BREATH ANALYSIS FOR THE NON-INVASIVE ASSESSMENT OF COLORECTAL CANCER

Geng-ping Lin

Department of Surgery and Cancer, Imperial College London,

St Mary's Hospital, London, UK

2023

A thesis submitted to the University of London for the degree of

Doctorate of Philosophy

Supervisors

Professor George B Hanna

Department of Surgery and Cancer, Imperial College London,

Mr Piers Boshier

Department of Surgery and Cancer, Imperial College London,

Dr Ilaria Belluomo

Department of Surgery and Cancer, Imperial College London,

STATEMENT OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any mediumor format. You may also create and distribute modified versions of the work.

This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to
others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

ABSTRACT

Colorectal cancer (CRC) is one of the leading malignancies worldwide and the early diagnosis of CRC is the key to better prognosis. Breath analysis which lacks standardization is a novel method potentially capable of non-invasive CRC detection and monitoring. The objective was to develop a VOC base model for diagnosis and monitoring of CRC. Secondary aims were to explore the different methodology including the influence of off-line sampling, perform multiplatform correlation of VOCs in human breath.

Initial studies determined the strengths and weaknesses of three different breath sampling devices. Findings showed that no single device has all the desired attributes of optimal method for breath collection system. The device chosen for future clinical studies because of the acceptable VOC capture, repeatability, and human factor analysis. A study of the comparability of the different VOC analytical methods were done to confirm the possibility of cross validating the VOCs. The results supported the implementation of an off- line thermal desorption-based platform for breath collection in multi-site studies for breath biomarker research. Analysis of room air VOCs provide baseline understanding of their ambient levels and potential contribution as contaminants. The background noise could be mostly ignored with a single background sample analysis.

Adapting these methods VOCs levels were examined in the tumour headspace and exhaled breath of patients with colorectal cancer. Tumour headspace analysis showed tumour and colonic mucosa might release different VOC which is detectable on mass spectrometry. Exhaled breath from CRC patients provided tentative evidence that selected exhaled VOCs were linked to CRC and that the compounds may change as a response to therapeutic intervention and disease status.

Further work is however, needed to refine the methodology for sample collection and handling to ensure reliability of results.

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 conducted work in collaboration with the author are appropriately credited. I authorise the library of the University of London to lend this thesis to other institutions or individuals.

Signed:

Name in print: Geng-ping Lin

Date:

PEER REVIEWED PUBLICATIONS

1. **Lin GP**, Vadhwana B, Belluomo I, Boshier PR, Španěl P, Hanna GB. Cross Platform Analysis of Volatile Organic Compounds Using Selected Ion Flow Tube and Proton-Transfer-Reaction Mass Spectrometry. J Am Soc Mass Spectrum. 2021 May 5;32(5):1215-1223. doi: 10.1021/jasms.1c00027. Epub 2021 Apr 8. PMID: 33831301.

2. Adam ME, Fehervari M, Boshier PR, Chin ST, **Lin GP**, Romano A, Kumar S, Hanna GB. Mass-Spectrometry Analysis of Mixed-Breath, Isolated-Bronchial-Breath, and Gastric-Endoluminal-Air Volatile Fatty Acids in Esophagogastric Cancer. Anal Chem. 2019 Mar 5;91(5):3740-3746. doi: 10.1021/acs.analchem.9b00148. PMID: 30699297.

PRESENTATIONS TO LEARNED SOCIETIES

1. Analysis of In Situ Colorectal Tumour Headspace for The Assessment of Cancer Specific Volatile Organic Compounds Release, APFCP-ASSR 2021, Taiwan, November 2021, oral presentation

2. Cross platform analysis of volatile organic compounds using selected ion flow tube-, and proton transfer reaction mass spectrometry, MSACL 2019 Salzburg, September 2019, poster presentation

3. Evaluation of off-line breath collection methods and assessment of utility for large scale clinical trials. Department of Surgery and cancer research afternoon cancer December 2016 poster presentation.

ACKNOWLEDGEMENTS

I am greatly indebted to my supervisors **Professor George Hanna** and **Dr Piers Boshier** and, **Dr. Ilaria Belluomo** for their continued guidance, encouragement, and support during the course of my study.

I would also like to express my gratitude to the following people who have helped me during the course of my research.

Professor David Smith – for guidance and advice during SIFT-MS studies
Professor Patrik Španěl – for guidance and advice during SIFT-MS studies
Dr. Andrea Romano – for guidance and advice during PTR-MS studies
Dr. Sung-tong Chin – for guidance and advice during GC-MS studies
Dr. Tanzeela Khalid – for assistance with device comparison studies
Dr. Georgia Woodfield – for assistance with device comparison studies
Mr. Malik Iltisham – for assistance with clinical sample collection
Miss Fatima Akbar – for assistance with clinical sample collection
Miss Prior Florence – for assistance with clinical sample collection

This PhD was supported by a studentship from Chang Gung Memorial Hospital, Linkou branch, Taiwan.

TABLE OF CONTENTS

Contents

STATEMENT OF ORIGINALITY 3

COPYRIGHT DECLARATION 4

ABSTRACT 5

DECLARATION 6

PEER REVIEWED PUBLICATIONS 7

PRESENTATIONS TO LEARNED SOCIETIES 8

ACKNOWLEDGEMENTS 9

TABLE OF CONTENTS 10

LIST OF FIGURES 13

LIST OF TABLES 14

1. BACKGROUND 17

1.1 Diagnosis and treatment for colorectal cancer 17

1.1.1 Epidemiology of Colorectal Cancer 17

1.1.2 Colorectal cancer detection 17

1.1.3 Treatment for colorectal cancer 18

1.1.4 Post-operative follow-up 21

1.2 Early detection of colorectal cancer 22

1.2.1 Faecal occult blood test (FOBT) 22

1.2.2 Direct assessment with colonoscopy 23

1.2.3 Developing methods for colorectal cancer screening 24

1.2.4 Comparison of the alternative method for early detection of colorectal cancer 25

1.3 Volatile organic compounds and cancer detection 26

1.3.1 VOCs within exhaled breath 26

1.3.2 Current application of VOCs in disease diagnosis 26

1.4 Systematic review of studies reporting volatile organic compounds as biomarkers of colorectal cancer 29
1.4.1 Introduction and aim of the systematic review 29
1.4.2 Methodology 29
1.4.3 Results 30
1.4.4 Discussion..31
1.5 Methods for volatile organic compounds analysis...37
1.5.2 Mass Spectrometry 37
1.5.3 Other analytical platforms 43
1.6 Current challenges with breath VOCs analysis..51
1.7 Summary of current knowledge and indications for future investigations 53
1.8 Overall objective of this thesis 54
2. METHOD DEVELOPMENT 56
2.1 Evaluation of off-line breath collection methods and assessment of utility for large scale clinical trials...57
2.1.1 Introduction 57
2.1.2 Method..58
2.1.3 Results 67
2.1.4 Discussion...81
2.1.5 Conclusion 83
2.2 Cross platform analysis of volatile organic compounds using selected ion flow tube - , proton reaction transfer- and gas chromatography mass spectrometry............................84
2.2.1 Introduction 84
2.2.2 Methods 85
2.2.3 Results 90
2.2.4 Discussion...98
2.3 Analysis of room air volatile organic compound concentrations within different hospital environments 102

2.3.1 Introduction 102

2.3.2 Methods 104

2.3.3 Results 106

2.3.4 Discussion 110

3. CLINICAL STUDIES INVESTIGATING THE ASSOCIATION BETWEEN VOCs AND COLORECTAL CANCER 115

3.1 Analysis of colorectal tumour headspace for the assessment of cancer specific VOC release 116

3.1.1 Introduction 116

3.1.2 Methods 117

3.1.3 Results 123

3.1.4 Discussion 129

3.1.5 Conclusion 133

3.2 Colorectal neoplasm and breath analysis 134

3.2.1 Introduction 134

3.2.2 Material and methods: 137

3.2.3 Results 144

3.2.4 Discussion 164

3.2.4.1 Conclusions 169

FINAL CONCLUSIONS 170

FUTURE WORK 174

REFERENCES 177

APPENDICES 201

LIST OF FIGURES

Figure 1 Flow chart of the systematic review. Explanation of the study selection screening .30

Figure 2 Gas chromatography mass spectrometry (GC-MS) 38

Figure 3 Mass spectrum of acetone (National Institute of Standards and Technology, NIST) 39

Figure 4 Proton transfer reaction mass spectrometry (PTR-MS)..40

Figure 5 Selected ion flow tube mass spectrometry (SIFT-MS) 43

Figure 6 ReCIVA breath sampler 48

Figure 7 HiSorb tubes and example of use 49

Figure 8 Breath sampling devices: 61

Figure 9 Overall net promotor score (NPS) for the three devices 80

Figure 10 Comparison of select volatile organic compounds analysed using the selected ion flow tube mass spectrometry and proton-transfer reaction time of flight mass spectrometry. 94

Figure 11 Comparison of select volatile organic compounds analysed with online and offline (thermal desorption) proton transfer reaction time of flight mass spectrometry. 95

Figure 12 Offline sample variability in abundant compound 97

Figure 13 Collection of cancer headspace samples..119

Figure 14 VOC elevated in colorectal cancer tissue headspace...127

Figure 15 VOC product ions association with colorectal cancer T Stage..............................128

Figure 16 Flow diagram of sample discarded 148

Figure 17 Significant difference from selected compounds 155

Figure 18 Custom breath sampling device..175

LIST OF TABLES

Table 1 comparison on the alternative method for pre-operative detecting colorectal cancer 25
Table 2 Summary of studies....................34
Table 3 Subject characteristics....................68
Table 4 VOCs detected within breath samples (in the order of the RT)....................69
Table 5 Variation in repeated collections (common abundant compounds)....................77
Table 6 Variation in the repeated collection (aldehydes)....................78
Table 7 Results of the Usability Metric for User Experience....................79
Table 8 Characteristics of target VOCs within breath analysed by SIFT-MS and PTR-ToF-MS .88
Table 9 Subject characteristics....................90
Table 10 Comparisons of online VOCs detection by SIFT-MS and PTR-ToF-MS....................92
Table 11 Comparisons of VOCs detection by online and offline (TD) PTR-ToF-MS. 93
Table 12 Target VOCs within exhaled breath quantified by the PTR-MS....................105
Table 13 Variation in ambient VOCs levels....................107
Table 14 Variation in outpatient clinic ambient VOCs levels....................108
Table 15 Variation in operating theatre waiting room ambient VOCs levels....................108
Table 16 Variation in endoscopy waiting room ambient VOCs levels....................109
Table 17 Variation in VOC laboratory ambient VOCs levels....................109
Table 18 Comparison of room and exhaled breath VOCs (expressed as fold change)....................110
Table 19 potential biomarker from previous study (repeated mentioned in different research) 116
Table 20 Data of the standard quality control....................121
Table 21 Characteristic of sample collection....................123
Table 22 comparison of known compounds or uncertain ions from PTR-MS....................126
Table 23 Remove patient 10 and 17....................131
Table 24 General characteristics of the patients recruited in the study....................144
Table 25 List of selected compounds....................146
Table 26 Comparison on the failure rate of the two methods....................150
Table 27 Comparison of the samples using different ReCIVA setting....................150
Table 28 characteristics of patients with qualified samples collected....................152

Table 29 Peak area count of selected VOC from the five study groups................................153
Table 30 Characteristics of patients with CRC lesions 158
Table 31 peak area counts according to cancer stage and cancer recurrence 160
Table 32 Characteristics of patients who provided serial breath samples 162

Section I

BACKGROUND

1. BACKGROUND

1.1 Diagnosis and treatment for colorectal cancer

1.1.1 Epidemiology of Colorectal Cancer

Colorectal cancer (CRC) is a major global health burden. The 2018 World Health Organization (WHO) reported CRC as the third most common cancer (1.8 million/year) and the second most common cause of cancer death (0.86 million/year) [1]. In the United Kingdom, CRC is ranked as the fourth most common cancer with 42,000 new cases and 16,000 deaths per year [2].

The development of CRC is multifactorial. Advanced age, male gender, high red meat or alcohol consumption, smoking, reduced physical activity and obesity are all recognised risk factors for CRC [3]. Whilst genetic factors play a part in the development of CRC, less than 25% of patients are found to have specific mutations that predispose to the development of this disease. Lynch syndrome and familial adenomatous polyposis (FAP) are the two well-defined hereditary disease related to CRC but these two diseases make up to less than 5% of CRC [4, 5]. Inflammatory bowel disease (IBD) is also found to be a risk factor for CRC, but its pathogenesis remains unclear [6, 7].

Outcomes of the American surveillance, epidemiology, and end results (SEER) program showed the overall survival rate for CRC is 65% at 5 years from diagnosis. The principal determinant of survival in CRC patients is disease stage at diagnosis. Whilst 5-year survival in patients with localised disease is reported to be 90%, survival declines to 14% in those with metastatic disease [8].

1.1.2 Colorectal cancer detection

Symptoms of CRC include change of bowel habit, weight loss, feeling of incomplete bowel emptying, and per-rectal bleeding. A minority of patients however present with these “red flag” symptoms. Around 10% of patients diagnosed with CRC are either asymptomatic or have vague non-specific symptoms that are commonly ascribed to benign gastrointestinal conditions [9, 10]. Therefore, only 39% of patients are found to have localised diseases at the

time of diagnosis. A significant proportion of patients (approximately 1 in 4) present as an emergency to acute hospital services, often with advanced metastatic disease [8].

The gold standard investigation for the diagnosis of CRC is considered to be colonoscopy with tissue biopsy. Diagnosis is confirmed through histopathological assessment of tissue samples acquired at the time of colonoscopy. Assessment of tissue samples helps to confirm the differentiation and cell type of cancer that are both important factors relevant to the treatment planning [5]. Ninety-six percent of all colorectal cancers are of the adenocarcinoma subtype arising from mucosal polyps [11].

Locoregional and distant assessment are also important before proceeding to further treatment. Computer tomography (CT) scan and magnetic resonance imaging (MRI) are the two most common imaging modalities used in CRC staging. CT can provide information including locoregional extension of the tumour and distant metastasis. MRI is particularly useful in the preoperative assessment of rectal cancer for accurate clinical stage which is important on the decision making of pre-operative radiotherapy. Positron emission tomography/computer tomography (PET/CT) can provide better indication of metastatic disease in some patients, but its routine use has not yet been proven to be beneficial to the survival of patients [12-15]. Carcinoembryonic antigen (CEA) is the only biomarker used in the routine care of CRC but is reserved for postoperative surveillance of disease recurrence as opposed to primary disease detection [16, 17].

1.1.3 Treatment for colorectal cancer

Surgical resection of the cancer lesion is the standard curative treatment for CRC. Additional local or systemic therapy with either radiotherapy or chemotherapy may also be used in selective patients [18-20].

1.1.3.1 Surgery

Surgery is the most widely used treatment for patients with localized CRC and in selected patients with locoregional disease [20, 21]. The goal of surgery is to completely resect the malignant neoplasm with clear proximal, distal, and circumferential margins. Excision of the

extended lymphatic drainage system related to the tumour is required for an accurate pathological staging and better prognosis [22]. Surgery can be performed either by open or laparoscopic methods. The oncological prognosis is similar for these two surgical approaches although laparoscopic surgery has been shown to be beneficial in terms of postoperative pain and shorter length recovery [23, 24]. For early-stage disease including *in situ* tumours and some T1 lesions, endoscopic mucosal resection (EMR) or endoscopic submucosal dissection (ESD) may be an alternative choice to surgical resection[25, 26]. However, based on the results of pathological review, further complete bowel resection may still be advised for some patients with higher risk of lymph node metastasis [27, 28].

Although distant metastasis is a poor prognostic factor in CRC, resection of the metastatic or recurrent lesion is still found to be beneficial certain patients with stage IV disease [29, 30]. With complete resection and an appropriate follow-up programme, the survival rate for those initially diagnosed distant disease may be similar to patients with localised. However, the proportion of patients who can receive salvage operations for recurrent cancer is low because of the delayed diagnosis of tumour recurrence [31-33].

In around one third of CRC patients, emergency surgery may be necessary for cancer related bowel obstruction or perforation. Bowel obstruction can be caused by the enlarged tumour itself or a tumour-related stricture of the bowel [34, 35]. Asymptomatic tumours can still invade deeply through serosae layer resulting in bowel perforation. In some patients, CRC may present as an emergency with bowel perforation with accompanying peritonitis or pelvic abscess [36]. In cases where there is found to be complete or partial obstruction of the colon other possible interventions include defunctioning colostomy, ileostomy, or colonic stent [37, 38].

1.1.3.2 Chemotherapy

Fluorouracil (5-Fu) based chemotherapy has been the most common agent given to CRC patients as adjuvant chemotherapy. In the past 30 years, new chemotherapy regimens including oxaliplatin, irinotecan, capecitabine, and biological therapies such as bevacizumab, and cetuximab have been approved for use in the treatment of CRC. The primary purpose of adjuvant chemotherapy in CRC patients is to decrease disease recurrence rates. In some

selected cases, neoadjuvant chemotherapy can be given to decrease the tumour burden prior to surgery [39]. While *en bloc* resection of the tumour offers the best oncological results of treatment, chemotherapy may decrease the volume of tumour and unresectable cancer may become resectable after neoadjuvant therapy. A similar principle can be applied to initial stage IV or recurrent disease.

Current guidelines for chemotherapy are based on the National Comprehensive Cancer Network (NCCN) and European Society for Medical Oncology (ESMO)[21, 40, 41]. It is still rare to see complete remission of cancer with chemotherapy alone, but lower recurrence rate or higher rate of complete resection were proven advantages of chemotherapy [39, 42-44].

Advances in genetics has meant that personalised treatment based on prediction of response to different chemotherapeutic regimens may become an important feature of patient care in the future. One example is the use of cetuximab, which is an epidermal growth factor inhibitor. This target therapy agent specifically works in patients with the wild type K-Ras gene [45-47]. Another important one is the assessment of the toxicity of fluoropyrimidines (5-FU). 5-FU toxicity results mainly from deficient uracil catabolism. This ^{13}C-uracil breath test could provide a valuable addition to the patients' standard of care [48].

1.1.3.3 Radiotherapy

Radiotherapy has mainly been used for neoadjuvant therapy in advanced rectal cancer for the purpose of local disease control. There are two principal methods of neoadjuvant radiotherapy, long course or short course, although no significant difference between the two methods has been reported in regards to oncological outcomes [49, 50]. Since radiotherapy is only used for local disease control, there is no influence on the distant metastasis or rates of disease recurrence. However, radiotherapy still decreases the local recurrence rate. Rarely, complete remission of rectal cancer can be seen after chemoradiotherapy. As a consequence, a *'watch and wait'* policy may be adopted for those patients found to have had a complete response to radiotherapy on initial assessment [51-54].

1.1.3.4 Palliative chemotherapy and radiotherapy

In patients with advanced disease that is deemed unsuitable for potentially curative therapy, chemotherapy and/or radiotherapy may be used in a palliative setting for disease control and to prolong survival. Radiotherapy has been proven to be helpful in reducing the tumour volume in unresectable tumours [55, 56]. Palliative chemotherapy has shown to be beneficial to prolong overall survival on initial unresectable CRC [57]. Rarely, resection may be feasible for patients who have good response to palliative chemoradiotherapy [58].

1.1.4 Post-operative follow-up

Post-operative follow-up after CRC resection is important to determine tumour recurrence. If identified early, recurrent disease may be amenable to salvage surgery which has proved to be beneficial in certain patients [20, 32]. The schedule of patient follow-up and modality of investigation varies considerably between different health services. Several investigations, including colonoscopy, CEA, CT, positron emission tomography with 2-deoxy-2-[fluorine-18] fluoro- D-glucose integrated with computed tomography (^{18}F-FDG PET/CT) may be performed regularly or if clinically indicated. Complete surveillance of the residual colon is required for every CRC patient because the possibility of metachronous lesions, especially for those patients whose pre-operative assessment is incomplete, particularly in cases where presentation was as an emergency.

Although intensive follow-up has been shown to improve the overall survival of CRC patients, there is no strong evidence to establish a single optimised program of surveillance [59-61]. Based on the pathological results and the patients' condition, the follow-up program may be adjusted to meet individual requirement.

1.2 Early detection of colorectal cancer

Better prognosis comes from earlier detection and treatment of CRC. However, in most cases early CRC is asymptomatic or is associated with on some minor symptoms such as bloody stool may be wrongly ascribed to common benign conditions such as haemorrhoids. There remains therefore a clinical need to develop methods to aid the earlier detection of CRC. Ideally such tests should be acceptable to patients, accurate and affordable. The United Kingdom's National Health Service (NHS) Bowel Cancer Screening Program (BCSP) was introduced in 2006 to improve CRC mortality by earlier detection of CRC. It is now offered to patients aged 60–74 years and involves a home-based stool test every other year. Follow-up colonoscopy would be arranged if there is a positive finding in stool test [62].

1.2.1 Faecal occult blood test (FOBT)

The FOBT has been designed as a screening test for CRC, through the detection of occult (hidden) blood in stool. The test is non-invasive and can be administered by patients in their own homes with samples being sent to a central laboratory for interpretation.

There are several methods looking into different properties in blood including antibodies, haem, globin, or porphyrins in clinical use for testing for occult blood in stool. The detection of DNA from the intestinal mucosal cell is another target for stool examination but is not currently used in routine practice.

Faecal immunochemical testing (FIT) which may also be named as immunochemical faecal occult blood test (iFOBT) focus on the detecting globin. Based on its cut-off point of globin detected, the sensitivity and specificity may be varied. From Meklin's review, diagnostic performance of FIT on CRC endpoint were assessed from twenty-four research projects. The pooled overall sensitivity and specificity of iFOBT tests were 0.86 (95% CI=0.78-0.93) and 0.85 (95% CI=0.81-0.88), respectively[63]. Another method for detecting stool occult blood is stool guaiac test (gFOBT). This method works as the heme component in haemoglobin has a peroxidase-like effect, rapidly breaking down hydrogen peroxide. In Meklin's review, gFOBT test results from pooled data revealed overall sensitivity and specificity of gFOBT tests for detecting CRC were 0.68 (95% CI=0.57-0.79) and 0.88 (95% CI=0.84-0.91), respectively.

Several studies have focussed on the comparison between FIT and gFOBT as mentioned above in the review article. From the results, FIT may be better than gFOBT on the detection of CRC, but larger scale studies on asymptomatic patients cohort are necessary to identify the preferable test in the future [64].

The American Cancer Society suggested that patients without additional risk of CRC should commence FOBT from 45 years old with a test interval of one or two years [65]. Further colonic survey will be advised if positive results are detected from the exam. In most European countries, the FOBT has been widely adopted with highest participation rates recorded in the Netherlands (68.2%)[66, 67].

Although FOBT has been widely accepted worldwide, its low specificity and high false positive rate remains an issue. Based on systemic reviews of FOBT, of those patients who have a positive test, 2-10% may ultimately be diagnosed with CRC whilst 20-30% are found to have colonic adenoma(s) [68, 69]. Although the participation rate of FOBT is high, the number of subjects having colonoscopy with positive results were lower than expected. In a study by Lurie and Welch, only 34% of patients who had recommended evaluation including colonoscopy or flexi-sigmoidoscopy with double contrast lower gastrointestinal enema [70]. Due to the low specificity of the FOBT, a considerable proportion of subsequent investigations including colonoscopy of CT virtual colonoscopy are ultimately unnecessary.

1.2.2 Direct assessment with colonoscopy

Colonoscopy has the highest sensitivity to detect CRC. Ideally, with adequate bowel preparation and performed by an experienced examiner, the sensitivity of colonoscopy for CRC detection is close to 100%. Therefore, direct colonoscopy has also been used as a screening test since 2002 [66, 71]. Unlike home based FOBTs, a trained endoscopist is necessary for colonoscopy. Further limitation of colonoscopy includes patient acceptability, risk of complications and requirement that it be performed in a medical setting. Therefore, the costs and risks associated with direct colonoscopy are greater than for other tests.

A major benefit of direct colonoscopy is the ability to not only directly visualise any bowel lesions, but also to acquire tissue biopsies for histological assessment.

1.2.3 Developing methods for colorectal cancer screening

There are a number of new methods being developed for early detection of CRC. A major challenge of such tests, which typically seek to determine unique biomarkers of CRC, has been to establish the reproducibility of findings.

1.2.3.1 Stool DNA

The detection of stool DNA has been proposed as an alternative method of screening for CRC [72, 73]. Like FOBT, it is a non-invasive test that can be performed by patients in their own home. Stool DNA testing looks to identify DNA alterations that are associated with cancer. At the present time however no single DNA markers have been reported as being diagnostic of CRC. Both Imperiale et al., and Ahlquist's et al., found that a multitarget panel may be most appropriate for cancer detection. Some mutant DNA found within the stools of patients with CRC and advanced adenomas, include K-Ras, p53, and APC genes; Bat-26, a microsatellite instability marker; and highly amplifiable DNA. However, increased sensitivity of stool DNA panels for CRC, was associated with a decrease in specificity. Further research investigating the possible combination of stool DNA and biomarker, as a potential hybrid test, is still required before adoption into routine clinical practice.

1.2.3.2 microRNA

MicroRNA from blood samples are a short sequence of RNA which may be used as a diagnostic biomarker for CRC [74-76]. Furthermore, detected microRNA can be used to guide target therapy for the development of new anti-oncological therapies [77]. However, the same challenges to adoption that are recognised for stool DNA also affect the wider application microRNA as a screening test for CRC. Agreement in the results between studies from different research groups is low and the repeatability of the test remains uncertain.

