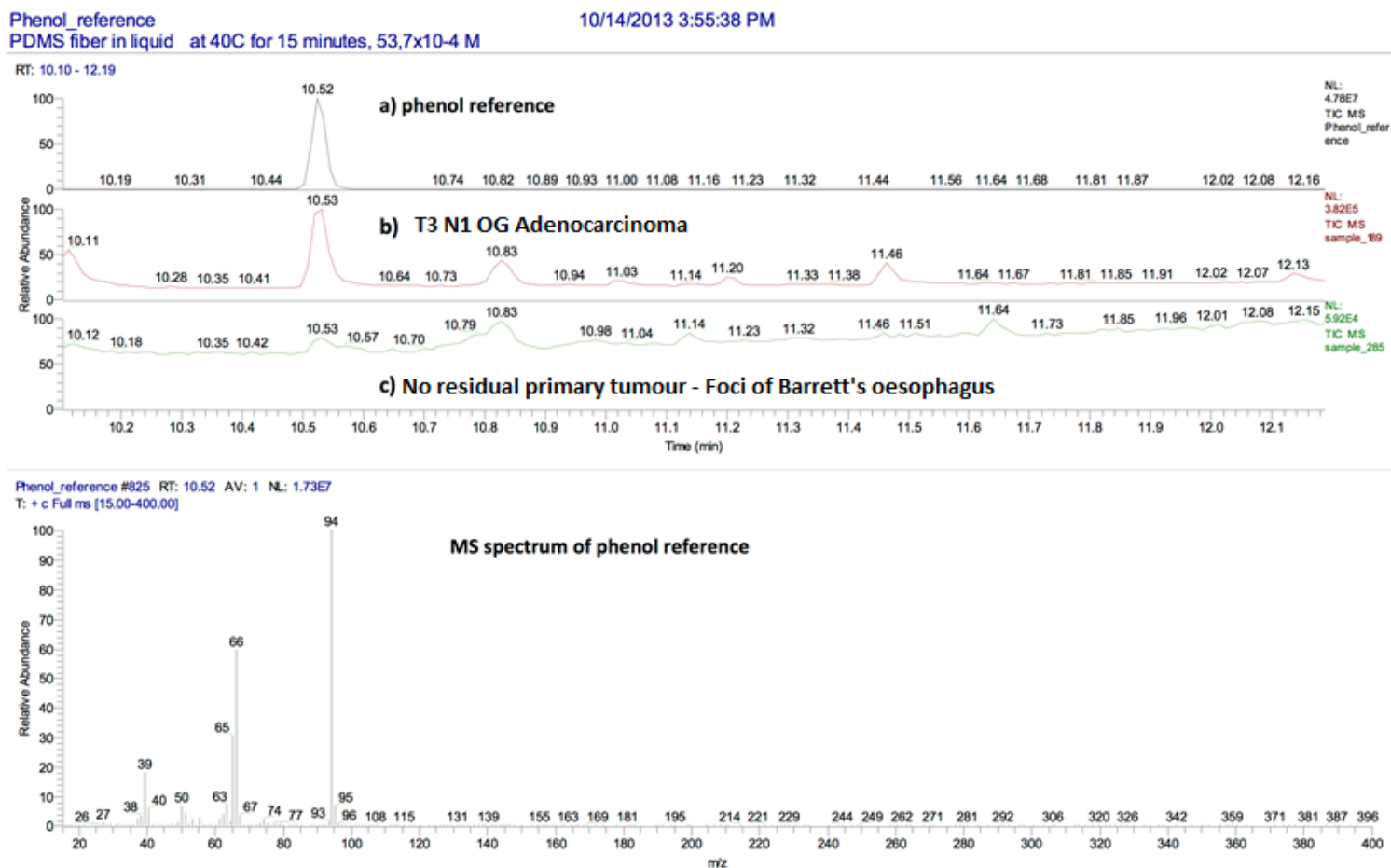
Solid phase microextraction, SPME, was used in combination with gas chromatography mass spectrometry, GC/MS, for qualitative analyses of the headspace of gastric content samples from patients with oesophago-gastric cancer. A SPME fibre coated by carboxen/polydimethylsiloxane (CAR/PDMS) purchased from Supelco (Bellafonte, PA, USA) was exposed to the headspace of gastric content for 45 minutes at 40°C. Extracted trace compounds were then analysed by inserting the SPME fibre into the injector of the Focus GC with ITQ 700 ion trap mass spectrometer (Thermo Electron Corporation) and heated to 200°C. A TR-1 column (30 m length, 0.25 mm internal diameter, 1 µm film of dimethyl polysiloxane) was used for separation using the following temperature programme: 13 min fixed at 38°C, 3°C/min ramp reaching 100°C, 5 min at 100°C, 4°C/min ramp to 150°C, 10°C/min to 200°C finished by a hold of 5 min at 200°C. The ion source was operated in the electron ionisation mode (70 eV, 20 µA) and the mass spectra were obtained using the ion trap operating in a scan mode (scan range from m/z 15-400 at a scan rate of 1 scan/s).