1.2.3.3 Circulating tumour DNA

Circulating tumour DNA (ctDNA) has been discussed for over 20 years [78, 79]. Some researchers showed its potential as a biomarker for both diagnosis of new CRC and predicting

cancer prognosis [79, 80]. Again, this approach to CRC detection shares many of the limitations faced by other emerging tests that have already been discussed.

1.2.4 Comparison of the alternative method for early detection of colorectal cancer

Although there are developing methods to detect CRC at earlier stage, still some defect with the above-mentioned method.

Methods	Sensitivity	Cost
Biomarker, carcinoembryonic antigen (CEA) [81, 82]	65%-74%	Low
Biomarker, carbohydrate antigen 19-9 (CA 19-9) [81, 82]	26%-48%	Low
Image, computer tomography colonoscopy (virtual colonoscopy) [83, 84]	89%	Medium
Image, double barium lower gastrointestinal image [70, 85, 86]	81%-83%	Medium
Image, magnetic resonance images (MRI)	Not appliable*	High
Image, positron emission tomography (PET) [87, 88]	95%	High
Blood, Circulating tumour cell [77]	Not appliable*	Low
Stool DNA test [65, 89]	92%	Low

Table 1 comparison on the alternative method for pre-operative detecting colorectal cancer.

* Not currently using for the detection of colorectal cancer. MRI is for tumour staging and circulating tumour is mainly used for post-operative follow up.

Currently, CT colonoscopy, PET/CT scan and multitarget stool DNA may provide good sensitivity for CRC. The cost of CT colonoscopy and PET/CT is high, and multitarget stool DNA test may be a time-consuming test. The acquire of stool sample is still the obstruction on the clinical practice. Considering the cost-effective and their practical for clinical use, a more convenient and cheap survey is necessary.

1.3 Volatile organic compounds and cancer detection

Volatile organic compounds (VOCs) are organic compounds with a low boiling point such that they are in the gas phase at ambient temperatures. VOCs may be produced from wide array of naturally occurring or industrial processes.

VOCs were first analysed in the context of air pollution [90, 91]. With advances in analytical platforms and sampling methods, more precise and efficient analysis of VOCs has been achieved, allowing for their detection in biological samples such as breath where they are found in low concentrations, typically parts-per-billion (ppb)[92, 93]. Analysis of VOCs has subsequently become a topic of interest in medical research, due to their association with human health and disease states.

Alteration of metabolism with associated abnormal cellular growth is a characteristic feature of disease states and cancer development [94-96]. Both normal physiology and disease states can lead to the production of VOCs as by-products of metabolic pathways. These volatile metabolites may circulate throughout the body within the blood stream and can be excreted in breath, urine, stool or from the skin. Previous studies have reported alteration in VOCs levels in a number of human diseases, including cancer [97-101].

1.3.1 VOCs within exhaled breath

The detection of VOCs within exhaled breath can be traced back to ancient Greece, with Hippocrates noting the distinctive breath odours of certain disease states including liver cirrhosis [99, 102, 103]. It was not however until the 1970's that scientists were able to accurately analyse breath using gas chromatography mass spectrometry (GC-MS). Linus Pauling in his seminal paper reported that human breath contained up to 250 different VOCs [104]. Subsequent studies have now identified more than 800 compounds within breath specimens [105].

1.3.2 Current application of VOCs in disease diagnosis

Although VOCs biomarker research remains in its infancy, some potential clinical applications have been supported by published studies. Considering the source and collection of VOCs,

breath sampling still has advantages over other sampling because of the lack of necessity for pre-sample preparation before analysis. Although there are many different analytical platforms for VOCs analysis, the aim of the analysis can be to identify and quantify compounds and see if there are specific trends on these VOCs. These specific VOCs have been found to be characteristic of some unique diseases.

1.3.2.1 Cancer

Previous studies have linked different cancers to a unique VOC signature. Initial studies focused on defining the VOC signature of lung and breast cancer [106, 107]. The sensitivity and specificity of detecting lung cancer with these VOC printing was between 0.53-1.00 and 0.72-1.00. Subsequently VOC research has extended to include other cancer sites including the gastrointestinal and urinary tract. Studies in oesophagogastric cancer, colorectal cancer, prostate cancer have revealed potential VOC biomarkers of these diseases. Whilst those studies suffered from a number of important limitations, including low patient numbers, different sampling method and lack of external validation, they still gave some encouraging results that analysing VOCs can be a possible way [98, 107-112]. The number of samples recruited in the previous studies varied from less than 30 to more than 100. Although these samples cannot give the answer directly, they are the foundation of future analysis. Also, from these studies, we can understand the importance of methodology.

1.3.2.2 Genetic and metabolic disorders

Some genetic disorders are related to deficiencies of enzymes which may result in over-accumulation or depletion of specific compounds. Diabetes mellitus or hepatic failure are two simple examples where there is accumulation of acetone and ammonia [113, 114]. Uraemia is a type of kidney failure defined by the presence of excessive nitrogenous waste products, such as urea, in the blood stream. Patients with uraemia may give breath urine like odour [115]. Scurvy is a disease caused by vitamin C deficiency, which is required for the synthesis of collagen. Scurvy patients produce sweat with a putrid odour[116]. Some other inherited lysosomal disorder related diseases such as maple syrup urine disease, or isovaleric acidaemia were also found to be potentially detectable with breath VOCs [117, 118].

1.3.2.3 Infectious disease

One possible origin of endogenous VOCs is as biproducts of by microbial species. Higher levels of methyl phenylacetate, methyl p-anisate, methyl nicotinate and o-phenylanisole were found in the breath of patients with active *mycobacterium tuberculosis* infection [119]. This may be a faster detection method than usual culture-based techniques. Besides, some other bacteria such as *Clostridium difficile* or *Campylobacter jejuni* can be found from the faecal headspace sample [120, 121].

1.3.2.4 Others

Other diseases including asthma, chronic obstructive pulmonary disease and, cardiovascular disease, or schizophrenia were also found to be detectable by their characteristic breath VOC signature [122-126].

Whilst many studies have identified an association between specific VOCs and disease states, further large-scale *in vitro* to *in vivo* studies are necessary to validate these findings.

1.3.3 Which came first?

While talking about the potential of VOCs as biomarker, there is a question. Could it be that people with specific disease modify their diet or lifestyle and, consequently, change their breath VOCs, and it is that is being detected, not biomarkers of cancer?

It is possible to change the VOCs with the modified diet or lifestyle, especially for some of the infectious disease which result in rapid changes of metabolism including higher consumption of energy and the breakdown of stored glycogen. However, in some of the cancer such as breast cancer, lung cancer or CRC, the most common initial symptom remains none symptom. For asymptomatic patients, it is less likely for them to modify their diet or lifestyle. Although it is not possible to exclude the possible correlation of disease and modification of diet, the impact might be minor at early stage. To give a solution to this question, a more detailed data record may be helpful.

1.4 Systematic review of studies reporting volatile organic compounds as biomarkers of colorectal cancer

1.4.1 Introduction and aim of the systematic review

Colorectal cancer (CRC) is one of the leading malignancies worldwide[1]. Previous studies have reported an association between different VOCs and CRC, offering the opportunity of a non-invasive method for detecting this disease.

The following systematic literature review is intended to identify and report the findings of all previous studies investigating the VOC signature of CRC.

1.4.2 Methodology

The systematic review was in accordance with the recommendations of the Cochrane library and MOOSE guidelines[127]. An electronic search (title and abstract) of the Embase and MEDLINE databases (1946 to 28th Aug 2023) was performed through the OVID platform. The following terms were used in the search strategy: breath, volatile organic compounds, VOC, colorectal cancer, and bowel cancer. All variations in spelling including truncated search terms using wild card characters and 'related articles' function was used in combination with the Boolean operators AND and OR. Reference lists of qualified articles were screened to include potentially relevant studies.

Two independent reviewers (GPL and PB) screened the titles and abstracts of all studies identified through database searching. Full texts of potentially relevant articles were reviewed for eligibility. Only original research articles published in English language were considered. Included studies identified named VOC biomarkers of colorectal cancer. Studies were excluded if they did not report named VOC biomarkers, and if they reported mixed cancers where results of each subtype could not be clearly separated. Review articles, conference abstracts, articles not written in the English language, animal and cell studies were excluded. A third reviewer (GH) was consulted where any disagreement in study inclusion arose.

1.4.3 Results

After excluding duplicates, 2126 published articles were identified from the electronic search. After screening and assessment for eligibility, 11 studies, which reported VOC profiles in 675 patients, were selected for inclusion in this review[98, 111, 128-140]. Two studies published by a Altomare et al., were considered as single study as they contained the same cohort of patients assessed both before[111] and after surgery [134]. Studies that used sensors or technologies such as Field Asymmetric Ion Mobility Spectrometry (FAIMS) were excluded as they did not present named VOC biomarkers of colorectal cancer.

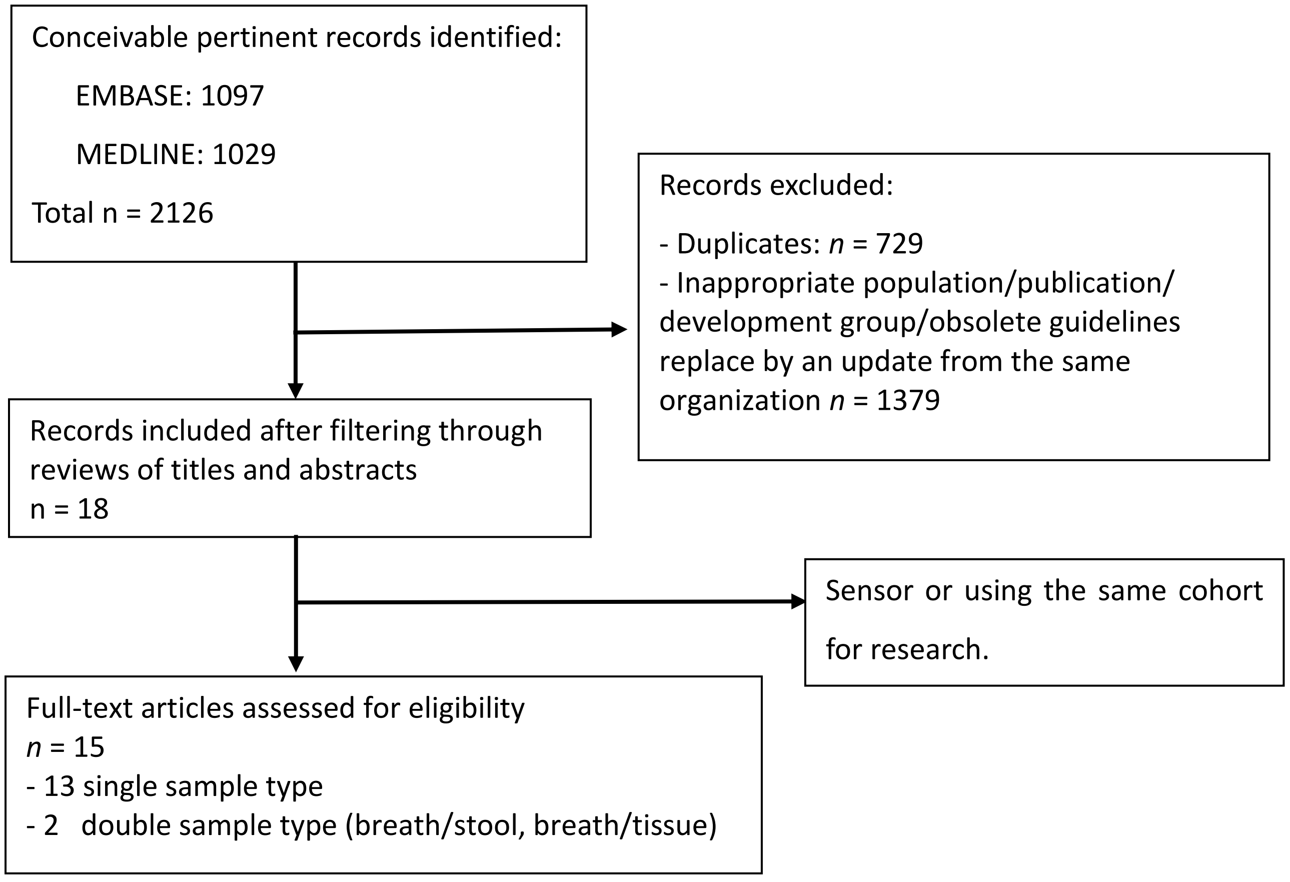


Figure 1 Flow chart of the systematic review. Explanation of the study selection screening

Details of included studies are presented in Table 1. Included studies analysed VOCs within a range of sample types: breath (n=6); tissue (n=2); faeces (n=2); blood (n=2), and urine (n=2). A single study analysed both the breath and tissue samples, and another study analysed both the breath and faeces. All studies used gas chromatography mass spectrometry (GC-MS) for

VOC analysis. Cross platform validation of findings was described by two studies that used selected ion flow tube mass spectrometry (SIFT-MS) in addition to GC-MS[130, 132]. Validation of findings within an independent patient cohort was performed by two studies[111, 132].

1.4.4 Discussion

Studies identified in this review showed the feasibility of VOCs as biomarkers for CRC. Although their sensitivity and specificity were not shown clearly in all papers, the available sensitivities were between 72%-100% and specificities were between 73%-98%. Whilst there was wide variation in the chemicals classes of potential CRC biomarkers identified by each study, common classes included: alcohols, aldehydes, alkanes (including cycloalkane or haloalkane, alkene), aromatic compounds, and fatty acids. Compared to other GI malignancy, some similarity including fatty acid, aldehyde could be found[109].

Despite the encouraging findings of this review, there remain some important challenges to the wider adoption of VOC based testing for CRC in clinical practice.

The first challenge is the variation in the results. Among all the compounds listed in the previous studies, only three compounds were found to be possible CRC biomarkers in different studies. Acetone and cyclohexanone showed different trend from cancer and control group whilst ethanol was noted to be higher in both studies (128,129,130,133,135). Since there was no defined method for breath sampling and analysis, it is possible to see the varied results. Besides, acetone and ethanol are the two common and abundant compounds to be found from human breath. Further investigation is still necessary for these results.

As mentioned before, the research on VOCs is still in its relative infancy and one of the major challenges is the standardization of the analytical process. From the review, although most of the studies utilised GC-MS for VOC detection, different methods including SPME or thermal desorption tube were used within these studies to trap VOC prior to analysis. Furthermore, different configuration of GC-MS instruments, including GC-column selection and flow rate may impact on sample separation and consequently VOC detection. Heterogeneity in these and other sampling methodologies within included studies means that it is not possible to

draw robust conclusions as to specific VOC associated with CRC. For the most part of the studies did not validate their findings within an independent patient cohort making it difficult to ensure their reproducibility.

While the common symptoms of CRC at a later stage may be tumour related bleeding or gastrointestinal symptoms including bowel obstruction or irregular bowel habit, it is less likely to confirm if blood in the bowel lumen or in the stool may alter the guts VOCs from the systematic review. If we do the cross-validation of the VOCs from faeces and blood in the systematic review, none of the VOCs could be found from both samples. It could be because of the breakdown of the blood in the bowel lumen and the VOCs were changed or the blood in the lumen was too little to change the VOCs. Further investigation including the research on the gut microbiota, cross-linking with stool occult blood test may give a more promising result.

Here in the systematic review, it is difficult to know any other symptom/sign from each patient or healthy volunteer. Some benign colorectal disease including constipation, diverticulum, or haemorrhoid were not mentioned in the systematic review. Normally, we define constipation as stool to stay in the bowel for a longer time. The cause of constipation varied. If not considering the mechanical obstruction related constipation such as cancer, dry and hard stool may be found from patients with constipation. While most of them may take some stool softener, we cannot totally neglect the impact of medication on the VOCs. A more detailed record including long term medication use may not to be taken into consideration as one factor in the final analysis. Diverticulum is mostly asymptomatic and found accidentally from colonoscopy or LGI series. Sometimes, we found patients with diverticulitis, but its related symptoms may help us to exclude the collection of samples. Finally, haemorrhoid is not a disease. Haemorrhoid is a normal tissue everyone has but not everyone has symptoms related to haemorrhoid. Although severe bleeding from haemorrhoid may lead to anaemia, it is impossible to tell if haemorrhoid is link to some specific VOCs.

VOCs may therefore be a promising biomarker for the detection of CRC; however, the standardization of the whole analytical process remains the major challenge. From this review, we can suggest that the concept that VOCs may be released in a unique pattern from CRCs,

but further research is required to understand the specific classes of VOCs that are dysregulated and to ensure the adequacy of the methods used in their analysis.

Table 2 Summary of studies

Author	Year	Country of origin	Sample type	Number of CRC patients	Number of controls	Disease stage	Analytical platform	VOC/VC biomarkers and their changes (tumour vs normal)	Independent validation	Sensitivity	Specificity	AUC-ROC	Ref.
Mezmale	2023	Latvia	Tissue headspace	50[1]	50[1]	I-IV	GC-MS	2-Butanone ↓ 1-Propanol ↑ Pyridine ↑ 2-Pentanone ↓ 2-Methyl-2-propanol ↓ Ethyl acetate, ↓ Isoprene ↑ 3-Methyl-1-butanol ↓ D-limonene, ↓ Methyl thiolacetate ↑ Tetradecane ↓ Dodecanal ↓ Tridecane ↓ 2-Ethyl-1-hexanol ↓ Cyclohexanone ↓	No	-	-	-	[128]
´Smiełowska	2023	Poland	Breath/faeces	15	20	-	TD-GC-MS	Breath: Heptanoic acid ↑ 2,6,10-Trimethyldodecane ↑ Stool: n-Hexane ↑ Dimethyl trisulfide, Skatole ↑ Both: Acetone ↑	No	100 (stool)[2]	100 (stool)[2]	1.00	[129]
Boulind	2022	UK	Urine	18	272	-	SIFT-MS, GC-MS	Carbon disulfide ↑ Acetone ↑ Ethanol ↑ unknown (NIST library best match for 2,2,6,6-Tetramethyl-4-ethyl-heptane) ↑ Dimethyldisulfide ↑ M-Xylene ↑ 4-Heptanone ↑ Benzenethiol ↑ Pyrrole ↑ 1,6-Dichloro-1,5-cyclooctadiene ↑ Biphenyl ↑ Phenol ↑ Dibenzofuran↑	No	83	82	0.913	[130]
De Vietro[3]	2020	Italy	Breath/tissue	7	20	III	TD-GC-MS	Benzaldehyde ↑ Ethylbenzene ↑ Indole ↑	No	-	-	-	[131]
Markar	2019	UK	Breath	50/25[4]	100/54[4]	I-IV	SIFT-MS, GC-MS	Propanal ↑	Yes	83	85	0.790	[132]
Amal	2016	Israel	Breath	65	122	I-IV	SPME-GC-MS	Acetone ↓ Ethanol ↑ Ethyl acetate ↑	No	-	-	-	[133]

								4- Methyl octane ↓					
Altomare[5]	2013	Italy	Breath	37/15[4]	41/10[4]	I-IV	TD-GC-MS	Nonanal	Yes	86	83	0.852	[111]
Altomare[5]	2015	Italy	Breath	48	32	I-IV	TD-GC-MS	4-Methyl-2-pentanone Decanal 2-Methylbutane 1,2-Pentadiene ↑ 2-Methylpentane 3-Methylpentane Methylcyclopentane Cyclohexane ↑ Methylcyclohexane ↑ 1,3-Dimethylbenzene ↑ 4-Methyloctane ↑ 1,4-Dimethylbenzene 4-Methylundecane Trimethyldecane	No	100	98	1.000	[134]
Wang	2014	China	Breath	20	20	I-III	SPME-GC-MS	Cyclohexanone ↑ 2,2-Dimethyldecane ↑ Dodecane ↑ 4-Ethyl-1-octyn-3-ol ↑ Ethylaniline ↑ Cyclooctylmethanol ↑ Trans-2-dodecen-1-ol ↑ 3-Hydroxy-2,4,4-trimethylpentyl 2-methylpropanoate ↑ 6-t-Butyl-2,2,9,9-tetramethyl-3,5-decadien-7- yne ↓	No	-	-	-	[135]
Peng	2010	Israel	Breath	20	22	I-IV	SPME-GC-MS	1,1'-(1-Butenylidene)bisbenzene ↑ 1,3-Dimethyl benzene ↓ 1-Iodononane ↑ [(1,1-Dimethylethyl)thio]acetic acid ↓ 4-(4-propylcyclohexyl)-4'-cyano[1,1'-biphenyl]-4-yl ester benzoic acid ↓ 2-Amino-5-isopropyl-8-methyl-1-azulenecarbonitrile ↓	No	-	-	-	[98]
Bond	2019	UK	Faeces	21	60	-	SPME-GC-MS	Propan-2-ol ↑ Hexan-2-one ↑ 3-Methylbutanoic acid ↑ Propan-2-yl butanoate ↑ Propan-2-yl pentanoate ↑ 1,4-Xylene ↑ Propan-2-yl propanoate ↑ 5-Methyl-2-propan-2-ylcyclohexan-1-ol ↓	No	88	85	0.860	[136]
Batty	2015	UK	Faeces	31[6]	31	-	SPME-GC-MS	Hydrogen sulphide ↑ Dimethyl sulphide ↑ Dimethyl disulphide ↑	No	72	78	-	[137]
Kim	2019	USA	Blood	30	30	I-IV	SPME-GC-MS	2,3,4- Trimethyl-hexane ↓ 3-Ethylhexane ↑	No	-	-	-	[138]

								2,4-Dimethylhept-1-ene ↑					
								3,5-Dimethyl-octane ↓					
								2,4,6-Trimethyl-1-nonene ↓					
Wang	2014	China	Blood	16	20	I-III	SPME-GC-MS	Phenyl methylcarbamate ↓	No	-	-	-	[139]
								Ethylhexanol ↓					
								6-t-Butyl-2,2,9,9-tetramethyl-3,5-decadien-7-yne ↑					
								1,1,4,4-Tetramethyl-2,5-dimethylenecyclohexane ↓					
Rozhentsov	2014	Russia	Urine	8	35	-	SPME-GC-MS	2,4-Dimethylfuran ↑	No	-	-	-	[140]
								1,6,7-Trimethylnaphthalene ↑					
								α,α,4-Trimethyl-(S)-3-cyclohexene-1-methanol ↑					
								β-Pinene ↑					
								2-Methyl-3-buten-2-ol ↑					
								5-(1-Methylethylidene)-1,3-cyclopentadiene↑					
								3,7-Dimethyl-(E)-1,3,6-Octatriene ↑					
								2-Methylhexanoic acid ↑					
								3-(2-Cyclohexylethyl)-6-cyclopentylhexylbenzene ↑					
								Pentanoic acid ↑					
								2,2,4-Trimethyl-3-carboxyisopropyl isobutyl ester ↑					
								1-Methyl-1-(2-hydroxyethyl)-1-silacyclobutane ↑					
								Vanillin, tert-Butyldimethylsilyl ether ↑					
								N,N'-ethylenebis(N)-nitroacetamide ↑					
								(Dibromomethyl)benzene ↑					

CRC, colorectal cancer patients. TD, thermal desorption. GC-MS, gas chromatography mass spectrometry. SIFT-MS, selected ion flow tube mass spectrometry. SPME, solid phase microextraction. [1]Cancer tissue and normal tissue were obtained simultaneously. [2]sensitivity and specificity were recorded based on the predictive model. [3]VOC biomarkers are those that were present in the breath and tissue headspace of colorectal cancer patients [4]Values are for model building / validation patient cohorts. [5]Only increased concentration of compounds mentioned in the article. Others were mentioned as variable. [6]Includes patients with high grade adenomas. The name of the marker compounds listed above were based on the typically used in International Union of Pure and Applied Chemistry (IUPAC) nomenclature for organic compounds.

1.5 Methods for volatile organic compounds analysis

The analysis of VOCs for the purpose of disease detection and monitoring remains an emerging field. At the present time there are no universally agreed guidelines for the collection and analysis of VOCs within clinical samples. In the majority of cases researchers have utilised mass spectrometry-based methods for the detection of VOCs. Mass spectrometry offers a versatile modality for the separation and detection of VOC. Different mass spectrometry techniques that have been used for the detection of VOC, have themselves unique features that are associated with specific advantages and disadvantages. Accordingly, there is no single instrument or technique that can be considered the optimum approach for VOC detection.

In the following section a number of the most commonly used mass spectrometry techniques used in the detection of VOCs is summarised.

1.5.2 Mass Spectrometry

Different mass spectrometry techniques offer distinctive capabilities in terms of VOC analysis. Analytical platforms are often distinguished base on their ability to identify and/or quantify VOCs either online or offline. The most widely used methods for VOC analysis are summarised below.

1.5.2.1 Gas chromatography mass spectrometry (GC-MS)

GC-MS is a combination of two analytical techniques, gas chromatography and mass spectrometry. The two parts work together to achieve separation and identification of compounds. The first GC-MS was described in 1959 [141] and over the past 60 years, technological advancement has simplified the use of GC-MS and made the technique

more accurate and versatile. Now, GC-MS has been used in drug detection, air pollution analysis, or clinical breath analysis [142, 143].

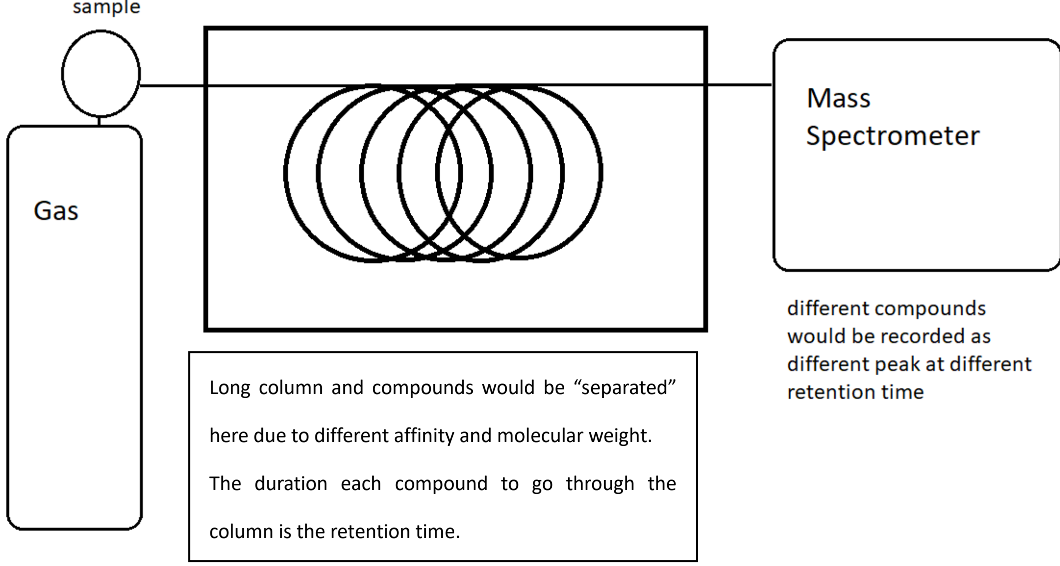


Figure 2 Gas chromatography mass spectrometry (GC-MS)

A schematic of a GC-MS instrument is shown in figure 2. Gas samples, such as breath, are introduced into the instrument via an inlet. The sample is transported by a carrier gas (typically helium) through the GC column where single components of the sample mixture are separated based on their molecular weight and chemical-physical properties, from which depends on the affinity for the column stationary phase. This affinity will determine the time taken for each component of the sample to pass through the GC column and is described as their retention time (RT). Separated compounds are subsequently transferred to a mass spectrometer where they are ionised into characteristic charged fragments that can be detected based on their mass-to-charge ratio (m/z). Results of the analysis are presented as a chromatogram which would be further compared to the mass spectrum (Figure 3) in a library.

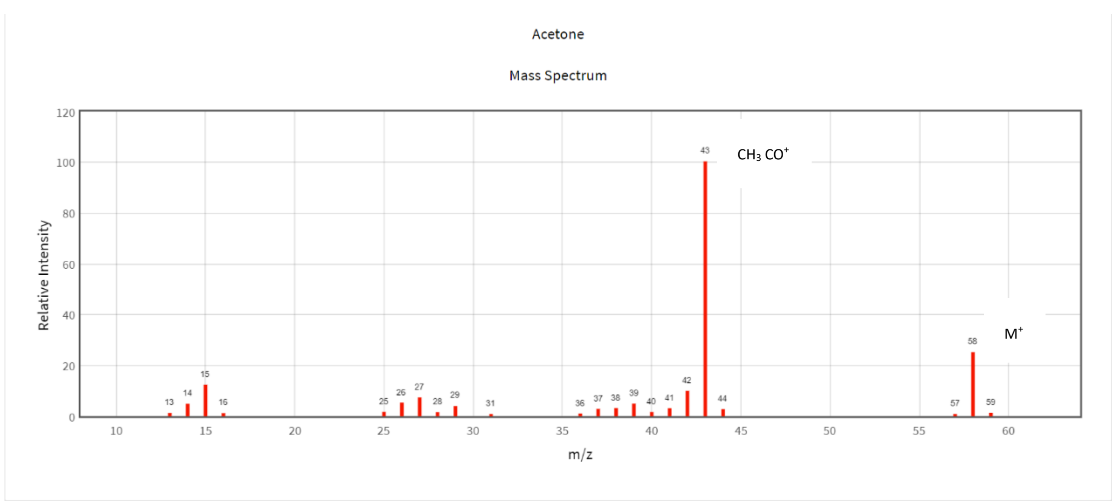


Figure 3 Mass spectrum of acetone (National Institute of Standards and Technology, NIST)

The mode in which the mass spectrometer within a GC-MS instruments analyses compounds eluted from the GC column can be changed depending on the aim of the analysis. The analysis of compounds may be conducted by either a full scan or selective ion monitoring (SIM) mode. Full Scan mode will monitor a range of masses known as mass-to-charge ratio (abbreviated m/z) and detect compound fragments within that range over a set time period. It is useful when identifying unknown compounds in a sample while the range of m/z is uncertain. SIM mode allows for detection of specific analytes with increased sensitivity relative to full scan mode. In SIM mode, only mass of interested would be calculated in a limited range. Typically, two to four ions are monitored per compound and the ratios of those ions will be unique to the analyte of interest. For targeted analysis, the sensitivity is higher in the SIM mode. In practical terms these two modes may be used to identify compounds with full scan mode and quantifying them with SIM mode aided by calibration curves.

Identification of compounds by GC-MS is achieved by comparing obtained spectra with spectra in on-line libraries. Whilst GC-MS provides an excellent modality for separating and identifying compounds within a sample, it is unable to provide direct quantification of those compounds. Further analysis with calibration curves using standard analytes of known concentrations is therefore required.

1.5.2.2 Proton transfer reaction mass spectrometry (PTR-MS)

PTR-MS was initially developed in 1995 for direct injection (on-line) monitoring of gas samples. PTR-MS instrument consists of an ion source that is directly connected to a drift tube and downstream mass spectrometer. Current application includes food science and air quality analysis. The analytical process is as summarised in the following schematic below (Figure 4).

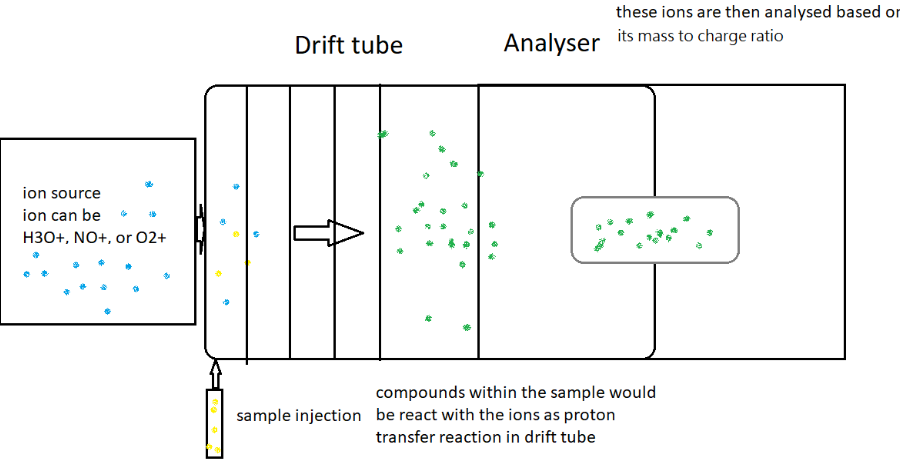


Figure 4 Proton transfer reaction mass spectrometry (PTR-MS)

Compounds react with precursor ions to produce characteristic product ions that are detected by the analyser.

An understanding of the ion molecular reaction between compounds and their precursor concentrations can be calculated from the rate of detection of their characteristic product ion(s).

Important advantages of PTR-MS are the lack of requirement for sample preparation as well as the low fragmentation and low energy associated with its ion molecular reactions. Because of the low fragmentation, it is easier to interpret the obtained results. Besides, the low energy requirement makes it possible to be built into a small space making the instrument more versatile and portable. However, there is one major disadvantage remaining. H_3O^+ was the only reagent ion in the initial design of PTR-MS, but the separation of isomers was poor. With the addition of two other reagent ions, NO^+ and O_2^+, the resolution improved but the affinity between compounds and ions is still the main issue affecting the results.

When applied to clinical practice, PTR-MS can offer the opportunity for point-of-care testing with continuous sampling and analysis. The cost of the PTR-MS instrument however limits wider use in clinical research and clinical practice. Although the use of different ions by PTR-MS makes the analysis of different compounds possible, one disadvantage is the difficulty of compound identification. This technique does not have a separation mechanism such as GC, so ions with same molecular weight and comparable properties, such as isomers, are indistinguishable. In addition, a specific library has not yet been developed.

1.5.2.3 Selected ion flow tube mass spectrometry (SIFT-MS)

Selected-ion flow-tube mass spectrometry (SIFT-MS) is a mass spectrometry technique for trace compounds that has similarities to PTR-MS.

The development of SIFT-MS was an early advancement of the VOC field and influence the design of other techniques such as PTR-MS. The technique was first developed in the 1970's to study molecule genesis within cold interstellar clouds of the Milky Way [144].

Analysis within SIFT-MS begins with the creation of precursor ions from a mixture of laboratory air and water vapor. In SIFT-MS analyses, H_3O^+, NO^+ and O_2^+ are used as precursor ions, and these have been chosen because they are known not to react significantly with the major components of air (e.g., nitrogen, oxygen, carbon dioxide), but do react with many trace gases. As shown in figure 5, the selected precursor ions are injected into a flowing carrier gas (typically helium) then travel along the reaction flow tube where gaseous samples entered at the same time. Within the flow tube chemical ionisation reactions occur between precursor ion and sample trace gases with the resultant production of characteristic product ions. The characteristic product ions of these reactions are subsequently detected by a downstream mass spectrometer. An in-build kinetics library permits direct quantification of target VOC when the instrument is used in in selective ion monitoring mode. Where desired the instrument can also be operated in a full scan mode although.

Like PTR-MS, there is no requirement for sample preparation and compounds can be quantified in units of absolute concentration. SIFT-MS is therefore ideally suited to online and real-time analysis of selected compounds within gas samples.

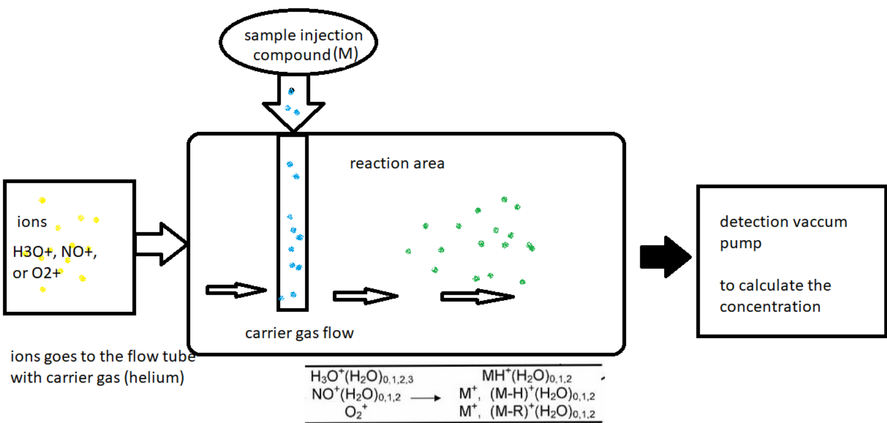


Figure 5 Selected ion flow tube mass spectrometry (SIFT-MS)

1.5.3 Other analytical platforms

Besides mass spectrometry-based methods that are describe above, there are other techniques that can be used to detect VOCs within gas samples.

1.5.3.1 Sensor based technologies

Two major gas sensors, the metal oxide semiconductor (MOS) sensor and the organic-based chemiresistive gas sensors, have been used in the detection of VOCs. MOS sensors are relative low cost, and possesses a short response time and long lifespan [145]. It can select VOCs including formaldehyde, benzene and naphthalene in ppb concentration [146]. Organic-based chemiresistive gas sensors depend on the change in the electrical resistance of the materials due to their chemical interaction with the analytes [147]. Current applications include TNT (Trinitrotoluene) detection [148]. Currently, the sensor is not a practical test in clinical research.

1.5.3.2 Electronic nose (eNose)

Dodd and Persaud developed the first model of electronic nose to describe smells [149]. Although the correlation between human odour impression and data from eNose was poor, it triggered the development of a new method to analyse non-odorant volatile compounds. However, the clinical application of eNose was interrupted due to poor intra-device repeatability, limited temporal stability. Poor selectivity due to the promiscuous nature of the chemical interactions between the sample and the sensors were also noted [150]. Although the use of eNose devices have been used to differentiate different disease state, wide application has been limited by their inability to define the identify of discriminatory VOCs [98, 151]. For this reason, several authors reporting the results of eNose analysis of breath have undertaken dual analysis with GC-MS [97, 98].

1.5.3.3 Field asymmetric ion mobility spectrometry (FAIMS)

FAIMS uses the application of a high-voltage asymmetric waveform in combination with a static waveform applied between two electrodes to separate ions at atmospheric pressure. The strength of the electric field is asymmetric and offset of the ion mobility occurs. The value of offset depends on the strength of the electric field and the structure of the ion. With different offset values, only some specific ions may go through a specific filter and are detected by the detector. Therefore, FAIMS can separate the ions and distinguish the identity of each ion.

Reports using FAIMS have revealed some advantages of this technique in clinical trials [152-155]. The application of FAIMS includes cancer, inflammation, or infectious

disease. Some studies were focussed on colorectal disease [156, 157]. The findings of from these limited studies reported the urinary VOC profile of CRC or and inflammatory bowel disease. Unlike many other mass spectrometry techniques, FAIMS is portable with a higher potential for point-of-care testing.

1.5.4 Method of breath sample collection

Breath sample collection varies in different studies. Ideally, sample collection and analysis should be as simple as the alcohol breath test. However, off-line sample analysis is still important in the design of clinical trials.

1.5.4.1 Direct sampling

PTR-MS and SIFT-MS are designed for direct injection of samples without requirement of pre-processing. The advantage of direct sample injection is less contamination or modification of the sample and real-time analysis of VOCs. As VOCs in clinical samples such as breath are often found at low concentrations, direct sample injection can prevent possible loss or transformation of compounds that may occur during collection and/or storage. This approach also has the opportunity to provide instantaneous results that is desirable in certain circumstances, such as clinical practice. At the present time, the expense and size of most PTR-MS and SIFT-MS instruments limits their wider use as on-line analytical platforms in clinical research. For this reason, direct sampling is not suitable for multi-centre clinical trials since the instrument needs to be located where the patients are recruited.

1.5.4.2 *offline sample analysis*

Although there are PTR-MS and SIFT-MS which can give results in seconds and gives results quantitatively, it cannot identify them directly (isomers for PTR-MS and selected compounds only for SIFT-MS). While there is no single mass spectrometry available for both identifying and quantifying compounds in a brief time, different platform analysis is suggested. Owing to the constraints of different analytical platforms there is often a requirement for indirect (offline) sample analysis. This is where samples are collected and stored for a period prior to being analysed. This may be because it is either not possible to bring the instrument to the point of sample collection or alternatively due to an instrument's inability to accommodate online sampling. Reliance on offline sampling has been particularly important in large scale clinical trials that are designed to examine VOCs with samples collected from many patients, often at various locations.

There are a range of methods available for the collection and storage of breath samples. Commonly, patients have been asked to exhale in breath sample bags that can be made from a range of materials, including Nalophan, Tedlar and Mylar [98, 158-161]. Samples held within sample bags can be analysed by direct injection (online) techniques by simply connecting the bag to the inlet of the instrument, as is possible with both PRT-MS and SIFT-MS. Previous studies have generally demonstrated that sample collection and analysis in this way is feasible with adequate repeatability and reliability [162]. A limitation of this approach is that a relatively brief time interval (<12hrs) is advised between sample collection and analysis to avoid sample loss and contamination.

To partially mitigate the recognised limitations of sample bag usage, alternative methods have been developed for the collection and storage of breath samples prior to offline sampling. Thermal desorption (TD) methods were developed in the 1970's as a way of concentrating VOCs onto a stainless-steel tube containing a sorbent phase, with the intention of providing a stable method of storage prior to analysis. Some recent studies showed that breath collected in TD tubes can be stored for several months, under proper conditions [163].

Typical materials used as sorbent solid phase include: Tenax, Carbograph and Polydimethylsiloxane. VOCs are retained in the TD tubes solid phase and then released at high desorption temperature. At the present time TD techniques are most used in conjunction with GC-MS analysis, although other techniques including PRT-MS have also been modified to permit this form of analysis.

Gas samples may be introduced to TD tubes through passive diffusions from an ambient environment or by active transfer using a manual or electronic pump.

Recently devices such as the ReCIVA® (Owlstone Medical Ltd, Cambridge, UK) have been developed in order to provide a more reliable method of exhaled breath capture onto TD tubes. A schematic of the ReCIVA device can be seen in figure 6. There are two computer-controlled pumps in the ReCIVA that regulate the flow of breath samples on to the TD tubes. The most important innovation for ReCIVA may be the carbon dioxide detector, which can distinguish the different phase of a breath cycle permitting activation of the pumps at the desired time to capture either an entire exhalation or only the end expiratory portion. The maximum TD tube loading for each sampling process is four. The ability to collect multiple sample simultaneous allows for the

acquisition of experimental repeats that form a critical component of quality control. In clinical trials, as discussed above, repeatability of sampling is important and the higher the similarity of the samples acquired, the more accurate of the results may be. Further discussion about the use of ReCIVA will be included in the method development chapter, 2.1.

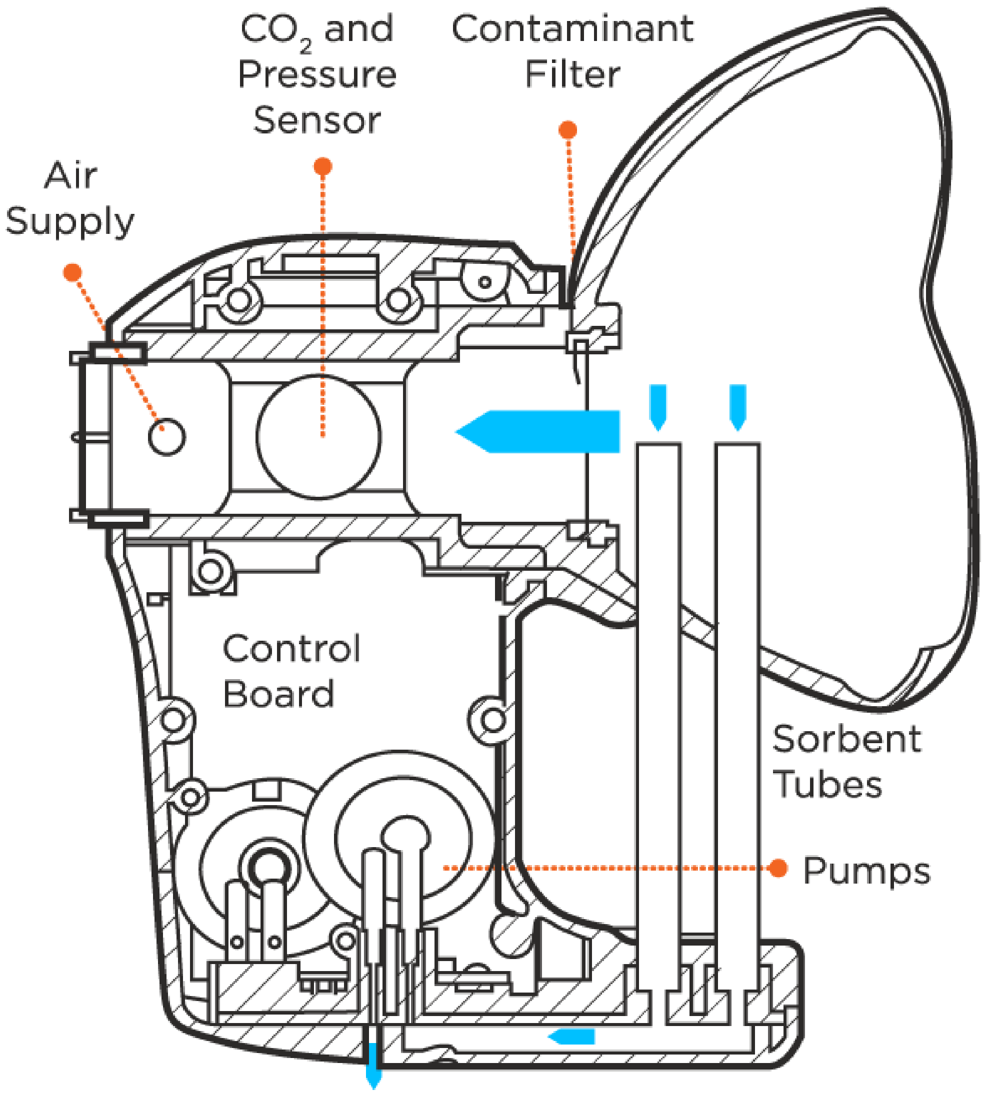


Figure 6 ReCIVA breath sampler

(Owlstone, https://www.owlstonemedical.com/)

Alternatively, HiSorb probes have been developed specifically for the analysis of headspace above a gas, liquid or solid (e.g., stool, tissue) sample. As shown in the figure 7, there is one segment in the HiSorb probe which is composed of a sorbent (e.g., polydimethylsiloxane (PDMS)) for the extraction of compounds. The probe is placed in the closed area environment which contains the sample. The duration and precise conditions required for Hisorb sample analysis are still under investigation and may be varied between each sample. After the sample collection, the probe is placed in an empty TD tube which can be analysed using a range of techniques.

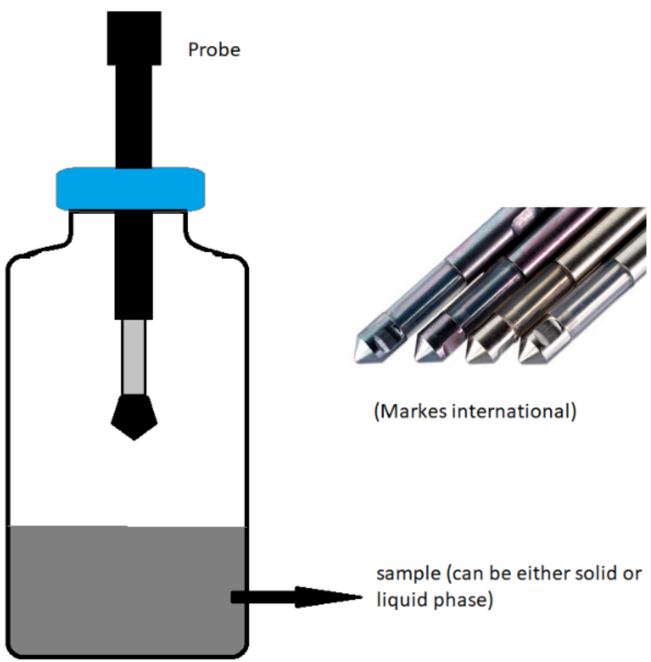


Figure 7 HiSorb tubes and example of use

An alternative method of VOC capture is solid phase microextraction (SPME). The SPME method is based on a coated fibre, which extracts different analytes (including both volatile and non-volatile) from different media, that can be in liquid or gas phase [164]. An image of liquid analysis by SPME is provide in figure 7. As with Hisorb, the SPME probe can be analysed using a range of techniques.

1.6 Current challenges with breath VOCs analysis

Despite growing evidence that VOCs are linked to a number of disease states, there remain several barriers to wider implementation of breath testing in clinical practice.

At the current time there are no universally agreed guidelines for standardized collection and analysis of breath samples. As a consequence, methodologies vary significantly between studies. Furthermore, the majority of studies lack appropriate quality assurance measure to ensure results are dependable. It has therefore often been difficult for findings to be externally validated. In addition, there remains a poor understanding of the factors responsible for methodological and subject-specific variability in VOC levels.

The origin of the majority of VOCs that can be detected within the breath is similarly unknown. It has therefore been difficult to provide a definitive biological rationale for most of those compounds that have been associated with disease states. To gain wider acceptance within clinical practice, the mechanistic drivers underlying observed changes in VOCs will need to be established.

Whilst there has been continued advancement of analytical strategies for VOC detection, their application in clinical practice remains a challenge due to their cost and complexity. Techniques remain primarily research tools that are poorly suited to broader usage within clinical practice. Finally, further work is needed to ensure analytical techniques have the required accuracy to reliably detect VOCs at the low concentrations found within exhaled breath samples. Furthermore, there needs to be a better understanding of the comparability of different VOC detection platforms, strengthening the value cross platform analysis and validation of findings.

The last question is the possibility of finding compounds from other specimen like stool or blood which is also none to little invasiveness and could be collected with a more well-established method. For the faeces, the major problem is the lack of standardized collecting method. It is less likely to collect faeces directly from the patient at clinic. Unlike stool occult blood test, while the preferable method is to decrease any factors related to sample collection and analysis. From the review in 1.4, there were two studies focusing on faeces and high variation could be found with the results. Blood sampling may offer other opportunities for the VOCs survey. Using liquid chromatography, we can find compounds from blood directly, but the acquirement of blood is still invasiveness. Therefore, blood or faeces may be preserved in a later stage if the aim of the study is to establish a none-invasiveness test.

1.7 Summary of current knowledge and indications for future investigations

CRC is one of the leading malignancies worldwide and the key to better prognosis is early detection and diagnosis whether it is primary lesion or relapsed disease.

As most CRCs are either asymptomatic or associated with minimal symptoms in their early stage, there is a need for acceptable, accurate and affordable tests to improve patient selection for definitive testing by colonoscopy. Furthermore, there remains an important unmet clinical need to improve the detection of recurrent CRC after potentially curative surgery. Follow-up after CRC resection varies in different country, but they share one similarity in that CEA is still the only biomarker in routine clinic use. However, the false negative rate of CEA is still high and relying on this single biomarker may result in delayed diagnosis. Therefore, colonoscopy CT or PET-CT is still required for an unknown cause of CEA elevation.

To overcome these challenges, research into other CRC biomarkers may provide an alternative strategy for CRC detection and assessment. Previous studies have established a link between selected VOCs in CRC. Critically VOC analysis has the potential to be entirely non-invasive.