Results & Discussion

Gastric content headspace analysis using a SPME GC-MS method for VOC analysis has not previously been reported. Hence, method development experiments were conducted to optimise adsorption of analytes from the headspace of gastric content. An exposure time of 45 minutes at a temperature of 40°C was the optimal conditions for headspace analysis of (unmodified) gastric content in this study. The VOC that appeared most frequently in the gastric content samples from oesophago-gastric cancer patients was phenol. In Figure B1 below, the phenol reference and 2 representative GC-MS spectra from gastric content samples are shown. Both spectra are from patients who underwent surgical resection of their oesophago-gastric cancer after completing 3 cycles of ECX neoadjuvant chemotherapy treatment (a standard management regimen). Figure B1b demonstrates a strong phenol peak in the patient with a T3 N1 adenocarcinoma. The patient in Figure B1c responded well to his chemotherapy treatment such that there was no residual tumour in the post-operative histology; in this case, there is a much weaker phenol signal. This study was

conducted in a small sample size but clearly identified phenol as a potentially important VOC in patients with OG cancer. In the study on gastric content using SIFT-MS (Section 4.1), phenol was increased in patients with OG cancer compared to controls but this did not reach statistical significance; the lack of statistical validation may be due to a higher proportion of the OG cancer patients in that study having early stage tumours. Alternatively, quantification of phenol using SIFT-MS may have been improved if the NO^+ precursor ion rather than H_3O^+ ion had been employed. Although these are both factors were accounted for in subsequent studies, this pilot study provided further evidence that phenol and its derivatives were likely to be influential VOCs in oesophago-gastric cancer.

Figure B1: GC-MS spectra demonstrating the phenol reference and analysis of the headspace of gastric content from two patients with OG adenocarcinoma



Appendix C

Summary of coefficients and constants derived from the binary logistic regression equation

	Gastric Adenocarcinoma versus normal upper gastro-intestinal tract	Oesophageal Adenocarcinoma versus normal upper gastro-intestinal tract	Gastric Adenocarcinoma versus non cancer controls¹	Oesophageal Adenocarcinoma versus non-cancer control²
Regression equation $\ln(p/(1-p)) = \text{coefficients} \times \text{VOCs} + \text{constant}$				
Butyric acid	n/a	0.03	n/a	0.02
Pentanoic acid	n/a	n/a	-0.06	-0.07
Hexanoic acid	-0.01	0.14	0.10	0.13
Phenol	-0.22	-0.23	-0.05	-0.06
Methyl phenol	-0.41	0.26	-0.14	0.07
Ethyl phenol	0.37	0.20	0.12	0.04
Butanal	0.06	-0.16	-0.38	-0.29
Pentanal	-0.57	-0.11	-0.06	0.08

Hexanal	0.26	0.26	-0.06	-0.03
Heptanal	-0.17	-0.51	-0.06	-0.13
Octanal	-0.31	-0.76	0.11	-0.15
Nonanal	0.92	0.51	-0.002	0.10
Decanal	1.25	0.97	0.49	0.39
Constant	-5.95	-5.87	-2.70	-3.20
R ²	0.85	0.80	0.63	0.59

1. Benign controls and normal upper gastro-intestinal tract.
2. Benign controls, Barrett's metaplasia and normal upper gastro-intestinal tract.
3. Gastric adenocarcinoma and oesophageal adenocarcinoma

Appendix D

Linear regression analysis of the influence of patient demographic and upper gastro-intestinal factors on the concentrations of VOCs:

A	Butanal		Pentanal		Hexanal		Heptanal		Octanal		Nonanal		Decanal	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Status	2.94	0.002	4.57	P<0.0001	4.14	0.002	4.43	P<0.0001	6.90	P<0.0001	8.82	P<0.0001	12.99	P<0.0001
<i>H.Pylori</i>	0.43	0.85	-0.36	0.90	-1.87	0.57	-0.34	0.91	-1.57	0.70	-0.91	0.85	-1.97	0.73
PPI	-0.59	0.50	-1.55	0.17	-0.57	0.64	-0.78	0.49	-0.88	0.57	-1.12	0.54	-1.44	0.50
Age	0.01	0.68	0.001	0.98	0.02	0.74	0.04	0.39	0.01	0.84	0.05	0.48	0.02	0.77
Sex	-1.29	0.17	-1.00	0.41	-1.02	0.45	0.15	0.90	-1.19	0.48	-0.72	0.72	-1.95	0.40
Smoking	0.23	0.69	-1.32	0.08	-0.51	0.54	-0.18	0.82	-0.52	0.62	-1.13	0.35	-0.50	0.97
Alcohol	-0.46	0.53	-0.50	0.60	-0.03	0.98	-0.61	0.53	-1.40	0.28	-1.53	0.32	-3.10	0.09

B	Butyric acid		Pentanoic acid		Hexanoic acid		Phenol		Methyl phenol		Ethyl phenol	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Status	19.57	0.001	5.31	0.04	22.89	P<0.0001	11.99	P<0.0001	8.61	P<0.0001	16.04	P<0.0001
<i>H.Pylori</i>	4.55	0.74	-2.35	0.72	13.84	0.19	0.35	0.95	-0.21	0.96	-0.74	0.94
PPI	-4.49	0.39	0.29	0.91	-1.08	0.79	-0.78	0.73	-0.64	0.68	-3.32	0.34
Age	0.15	0.46	0.01	0.94	0.10	0.54	0.04	0.67	0.02	0.77	0.08	0.57
Sex	-4.03	0.48	-2.75	0.30	-3.62	0.41	-4.65	0.06	-2.16	0.21	-5.02	0.18
Smoking	-3.02	0.39	1.52	0.36	-1.71	0.53	-1.26	0.40	-0.27	0.78	-2.61	0.26
Alcohol	-5.98	0.18	-1.86	0.37	0.18	0.96	-1.18	0.54	-0.50	0.71	-1.46	0.62

C	Methanol		Acetone		Ammonia		Isoprene	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
Status	-2.83	0.89	-61.34	0.64	90.17	0.17	18.13	0.04
<i>H.Pylori</i>	-9.92	0.84	167.12	0.61	335.48	0.04	6.58	0.76
PPI	9.47	0.61	-101.67	0.41	119.56	0.06	-9.55	0.24
Age	0.01	0.99	2.77	0.56	1.55	0.52	0.13	0.67
Sex	-5.33	0.79	-123.57	0.36	-148.05	0.03	-2.56	0.78
Smoking	-18.63	0.14	-141.22	0.09	-70.61	0.09	-2.46	0.65
Alcohol	22.91	0.15	176.55	0.09	46.48	0.38	-3.55	0.61

Selected Ion Flow Tube-MS Analysis of Headspace Vapor from Gastric Content for the Diagnosis of Gastro-Esophageal Cancer

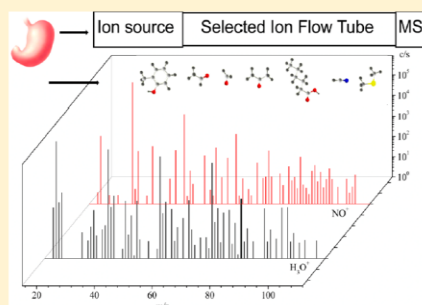
Sacheen Kumar,^{†,‡} Juzheng Huang,^{†,‡} Julia R. Cushnir,[†] Patrik Španěl,[¶] David Smith,[§] and George B. Hanna^{*,†}

[†]Department of Surgery and Cancer, Imperial College London, 10th Floor QEQM Wing, St Mary's Hospital, London W2 1NY, United Kingdom

[¶]J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 18223 Prague 8, Czech Republic

[§]Institute for Science and Technology in Medicine, Keele University, Guy Hilton Research Centre, Thornburrow Drive, Hartshill, Stoke-on-Trent ST4 7QB, United Kingdom

ABSTRACT: Gastric content is a complex biofluid within the human stomach which has an important role in digestive processes. It is believed that gastric content may be a contributory factor in the development of upper gastro-intestinal diseases. In this work, selected ion flow tube mass spectrometry (SIFT-MS) has been applied to the quantification of volatile organic compounds (VOCs) in the headspace vapor of gastric content samples, which were retrieved from three groups of patients, including those with gastro-esophageal cancer, noncancer diseases of the upper gastro-intestinal tract, and a healthy cohort. Twelve VOCs have been investigated in this study; the following 7 VOCs, acetone, formaldehyde, acetaldehyde, hexanoic acid, hydrogen sulphide, hydrogen cyanide, and methyl phenol, were found to be significantly different between cancer and healthy groups by the Mann–Whitney U test. Receiver operating characteristics (ROC) analysis was applied for the combined VOCs of acetaldehyde, formaldehyde, hydrogen sulphide, and methyl phenol to discriminate cancer patients from healthy controls. The area under the curve (AUC) was 0.9. This result raises the prospect that a VOC profile rather than a single biomarker may be preferable in the molecular-orientated diagnosis of gastro-oesophageal cancer, and this warrants further investigation to assess its potential application as a new diagnostic test.



Volatile organic compounds (VOCs) emitted from the human body have been of interest to researchers for over four decades. In 1971, Nobel-prize winner Linus Pauling demonstrated that breath is a complex mixture containing about 250 VOCs; he also showed that urine vapor has over 280 VOCs.¹ Since that time, research in this field has focused on identifying VOCs associated with certain disease states, the ultimate translational application for this research being the quantification of VOCs as a noninvasive diagnostic and therapeutic monitoring tool in medicine. Several chemical analytical techniques have been employed to study VOCs in disease, the most common being gas chromatography mass spectrometry (GC/MS). One of the most common disease states that has been studied in relation to VOC analysis is cancer. Phillips et al. studied the exhaled breath of lung cancer patients and controls and reported that alkanes, alkane derivatives, and benzene derivatives were able to differentiate those with and without lung cancer.² GC/MS coupled with sample collection and concentration techniques, in particular solid phase microextraction, SPME, has further enhanced the isolation and identification of VOCs from mixed samples. Using SPME/GC/MS to analyze urine, Silva et al. showed that

concentration of dimethyl disulfide was statistically significantly lower in cancer patients compared to healthy controls, while concentrations of ketones, phenols, and benzene derivatives were found to be higher in the cancer group when compared to the healthy control group.^{3,4} Xue et al. used SPME/GC/MS to investigate the blood VOCs linked with liver cancer. Nineteen blood VOCs were found to be different between the liver cancer and the control groups ($p < 0.05$). Of these 19 compounds, 3 (hexanal, 1-octen-3-ol, and octane) were found to be present at much higher concentrations in blood of the liver cancer group compared to the control group.⁵ These studies all demonstrate promising associations between certain VOCs and cancer states in a variety of biological samples.

However, there are several important considerations for the application of GC/MS in the analysis of biological samples. It remains an offline technique requiring appropriate GC-column selection and preanalysis sample preparation. Analysis times are lengthy and VOCs with low molecular weight (some of which

Received: August 20, 2012

Accepted: October 4, 2012

Selected Ion Flow Tube Mass Spectrometry Analysis of Volatile Metabolites in Urine Headspace for the Profiling of Gastro-Esophageal Cancer

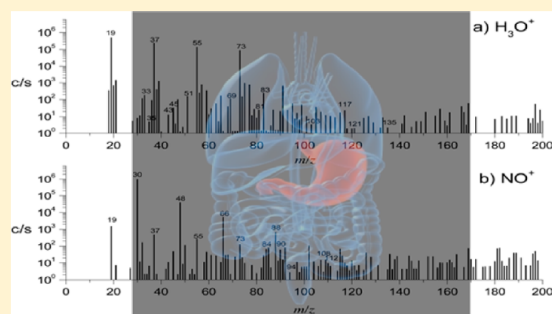
Juzheng Huang,^{†,||} Satcheen Kumar,^{†,||} Nima Abbassi-Ghadi,[†] Patrik Španěl,[‡] David Smith,[§] and George B. Hanna^{*,†}