Previous studies investigating the role of VOCs in CRC detection have been limited by their small sample size and heterogeneous methods resulting in significant variability in results. An improved understanding of the influence of methodological variability may support a wider acceptance of VOC analysis in clinical practice. Such work could provide a basis for standardizes and quality-controlled VOC analysis.

1.8 Overall objective of this thesis

Hypothesis: colorectal cancers is associated with a unique profile of VOCs within exhaled breath, which can be exploited for the purpose of early detection and disease monitoring.

Primary aim

i To develop a VOC base model for diagnosis and monitoring of colorectal cancer

Secondary aims

ii To explore the influence of off-line sampling methodology on the detection of exhaled VOCs

iii To perform multiplatform correlation of VOCs in human breath

iv To determine the VOCs detected in colonic gas

v To determine the influence of treatment modality and disease recurrence on the VOC signal induced by CRC

With the literature review, we could understand that the high variation from each result may be related to the varied methods. Therefore, the primary aim of the thesis is to develop a clinically feasible method on the collection and analysis of breath samples. With the developed method, we can proceed to the clinical trial for further breath sampling from CRC patients. The secondary aim is to establish a model on the research of the human breath. With the modified method based on the direct sample analysis, we may find possible CRC biomarker from breath with a reasonable, and clinically feasible methods.

Section II

Method Development

2. METHOD DEVELOPMENT

Unlike blood or urine samples, breath samples cannot be visualised directly and there is no standardize method for sample collection and analysis. In this session, the process of method development will be described in three aspects including the evaluation of breath collection device, cross-platform sample correlation and assessment of background noise through room air sample analysis. These studies are intended provide a greater understanding of factors responsible for variability in VOC analysis, which will in turn inform a standardized methodology for subsequent clinical studies.

2.1 Evaluation of off-line breath collection methods and assessment of utility for large scale clinical trials.

2.1.1 Introduction

Numerous studies have reported that there are apparently VOCs profiles associated with disease. These VOCs within human breath and body fluids can be used to detect disease states, including cancer [109, 112, 133, 165-167]. However, these compounds found from different studies for the same disease could be varied not only the composition of the compounds but also for the concentration of these compounds (128,129,130,133,135). The detection of VOCs within the exhaled breath of diseased patients is based on the hypothesis that the disease process induces local changes in metabolic pathways that lead to the release of VOCs in the systemic circulation and hence to the lungs where they are partially excreted [168]. Similarly, these compounds may be released into urine, sweat, alimentary tract secretions or kept within the blood stream. Ideally, we can find similar compounds in the headspace of urine, blood, or stool specimens, but there is no current study support this hypothesis.

Collection and analysis of exhaled breath remains the preferred approach to determine disease-specific changes in VOCs. Breath testing is non-invasive and is almost universally acceptable to patients. Unlike other samples (e.g., urine, blood) that require vaporisation of VOCs before they can be trapped and analysed, breath samples do not typically required pre-processing. Limitations of breath testing however include the often-large variability in exhaled VOC levels that is presumed to be a consequence of both subject-specific variability (influenced by diet, medication, oral bacteria, disease) and methodological variability (influenced by collection and analytical methodology).

Such variability has contributed to poor repeatability of studies investigating disease-specific VOC biomarkers [97, 169]. A review of VOC biomarkers in CRC in the chapter 1.4 highlighted various sources of samples (breath, blood, urine, faeces, or tissue), different analytical platforms, and different putative biomarkers. Whilst findings suggest a potential benefit of VOCs as biomarkers of colorectal cancer, there is currently no consensus of evidence regarding the appropriate methodology for sampling or the chemical targets. Clearer evidence and consensus guidelines are therefore required before wider acceptance of VOCs analysis in clinical practice.

Standardized breath sampling methods should be the foundation of the analytical process. This issue has been the subject of much debate in the field in recent years [170]. A number of breath sampling devices have been developed and used in clinical trials. Such devices typically facilitate the transfer of exhaled breath from a sampling manifold onto TD tubes that can store breath samples for up to 30 days. These devices vary however in the method by which they sample breath and transfer it to TD tubes. At the current time little is known about the performance of these devices in terms of VOC retrieval, sample contamination and repeatability. An understanding of these factors is important in ensuring standardized high-quality data collection within clinical trials.

The aim of this study is to evaluate three different methods of offline breath collection and transfer to thermal desorption tubes. Objectives were to determine their performance, reproducibility, and utility.

2.1.2 Method

Study population

Twenty healthy volunteers within the Department of Surgery and Cancer (Imperial College London) were recruited to this study. These twenty volunteers were asked to provide breath samples using each of the three devices described below after well explained of the sampling protocol and acquired their agreement with the consent form.

Ethical Approval

NHS Health Research Authority (NRES Committee London –Camden and Islington) approval was gained on July 16, 2014 (REC reference 14/LO/1136). Trials registration number: UKCRN18063.

Breath sampling devices

Three different devices were compared: BioVOC (Markes International Ltd., Llantrisant, United Kingdom), Breath concentrator (UWE,UK) [171], and; ReCIVA (Owlstone Ltd, Cambridge, United Kingdom) (Figure 8). Selection of these three devices was based on: (i) previous experience using them; (ii) their ability to transfer breath directly to TD tubes without requirement for sample bags, and (iii) the potential for their future use for breath sample collection in large-scale clinical trials. Although we chose three different devices here, it does not mean that we will only use these devices. The devices we use is to discover the biomarker. Once we find the biomarker, we may use a completely different technology to detect these biomarkers. It is like the COVID-19 test before. Before, to confirm the diagnosis of COVID-19 may be a complicated process on the detection of DNA. With the better understanding of the virus, we can use some other rapid test with acceptable sensitivity and specificity. Therefore, the accuracy may be a crucial factor than the cost of time or money for each device.

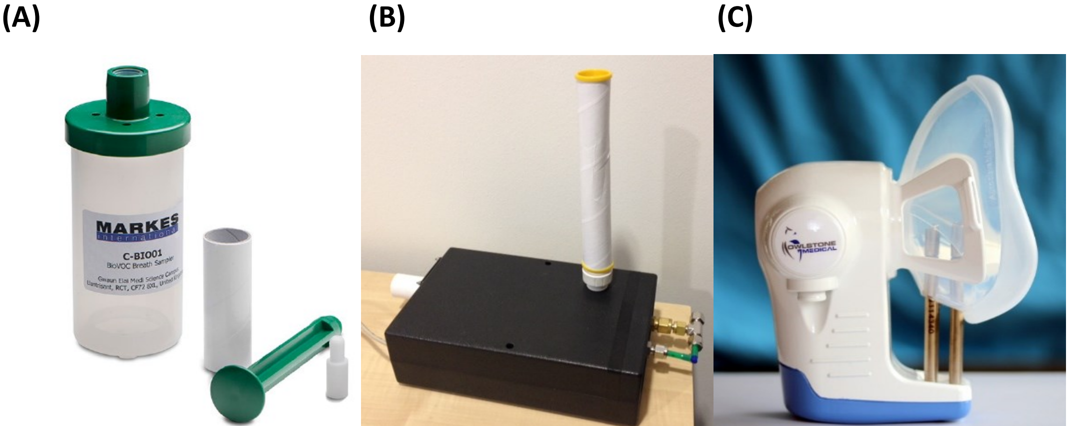


Figure 8 Breath sampling devices:

(A) *BioVOC (Markes International),* ***(B)*** *Breath concentrator, and* ***(C)*** *ReCIVA (Owlstone)*

BioVOC (Markes international Ltd., Lantrisant, United Kingdom)

The BioVOC device is lightweight, compact, and relatively simple to use. Samples are collected by asking subjects to exhale directly into the device via a disposable mouthpiece. The device is able to sample 129ml end tidal breath. After a breath sample has been delivered to the device, a plunger mechanism can be manually activated in order to expel the sample from the device into a TD tube that is inserted into its inlet. Sampling of breath can be repeated to increase the volume of breath transferred to TD tubes. While BioVOC is designed as a single-use device, the cost of sampling a single patient was more than £30.

In this study, subjects were asked to provide two separate breath samples (total sample 258ml) that were sequentially loaded into a single TD tube.

Breath concentrator

This bespoke device was design by researchers working within the field of breath analysis [171]. Samples were collected by asking subjects to repeatedly exhale into the

device via a disposable mouthpiece. Computer software is used to sample the end tidal (alveolar) portion of each breath. Breath collection depends on the exhalation flow rate. Once the exhalation rate exceeds a user-defined threshold an internal pump is activated to extract a sample of breath that is loaded on to a TD tube. The device is designed for repeated use with the only consumable being an inexpensive disposable mouthpiece. More detailed design regarding to breath concentrator is described in appendix 1. Since there is only one disposable mouthpiece necessary for each sampling, the cost might be lower than £2 to sample a single patient. However, breath concentrator is a self-designed sampler and is not commercially available. Considering the large-scale clinical trial in the future, three or four more samplers may be required if we would like to take samples from different hospital. Overall cost of each sample with breath concentrator for clinical trial may be difficult to calculate.

In this study a 250ml breath sample was collected using the breath concentrator device and transferred directly on to a TD tube.

ReCIVA

This commercially available device utilises a disposable mask and reusable handheld pump system to coordinate the simultaneous collection of breath samples into four TD tubes. Sampling is controlled through detection of carbon dioxide (CO_2) within each exhalation. Once a predefined CO_2 threshold is reached, sample pumps within the device are activated, transferring breath to TD tubes. The device requires a source of VOC-free 'scrubbed' air as well as a computer to coordinate sample collection. Including the disposable mask, this method has relatively high start-up and running costs. The cost for each sample taken with ReCIVA may be £20 for the mask and the sampler.

In this study 250 ml of whole breath was collected on to four TD tubes (250ml each) at a rate of 400ml/min. The four TD tubes were collected at the same time and further analysis as the repeatability and intra-device analysis would be done.

Breath sampling protocol

Three exhaled breath samples were collected using the three breath sampling devices described above. For breath concentrator and BioVOC, simply three repeat samples were taken. For ReCIVA itself, some different setting was applied while it can collect four tubes at the same time. The first sample taken with ReCIVA may be four tubes at the same time, but we collected only single tube with ReCIVA for the 2nd and 3rd sample taken. The first sample taken was to measure if the four tubes were similar. This may be important for future clinical trial if we can decrease time consumed for sample taking. All sample were collected on the same day during a single period of sampling. The order in which breath samples were collected with each device was randomised using sealed envelopes. Repeatability of each device was further analysed through comparison of the three sets of samples. Following the protocol, twelve TD tubes may be required for a single examinee. We used 24 TD tubes (12 x 2) for sample collection and the tubes used for each volunteer were randomly selected.

Breath sample analysis

Prior to sampling, TD tubes were conditioned using TC-20 (Markes International Ltd., Lantrisant, United Kingdom) with settings: 20 psi pressure of nitrogen gas under 300°C for 70 minutes.

Exhaled breath samples stored within TD tubes were analysed using GC-MS. The TD tubes were desorbed using a Markes TD-100 TD unit (Markes International Ltd, Llantrisant, UK) using a two-stage desorption programme, applying a constant flow of helium at 50 ml min^{-1}. In the primary desorption stage, TD tubes were dry purged for 3 min and heated at 280 °C for 10 min. In the secondary desorption stage, the cold trap (U-T12ME-2S, Markes International Ltd, Llantrisant, UK) was rapidly heated (99 °C min^{-1}) from 10 °C to 290 °C. VOCs were transferred from the TD unit to the GC by means of a capillary line heated at 140°C.

GC-MS analysis was performed using an Agilent 7890B GC with 5977A MSD (Agilent Technologies Ltd, Santa Clara, USA) equipped with a ZB-642 capillary column (60 m × 0.25 mm ID × 1.40 µm df; Phenomenex Inc., Torrance, USA) with helium used as the carrier gas (1.0 ml min^{-1} flow rate). The GC column temperature programme was set as follows: 4 min at 40 °C, ramped up to 100 °C at 5 °C min^{-1} with a 1 min hold, ramped up to 110 °C at 5 °C min^{-1} with a 1 min hold, ramped up to 200 °C at 5 °C min^{-1} with a 1 min hold and finally ramped up to 240 °C at 10 °C min^{-1} with a 4 min hold. The MS transfer line temperature was 240 °C and EI source conditions were 70 eV at 230 °C. Mass acquisition was conducted in the range 20–250 m/z with a rate of approximately six scans s^{-1}.

Prior to analysis, TD tubes were spiked with the internal standard Toluene D-8 that was added manually to each sampled tube. The purpose of adding toluene D-8 is to ensure the quality of analysis in GC-MS. A constant flow of nitrogen was given during the spiking. The setting of flow rate was 20 psig. 0.5 µL toluene D-8 was injected into each sample tube before analysed in GC-MS. Because all the samples were analysed with

GC-MS, it was not necessary to quantify the concentration of each compound with the fixed amount of internal standard added.

Data analysis

VOC analysis was completed using the Agilent Mass Hunter Qualitative Analysis software (Agilent Technologies Ltd, Santa Clara, USA) and VOC identification was completed using the National Institute of Standards and Technology (NIST) Mass Spectral Database version 2.0 (NIST, Boulder, USA). Using the Mass Hunter software, the lower limit of peak area count detection was set at 10^4. Whilst the detection limit for compounds analysed by GC-MS may be different based on their structure and affinity, 10^4 was considered to be an appropriate threshold for most compounds that are detectable within human breath.

The compounds and contaminant collected using the three devices were compared using the Friedman's test. Repeatability was analysed by calculating the standard deviation and variation of each compound from each device individually. The equation for standard deviation is as following:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(xi - \mu)^2}$$

In the equation, μ is the average of the X1 to XN.

The equation for coefficient variation was calculated as

$$\frac{standard\ deviation\ (\sigma)}{mean\ (\mu)}$$

Due to the lack of reference concentration of compounds from human breath, it was not possible to determine the accuracy of the results directly.

Human factors analysis of sampling devices

After using each device, subjects were requested to complete a short questionnaire to evaluate their experience with each device. Participant's experience was assessed as follows:

i) The Usability Metric for User Experience (UMUX-LITE) was used to measure participants satisfaction in using the devices [172-174];
ii) The Net promoter Score (NPS) was used to identify participants intention to promote or to prevent the use of each device [175];
iii) The pairwise comparison was used to qualitatively assess the participant preferences of each device against the others in terms of perceived ease of use.

The whole questionnaire is attached in the appendix 2.

2.1.3 Results

Characteristics of the twenty volunteers enrolled in this study are presented in Table 3. Most subjects had eaten within 4 hours of breath sampling. all the twenty healthy volunteers were the research workers at St. Mary's hospital. Although we cannot guarantee that they all have no not-known underlined disease, at least no active disease mentioned while doing sampling. Most of them works at the VOCs research group or the medical educational research group. Four of the volunteers needed to do night shift but none of them had night shift on the day before sampling. To simulate clinical study in the future, we did not request a strict fasten time. Also, sample would be taken in the morning or in the afternoon as for future sampling method.

VOC detection

Table 4 shows eighty-seven compounds detected by GC-MS within exhaled breath samples from these twenty healthy volunteers.

Compounds and contaminant

Initial comparisons were made to compare VOC detection and relative abundance (peak area count) in breath samples collected using each of the three sample devices (Table 4). There were eighteen potential contaminants detected in breath samples. Comparison of potential contaminants from the sampling pathways were also assessed. Excluding the internal standard Toluene-D8, the relative abundance of 17 compounds (including 5 potential contaminants) were not significantly difference in breath collected with each device ($P>0.05$).

Table 3 Subject characteristics

Gender	(Male/Female)	9/11
Age	(years)	33.7 ± 5.85 (Range 25-52)
Ethnic origin	Asian	7
	Caucasian	7
	Mixed or African	6
Smoking	Non smoker	19
	Smoker	1 (e-cigarette)
Alcohol	0 units/week	6
	1-14 units/week	14
	>14 units/week	0
Last oral intake	< 1 hour	3
	1-4 hours	15
	> 4 hours	2

Table 4 VOCs detected within breath samples (in the order of the retention time and the chemical group)

	Chemical group	Breath concentrator		BioVOC		ReCIVA		P
		Median	IQR	Median	IQR	Median	IQR	
Isoprene	Abundant compound	15913800	[9525161 - 18832014]	1766801	[1224055 - 2651746]	4779060	[3970548 - 10439388]	<0.001
Acetone		21901718	[14849847 - 27918493]	3848862	[2885580 - 5526904]	11394349	[8335275 - 16023699]	<0.001
Methanal	Aldehydes	1034539	[748478 - 1398760]	239383	[134396 - 269522]	558336	[418988 - 855334]	<0.001
Propanal		71718	[61563 - 90694]	36181	[30238 - 58583]	55713	[42968 - 78902]	<0.001
Hexanal		98873	[83122 - 146030]	72308	[56302 - 97017]	89385	[78342 - 126817]	<0.001
Heptanal		112523	[100290 - 191353]	87272	[70965 - 174982]	117627	[76370 - 135736]	0.086
Octanal		348619	[237012 - 502308]	330715	[191814 - 523873]	474950	[341774 - 601323]	0.026
Nonanal + acetophenone		1248478	[804608 - 2171625]	1109429	[734515 - 1514520]	974926	[617427 - 1562749]	0.091
Decanal		1200620	[750265 - 2264576]	1193577	[638250 - 2029607]	980411	[616610 - 1689349]	0.247
Undecanal		358431	[324170 - 533672]	189191	[131071 - 313246]	174117	[136049 - 288614]	0.002
Dodecanal		606880	[406187 - 707992]	489306	[350175 - 553136]	419897	[388554 - 540346]	0.116
Methanol	Alcohols	13372414	[12272451 - 17191258]	10947285	[8311341 - 12426494]	13726705	[11200702 - 15004587]	0.001
Ethanol		1705562	[1012548 - 2470274]	387869	[262420 - 662416]	746167	[627993 - 1159055]	<0.001
1-propanol		309317	[193159 - 547272]	49727	[36042 - 72695]	74338	[46586 - 103327]	<0.001
Isopropanol		320368	[283587 - 411702]	68577	[45221 - 92160]	260785	[181266 - 551514]	<0.001
1-butanol		55354	[42513 - 70291]	41603	[28262 - 52985]	83015	[59523 - 104543]	<0.001
1,1'-biphenyl		112073	[96333 - 142778]	107881	[88739 - 125691]	111700	[85879 - 128821]	0.282
Methyl formate	Esters	394251	[343127 - 513646]	146517	[98319 - 188984]	300397	[182865 - 376798]	<0.001
Acetic acid, methyl ester		882404	[655267 - 1114566]	227849	[149267 - 445888]	530576	[358124 - 729666]	<0.001
Carbonic acid, dimethyl ester		154052	[127417 - 245313]	49297	[34122 - 68944]	129521	[95603 - 198504]	<0.001
Benzoic acid, methyl ester		266545	[182488 - 383459]	107876	[71545 - 182232]	184064	[126708 - 261382]	<0.001
Benzonitrile		131291	[74835 - 168987]	50502	[43071 - 68485]	74307	[55865 - 95651]	<0.001

Glycerol 1,2-diacetate		33007559	[29478099 - 37416892]	525406	[178049 - 1237048]	171918	[109745 - 667712]	<0.001
Isopropyl myristate		109265	[88916 - 155280]	134073	[114918 - 186706]	129397	[91592 - 187730]	0.211
2-Methylpropane	Hydrocarbons	269272	[205690 - 486107]	68565	[57992 - 135507]	154068	[120474 - 231349]	<0.001
Isobutene		196142	[138467 - 325554]	149038	[119523 - 205715]	181760	[135911 - 229776]	0.002
Butane		224360	[165619 - 425758]	61527	[37081 - 85084]	143641	[57418 - 196001]	<0.001
2-Methylbutane		301391	[186316 - 414217]	94427	[41052 - 152326]	146880	[109470 - 214972]	<0.001
Pentane		320270	[229747 - 593578]	140496	[69850 - 247116]	170889	[108362 - 438850]	<0.001
n-Hexane		3205738	[1393378 - 5371928]	1320285	[416375 - 2249622]	1531956	[784964 - 3759221]	<0.001
Decane		96807	[80507 - 137615]	56350	[41950 - 76526]	266906	[222136 - 332161]	<0.001
Undecane		221613	[68065 - 262156]	42835	[36810 - 222589]	156055	[71486 - 254179]	<0.001
Tetradecane		3429716	[2576596 - 4093292]	132100	[118738 - 187817]	183950	[149870 - 252823]	<0.001
Hexadecane		1378051	[1066915 - 1774383]	155583	[127169 - 231515]	169355	152211 - 181910]	<0.001
Benzene	Aromatic	538446	[413018 - 656163]	431173	[338502 - 541541]	494804	[347556 - 627949]	0.157
Toluene-D8 (Internal standard)	compounds	**42193935**	**[36379915 - 47308795]**	**41635829**	**[35588446 - 47290320]**	**45052027**	**[37856193 - 49012942]**	**0.350**
Toluene		245745	[197496 - 345045]	153678	[126183 - 218252]	246503	[195428 - 311852]	<0.001
Benzaldehyde		439696	[319108 - 573886]	391144	[331228 - 425925]	396908	[302651 - 460009]	0.007
Limonene		745701	[461603 - 1471145]	141024	[70518 - 212918]	241344	[179494 - 515717]	<0.001
Eucalyptol		110707	[70818 - 147979]	22696	[19323 - 36563]	57364	[47789 - 79205]	<0.001
Xylene		115253	[86545 - 158280]	66761	[54604 - 90556]	110539	[84244 - 148405]	<0.001
Xylene combined to styrene 26.28		70977	[52063 - 85853]	45232	[31711 - 47464]	64581	[52917 - 90951]	<0.001
Acetic acid	Fatty acid	320233	[169199 - 434935]	105724	[67943 - 152465]	178599	[138329 - 308490]	<0.001
Propanoic acid		46775	[28495 - 93863]	12030	[5212 - 14812]	14197	[12686 - 26762]	<0.001
Sevoflurane	Others	395979	[188151 - 536816]	109998	[70884 - 618735]	154520	[127111 - 527523]	0.026
Acrolein		115195	[89443 - 150051]	81987	[55186 - 113138]	86055	[64327 - 123896]	0.009
Dimethyl sulphide		326335	[186935 - 376520]	40840	[28444 - 74746]	130247	[92605 - 187501]	<0.001
Isoflurane		53694	[20548 - 146257]	23953	[12287 - 141380]	33701	[24708 - 130197]	0.021

Carbon disulphide		242873	[126813 - 899014]	203933	[135962 - 1196726]	217571	[93795 - 3644593]	0.465
3-methyl-2-butanone		32181	[20701 - 42925]	28962	[19354 - 42874]	35152	[24236 - 43063]	0.144
1,3,5-triazine		27462	[17073 - 59457]	10916	[0 - 12866]	19025	[12585 - 28395]	<0.001
Allyl methyl sulphide		128204	[72667 - 249850]	17327	[12176 - 33361]	46476	[24639 - 99776]	<0.001
2-pentanone		154660	[110414 - 193107]	59148	[44310 - 69924]	85343	[77744 - 101583]	<0.001
1-(methylthio)-propane		118524	[72710 - 155228]	41082	[28429 - 45192]	66046	[59128 - 72935]	<0.001
3-methyl-thiophene		178572	[113184 - 289901]	31793	[10966 - 48248]	50769	[35085 - 134399]	<0.001
Furfural		28748	[18655 - 32448]	10513	[0 - 15462]	15904	[12523 - 18646]	<0.001
Ethylbenzene		73870	[56537 - 90439]]	44760	[35465 - 69956]	162692	[115451 - 187746]	<0.001
3-methylcyclopentyl acetate		36253	[29441 - 49197]	28385	[20584 - 50112]	66954	[47861 - 83182]	<0.001
N,N-Dimethylacetamide		26301	[18951 - 42596]	15928	[7817 35360]	30944	[23706 - 48175]	<0.001
N,N-diethyl-formamide		83012	[53998 - 336141]	53481	[30381 - 72602]	101860	[60575 - 250356]	<0.001
6-methyl-5-hepten-2-one		118857	[83924 - 178082]	74876	[48809 - 152930]	76402	[57651 - 115264]	0.091
2-ethyl-1-hexanol		266185	[195778 - 359660]	175379	[130885 - 246547]	198779	[167250 - 264045]	0.003
N,N-diethyl-acetamide		52231	[30499 - 63401]	33838	[21740 - 49792]	51084	[35189 - 68894]	<0.001
Phenol		792703	[573456 - 1155668]	150878	[91662 - 177586]	205818	[118370 - 309587]	<0.001
Dodecane		558557	[364087 - 750557]	69351	[54737 - 93595]	227909	[184410 - 258176]	<0.001
Tridecane		107277	[85665 - 127438]	64874	[44976 - 80860]	73456	[64409 - 95106]	<0.001
Methenamine		1343577	[645944 - 2717041]	30806	[26107 - 51597]	666400	[215049 - 1429899]	<0.001
Anethole		43192	[34858 - 68663]	35987	[31422 - 39421]	36199	[29541 - 51918]	0.212
N,N-dimethyl-1-dodecanamine		157318	[119084 - 177971]	120015	[95849 - 144429]	134189	[118972 - 182967]	0.760
1,1,3,3,3-Pentafluoropropene	Contaminant	19423	[13515 - 47192]	20433	[10912 - 40841]	6211	[0 - 62979]	0.816
1,1,1,2,2-Pentafluoroethane		322509	[148706 - 647630]	250101	[113964 - 1616217]	386286	[134548 - 1623991]	0.949
Trimethylsilyl fluoride		55122	[46612 - 66973]	44334	[38849 - 53664]	615953	[450220 - 920481]	<0.001
Methanol, TMS derivative		69243	[53169 - 126350]	37104	[24585 - 59856]	2937539	[2398420 - 4064365]	<0.001
Silanol, trimethyl-		726458	[591097 - 1109864]	318900	[234985 - 380317]	27960612	[24227244 - 31973604]	<0.001

Disiloxane, hexamethyl-	224170	[132898 - 1061131]	47957	[20737 - 101749]	46724581	[41041646 - 50468850]	<0.001
Diglycolic acid, di(pentafluorobenzyl) ester	10893	[0 - 12719]	0	[0 - 13438]	26763	[20598 - 36551]	<0.001
2-Propanol, 1,1,1,3,3,3-hexafluoro-	84447	[58479 - 107749]	61663	[41960 - 112468]	55131	[39445 - 111397]	0.115
Disiloxane, 1-ethenyl-1,1,3,3-tetramethyl-3-(2-propenyl)	23362	[12075 - 53620]	19031	[11917 - 24577]	2864845	[1200001 - 6388109]	<0.001
Di-trimethylsilyl peroxide	36242	[13160 - 59180]	20298	[13288 - 33881]	242797	[96771 - 316375]	<0.001
1,1,1,3,3,3-Hexafluoro-2-(trifluoromethyl)-2-propanol	13662	[10506 - 16386]	10050	[0 - 11351]	10476	[0 - 13337]	0.087
Disiloxane, ethylpentamethyl-	14950	[12412 - 19920]	5556	[0 - 13401]	825870	[499368 - 966385]	<0.001
Thiocyanic acid, methyl ester	22240	[0 - 114532]	14247	[0 - 49555]]	11866	[0 - 108081]	0.249
Acetamide, 2-fluoro-	44216	[35302 - 77215]	23472	[13251 - 27448]	34892	[27077 - 42915]	0.002
Cyclotrisiloxane, hexamethyl-	1163542	[1036515 - 1473194]	802608	[635124 - 976580]	2247328	[1997809 - 2790109]	<0.001
Trisiloxane, octamethyl-	24432	[13813 - 52346]	11440	[0 - 13303]	9349283	[4688260 - 12078791]	<0.001
Trimethylaluminum	13418	[12277 - 16865]	10547	[0 - 12960]	15080	[13251 - 18902]	0.002
2'-Hydroxy-5'-methylacetophenone, TMS derivative	18587	[11234 - 22377]	0	[0 - 14614]	318842	[259568 - 524405]	<0.001

Quality of sample collection and analysis

Toluene D-8 was used as an internal standard to ensure the stability of GC-MS analysis of TD tubes. No significant difference in the peak area of the internal standard Toluene D-8 was observed (P=0.350).