[†]Department of Surgery and Cancer, Imperial College London, 10th Floor QEOM Wing, St Mary's Hospital, London W2 1NY, United Kingdom

[‡]J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 18223 Prague 8, Czech Republic

[§]Institute for Science and Technology in Medicine, Keele University, Guy Hilton Research Centre, Thornburrow Drive, Hartshill, Stoke-on-Trent ST4 7QB, United Kingdom

ABSTRACT: Urine is considered an ideal biofluid for clinical investigation because it is obtained noninvasively and relatively large volumes are easily acquired. In this study, selected ion flow tube mass spectrometry (SIFT-MS) has been applied for the quantification of volatile organic compounds (VOCs) in the headspace vapor of urine samples, which were retrieved from three groups of patients with gastro-esophageal cancer, noncancer diseases of the upper gastro-intestinal tract, and a healthy cohort. Eleven VOCs have been investigated in this study. The concentrations of seven VOCs—acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulfide, methanol, and phenol—were found to be significantly different between cancer, positive control, and healthy groups using the Kruskal–Wallis test. The concentrations of acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulfide, and methanol were increased in the cancer cohort compared with healthy controls while the concentration of phenol decreased. The differences in the concentrations of ethanol, propanol, methyl phenol, and ethyl phenol were not significant between cancer and control groups. Receiver operating characteristics (ROC) analysis was applied for a combination of six VOCs (acetaldehyde, acetone, acetic acid, hexanoic acid, hydrogen sulfide, and methanol) to discriminate cancer patients from noncancer controls. The integrated area under ROC curve is 0.904. This result indicates that VOC profiling may be suitable in identifying those at high risk of gastro-esophageal cancer. Therefore, further investigations should be undertaken to assess the potential for VOC profiling as a new screening test in gastro-esophageal cancer.



We have previously investigated volatile organic compounds (VOCs) in the headspace vapor of gastric content in patients with gastro-esophageal cancer and controls using selected ion flow tube mass spectrometry (SIFT-MS).¹ Seven VOCs, (viz., acetone, formaldehyde, acetaldehyde, hexanoic acid, hydrogen sulfide, hydrogen cyanide, and methyl phenol) were identified that were significantly different between cancer patients and the healthy control group.¹ These results identified potential VOCs that may be important in the diagnosis of gastro-esophageal cancer. However, the retrieval of gastric content as part of a diagnostic test is not a straightforward prospect. Gastric content can be accessed by only invasive procedures. In order to develop a VOCs-based diagnostic test for gastro-esophageal cancer, it would be preferable to utilize a biofluid that can be obtained non-invasively.

Urine is a biofluid, which can provide good information regarding the physiological state of a person. The analysis of urine is a quick, cost-effective, routinely employed diagnostic test for the assessment of infection, diabetes mellitus, and pregnancy. However, the chemical complexity of urine was first highlighted in 1971 by Linus Pauling, who demonstrated that urine vapor contains over 280 VOCs.² In subsequent years, researchers have focused on identifying VOCs in specific disease states; the ultimate aim of such research being the development of VOC profiling as a clinical tool in disease management. Urine is considered an ideal biofluid for analysis because it is obtained by noninvasive sampling and relatively large volumes are available.

Received: January 11, 2013

Accepted: February 19, 2013

Published: February 19, 2013

Selected Ion Flow Tube Mass Spectrometry Analysis of Exhaled Breath for Volatile Organic Compound Profiling of Esophago-Gastric Cancer

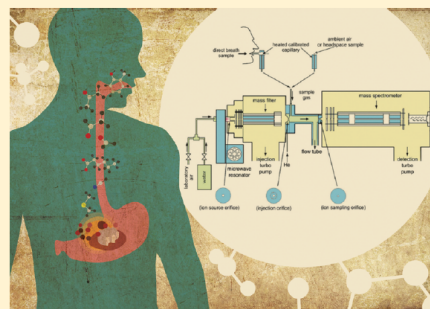
Sacheen Kumar,^{†,‡} Juzheng Huang,^{†,‡} Nima Abbassi-Ghadi,[†] Patrik Španěl,[¶] David Smith,[§] and George B. Hanna^{*,†}