Abundant compounds: acetone, isoprene, methanol, and ethanol

Peak area count of common abundant breath compounds was highest in samples collected using the breath concentrator (P<0.05), although methanol levels were equivalent in ReCIVA samples. Breath samples collected using the BioVOC consistently had the lowest levels of these abundant VOCs.

Fatty acids

Short chain fatty acids were previously found to be potential volatile biomarkers of upper gastrointestinal cancers [109, 176]. Based on Kumar's findings, elevated concentration of Butyric acid, Pentanoic acid, and Hexanoic acid from breath was noted with SIFT-MS for gastroesophageal adenocarcinoma. While most of the CRC are adenocarcinoma, we tried to analyse short chain fatty to understand if this method can be clinically practical. GC-MS analysis of breath samples identified only two fatty acids: Acetic acid and Propanoic acid. Again, concentrations were highest in samples collected using the breath concentrator and lowest samples collection using the BioVOC.

Aldehydes

Previous studies suggested that aldehydes (C3 to C10) may be important upper gastrointestinal cancer biomarkers [109]. Also, propanal was found to be possible biomarker from Markar's study[132]. As mentioned in the fatty acid, we tried to

analyse possible cancer related compounds mentioned before, especially those diagnosed adenocarcinoma. Pentanal and butanal were excluded from the compound list because of their low peak area count. Three (heptanal, nonanal, and decanal) aldehydes showed similar peak area counts in samples collected using the three devices.

Other VOCs

Other VOCs including, carbon disulphide, benzene, 3-methyl-2-butanone, 6-methyl-5-hepten-2-one, anethole, dodecanal (aldehyde), 1,1'-biphenyl, N,N-dimethyl-1-dodecanamine, and isopropyl myristate showed comparable results in all three devices. As with other chemical classes VOC levels tended to be higher in breath samples collected using the breath concentrator.

Contaminants

Contamination may come from the sampling process or analytical platform. In this section, the specific contaminants associated with each device are reported. VOCs detected within breath samples were compared to an existing database of 872 VOCs detected in healthy humans [105]. It was noted that none of the participants in this study reported that they were suffering from any specific active illness. Compounds not previously reported in humans were considered to be contaminants. These contaminants (n=18, Table 3) were divided into two groups: (i) siloxane-related compounds and halogenated compounds (containing bromide, fluorine, and chlorine) which are typically not found within the exhaled breath of healthy subjects.

Siloxane-related compounds

Of those contaminants identified, ten were siloxane-related compounds. These compounds were found in higher levels in samples collected using the ReCIVA device and potentially signify contamination from the disposable mask. Most of the siloxane related compounds share some similarity including low thermal conductivity, chemical reactivity, and toxicity [177]. For its characteristic, siloxanes in ReCIVA device may not interfere in spectral analyses.

Halogenated compounds

Five of the remaining eight contaminants showed similar levels in the three devices ($P>0.05$). The peak area count for these compounds varied from < 10000 to >$10x10^7$. These contaminants are believed to derive from the analytical process including: the GC-MS, the tube conditioner, and ambient room air.

Repeatability of the three devices

Tables 5 and 6 show the variation (standard deviation/average) of the three repeat sample collections in common abundant breath compounds.

For abundant compounds (acetone, isoprene, and ethanol) higher variation was observed for samples collected with BioVOC (Table 5). The ReCIVA device tended to demonstrate the least variation (Table 6). High variation can be found with the four aldehydes analysed in the experiment. This finding likely reflects the lower relative peak area count for aldehydes.

Intra-device comparison

As the ReCIVA device is able to load four TD tubes with breath simultaneously it was also possible to assess the variation between samples collected at the same time. The

results were shown in the Table 5 and 6 as ReCIVA (II). Comparing to ReCIVA (I), only subject 3 showed higher variation in the common compounds. Considering the possibility of failed sampling in a single subject, the repeatability in common compounds is acceptable.

Table 5 Variation in repeated collections (common abundant compounds)

	BioVOC			Breath concentrator			ReCIVA (I)			ReCIVA (II)		
Subject	**Acetone**	**Isoprene**	**Ethanol**	**Acetone**	**Isoprene**	**Ethanol**	**Acetone**	**Isoprene**	**Ethanol**	**Acetone**	**Isoprene**	**Ethanol**
1	0.16	0.17	0.12	0.06	0.12	0.52	N/A	N/A	N/A	0.09	0.11	0.64
2	0.18	0.06	0.06	0.00	0.06	0.10	N/A	N/A	N/A	0.12	0.19	0.36
3	0.17	0.25	0.16	0.05	0.10	0.05	0.25	0.29	0.73	0.83	0.86	0.77
4	0.32	0.35	0.28	0.04	0.11	0.10	0.06	0.42	0.23	0.19	0.19	0.09
5	0.39	0.43	0.25	0.04	0.05	0.26	0.13	0.20	0.29	0.28	0.34	0.12
6	0.27	0.30	0.18	0.02	0.13	0.25	0.00	0.07	0.40	0.39	0.45	0.03
7	0.15	0.16	0.12	0.09	0.11	0.29	0.06	0.15	0.61	0.21	0.29	0.13
8	0.16	0.27	0.19	0.11	0.09	0.27	0.13	0.29	0.16	0.12	0.14	0.15
9	0.11	0.12	0.34	0.04	0.03	0.22	0.13	0.15	0.22	0.21	0.31	0.07
10	N/A	N/A	N/A	0.17	0.11	0.23	0.08	0.14	0.08	0.21	0.17	0.15
11	0.21	0.44	0.27	0.03	0.09	0.18	0.03	0.02	1.41	0.19	0.19	1.62
12	0.17	0.30	0.42	0.02	0.19	0.23	0.35	0.16	0.20	0.26	0.33	0.11
13	0.87	0.94	0.87	0.91	1.09	0.90	0.45	0.29	0.31	0.23	0.23	0.12
14	0.03	0.20	0.10	0.14	0.07	0.25	0.13	0.21	0.39	0.13	0.20	0.08
15	0.38	0.44	0.30	0.27	0.24	0.51	0.06	0.18	0.27	0.16	0.20	0.14
16	0.27	0.49	0.53	0.15	0.18	0.32	0.02	0.30	0.13	0.21	0.21	0.13
17	0.72	0.48	0.56	0.77	0.87	0.91	0.49	1.06	0.43	0.15	0.15	0.19
18	0.20	0.19	0.11	0.06	0.17	0.19	0.23	0.49	0.24	0.14	0.19	0.18
19	0.24	0.13	0.30	0.05	0.07	0.14	0.13	0.14	0.17	0.08	0.06	0.29
20	0.16	0.27	0.06	0.05	0.19	0.25	0.12	0.46	0.15	0.09	0.05	0.25

Data shown are standard deviation (of peak area count) over mean (of peak area count) N/A: data not available due to failure of analysis of one or more samples. ReCIVA (I), comparison for repeated breath samples (for each repeat the TD was taken from the same position of the ReCIVA device). ReCIVA (II), comparison of 4 samples collected simultaneously.

Table 6 Variation in the repeated collection (aldehydes)

	BioVOC				Breath concentrator				ReCIVA (I)				ReCIVA (II)			
subject	Propanal	Heptanal	Octanal	Nonanal	Propanal	Heptanal	Octanal	Nonanal	Propanal	Heptanal	Octanal	Nonanal	Propanal	Heptanal	Octanal	Nonanal
1	0.28	0.51	0.71	0.80	0.07	0.28	0.36	0.34	N/A	N/A	N/A	N/A	0.12	0.26	0.12	0.41
2	0.57	0.40	0.43	0.42	0.51	0.18	0.10	0.18	N/A	N/A	N/A	N/A	0.12	0.32	0.24	0.44
3	0.23	0.15	0.22	0.38	0.08	0.24	0.35	0.40	0.22	0.22	0.34	0.47	0.76	0.79	0.72	0.68
4	0.15	0.55	0.73	0.80	0.36	0.37	0.56	0.57	0.27	0.36	0.26	0.29	0.14	0.35	0.17	0.38
5	0.30	0.13	0.29	0.40	0.38	0.08	0.14	0.17	0.30	0.08	0.10	0.43	0.25	0.08	0.21	0.45
6	0.17	0.24	0.34	0.33	0.12	0.20	0.26	0.27	0.14	0.44	0.11	0.49	0.02	0.03	0.26	0.19
7	0.32	0.53	0.91	0.93	0.20	0.69	0.82	0.77	0.27	0.61	0.53	0.55	0.32	0.19	0.27	0.65
8	0.16	0.17	0.53	0.58	0.24	0.64	0.92	0.68	0.38	0.11	0.12	0.02	0.19	0.14	0.12	0.28
9	1.65	0.30	0.55	0.64	0.35	0.05	0.26	0.44	0.33	0.19	0.35	0.67	0.39	0.14	0.54	0.39
10	N/A	N/A	N/A	N/A	0.28	0.08	0.14	0.23	0.12	0.30	0.20	0.36	1.75	0.51	0.16	0.59
11	0.49	1.04	1.37	1.50	0.20	0.23	0.28	0.25	0.29	0.24	0.07	0.42	0.66	0.23	0.13	0.10
12	0.21	1.02	1.07	1.04	0.13	0.79	0.79	0.85	1.63	0.44	0.24	0.70	0.46	0.44	0.12	0.74
13	1.07	1.07	1.27	1.24	0.76	0.64	0.43	0.40	0.06	0.59	0.67	0.70	0.18	0.72	0.56	0.97
14	0.43	0.45	0.63	0.99	0.64	0.95	0.96	1.11	0.50	0.34	0.19	0.58	0.10	0.39	0.08	0.35
15	0.15	0.47	0.40	0.45	0.29	0.01	0.08	0.18	0.38	0.25	0.24	0.43	0.47	0.51	0.17	0.45
16	0.94	0.28	0.48	0.49	0.96	0.54	0.48	0.46	0.59	0.69	0.70	1.35	0.47	0.43	0.12	0.54
17	0.35	0.23	1.06	0.40	0.38	0.32	0.04	0.01	0.37	0.15	0.68	0.15	0.22	0.18	0.15	0.21
18	0.41	0.59	0.65	0.50	0.21	0.82	0.76	0.62	0.04	1.53	1.46	1.59	0.26	0.54	0.34	0.73
19	0.69	0.38	0.64	0.42	0.18	0.15	0.29	0.42	0.22	0.31	0.28	0.66	0.79	1.12	1.11	1.11
20	0.22	0.39	0.26	0.35	0.20	0.17	0.03	0.09	0.09	0.46	0.27	0.43	0.32	0.32	0.16	0.45

Data shown are in standard deviation (of peak area count) over mean (of peak area count)

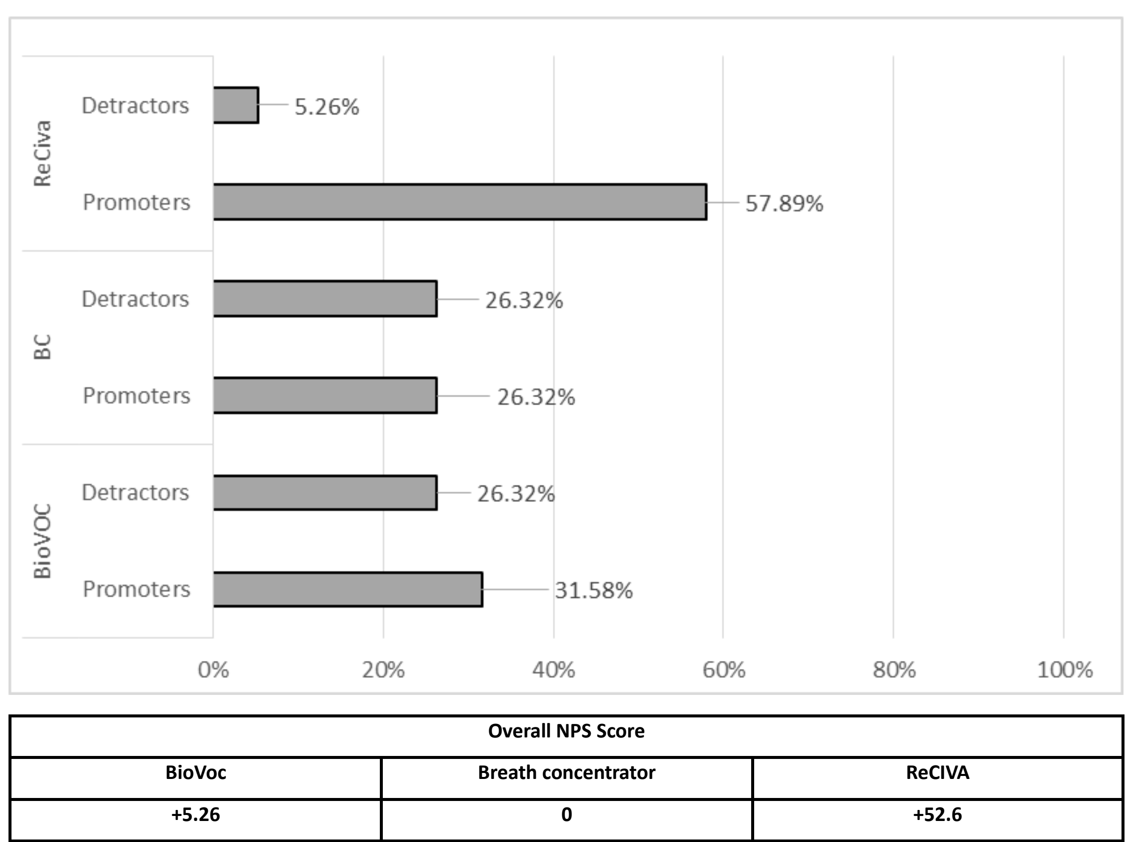
N/A: data not available due to failure of analysis of one or more samples. ReCIVA (I), comparison for repeated breath samples (for each repeat the TD was taken from the same position of the ReCIVA device). ReCIVA (II), comparison of 4 samples collected simultaneously

Human factors analysis of sampling devices

Results of UMUX-LITE (Table 7) calculated by using the standard curved grading (SCG) scale of usability, show that participants were generally more satisfied using the ReCIVA (level B) compared to other the devices (Level C)[174, 178]. Responses to the NPS demonstrated a strong willingness of subjects to promote the use of ReCIVA, compared to BioVOC, and breath concentrator which received no intention to promote (Figure 9). Percentage of participants who are willing to promote the use (Promoters) or express the intention to act against (Detractors) the use of each device. Overall NPS scores are calculated by subtracting the number of detractors to the number of promoters. The pair-wise comparison revealed that 47.4% of participants perceived ReCIVA more user friendly compared to BioVOC and breath concentrator. The least user- friendly device was the breath concentrator with 18.2% of the preferences.

Table 7 Results of the Usability Metric for User Experience

	BioVOC	**Breath concentrator**	**ReCIVA**
Perceived Usability	70.7%	67.5%	73.2%
SCG	C	C	B-



Overall NPS Score		
BioVoc	**Breath concentrator**	**ReCIVA**
+5.26	**0**	**+52.6**

Figure 9 Overall net promotor score (NPS) for the three devices

2.1.4 Discussion

Establishing reliable standardized protocols for clinical breath testing remains a major challenge. Central to overcoming this challenge is the ability to optimise the method of breath sample collection and storage. This is the first study to directly compare different methods of breath sample collection on to TD tubes, which are the preferred modality for long term VOC storage. Three devices were compared, including two that are commercially available. Factors that were assessed included: VOC retrieval; potential contaminants; repeatability of sampling; human factors. Findings indicate that each sampling device has its own strengths and weaknesses that influence their suitability for use in large-scale clinical trials.

Breath samples collected using each device were found to contain an appropriate number of VOCs. The number of VOCs that are detectable remains important in phase 1 biomarker discovery studies that rely on untargeted analysis of the volatilome. It is however important to note that other factors impact on the ability to detect different chemical classes of VOCs, including patient health status, the sorbent within TD tubes and characteristics of the analytical platform. In the current study those factors remained constant. The numbers of VOCs detected from samples collected from each device were broadly similar. Concentrations of VOCs were generally higher in samples collected using the breath concentrator. This device specifically samples end tidal breath where the initial portion of 'dead space' air is excluded. For those VOCs that are derived from the systemic circulation and exchanged within the lower airways and alveoli (e.g., Isoprene), this sampling method is anticipated to yield higher concentrations.

Another factor influencing the number and type of VOCs detected include the presence of contaminants. This study identified a number of contaminants that may have originated both from the breath sampling devices and from the subsequent method of analysis by GC-MS. Exogenous siloxane related compounds were particularly linked to the ReCIVA device and its single use face mask that is required for breath sampling. Recognition of such contaminants is important to ensure that they are excluded from datasets prior to making any comparisons based on clinical status.

The repeatability of breath sampling using each device was also assessed. Again, results tended to show the highest repeatability was for the breath concentrator device. The ReCIVA device however offers the unique ability to collect breath simultaneously on to four TD tubes. Comparison of results with those from experiments where breath samples were collected sequentially show comparable results, suggesting inherent internal variability in sample collection which is likely to reflect variation in the performance of the sampling pumps within the ReCIVA device.

The human factor analysis showed that the ReCIVA device was preferred by users. Besides analysis of the samples, it is important to acknowledge the acceptance of sampling methodology amongst users. Higher approval of a device might impact on its use within a large clinical trial.

While the aim of the study is to evaluate the different device to find the possible sampler for future clinical trial, some crucial factors shall be taken into consideration. Only ReCIVA may require longer time to take samples, but the exact time consumed may be depended on the setting of sampling program. Besides, ReCIVA may cost more than breath concentrator for sampling because of the mask but it is commercially available. While it is difficult to count the cost of getting another breath concentrator for clinical trial, it is less likely to choose breath concentrator as main sampler for clinical trial.

The fasten and sampling time may be varied from each volunteer. Here in the study, we checked the VOCs detected by each device and their potential for repeated sampling. We cannot deny that unstandardized sampling may change the individual VOCs pattern, but it may be only minor impact to the result. While the samples were taken at the same time, the three devices would take similar samples. Since the comparison were done with the subject mentioned above, the changes from fasten or the sampling time may be erased from database. The comparison was based on the relative value but not the absolute value. As long as we ensure the quality of Sampling in the study, there should be little bias from sampling time and fasten time.

Acknowledging the strengths and limitations of each device is a crucial step in deciding their utility within clinical trials. Whilst the breath concentrator tended to exhibit desirable characteristics in regard to the reliable capture and storage of breath on to TD tubes, it remains largely a bespoke research tool with no current commercially available platform. In comparison the ReCIVA device demonstrates broadly acceptable performance as a breath sampler with the benefit of being able to collect four samples (TD tubes) simultaneously using a commercially available platform. The ability to collect four samples is particularly important for phase 1 discovery studies were broad, multi-platform analysis, may be desirable.

This study has a number of limitations, including a relatively small study population. Furthermore, the study population were relatively young and healthy individuals who represent a different population to that for which breath testing is generally intended. Accordingly, there may be a different profile of VOCs and acceptability in other populations. The analytical strategy chosen for this study was GC-MS, which permits broad untargeted analysis of VOCs. As was the case in the current study, the choice of GC-MS column and experimental method may impact the ability to detect certain VOCs, including short chain fatty acids that have been linked to a number of human disease states. Subsequent studies are therefore planned to explore cross-platform variation in VOC detection with additional comparison between on-line (direct injection) and off-line (TD tube) analysis.

2.1.5 Conclusion

In conclusion this is the first study to examine the characteristics of different devices intended for off-line collection of exhaled breath. Results suggest that for future large-scale clinical trials the ReCIVA breath samplers offers an acceptable performance in terms of VOC collection, retrieval and repeatability coupled with availability and acceptability. Findings however indicate that there remains scope to further optimise the method of breath sample collection and storage.

2.2 Cross platform analysis of volatile organic compounds using selected ion flow tube - , proton reaction transfer- and gas chromatography mass spectrometry

2.2.1 Introduction

VOCs within exhaled breath are typically found in low concentrations, therefore accurate and reliable measurements are required with a high degree of precision and reproducibility. GC-MS has been traditionally viewed as the standard analytical platform in volatilomics research owing to its established chromatographic separation and accurate identification of analytes using spectral libraries[179]. Advances in technology and gas phase ion chemistry has led to the development of alternative and more versatile MS techniques including proton transfer reaction time of flight mass spectrometry (PTR-ToF-MS) and select SIFT-MS [176, 180, 181]. These platforms offer the option of on-line VOC analysis without requiring breath collection and storage. Recently a novel workflow for TD coupled to PTR-MS has been developed allowing for longer breath sample storage and high throughput automated analysis [182].

The increasingly varied array of MS methods used for the analysis of exhaled breath presents a particular challenge as a lack of methodological standardisation and correlation between different techniques means that it is not possible to establish comparability between datasets [3,9]. This study aims to assess the inter- and intra-variability across two widely used MS platforms; SIFT-MS and PTR-ToF-MS, for the detection of target breath VOCs in both online and offline sampling techniques.

2.2.2 Methods

Participants

Exhaled breath was collected from twenty-six healthy adult volunteers. Inclusion criteria were over 18 years of age and the absence of systemic disease. Informed written consent was obtained from each participant (NRES Committee London Camden and Islington; REC number: 14/LO/1136). Exhaled breath collection was performed between 9am and 5pm over a two-week period without restriction of participants' diet or deviation from normal daily activities.

Breath sampling

Breath was collected by asking subjects to exhale into a single use six litre double thickness Nalophan® (Kalle UK Ltd., Witham, UK) sample bag via a sterile mouthpiece[158]. Subjects were typically required to provide two sequential exhalations to fill the sample bag with an adequate volume of breath. Following collection, breath samples were immediately placed in an incubator at 37°C for 30 minutes [158].

Sample analysis

Breath samples were analysed using PTR-ToF-MS (PTR-ToF 1000, Ionicon Analytik GmbH, Innsbruck, Austria) and SIFT-MS (VoiceUltra 200, Syft Technologies, Anatune, Cambridge, UK). A detailed description of these analytical techniques is provided in Section I and elsewhere [183-185]. The analytical strategy was devised in order to assess the inter- (PTR-ToF-MS vs. SIFT-MS) and intra- (PTR-ToF-MS) platform variability of exhaled breath analysis.

Inter-platform variability

Inter-platform variability of breath analysis was assessed through direct-injection of collected breath samples by PTR-ToF-MS and SIFT-MS. This was achieved by securing a three-way connector to sample bags containing breath. Breath was simultaneously analysed by each instrument through one of the two sampling ports. Additional sample preparation was not required.

Optimised instrument conditions for direct-injection analysis of breath VOCs using PTR-ToF-MS to minimise fragmentation and high protonation rates have previously been established [186]. The following reaction conditions were used: drift tube temperature of 110°C, pressure 2.30 mbar, voltage 350 V, with a reduced electric field (E/N) at relatively low value of 84 Td. An instrument inlet flow rate of 200 ml per minute was applied. A single reagent ion, hydronium ion (H_3O^+) was selected.

For SIFT-MS, a sixty second multi-ion monitoring mode was used for targeted analysis of VOCs within breath samples. VOCs of interest were analysed consecutively using four previously optimized methods to prevent unwanted interactions. The SIFT-MS uses a flow rate of 25 ml per minute, with a sampling time of 15 seconds, and temperature control between 10-30 degrees. Simultaneous reactions occurred with three precursor ions, and resultant product ions were assessed for each ionisation. For final analysis, H_3O^+ was selected for all reactions, except for 1-butanol for which NO^+ was used.

Intra-platform variability

Intra-platform variability of breath analysis was assessed through comparison of direct-injection and thermal desorption (TD) PTR-ToF-MS. Analysis by direct-injection was completed as described in the previous section. Breath (500ml) was subsequently aspirated at a rate of <10ml/second from sample bags onto TD tubes containing 200mg Tenax and 100mg Carbograph5 (Markes International Ltd, Llantrisant, UK) using the Easy-VOC® sampler (Markes International Ltd, Llanstrisant, UK). From each sample bag, duplicate samples were collected and stored within TD tubes. Breath sample volume was optimised in a previous study [187]. Prior to use, TD tubes were conditioned at 330°C for 40 minutes in a stream of nitrogen of 50 ml/min (TC-20 tube conditioner, Markes International Ltd). TD tubes that contained breath samples were stored at -80°C prior to analysis.

Breath specimens stored within TD tubes were analysed with PTR-ToF-MS coupled with a TD autosampler (TD100-xr, Markes Ltd, Llantrisant UK) by means of a customized automated

interface, details of which have been published previously [10]. TD tubes were desorbed for 10 minutes at 280°C using a nitrogen flow of 130 sccm in a one-stage desorption process. Subsequently, VOCs were transferred from the TD autosampler to the PTR-ToF-MS through an inlet formed of polyether ether ketone (PEEK) tubing maintained at 110°C, with a flow rate of 130 ml per minute. The instrument is equipped with a commercial Select Reagent Ion (SRI) feature, using a single reagent ion. The H_3O^+ precursor ion was used in the untargeted analysis of duplicate breath samples.

Off-line sample variability

Off-line sample variability was assessed through comparison of thermal desorption tubes analysed on TD-GC-MS and TD-PTR-MS. Sample collection and analysis for TD-PTR-MS was the same as above mentioned. The TD tubes were desorbed using a Markes TD-100 TD unit (Markes International Ltd, Llantrisant, UK) using a two-stage desorption programme, applying a constant flow of helium at 50 ml min^{-1}. In the primary desorption stage, TD tubes were dry purged for 3 min and heated at 280 °C for 10 min. In the secondary desorption stage, the cold trap (U-T12ME-2S, Markes International Ltd, Llantrisant, UK) was rapidly (99 °C min^{-1}) heated from 10 °C to 290 °C. VOCs were transferred from the TD unit to the GC by means of a capillary line heated at 140°C. GC-MS analysis was performed using an Agilent 7890B GC with 5977A MSD (Agilent Technologies Ltd, Santa Clara, USA) equipped with a ZB-642 capillary column (60 m × 0.25 mm ID × 1.40 µm df; Phenomenex Inc., Torrance, USA) with helium used as the carrier gas (1.0 ml min^{-1} flow rate). The GC column temperature programme was set as follows: 4 min at 40 °C, ramp to 100 °C at 5 °C min^{-1} with a 1 min hold, ramp to 110 °C at 5 °C min^{-1} with a 1 min hold, ramp to 200 °C at 5 °C min^{-1} with a 1 min hold and finally ramp to 240 °C at 10 °C min^{-1} with a 4 min hold. The MS transfer line temperature was 240 °C and EI source conditions were 70 eV at 230 °C. Mass acquisition was carried out in the range 20–250 m/z with a rate of approximately 6 scans s^{-1}. The comparison was carried out as concentration on PTR-MS versus peak area count from GC-MS.

VOC selection

Details of VOCs selected for analysis, including their analytical precursor and product ions, are presented in Table 8. VOCs were selected based on either their known relative abundance within the human breath (e.g. acetone, isoprene and ethanol) or their previously documented associated with human disease [184, 188, 189]. Acetone was used as an internal control for each method to ensure the 'quality' of breath analysed.

Table 8 Characteristics of target VOCs within breath analysed by SIFT-MS and PTR-ToF-MS

	Formula	**Molecular weight**	**Precursor ion**		**Product ion (m/z)**
			SIFT-MS	**PTR-ToF-MS**	
Ketone					
Acetone	C_3H_6O	58	H_3O^+	H_3O^+	59
Diene					
Isoprene	C_5H_8	68	H_3O^+	H_3O^+	69
Short chain fatty acids	CH_3COOH	60	H_3O^+	H_3O^+	61
Acetic acid					
Propanoic acid	CH_3CH_2COOH	74	H_3O^+	H_3O^+	75
Butyric acid	C_4H_8O	88	H_3O^+	H_3O^+	89
Pentanoic acid	$C_5H_{10}O_2$	102	H_3O^+	H_3O^+	103
Hexanoic acid	$C_6H_{12}O_2$	116	H_3O^+	H_3O^+	117
Alcohols					
Methanol	CH_3OH	32	H_3O^+	H_3O^+	33
Ethanol	C_2H_5OH	46	H_3O^+	H_3O^+	47
1-butanol	C_4H_9OH	74	NO^+	H_3O^+	73, 75
1-pentanol	$C_5H_{12}OH$	71	H_3O^+	H_3O^+	71

Quality control measures

Quality control checks were employed daily to ensure accuracy of the outputs. A permeation unit (ES 4050P, Eco Scientific, Stroud, Gloucestershire UK) provides a steady flow directly to the PTR-ToF-MS for characterisation of impurities, accuracy, fragmentation rates and resolution of the three precursor ions. Thermal desorption tubes are directly loaded from the permeation unit and

analysed sequentially on the TD-PTR-ToF-MS to ensure the stability of the system. The SIFT-MS requires confirmation of the flow tube pressure at 102.5 ± 2.0 mTorr when calibrant flow is introduced, followed by a series of internal validation checks.

Data analysis

PTR-ToF-MS data was extracted using PTR-MS Viewer version 3.2.8.0 (Ionicon Analytik) and analysed using in-house generated scripts written using R programming language. LabSyft 1.7 Software Suite (Syft Technologies, Anatune, Cambridge, UK) was used to analyse SIFT-MS data. Statistical analysis with Spearman correlation coefficients, simple linear regression and Bland-Altman analysis were performed to validate the relationship and agreement between analytical platforms using SPSS (IBM SPSS version 26.0, Armonk, NY: IBM Corp).

2.2.3 Results

Participant characteristics are presented in Table 9. All 26 participants (15 males; mean age 32.6 ± 3.6 years) reported no significant co-morbidities, and only one was an active smoker.

Table 9 Subject characteristics

	Number of participants (%)
Gender	
Male	11 (42)
Female	15 (58)
Age (mean ± SD)	32.6 ± 3.6
Race	
White	18 (69)
Black	5 (19)
Asian	3 (12)
Smoking status	
Non-smoker	25 (96)
Current smoker	1 (4)
Last food intake	
<1 hour	5 (19)
1-4 hours	14 (54)
>4 hours	7 (27

Inter-platform variability

Details of inter-platform variability in VOC analysis are presented in Table 10 and Figure 10. Abundant breath metabolites such as acetone and isoprene demonstrated a strong positive correlation and linear regression patterns. Acetone (m/z 59) concentrations were comparable at 783 ppbv and 752 ppbv (rho=0.97, p<0.01) when analysed by SIFT-MS and PTR-ToF-MS respectively. SIFT-MS (H_3O^+ precursor ion) recorded higher isoprene concentrations compared to PTR-ToF-MS (192 ppbv vs. 92 ppbv; rho=0.89, p<0.01). Isoprene displayed a strong linear correlation when measured by SIFT-MS using both H_3O^+ and NO^+ precursor ions. For the remaining VOCs that were assessed, both on-line analytical platforms show a high degree of agreement with the majority of values within 95% of the limits of agreement (2 SD) as determined

by Bland-Altman analysis. Pentanoic acid and 1-butanol were not however observed to correlate when analysed by both instruments.

For short chain fatty acids, butyric-, acetic- and propanoic acid concentrations were detected at up to 2-fold higher levels with the SIFT-MS. Hexanoic was detected at 6.5-fold higher concentrations, albeit with a correlation of 0.97. A distinct correlation was observed with hexanoic acid (m/z 117) and acetone (m/z 59) indicating there may be contribution from the acetone dimer [MH+M]. Pentanoic acid presented a statistically significant relationship with a moderate level of correlation (rho=0.41, p=0.04), which may be due to low values below the limits of quantification of PTR-ToF-MS (<1 ppbv) [185]. Repeated calculations utilising only results with concentrations above the limit of quantification (n=6) produced strong correlations consistent with the remainder of the fatty acid group. Pentanoic acid concentrations were 8-fold higher with the SIFT-MS which may reflect background instrument interference. This has been studied and adjusted for in a recent paper investigating atmospheric VOCs [190].

A single compound, 1-butanol, showed a weak correlation (rho=0.32, p=0.12) with NO^+ ionisation. The outcome of SIFT-MS H_3O^+ ionisation produced a concentration higher than expected (60 ppbv) which is likely a contribution from the ^{18}O isotopologue (m/z 73). For inter-platform comparisons of 1-butanol levels, NO^+ data should therefore be considered.

Table 10 Comparisons of online VOCs detection by SIFT-MS and PTR-ToF-MS

	Reagent Ion		SIFT-MS		PTR-ToF-MS		Median ratio	Online SIFT-MS vs. PTR-ToF-MS		Linear regression		
	SIFT-MS	PTR-MS	Median	IQR	Median	IQR	(SIFT: PTR)	Spearman's Rho	P value	R^2	Y intercept	Slope
Acetone	H_3O^+	H_3O^+	783	(553-1305)	752	(646-1602)	1.04	0.97	<0.001	0.98	110	0.76
Isoprene	H_3O^+	H_3O^+	192	(150-232)	92	(78-129)	2.07	0.89	<0.001	0.81	46.8	1.45
Isoprene	NO^+	H_3O^+	174	(135-224)	92	(78-129)	1.88	0.93	<0.001			
Short chain fatty acids												
Acetic acid	H_3O^+	H_3O^+	50	(39–77)	25	(15-34)	2.03	0.85	<0.001	0.74	17.5	1.58
Propanoic acid	H_3O^+	H_3O^+	16	(12- 35)	7.6	(4.25 - 19.04)	2.11	0.90	<0.001	0.89	8.7	1.2
Butyric acid	H_3O^+	H_3O^+	60	(42-75)	39	(30-51)	1.56	0.91	<0.001	0.82	-1.9	1.6
Pentanoic acid	H_3O^+	H_3O^+	6.0	(1.8-6.7)	0.71	(0.50-0.90)	8.38	0.41	0.04	0.08	4.9	1.5
Hexanoic acid	H_3O^+	H_3O^+	5.1	(4.2-9.2)	0.8	(0.5-1.7)	6.41	0.97	<0.001	0.91	3.2	2.3
Alcohols												
Methanol	H_3O^+	H_3O^+	322	(198-464)	194	(105-265)	1.66	0.90	<0.001	0.93	4.2	1.8
Ethanol	H_3O^+	H_3O^+	133	(85-290)	108	(80-231)	1.23	0.98	<0.001	0.97	32.4	0.93
Ethanol	NO^+	H_3O^+	235	(161-442)	108	(80-231)	2.18	0.97	<0.001			
1-butanol	NO^+	H_3O^+	10	(9.3-12)	7.6	(4-19)	1.38	0.32	0.12	0.05	10.2	0.02
1-pentanol	H_3O^+	H_3O^+	16	(14-18)	39	(30-51)	0.41	0.84	<0.001	0.37	6.75	0.24
1-pentanol	NO^+	H_3O^+	46	(37-68)	39	(30-51)	1.20	0.79	<0.001			

Table 11 Comparisons of VOCs detection by online and offline (TD) PTR-ToF-MS.

	Reagent Ion		TD-PTR-MS		Online PTR-MS		Median ratio	PTR TD vs online			Linear regression	
	TD-PTR	Online-PTR	Median	IQR	Median	IQR	(TD : Online)	Spearman's Rho	P value	R^2	Y intercept	Slope
Acetone	H_3O^+	H_3O^+	273	(204-562)	823	(624-1945)	0.33	0.97	<0.001	0.97	1.62	0.49
Isoprene	H_3O^+	H_3O^+	35	(24-44)	95	(75-129)	0.37	0.92	<0.001	0.85	6.25	0.66
Short chain fatty acids												
Acetic acid	H_3O^+	H_3O^+	29	(23-41)	25	(17-36)	1.14	0.68	<0.001	0.55	45	0.64
Propanoic acid	H_3O^+	H_3O^+	5	(4-10)	9	(5-20)	0.59	0.93	<0.001	0.94	6.65	0.57
Butyric acid	H_3O^+	H_3O^+	14	(12-18)	38	(30-52)	0.36	0.96	<0.001	0.92	6.79	0.58
Pentanoic acid	H_3O^+	H_3O^+	1.2	(0.91-1.3)	0.72	(0.56-0.97)	1.58	0.46	0.06	0.04	2.04	0.62
Hexanoic acid	H_3O^+	H_3O^+	1.3	(0.97-2.1)	0.83	(0.61-2.3)	1.60	0.86	<0.001	0.82	1.27	1.35
Alcohols												
Methanol	H_3O^+	H_3O^+	16	(14-18)	203	(117-283)	0.08	0.72	<0.001	0.72	17.7	0.08
Ethanol	H_3O^+	H_3O^+	58	(40-90)	134	(80-283)	0.43	0.97	<0.001	0.98	58.3	0.45
1-butanol	H_3O^+	H_3O^+	5	(4-8)	7	(5-15)	0.67	0.90	<0.001	0.69	6.76	0.56
1-pentanol	H_3O^+	H_3O^+	14	(12-18)	39	(30-52)	0.36	0.96	<0.001	0.92	6.79	0.58

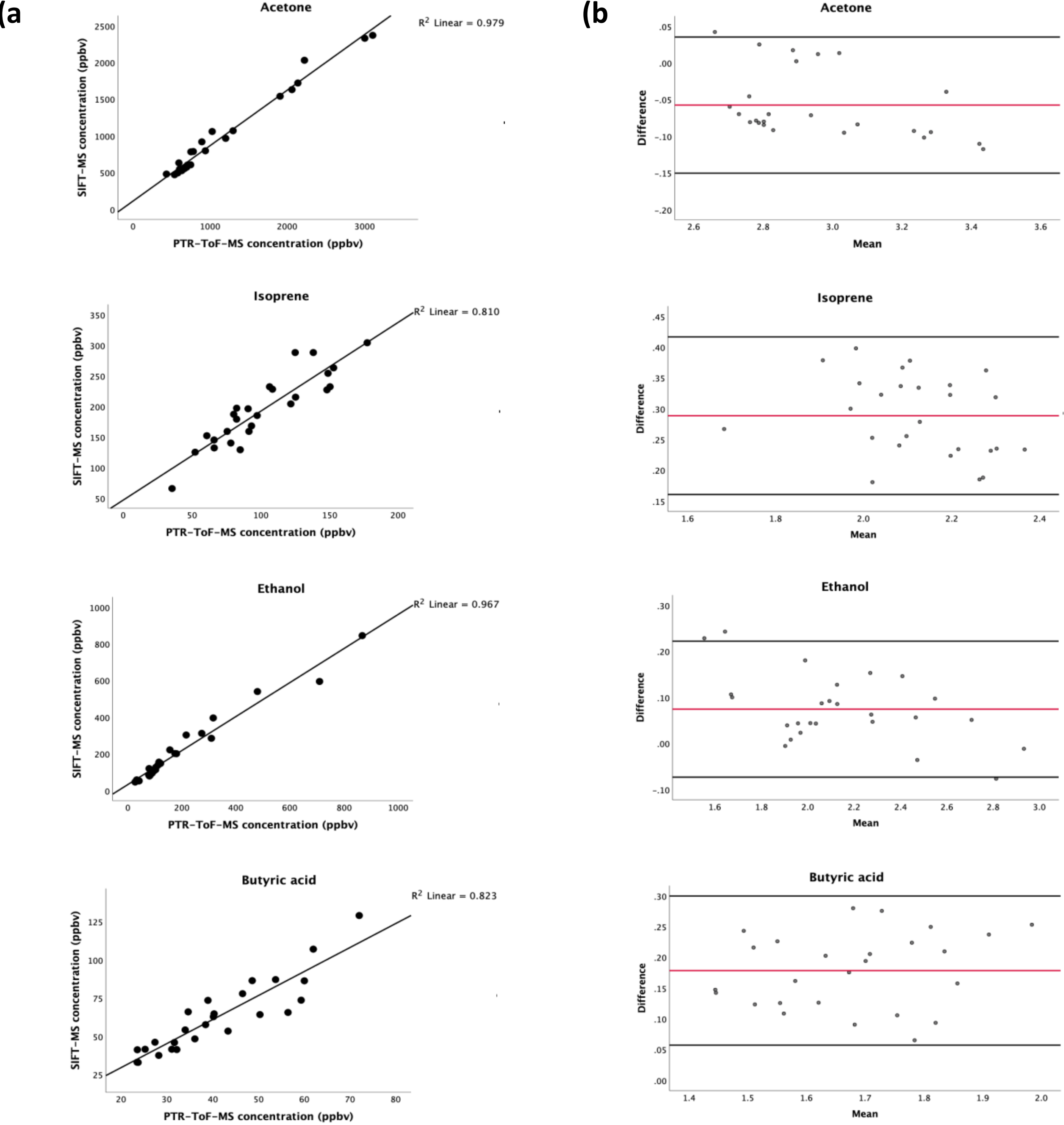


Figure 10 Comparison of select volatile organic compounds analysed using the selected ion flow tube mass spectrometry and proton-transfer reaction time of flight mass spectrometry.

(a) Linear regression plots and (b) Bland-Altman plots constructed with logarithmic transformation (log10) with mean of the difference (red line) and 95% limits of agreement using +/- SD (black lines).

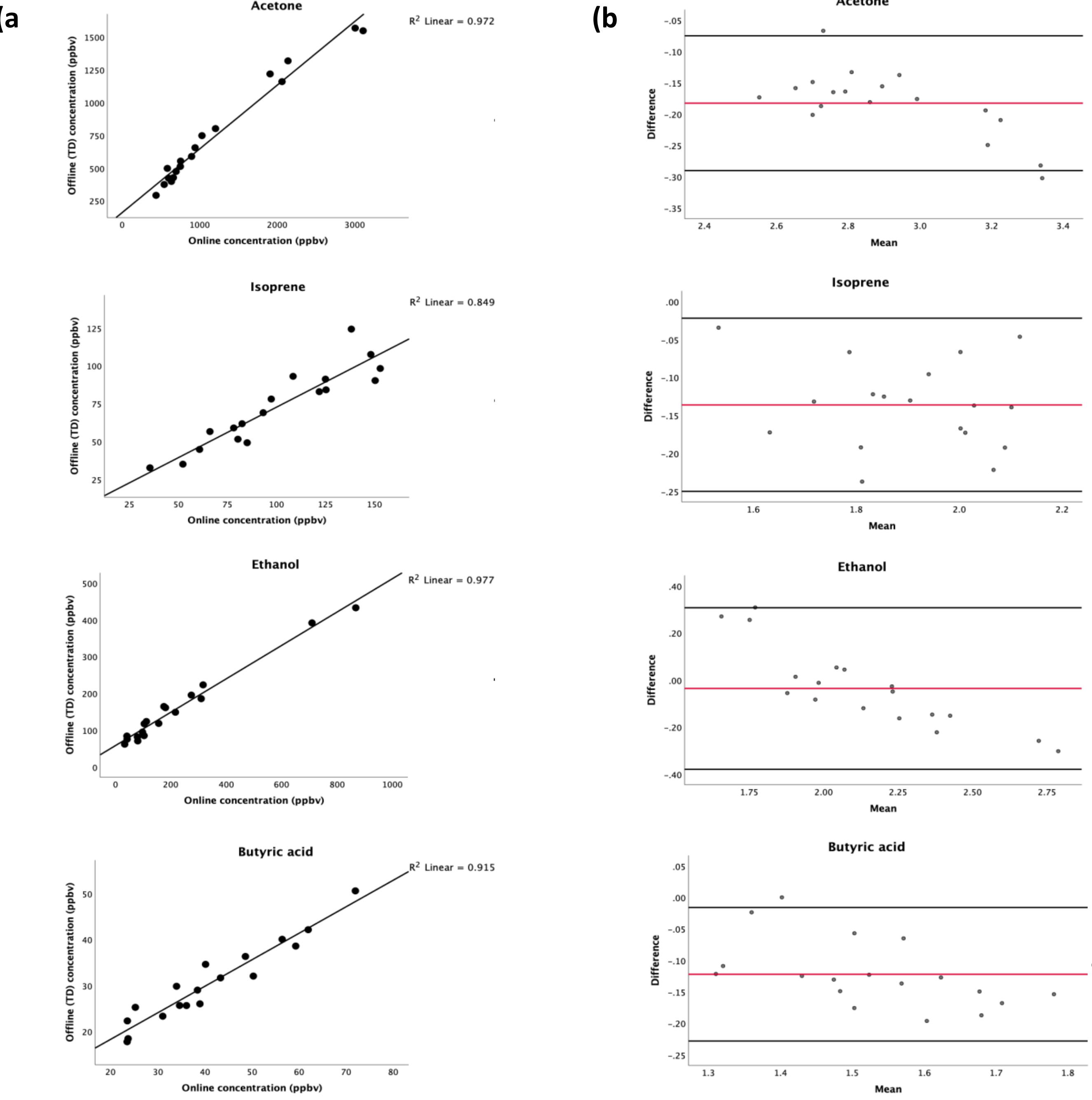


Figure 11 Comparison of select volatile organic compounds analysed with online and offline (thermal desorption) proton transfer reaction time of flight mass spectrometry.

(a) Linear regression plots and (b) Bland-Altman plots constructed with logarithmic transformation (log10) with mean of the difference (red line) and 95% limits of agreement using +/- SD (black lines).

Intra-platform variability

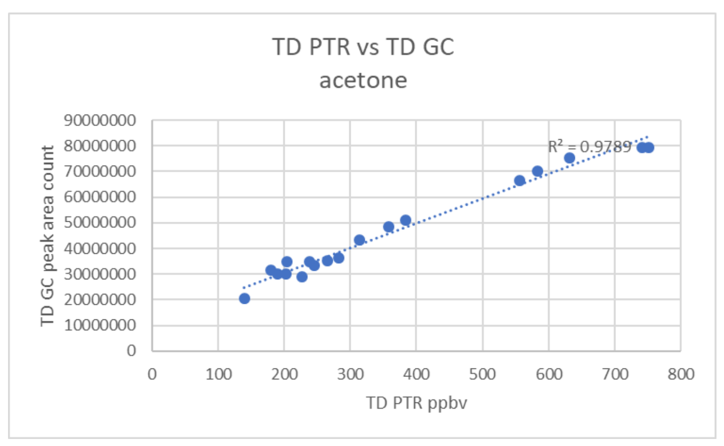
This study assessed the concentration of breath metabolites using direct-injection and offline (TD) techniques on a single platform, PTR-ToF-MS. Concentrations of acetone and isoprene within TD breath samples were lower than those detected by direct-injection (Table 11 and Figure 11). They did however demonstrate strong positive linear correlations (acetone rho=0.97; isoprene rho=0.92; $p<0.01$) and a high degree of agreement with 95% of all values falling within the 95% limits of agreement.

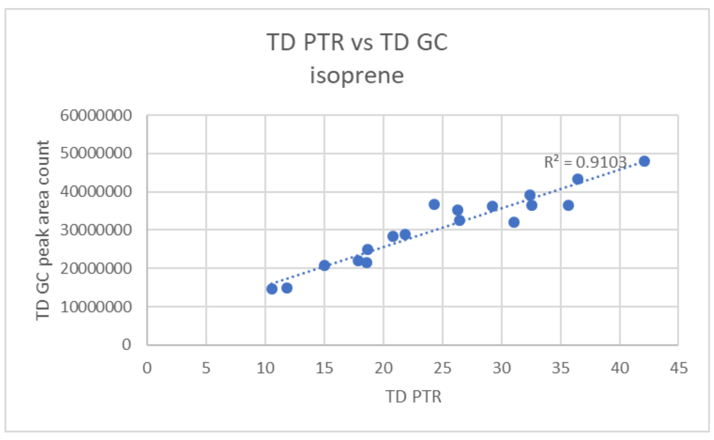
All short chain fatty acids showed a strong positive correlation. Acetic acid concentrations were over 2-fold higher with TD analysis. Whilst hexanoic and pentanoic acids showed a modest correlation, online analysis revealed concentrations below the limit of quantification. Observed higher concentrations of these VOCs within TD tubes may represent a background interference.

A strong positive correlation was observed for ethanol, 1-butanol and 1-pentanol analysed by both methods. Although a similar relationship was observed for methanol, TD analysis resulted in concentrations almost 85% lower than that of online sampling, potentially reflecting sample loss in the transfer process, or ineffective capture onto the sorbent material.

Off-line sample variability

The comparison of off-line samples on GC-MS and PTR-MS was to compare peak area count from GC-MS and concentration from PTR-MS. Because of the different parameters, it cannot be a formal comparison. Although the comparison was based on different parameters, a strong correlation could still be found with three major compounds from breath (Figure 12). The remainder of the compounds included in the study did not demonstrate a good correlation.