[†]Department of Surgery and Cancer, Imperial College London, 10th Floor QEOM Wing, St Mary's Hospital, London, W2 1NY United Kingdom

[¶]J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 18223 Prague 8, Czech Republic

[§]Institute for Science and Technology in Medicine, Keele University, Guy Hilton Research Centre, Thornburrow Drive, Hartshill, Stoke-on-Trent, ST4 7QB United Kingdom

ABSTRACT: Exhaled breath analysis of volatile organic compounds (VOCs) has great potential in terms of disease diagnosis and measuring physiological response to treatment. In this study, selected ion flow tube mass spectrometry (SIFT-MS) has been applied for the quantification of VOCs in the exhaled breath from 3 groups of patients, viz., those with esophago-gastric cancer, noncancer diseases of the upper gastro-intestinal tract, and a healthy upper gastrointestinal tract cohort. A total of 17 VOCs have been investigated in this study. The concentrations of 4 VOCs, hexanoic acid, phenol, methyl phenol, and ethyl phenol, were found to be significantly different between cancer and positive control groups using the Mann–Whitney U test. Receiver operating characteristics (ROC) analysis was applied for a combination of 4 VOCs (hexanoic acid, phenol, methyl phenol, and ethyl phenol) to discriminate the esophago-gastric cancer cohort from positive controls. The integrated area under the ROC curve (AUC) is 0.91. The results highlight the potential of VOC profiling as a noninvasive test to identify those with esophago-gastric cancer.



The analysis of exhaled breath through the detection and accurate quantification of volatile organic compounds (VOCs) has great potential in terms of disease diagnosis and measuring physiological response to treatment. The major appeal of breath analysis for both clinicians and patients is that it is a noninvasive technique and painless and samples can readily be provided. The development of clinical breath tests for specific diseases may also potentially reduce the number of invasive procedures undertaken and result in fewer hospital visits. As well as the potential financial savings, these indirect benefits are particularly advantageous in certain patients groups, including young children and the elderly. However, there remain limited medical applications of human breath analysis. The ¹³C urea breath testing for *H. pylori* bacterium, hydrogen breath test for small bowel bacterial overgrowth, and the measurement of exhaled nitric oxide in asthma represent the few routinely employed clinical breath tests.^{1–3} Hence, there remains much scientific interest and research effort in developing breath analysis for other diseases.

Cancer remains one of the leading causes of death worldwide accounting for 7.6 million deaths in 2008.⁴ Between 1979 and 2008, the European age-standardized mortality rates for all malignant neoplasms fell by 20% from 219 to 176 per 100,000

population.⁵ Screening programs and disease awareness coupled with treatment advances have contributed to the decrease in cancer mortality rates, although the incidence continues to increase. However, it is estimated that a further 30–40% of these deaths could be prevented.⁶ In particular, esophago-gastric cancer has a poor long-term survival. Between 2002 and 2008, the relative five-year survival rate in the USA for regional esophageal cancer was 20% and for distant esophageal cancer was 3%.⁷ Only 20% of these patients are suitable for potentially curative treatment at first presentation due to a paucity of “red-flag” symptoms at the early stages of the cancer. These statistics highlight the importance of developing new methods for early disease detection in order to improve esophago-gastric cancer-related outcomes. The rationale for development of a non-invasive test would be to identify “at-risk” patients, who could subsequently be sent for esophagoduodenoscopy (EGD), a camera test which allows both direct visualization and tissue biopsy of the upper gastro-intestinal tract. The ideal biological

Received: April 8, 2013

Accepted: May 9, 2013

Published: May 9, 2013

Investigation of C3–C10 aldehydes in the exhaled breath of healthy subjects using selected ion flow tube-mass spectrometry (SIFT-MS)

Juzheng Huang¹, Sacheen Kumar¹ and George B Hanna

Department of Surgery and Cancer, Imperial College London, 10th Floor QEOM Wing, St Mary's Hospital, London W2 1NY, UK

E-mail: g.hanna@imperial.ac.uk

Received 10 January 2014, revised 28 March 2014

Accepted for publication 7 April 2014

Published 4 September 2014

Abstract

Aldehydes have attracted great scientific and clinical interest as potential disease biomarkers. We have investigated selected ion flow tube-mass spectrometry (SIFT-MS) in detecting and quantifying C3 to C10 saturated aldehydes (propanal, butanal, pentanal, hexanal, heptanal, octanal, nonanal and decanal) from the exhaled breath of 26 healthy human volunteers. To assess the reliability of the Nalophan[®] bag sampling method employed, the water level in the breath sample was measured up to 4 h after collection and showed no significant degradation. Propanal was found to be the most abundant aldehyde in the exhaled breath of healthy volunteers. For the C4–C10 aldehydes, their median concentrations were all less than 3 ppbv, demonstrating only trace quantities are present in the exhaled breath of the 26 healthy volunteers.