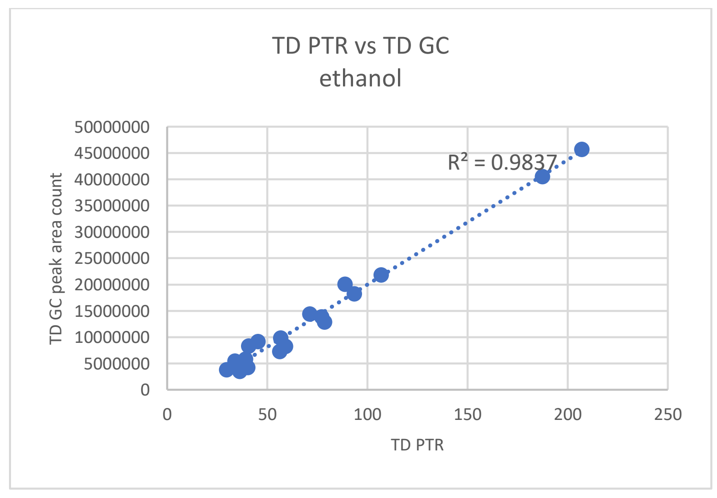


Figure 12 Offline sample variability in abundant compound

2.2.4 Discussion

The emergence of an increasingly broad array of MS techniques capable of analysing VOCs within exhaled breath has made it necessary to establish an understanding of how methods and datasets may be compared and potentially aggregated. Such information is vital as it will support cross-platform validation of clinically significant findings and help to provide insight into potential sources of variability. The principal finding of this study were: (i) a typically strong linear correlation and agreement between PTR-ToF-MS and SIFT-MS methods for abundant and trace compounds analysed by direct-injection;(ii) a tendency for concentrations of VOCs analysed by SIFT-MS to be higher than those measured by PTR-ToF-MS; (iii) an acceptable correlation between direct-injection and TD PTR-ToF-MS methods but with evidence of loss or contamination for selected compounds, and; (iv) an acceptable correlation between TD PTR-ToF-MS and TD GC-MS for abundant compounds.

In the context of clinical breath research, PTR-ToF-MS and SIFT-MS offer real-time and online quantification of trace analytes, eliminating the need for sample preparation and compound separation. These benefits are offset by the requirement for sampling to occur in close proximity to the instrument in order to allow patients to either exhale directly into the instrument or, more frequently, into a sample bag that can be promptly analysed. Previous studies have suggested that the stability of breath stored within sample bags, such as the ones used in this study, is up to 24 hours [158, 191, 192]. Incubation of sample bags at 37°C does however allow breath samples to remain at physiological temperatures prior to analysis. Observed correlations obtained in exhaled VOC concentrations in this study would appear to infer the stability of breath in sample bags at least for short periods of time.

Abundant and trace analytes within exhaled breath were shown to be detected reliably with strong linear correlations and agreement between direct-injection PTR-ToF-MS and SIFT-MS methods. Abundant compounds such as acetone and isoprene demonstrated a strongly positive relationship with correlation coefficients of 0.97 and 0.89, respectively. Other VOCs that have

been linked to human disease also showed strong linear correlations. The narrow limits of agreement determined by Bland-Altman analysis indicates that these platforms produce reliable and related outcomes. Absolute concentrations of VOCs detected by direct-injection tended to be higher for SIFT-MS. Both PTR-ToF-MS and SIFT-MS employ soft chemical ionisation techniques which reduces fragmentation rates and improves detection of analytes. However, the PTR drift tube uses reagent ions with higher kinetic energy resulting in greater fragmentation than SIFT-MS. In comparison, SIFT-MS utilises a flow tube which carries trace gases via the helium flow at a slower speed. The majority of correlated chemical ionisation reactions for both platforms used the H_3O^+ reagent ion, resulting in protonated product ions and water adducts (MH^+, MH^+H_2O). The PTR-ToF-MS method achieves higher resolution with the ToF analyser, and uses single ionisation compared to the SIFT-MS which utilises three simultaneous precursor ions [184]. In addition, absolute concentrations are calculated using rate coefficients which are different for both platforms. SIFT-MS uses a coefficient from the in-built kinetics library that is unique to each identified compound. Kinetic rate constants have been derived from accurate calculations using ion velocities in the flow tube and known flow rates of the carrier and sample gases. In contrast, PTR-ToF-MS uses a single approximate coefficient [184]. This may contribute to the observed variability in VOC concentrations. The influence of instrument artefact may also be a factor [184]. Finally, it is noteworthy that both H_3O^+ and NO^+ reagent ions of the SIFT-MS produced comparable correlations with the H_3O^+ PTR-ToF-MS data. This is important as studies will continue to utilise both H_3O^+ and NO^+ precursor ions for tentative VOC detection, especially for compounds where product ions share the same or similar mass-to-charge ratio.

In spite of its advantages, direct-injection analysis of breath is not always feasible especially where large scale population testing is desired. Also, direct injection cannot be used for GC-MS identification of separated compounds. TD analysis has therefore been used in multi-centre trials and presents a reliable and transferrable method of breath collection and storage. As shown in the current study when compared to direct sample injection, TD can result in variable VOC concentrations. Acetone, isoprene, and butyric acid concentrations were 25-30% lower, whilst methanol was almost 85% lower in TD samples. Conversely, elevated concentrations of acetic-,

hexanoic- and propanoic acids were detected within TD tube samples. Explanation for these variations is likely to be multifactorial and compound dependent. Transfer of breath from sample bags to the TD tubes risks sample degradation through either the loss of highly volatile compounds or contamination from transfer lines. Further losses and contamination may also occur within the TD tubes. This study has identified potential background signals from the TD tubes even after robust conditioning protocols. This is a recognised phenomenon that may have important implications for specific breath biomarkers such as short chain fatty acids. Whilst this study was not designed to optimise these processes it does offer valuable insight into their importance. Introduction of standardized and quality-controlled processes that are intended to mitigate these effects at least partly should be the focus of further investigation.

Although it was not a formal comparison, there were good correlations among the three abundant compounds from breath specimens. One possible application for this result is quality control of the sample. Acetone is a major compound from breath and its concentration is between 1ppm to 1250 ppm [193]. Although we cannot rule out the possibility of specific disease to decrease the formation of acetone or accelerate the consumption of acetone, it is still reasonable to use acetone as quality control for samples. It would be easy to transfer peak area count to absolute concentration since they are highly correlated. Because it is not possible to directly visualise samples collected to TD tube, an excellent quality control is necessary to minimise any bias. By calculating the concentration of acetone on GC, we can therefore potentially exclude any poor-quality samples.

The current study has a number of acknowledged limitations. This study relied on the analysis of breath samples from human subjects. Samples were therefore representative of those collected within clinical trials. By its nature, human breath constitutes an overly complex biological matrix of VOCs, ambient gases (e.g., N_2, O_2, CO_2 and Ar) and water vapour. Whilst highly desirable, especially for studies such as this one, it has not yet been possible to establish a synthetic breath standard within which the concentrations of its components are known and controlled. Furthermore, without creating accurate external calibration, such as calibration curves, it is not

possible to ascertain the precise concentrations of VOCs within breath samples. The identification of compounds by both PTR-ToF-MS and SIFT-MS presents a challenge in the discrimination of isobaric compounds. A more comprehensive platform such as the GC-MS may be implemented in the workflow, however, this technique requires sample preparation and compound separation with a longer analytical time frame and more detailed data extraction processes.

Breath analysis continues to be an attractive strategy for the non-invasive detection of disease. The analytical value and performance of the SIFT-MS and PTR-ToF-MS in breath biomarker research has been established. This study provides evidence of a strong linear correlation between these instruments for the quantification of the majority of assessed VOCs. Bland-Altman analysis concluded a high degree of agreement between both platforms with many of the outcomes falling with in the 95% limits of agreement. These findings give new insight into cross platform analysis of VOCs that is relevant to both ongoing and future clinical trials.

2.3 Analysis of room air volatile organic compound concentrations within different hospital environments

2.3.1 Introduction

An inherent challenge to breath analysis is the influence of VOC contaminants that may arise from a multitude of exogenous and endogenous sources. Previous chapters have considered contaminants arising from both the sampling manifold and analytical platform. VOCs inhaled within ambient air can also be an exogenous source of contamination. Previous studies have identified variable levels of selected VOCs within different clinical environments. Those VOC that occurs in prominent levels within inhaled ambient air may not be suitable as disease biomarkers. A common of example is ethanol that is found in high concentrations within hospitals, presumably due to the ubiquitous use of alcohol hand sanitized and other cleaning products.

Previous authors have commented on the possible implications of background contamination of breath samples. Schubert et al., previously suggested that high concentrations of inhaled VOCs may disproportionally affect their alveolar concentration gradients, in turn influencing their excretion within breath [194]. The same authors suggested that only VOCs whose inspiratory levels were <5% of their expiratory levels could be confidently analysed within exhaled breath [195]. Implementation of this recommendation would however preclude to the analysis of a substantial proportion of VOCs that are considered to be of clinical interest. Alternative strategies that have been proposed to mitigate the effects of ambient VOC contaminants, include background subtraction and pre-test inhalation of 'scrubbed' air. Whilst background subtraction is straightforward to perform it may not fully account for the effects of inspiratory and expiratory VOC levels as these factors so do exhibit collinearity in relation to blood concentrations. Breathing scrubbed air, although a more meticulous method of mitigating the effects of inspiratory VOC levels, is laborious and impractical for use in routine clinical practice. It is also recognised that scrubbed air and the apparatus and masks used to deliver it to patients contain background levels of many VOCs that can themselves contribute to sample contamination (Please refer to Chapter 2.1).

To better understand the potential effects of ambient VOCs on breath biomarkers it is first necessary to assesses their levels within the environments in which sampling is proposed. In environments in which VOCs of interest are found to be in high concentrations, it may be necessary to either avoid sampling in that location or to modify sampling methods to reduce the potential for contamination.

The current study therefore intends to examine longitudinal variation in the levels of common VOCs in different hospital environments in which future clinical studies are planned.

2.3.2 Methods

Sampling locations

Based on clinical locations proposed for sampling in future clinical trials three hospital environments were chosen in which to measure ambient VOC levels: (i) outpatient clinics; (ii) endoscopy department waiting areas, and; (iii) operating theatre waiting areas. The order in which samples were collected from each location was random. Ambient air samples were also collected from the VOC laboratory as this is the location where all samples are processed and analysed.

Sample collection

Ambient room air samples (500ml) were collected on to TD tubes (Markes Ltd, Llantrisant UK) from each location using a manual pump (EasyVOC, Markes Ltd, Llantrisant UK).

Samples were collected from each location both in the morning (09:00 to 10:00 am) and afternoon (16:00 to 17:00 pm) over a period of sequential working days (Monday to Friday) during a two-week period.

Sample analysis

Samples were analysed by TD-PTR-ToF-MS (PTR-ToF 1000, Ionicon Analytik GmbH, Innsbruck, Austria) which is equipped with a commercial Select Reagent Ion (SRI) feature, using a single reagent ion H_3O^+. The PTR-ToF-MS instrument is coupled with a TD autosampler (TD100-xr, Markes Ltd, Llantrisant UK) by means of a customized interface. Analysis was performed with a one-stage desorption method. TD tubes were desorbed for 10 minutes at 280°C using 130 sccm of nitrogen. Subsequently, VOCs were transferred to the PTR-ToF-MS through an inlet formed of polyether ether ketone (PEEK) tubing maintained at 110°C, with a flow rate of 130 sccm. Drift tube analysis conditions were equivalent to the direct sampling analysis. An untargeted approach was taken for analyte quantification.

Target VOCs

Target VOCs analysed within collected samples are listed in Table 12. Compound selection was based on either their known abundance in exhaled breath (e.g. acetone, isoprene) and/or their prior identification as potential cancer biomarkers [133, 166, 181, 196-199].

Table 12 Target VOCs within exhaled breath quantified by the PTR-MS.

Chemical compound	Formula	Molecular weight	Selected precursor ion
Ketone			
Acetone	C_3H_6O	58.080	H_3O^+
Diene			
Isoprene	C_5H_8	68.12	NO^+
Alcohol			
Ethanol	C_2H_5OH	46.07	H_3O^+
Methanol	CH_3OH	32.04	H_3O^+
Butanol	C_4H_9OH	74.12	H_3O^+
Short chain fatty acid			
Acetic acid	CH_3COOH	60.052	H_3O^+
Propanoic acid	CH_3CH_2COOH	74.08	H_3O^+
Butyric acid	C_4H_8O	88.11	H_3O^+
Pentanoic acid	$C_5H_{10}O_2$	102.13	H_3O^+
Hexanoic acid	$C_6H_{12}O_2$	116.1583	H_3O^+
Aromatic			
Phenol	C_6H_5OH	94.11	NO^+

Data analysis

TD-PTR-ToF-MS data was extracted using PTR-MS Viewer version 3.2.8.0 (Ionicon Analytik) and analysed using in-house generated scripts written using R programming language. All the results are further presented as median and their range. Further calculations of the variation were done with the dataset.

Variation was determined as follows:

$$variation = \frac{standard\ deviation}{average}$$

2.3.3 Results

Comparison between clinical environments

Variation in ambient VOC levels between different clinical environments are presented in Table 13. VOC levels were broadly equivalent within each of the three clinical environments. Acetone, methanol, ethanol, and acetic acid were found in each clinical environment at levels greater than ten parts per billion (ppb).

Comparison of morning and afternoon room air samples

Variation in VOCs levels detected in the morning and afternoon within each clinical environment are presented in Tables 14-17. In each of the three clinical environments ethanol levels were higher in the afternoon. Remaining VOCs demonstrated no significant variation between morning and afternoon samples. Ethanol levels also demonstrated the highest variability of all VOCs both in morning and afternoon samples.

Comparison between room air and breath samples

A comparison between ambient and exhaled breath levels of target VOCs is presented in Table 13. Of the VOCs assessed ambient acetone, isoprene, and butyric acid levels with were ~10% of levels observed in exhaled breath (Table 18).

Table 13 Variation in ambient VOCs levels

	Outpatient clinic		Operating theatre waiting room		Endoscopy waiting room		VOC laboratory		Exhaled breath	
	Median	**Range**	**Median**	**Range**	**Median**	**Range**	**Median**	**Range**	**Median**	**Range**
Acetone	18.76	12.50-27.14	18.73	12.89-30.63	18.36	12.36-6-.13	22.59	16.00-35.84	265.09	130.28-787.49
Isoprene	2.68	1.79-4.01	2.37	1.56-4.37	2.51	1.47-4.74	2.90	2.07-4.25	26.03	10.26-42.44
Methanol	13.27	10.67-21.65	14.54	10.48-22.23	14.76	11.08-20.61	17.35	12.88-18.21	15.07	7.51-38.52
Ethanol	93.60	44.52-247.45	82.70	37.63-373.60	76.14	34.55-300.36	43.79	27.75-69.07	59.27	28.35-208.21
Acetic acid	20.98	12.63-32.46	22.83	13.17-29.90	22.86	11.83-33.77	22.65	12.26-36.81	29.01	12.95-54.70
Propanoic acid	4.98	2.22-8.71	4.67	3.06-8.56	4.95	2.13-6.39	4.81	2.37-7.04	5.31	2.89-31.25
Butyric acid	1.58	1.03-1.94	1.47	0.97-2.13	1.52	0.88-2.07	1.54	1.02-2.11	14.05	8.24-25.04
Pentanoic acid	0.84	0.57-1.21	0.95	0.52-1.28	0.95	0.62-1.49	0.88	0.56-1.19	1.15	0.55-3.98
Hexanoic acid	1.22	0.57-1.91	1.22	0.88-2.51	1.13	0.79-2.73	1.31	1.01-1.84	1.29	0.65-7.82
Phenol	5.56	1.69-7.22	5.69	1.86-9.14	6.31	1.99-9.08	5.12	2.22-8.67	1.21	0.49-3.89

Table 14 Variation in outpatient clinic ambient VOCs levels

	Morning			Afternoon		
	Median	**Range**	**Variation**	**Median**	**Range**	**Variation**
Acetone	18.31	12.50 - 22.84	0.15	18.76	13.42 - 27.14	0.2
Isoprene	2.76	1.79 - 4.01	0.26	2.62	1.82 - 3.79	0.23
Methanol	13.17	10.79 - 21.65	0.22	14.73	10.67 - 20.80	0.21
Ethanol	91.67	46.94 - 230.10	0.5	115.94	44.52 - 247.45	0.52
Acetic acid	21.58	12.63 - 32.46	0.28	20.02	15.20 - 31.73	0.24
Propanoic acid	4.70	2.22 - 5.59	0.25	5.01	2.71 - 8.71	0.3
Butyric acid	1.58	1.03 - 1.94	0.24	1.56	1.19 - 1.81	0.15
Pentanoic acid	0.84	0.57 - 1.21	0.28	0.84	0.71 - 1.18	0.16
Hexanoic acid	0.89	0.57 - 1.71	0.4	1.23	0.90 - 1.91	0.2
Phenol	5.46	1.69 - 7.22	0.35	5.56	2.77 - 7.06	0.22

Table 15 Variation in operating theatre waiting room ambient VOCs levels

	Morning			Afternoon		
	Median	**Range**	**Variation**	**Median**	**Range**	**Variation**
Acetone	18.47	13.56 - 25.03	0.2	20.28	12.89 - 30.63	0.25
Isoprene	2.47	1.56 - 4.37	0.29	2.37	1.89 - 3.10	0.13
Methanol	14.73	12.47 - 22.23	0.18	14.40	10.48 - 19.27	0.18
Ethanol	60.39	37.63 - 292.57	0.84	98.43	69.37 - 373.60	0.66
Acetic acid	24.64	13.17 - 29.90	0.21	19.13	16.40 - 26.84	0.17
Propanoic acid	5.05	3.06 - 6.65	0.22	4.14	3.47 - 8.56	0.3
Butyric acid	1.52	0.97 - 2.13	0.21	1.40	1.16 - 1.82	0.15
Pentanoic acid	0.99	0.52 - 1.28	0.22	0.87	0.67 - 1.24	0.19
Hexanoic acid	1.22	0.88 - 1.85	0.25	1.19	0.97 - 2.51	0.34
Phenol	5.93	1.86 - 7.36	0.27	5.15	3.48 - 9.14	0.29

Table 16 *Variation in endoscopy waiting room ambient VOCs levels*

	Morning			Afternoon		
	Median	**Range**	**Variation**	**Median**	**Range**	**Variation**
Acetone	17.57	12.51 - 60.13	0.59	18.51	12.36 - 23.69	0.18
Isoprene	2.45	1.66 - 4.74	0.3	2.76	1.47 - 4.00	0.29
Methanol	15.42	11.08 - 20.59	0.17	14.49	11.83 - 20.61	0.17
Ethanol	61.44	36.36 - 229.05	0.69	85.71	34.55 - 300.36	0.79
Acetic acid	22.32	15.33 - 31.75	0.18	24.30	11.83 - 33.77	0.23
Propanoic acid	4.62	2.54 - 6.07	0.24	4.98	2.13 - 6.39	0.24
Butyric acid	1.48	0.98 - 1.62	0.13	1.58	0.88 - 2.07	0.23
Pentanoic acid	0.94	0.67 - 1.20	0.17	0.97	0.62 - 1.49	0.27
Hexanoic acid	1.09	0.79 - 1.80	0.22	1.30	0.99 - 2.73	0.37
Phenol	6.02	2.38 - 8.61	0.35	6.70	1.99 - 9.08	0.31

Table 17 Variation in VOC laboratory ambient VOCs levels

	Morning			Afternoon		
	Median	**Range**	**Variation**	**Median**	**Range**	**Variation**
Acetone	19.90	16.33-24.27	0.17	22.83	16.00-35.84	0.28
Isoprene	2.40	2.07-3.40	0.21	3.12	2.34-4.25	0.19
Methanol	16.72	12.88-17.69	0.12	17.90	15.29-18.21	0.07
Ethanol	38.19	27.75-46.48	0.18	50.43	32.48-69.07	0.23
Acetic acid	20.69	12.26-36.81	0.41	22.65	15.28-29.78	0.20
Propanoic acid	4.35	2.37-6.94	0.39	4.81	3.93-7.04	0.20
Butyric acid	1.38	1.02-2.11	0.28	1.60	1.24-1.91	0.14
Pentanoic acid	0.89	0.56-1.19	0.25	0.88	0.67-1.10	0.16
Hexanoic acid	1.27	1.01-1.55	0.15	1.33	1.06-1.84	0.18
Phenol	4.21	2.22-7.75	0.46	5.78	4.14-8.67	0.25

Table 18 Comparison of room and exhaled breath VOCs (expressed as fold change)

	Outpatient clinic	Operating theatre	Endoscopy	VOC laboratory	Exhaled breath
Acetone	0.07	0.07	0.07	0.09	1.00
Isoprene	0.10	0.09	0.10	0.11	1.00
Methanol	0.88	0.96	0.98	1.15	1.00
Ethanol	1.58	1.40	1.28	0.74	1.00
Acetic acid	0.72	0.79	0.79	0.78	1.00
Propanoic acid	0.94	0.88	0.93	0.91	1.00
Butyric acid	0.11	0.10	0.11	0.11	1.00
Pentanoic acid	0.73	0.83	0.83	0.77	1.00
Hexanoic acid	0.95	0.95	0.88	1.02	1.00
Phenol	4.60	4.70	5.21	4.23	1.00

2.3.4 Discussion

This study provides a description of the longitudinal variation in selected VOCs within three hospital environments relevant to planned clinical studies. Findings indicate that concentrations of ambient VOCs are broadly equivalent within each clinical environment. Furthermore, for the majority of VOCs, concentrations do not appear to vary significantly both within the same and consecutive days.

In keeping with previous studies ethanol was found in high ambient levels within each hospital environment [200]. Ethanol showed high variation not only within a single day but between consecutive day of the. Ethanol is a common constituent within hand sanitisers and cleaning productions found within hospitals. The observation of higher ethanol concentrations in afternoon samples would appear to suggest that levels accumulate during the day, possibly due to the presence of patients within the clinical space and concurrent use of ethanol-based cleaning products. Findings would suggest that ethanol remains an unsuitable disease biomarker within breath samples due to the potential influence of ambient contamination.

In accordance with previously developed quality control pathways developed within our group, levels of acetone and isoprene within breath samples collected in TD-tubes should exceed 50 ppb and 5 ppb, respectively. With the exception of a single acetone level recorded from the endoscopy department waiting area, ambient levels of acetone and isoprene did not exceed the recommended lower levels for breath. It would appear sensible if breath sample collected. It can also be evidence that the quality control of breath sample analysed on TD-PTR-MS.

Remaining VOCs typically occurred at trace (<10ppb) levels and exhibited minimal to moderate variability over repeated measurements.

To assess the impact of room air on breath samples, further comparison on human breath and ambient VOC levels was performed. Although the method for sample collection was different in the two experiments, trace compounds including alcohols, phenols and long chain fatty acid showed minor difference from room air and breath samples. The initial reason fatty acid was selected as target compound was that fatty acid was found to be a potential breath biomarker of CRC. Although these trace compounds showed similar concentration from healthy control and room air, the relative elevation of these compounds from cancer patients could still be evidence of the presence of disease.

In this study, EasyVOC was used for room air sampling. Unlike ReCIVA, it is a manual pump that aspirates ambient air directly on to a TD tube, therefore reducing the potential for instrument-related contamination. Previous studies have reported the EasyVOC to be a simple and effective method for collection of cell culture headspace [201, 202]. The current study would appear to corroborate those studies, endorsing the EasyVOC as a reliable gas sampling device.

Limitations of this study include the relatively short duration of sampling and the limited selection of VOCs assessed. It was, however, the intention of this study to provide an overview of VOC variability in clinical environments pertinent to planned clinical studies. It was important to determine whether targeted VOCs varied significantly between clinical locations and over time.

Such information was intended to inform the design of future studies. This study did not seek to directly compare ambient VOC levels with those detected within the exhaled breath of patients and staff present in each of the clinical areas. That approach would have provided greater insight into the relative contribution of ambient VOCs to levels detected within breath.

In conclusion this study provides useful information relating to the concentrations of clinically relevant ambient VOC levels within different hospital environments. With the exception of ethanol, ambient VOC levels did not appear to vary significantly both over time and between different areas. There does not therefore appear to be a necessity to restrict the timing or place of breath sampling in future clinical studies. It would however seem prudent to ensure that ambient VOC levels are regularly measured in areas where breath sampling occurs in order to provide a reference point to assess the potential influence of environmental contamination.

Section III

Clinical application

3. CLINICAL STUDIES INVESTIGATING THE ASSOCIATION BETWEEN VOCs AND COLORECTAL CANCER

Previous studies have reported an association between VOCs and colorectal cancers, which offers the potential of a future non-invasive test to both detect and monitor this disease (Section 1.4 Systemic review of studies reporting volatile organic compounds as potential biomarkers of colorectal cancer).

This section intends to further explore the unique VOC signature of colorectal cancer in human subjects. Studies will examine the exhaled breath profile of colorectal cancer patients both before and after surgery, do decern both the effects of the tumour and its treatment of VOC levels. Additional studies will seek to explore the origin of cancer specific VOCs through the analysis of in-situ colonic tumours at the time of resection. VOC analysis will be conducted using a robust quality-controlled methodology.

3.1 Analysis of colorectal tumour headspace for the assessment of cancer specific VOC release

3.1.1 Introduction

Previous studies that have investigated an association between VOCs and CRC have typically analysed samples that are reflective of the systemic metabolic milieu, including: exhaled breath, blood, urine, stool, or cell culture [120, 136, 138, 181, 202, 203]. The mechanisms by which VOCs are produced in cancer tissues and released from the body remains however largely unknown. It is presumed that the tumour is the source of aberrant VOC production in CRC and that VOCs travel via the bloodstream to organs such as the lungs and kidneys where they may be partially excreted. Although some of the compounds may be changed at the excretion system, still some similarities or some clues related to the newly formed compounds could be found.