Keywords: aldehydes, breath analysis, SIFT-MS

(Some figures may appear in colour only in the online journal)

1. Introduction

The detection and quantification of volatile organic compounds (VOCs) within exhaled breath is an evolving scientific field which has great potential for clinical diagnostics. The major appeal of human breath analysis lies in the fact that it is a non-invasive technique, which could eventually be used in the physician's consultation room. Moreover, the development of breath testing for therapeutic monitoring of diseases could improve patient experience through a reduction in the number of invasive procedures required and fewer hospital visits. Hence, there remains continued scientific interest and research effort in developing breath analysis into a clinical reality.

It is known that human breath is a complex mixture containing over 250 VOCs [1]. In medicine, ¹³C urea breath testing for *Helicobacter pylori* bacterium and the measurement of exhaled nitric oxide as a biomarker of airway inflammation in asthma represent the only routinely employed clinical tests [2, 3]. In recent years, the potential of breath analysis has been studied in several diseases including lung cancer, cystic fibrosis, breast cancer and end-stage renal disease [4–7]. These studies have employed various chemical analytical techniques and have demonstrated some very promising VOC-disease associations.

Gas-chromatography mass spectrometry (GC-MS) is the most commonly employed analytical instrument for the detection of VOCs from exhaled breath. However, it remains an analytical technique which requires capture devices, sample pre-concentration, column calibration for all desired analytes as well as lengthy sample processing times. GC-MS extraction

¹ JH and SK are joint first authors.

Mass Spectrometric Analysis of Exhaled Breath for the Identification of Volatile Organic Compound Biomarkers in Esophageal and Gastric Adenocarcinoma

Sacheen Kumar, MRCS,* Juzheng Huang, PhD,* Nima Abbassi-Ghadi, MRCS,* Hugh A. Mackenzie, MRCS,* Kirill A. Veselkov, PhD,* Jonathan M. Hoare, PhD, FRCP,† Laurence B. Lovat, PhD, FRCP,‡ Patrik Španěl, PhD,§ David Smith, PhD, FRS,¶ and George B. Hanna, PhD, FRCS*

Objective: The present study assessed whether exhaled breath analysis using Selected Ion Flow Tube Mass Spectrometry could distinguish esophageal and gastric adenocarcinoma from noncancer controls.

Background: The majority of patients with upper gastrointestinal cancer present with advanced disease, resulting in poor long-term survival rates. Novel methods are needed to diagnose potentially curable upper gastrointestinal malignancies.

Methods: A Profile-3 Selected Ion Flow Tube Mass Spectrometry instrument was used for analysis of volatile organic compounds (VOCs) within exhaled breath samples. All study participants had undergone upper gastrointestinal endoscopy on the day of breath sampling. Receiver operating characteristic analysis and a diagnostic risk prediction model were used to assess the discriminatory accuracy of the identified VOCs.

Results: Exhaled breath samples were analyzed from 81 patients with esophageal (N = 48) or gastric adenocarcinoma (N = 33) and 129 controls including Barrett's metaplasia (N = 16), benign upper gastrointestinal diseases (N = 62), or a normal upper gastrointestinal tract (N = 51). Twelve VOCs—pentanoic acid, hexanoic acid, phenol, methyl phenol, ethyl phenol, butanal, pentanal, hexanal, heptanal, octanal, nonanal, and decanal—were present at significantly higher concentrations ($P < 0.05$) in the cancer groups than in the noncancer controls. The area under the ROC curve using these significant VOCs to discriminate esophageal and gastric adenocarcinoma from those with normal upper gastrointestinal tracts was 0.97 and 0.98, respectively. The area under the ROC curve for the model and validation subsets of the diagnostic prediction model was 0.92 ± 0.01 and 0.87 ± 0.03 , respectively.

Conclusions: Distinct exhaled breath VOC profiles can distinguish patients with esophageal and gastric adenocarcinoma from noncancer controls.

Keywords: breath analysis, esophageal cancer, gastric cancer, mass spectrometry, volatile organic compounds

(Ann Surg 2015;00:1–10)

From the *Department of Surgery and Cancer, Imperial College London, St Mary's Hospital, London, United Kingdom, †Department of Medicine, Imperial College London, St Mary's Hospital, London, United Kingdom; ‡Department of Surgery and Interventional Science, National Medical Laser Centre, University College London, London, United Kingdom; §J. Heyrovsky Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Prague, Czech Republic; and ¶Institute for Science and Technology in Medicine, Keele University, Guy Hilton Research Centre, Hartshill, United Kingdom.

Disclosure: Supported by the Imperial National Institute for Health Research Biomedical Research Centre, Rosetrees Trust and The Stonegate Trust (M261CD1), and the Imperial College Junior Fellowship Research Program. The authors declare no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.annalsofsurgery.com).

Reprints: George B. Hanna, PhD, FRCS, Department of Surgery and Cancer, Imperial College London, 10th Floor QEOM Wing, St Mary's Hospital, London, W2 1NY UK. E-mail: g.hanna@imperial.ac.uk

Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 0003-4932/15/00000-0001

DOI: 10.1097/SLA.0000000000001101

Esophageal and gastric malignancies account for 15% of cancer-related deaths globally.¹ Only one-third of patients diagnosed with esophagogastric (EG) cancer are considered suitable for treatment with curative intent.² Between 2003 and 2009, the 5-year relative survival rate in the United States for esophageal cancer was 17.9% and for gastric cancer was 27.7%.³ The lack of “alarm” symptoms until the disease is at a more advanced stage contribute to these poor statistics. However, diagnosis of these cancers at a treatable stage is associated with a significant survival benefit. Treatment of early esophageal adenocarcinoma has a reported overall 5-year survival rate of 83% to 96%.⁴ However, in Western countries, an endoscopy-based screening approach is not justified given the low incidence of significant pathology and associated health care costs [estimated at ~\$88,000 for detecting each upper gastrointestinal (UGI) malignancy].⁵ Alarm symptoms also have low predictive value for underlying malignancy, despite forming a major part of urgent referral guidance for primary care physicians.⁶ Thus, there remains an important clinical need to develop novel methods for early disease detection in esophageal and gastric cancer.

Breath analysis is a diagnostic modality with few routine applications in clinical practice. Notable examples include the measurement of ¹³C urea for *Helicobacter pylori* bacterium, hydrogen breath testing for small bowel bacterial overgrowth, and exhaled nitric oxide in asthma.^{7–9} However, human breath is also a complex gaseous biological sample containing more than 250 volatile organic compounds (VOCs).¹⁰ In recent years, the role of VOCs in cancer has been investigated using both innovative and conventional methods. Sonoda et al¹¹ demonstrated that canine olfaction of VOCs within exhaled breath and stool samples could discriminate patients with colorectal cancer from controls with good accuracy. The study reinforced the theory that cancer-related VOCs exist, and investigation of these compounds can be undertaken using modern chemical analytical techniques. Several studies employing Gas-Chromatography Mass Spectrometry (GC-MS) have identified potential VOCs of interest within exhaled breath in cancer.^{12–15} However, GC-MS is an off-line analytical technique that requires breath adsorption devices (eg, Solid Phase Micro-extraction) and column calibration for desired analytes. GC-MS is also less preferable to quantify some important trace compounds that are present in exhaled breath and has lengthy processing times. Technological advances have resulted in the development of novel chemical analytical techniques including proton transfer reaction mass spectrometry and Selected Ion Flow Tube Mass Spectrometry (SIFT-MS), both of which offer real-time quantification of VOCs.^{16,17} SIFT-MS is a valuable method for ambient gas analysis that utilizes a soft ionization method, which can accurately analyze the trace organic compounds present in air and other complex matrices including exhaled breath and the headspace of blood, urine, and cell and bacterial cultures. The samples can be analyzed in real time without modifying or disturbing the medium, which can occur, for example, when collecting and concentrating the gas phase trace