In section 1.4, we reviewed the possible biomarkers for CRC found in different source of samples. Only three VOCs were found to be potential biomarker in different research (Table 19). Further investigation for the similarities from previous study may be beneficial on the narrowing down the area of research. besides of the compounds found from different research, some other compounds including aldehyde, short chain fatty acid were also mentioned in previous study.

Table 19 potential biomarker from previous study (repeated mentioned in different research)

Name	Source of sample
Acetone	Breath, urine, faeces
Ethanol	Breath, urine
Cyclohexane	Tissue headspace, breath
Propanal	*Breath, from current laboratory*

Researchers who have examined VOCs release in association with gastroesophageal cancers found cancer specific VOCs were significantly higher in the headspace above tumour biopsy samples compared to healthy tissue [202, 204]. Those same compounds were also found to be enriched within the upper gastrointestinal luminal air of cancer patients compared to controls. VOC levels detected within the gastrointestinal tract adjacent to the tumour were significantly

higher than recorded within exhaled breath, suggesting that the tumour was the source of their production.

Previous studies were focusing on the faeces which may release VOCs to its headspace [120, 137]. While the VOCs from faeces may be the mixture of stool, colonic tissue, and the microbes. It is difficult to purify the VOCs from lower GI tract. to simplify the study, another simple but feasible source of VOCs is considered. No previous studies have examined VOC release in direct association with CRC tissue, but in vitro cell culture experiments showed differences between CRC and normal cell lines [201, 202]. Whilst it is possible to sample endoluminal air during endoscopy, such as described by Adam *et al*. for upper gastrointestinal tumours, this method is less effective during colonoscopy [204]. In pilot studies we found that the luminal gas sampling was limited by the length of the colonoscope as well as limited space within the colon. These factors made it more common for the sample line to become blocked rendering sampling ineffective. An alternative approach is to analyse the local headspace of colonic tissue at the time of surgical resection. An added advantage of this method is that tumour and normal mucosa can be identified and analysed separately.

The purpose of this study was to investigate the productions of VOCs from CRC tissue through direct in-situ analysis of tumour headspace and adjacent 'normal' mucosa. By determining the relative abundance of VOCs associated with both tumour and healthy tissue it may be possible to better understand their source of origin.

3.1.2 Methods

Patients:

Patients undergoing elective resection of CRC (colectomy or anterior resection) at Imperial College Healthcare NHS Trust during the period July 2019 to January 2020 were eligible for inclusion in this study. Patients were required to provide written informed consent prior to enrolment. Patients without CRC and patients who were unable to provide informed written consent were excluded.

Sample collection:

Study participation did not alter patients routine perioperative (anaesthetic or surgical) care. On resection of the surgical specimen, it was immediately (within 10 minutes) prepared for sampling either in the operating theatre or histopathology department. The bowel was opened longitudinally away from the tumour to exposure the tumour (Figure 22). Any residual stool or blood/mucus was wiped from the mucosal surface with gauze or tissue.

Headspace gas above the tumour and adjacent macroscopically normal colonic mucosal was collected using a handheld manual precision pump (Easy-VOC, Markes International, Llantirsant, UK) and polystyrene sample pot (Figure 22(B)). Headspace was refreshed with entrained room air through a separate opening to the sample port. Aspirated headspace gas was collected and stored within thermal desorption tubes (Markes International, Llantirsant, UK). In total 500ml headspace was sampled from both the tumour and normal mucosa. A separate sample of ambient room air (500ml) was also acquired at the time of sampling.

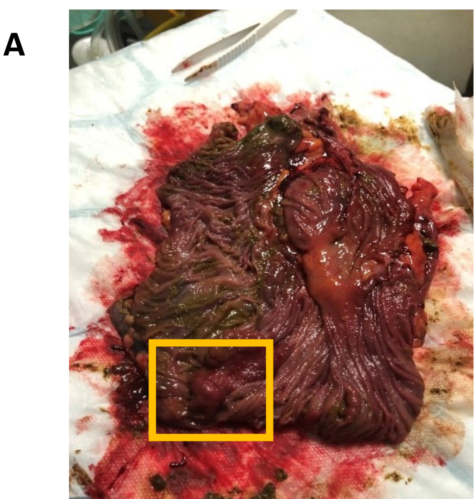
A  B

Figure 13 Collection of cancer headspace samples

Tumour headspace sampling method. The resected specimen (large intestine) was open for tumour exposure. (A) The tumour is located near caecum. (B) Headspace gas collection method.

Sample analysis

Sample analysis was conducted using PTR-ToF-MS. Samples were analysed by PTR-ToF-MS (Ionicon Analytik GmbH, Innsbruck, Austria) equipped with a commercial Select Reagent Ion (SRI) feature, using one of three reagent ions H_3O^+, NO^+, O_2^+. The PTR-ToF-MS instrument was coupled with a TD autosampler (TD100-xr, Markes Ltd, Llantrisant UK) by means of a customized interface [205]. Analytical conditions and functioning of the TD-PTR-ToF-MS were previously described. Analysis was performed with a one-stage desorption method; tubes were desorbed for 10 minutes at 280°C using 130 sccm of nitrogen; VOCs were then transferred to the PTR-ToF-MS through an inlet made by polyether ether ketone (PEEK) tubing kept at 110°C, with a flow rate of 130 sccm. Drift tube analysis conditions were the following: temperature 110 °C, pressure 2.30 mbar, and voltage 350 V, resulting in an E/N of 84 Td. The optimization of the PTR-ToF-MS reaction

conditions was carried out as previously detailed [206]. Similar instrumental parameters were applied for the direct sampling analysis, with an inlet flow rate of 200 sccm.

Quality control of the sample analysis

All the sampling tubes were conditioned with TC20® (Markes International) on the sampling day or one day prior to sample taking. Tubes were sealed (airtight) before and after sampling with brass caps. Where possible samples were analysed on the day of collection. If sample analysis on the day of sampling was not possible, TD tubes were stored at -80°C.

PTR-ToF-MS Quality control. Two types of quality checks were performed every day to ensure a constant analytical performance of the PTR-ToF-MS [206]. The first test is the *instrument quality control* evaluating instrument reproducibility with the three ionization modes against four parameters: impurities, fragmentation, mass resolution and accuracy. The permeation unit was connected to the PTR-ToF-MS inlet through a PEEK tube union connection. Measurement was carried out for five minutes for each ionization mode. Accuracy was evaluated through quantification of a benzene certified standard permeation tube (Kin-Tek Analytical Inc., La Marque TX). The second quality control (*Thermal desorption quality control*) method evaluated the recovery of VOCs from desorbed TD tube. TD tubes were loaded with standard mix from the permeation unit with a previously described method [205]. The standard mix was composed by benzene (63 ppb), phenol (90 ppb), butyric acid (20 ppb), pentanoic acid (5 ppb), hexanoic acid (5 ppb), decanal (4 ppb) and butanal (5 ppb), generated by a permeation unit (ES 4050P, Eco Scientific, Gloucestershire UK), that provided a constant VOC concentration, maintained at 30°C and with a nitrogen flow of 0.9 L/minute. Concentration of VOCs calculated from TD tubes were compared to the concentration calculated from the permeation unit through direct measurement obtained during the first quality control test. Details of tube loading, average recovery, and relative standard deviation (RSD) are summarized in Table 20. The first quality check evaluated the performance of the PTR-ToF-MS instrument, while the second assessed the performance of the TD unit coupled with the PTR-ToF-MS.

Table 20 Data of the standard quality control

Loading concentrations from the permeation unit, mean recovery % and RSD % for each of the standards analysed with the three primary ions. All the data referred to the period of the study.

		Benzene	Butanal	Decanal	Phenol	Butyric Acid	Pentanoic Acid	Hexanoic Acid
Loaded (ppb)		62.5	5	4	90	20	5	5
PTR H_3O^+	Mean Recovery %	93.7			101.2	87.7	95.5	90.7
	RSD %	9.8			8.8	12.9	14.9	17.0
PTR NO^+	Mean Recovery %	80.5	87.6	111.2	88.7			
	RSD %	12.4	10.4	30.8	13.1			
PTR O_2^+	Mean Recovery %	81.2			91.1			
	RSD %	9.4			12.2			

Data processing

TD-PTR-MS data were extracted using PTR-MS Viewer version 3.2.8.0 (Ionicon Analytik, Austria) and analysed using in-house generated scripts written with R programming language [207]. VOC ions included in the script were selected using a resolution of 0.01 mass ratio, to obtain a better separation of ions with similar molecular weight. The separation was obtained manually dividing the peaks obtained from the time-of-flight analysis, based on the instrument mass resolution power given by the high resolution of the time-of-flight analyser. The R script generated in-house gave a direct quantitative output, as ppb concentration for each VOC.

The identification of the compounds found with PTR analysis presented some limitations. This technique does not have any chromatographic separation; therefore, separation of isobars and isomers result impossible. Identification of VOC obtained with PTR analysis is only tentative since there could be the overlap of different product ions. The process of tentative compound identification was based solely on the obtained product ion. Each VOC react with a reagent ion

(H_3O^+) to produce a characteristic product ion. Based on the knowledge of the specific chemical reaction of VOCs and reagent ions, we tried to identify the compounds obtained by our analysis.

Statistical analysis

The concentrations of reported VOCs are presented as median (interquartile range). Comparison of VOC concentration between tumour and normal mucosa was performed using the Wilcoxon signed-rank test. IBM SPSS version 26 was used on the data analysis.

3.1.3 Results

In total twenty tumour headspace samples were collected from nineteen patients. One patient (No. 11) was diagnosed with synchronous tumours of the transverse and ascending colon. Details of tumour characteristics are presented in Table 21.

Table 21 Characteristic of sample collection

Patient	Tumour location	Pathological T stage	Pathological stage	Max tumour diameter (mm)	Tumour Differentiation	Neoadjuvant therapy
1	Ascending colon	2	I	40	Moderate	No
2	Caecum	4	IV	60	Poor	Yes
3	Ascending colon	3	II	33	Poor	No
4	Rectum	2	III	49	Moderate	No
5	Sigmoid colon	4	III	45	Well/moderate	No
6	Sigmoid colon	1	I	14	Moderate	No
7	Rectum	3	II	-	Well/moderate	Yes
8	Rectum	3	II	21	Poor	Yes
9	Sigmoid colon	3	II	36	Moderate	No
10	Sigmoid colon	0*	I	-	Poor	No
11	transverse	2	III	28	Moderate	No
11	Ascending colon	1	I	25	Moderate	No
12	Sigmoid colon	3	II	49	Moderate	No
13	Rectum	4	II	44	Moderate	No
14	transverse	4	III	55	Moderate	No
16	Sigmoid colon	3	II	60	Poor	No
17	Sigmoid colon	0*	III	2.5**	Poor	No
18	Rectum	3	IV	-	Poor	Yes
19	Rectum	3	III	30	Moderate	No
20	Caecum	1	III	24	Moderate	No

* Patients # 10 and # 17 had endoscopic resection of the polyp tumour and were diagnosed T1 lesion for the resected polyps. There was no residual tumour found with the resected colonic specimen. Therefore, the pathological T stage was 0.

** Piecemeal resection was done with the polyp cancer and 2.5 mm was the biggest diameter could be measured with the specimen.

Variation in VOC levels between room air, tumour and normal mucosal headspace are presented in Table 22. Because all the tubes were analysed with H_3O^+, some possible cancer related compounds including aldehydes cannot be optimally detected. While it is a small sample size exam, we could not include all the possible cancer compounds in the study. Besides, we cannot make sure if taking two samples from the same site would give the same result. Therefore, some compounds such as propanal, which was found to be possible CRC biomarker compound was discarded in the study.

Results show clear differences between room air, headspace of cancer or normal mucosa. This can be seen as confirmation that that colonic mucosa can produce and release VOCs directly. Further comparison was done between cancer and normal mucosa only. Four compounds, hexanoic acid, benzene, 2-butanone, and propofol showed significant difference. Among these four compounds, the concentration was low (mostly less than 1 ppbv) with hexanoic acid and propofol. Although the difference was not clear between cancer and normal mucosa, it may be related to the number of the samples.

From Figure 14, selected compounds showed a difference between room air, cancer, and normal mucosa. Typically, the concentration of these selected compounds from room air was <10 ppb. Higher concentration of these compounds was noted from cancer headspace than normal mucosa tissue, but they both release some compounds directly. Although in the small sample size study cannot reveal significant difference, further sample collection may be beneficial for better understanding about the VOCs colonic mucosa may release.

For those VOCs that were identified to be increased from colonic mucosa, further comparison was made based on the depth of tumour invasion (T stage, Figure 25). Although significant difference between cancer and normal mucosa could only be seen in four compounds only, concentration of some compounds (ions) increased with T stage. Similarly, the number of samples collected may be the key factor.

The effects of pathological stage, max diameter of tumour, and tumour differentiation on VOC levels were also reviewed, but no strong correlation could be obtained with the above characteristic.

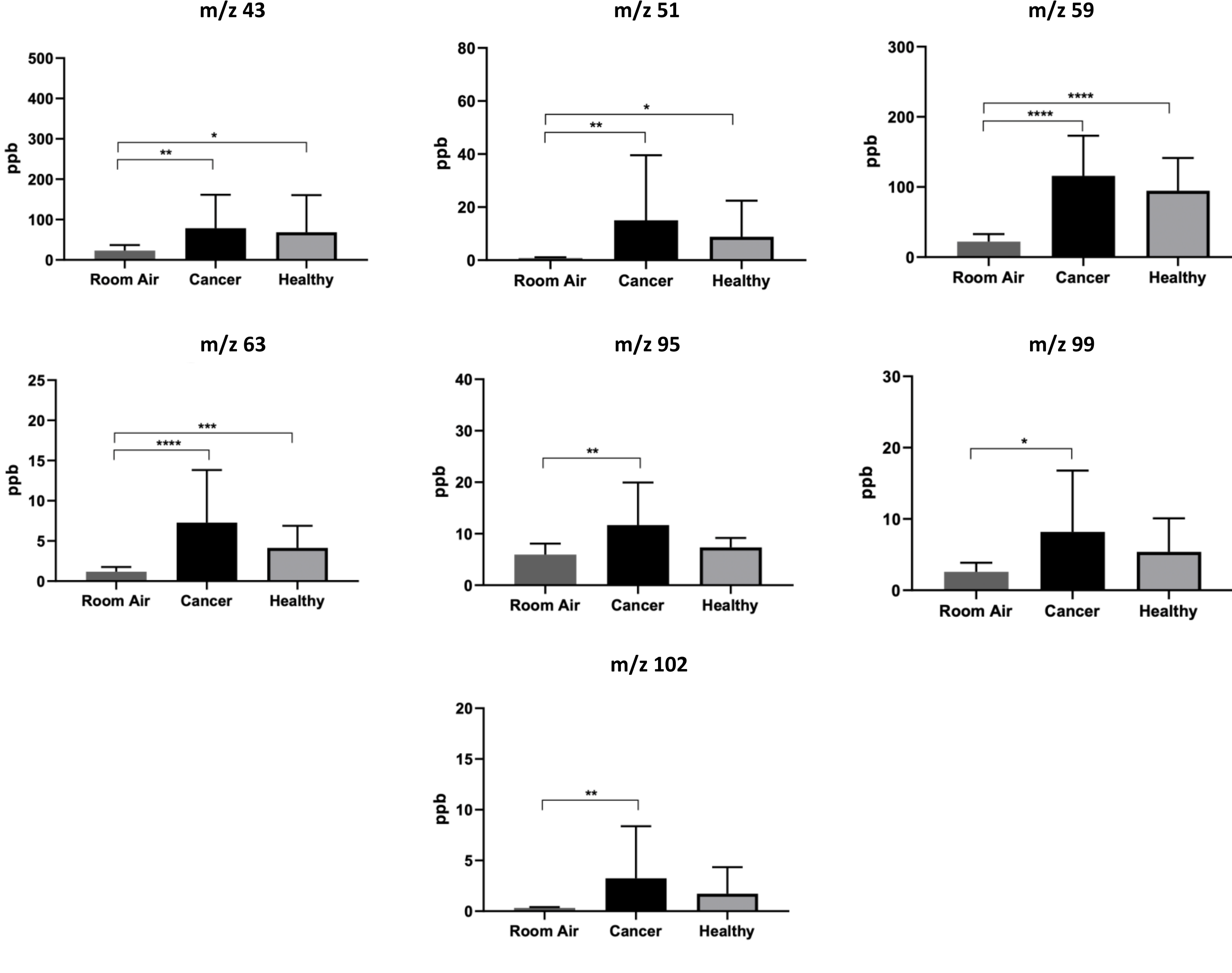
Table 22 comparison of known compounds or uncertain ions from PTR-MS

	Room air		Cancer		Control			
	median	IQR	Median	IQR	median	IQR	P1	P2
Acetone	19.7	7.1	123.9	59.7	87.3	58.7	0.000	0.228
Acetic acid	20.3	4.2	25.4	8.0	20.3	4.2	0.386	0.386
Propanoic acid	5.7	2.2	6.2	2.1	5.7	2.2	0.493	0.493
Butyric acid	1.6	0.8	2.3	0.8	2.0	1.0	0.637	0.783
Pentanoic acid	1.0	0.6	1.4	0.8	1.1	0.6	0.392	0.627
Hexanoic acid	0.8	0.4	1.3	0.7	1.1	0.4	0.001	0.004
Ethanol	48.6	256.2	151.5	231.5	102.9	183.0	0.468	0.524
Benzene	2.0	1.1	3.0	3.6	3.0	1.5	0.013	0.025
Phenol	6.3	2.8	8.0	4.9	7.4	2.8	0.035	0.063
Methyl phenol	1.1	0.4	1.7	1.8	1.5	1.1	0.027	0.149
Ethyl phenol	0.7	0.5	1.0	0.6	0.9	0.6	0.017	0.064
Xylene	4.4	25.7	6.8	36.2	5.0	29.7	0.279	0.245
Acetaldehyde	46.9	21.4	50.5	31.2	50.1	18.0	0.663	0.997
Acetonitrile	1.7	1.0	2.9	1.9	2.5	0.8	0.172	0.494
2-butanone	4.0	2.0	5.5	2.2	4.5	0.9	0.016	0.016
Hexanone	3.5	1.6	4.4	171.6	4.6	152.9	0.042	0.237
Isoprene	3.0	1.2	3.9	1.5	3.0	1.5	0.046	0.325
Methanol	16.4	8.2	19.4	13.8	18.7	13.2	0.618	0.775
Pinene (alpha)	1.0	0.9	1.6	1.9	1.3	1.2	0.089	0.173
Propofol	0.2	0.1	0.3	0.4	0.2	0.3	0.007	0.038
Toluene	3.0	1.5	4.0	1.9	3.4	2.2	0.132	0.102
m.z. 27	5.6	8.5	10.0	18.2	9.9	8.2	0.213	0.676
m.z. 43	17.8	12.9	34.9	38.9	32.0	28.3	0.18	0.763
m.z. 51	0.9	0.4	1.5	18.4	1.4	14.1	0.038	0.251
m.z. 63	1.0	0.7	3.5	6.5	3.7	4.4	0.01	0.074
m.z. 80	2.7	2.7	3.7	4.4	3.0	2.28	0.232	0.502
m.z. 99	2.26	1.4	3.4	7.5	3.0	5.7	0.014	0.156
m.z. 102	0.3	0.1	0.6	3.3	0.6	3.1	0.033	0.181
m.z. 147	0.3	0.3	0.6	41.3	0.4	29.0	0.039	0.252
m.z. 149	1.4	2.1	3.3	456.3	2.6	261.4	0.034	0.413

P1 = P value from comparison of all three sample sources.

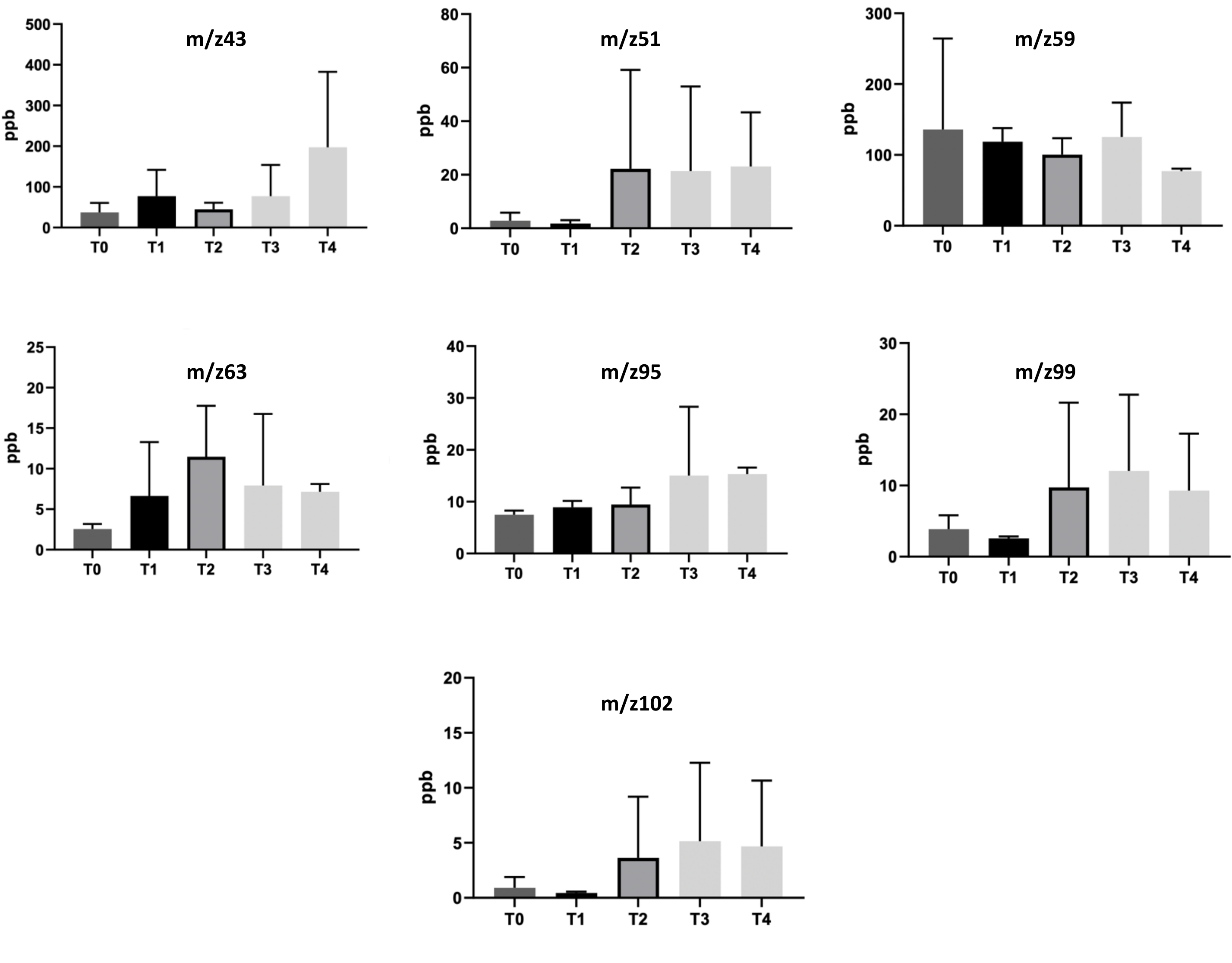
P2 = P value from comparison of cancer and normal mucosa

Figure 14 VOC elevated in colorectal cancer tissue headspace



(m/z 59 is acetone, m/z 95 is phenol)

Figure 15 *VOC product ions association with colorectal cancer T Stage*



(m/z 59 is acetone, m/z 95 is phenol)

3.1.4 Discussion

Previous studies examining the VOC profile of CRC have typically relied upon the analysis of VOCs within breath, stool, and urine. Whilst such studies have provided a link between CRC and aberrant VOC production, they provide limited insight in the origin of target VOCs. It is presumed that the tumour itself produced VOC that are released into the systemic circulation before being subsequently excreted within breath and urine.

The current study provides new evidence that colorectal tumours can be the source of increased VOC production. Consistent with the findings of previous studies that have examined VOCs in relation to CRC, no single VOC was identified as being diagnostic of CRC. A panel of both named and un-named VOCs were detected in higher levels above CRC tissue compared to adjacent 'healthy' mucosa. The three compounds, hexanoic acid, benzene, 2-butanone were observed in higher concentrations above colonic tissue. A fourth compounds, propofol, was also seen in higher levels in CRC tissue, however as this compound is of known exogenous origin it cannot be considered a disease biomarker. Other VOCs including phenol, ethyl phenol acetone, isoprene, and hexanone, and ions (51, 63, 99, 102, 147, and 149) also showed non-significant differences between cancer and normal mucosa. These compounds could be produced by the colonic mucosa and released directly into bowel lumen.

The four compounds showed significant difference were hexanoic acid, benzene, 2-butanone, and propofol. This is the first time we found benzene as a possible biomarker from literature review (section 1.4). However, the median concentration of benzene was the same in caner and normal mucosa. Besides, the concentration was around 3.0 ppbv. Further investigation with bigger sample size may reveal the character of benzene as potential CRC biomarker. Propofol is one of the widely used anaesthetic agent. For almost all the colorectal surgery, intubation for ventilation is necessary and most of the anaesthetic agents would be given through intravenous fluid and endotracheal tube. Propofol may be used initially for rapid intubation. The dose of propofol may be varied based on the body weight. The dose would be slightly different from each patient. However, the metabolism of propofol would be fast and the half-life is 40 minutes [208].

Considering its rapid metabolism and the low concentration in the study, it is difficult to tell if propofol is a potential biomarker. 2-butanone was also found as a potential biomarker in another tissue headspace study (128). Further investigation is recommended.

Previous studies have suggested hexanoic acid and phenol are possible biomarkers of upper gastrointestinal adenocarcinoma [109]. Hexanoic acid was also found higher concentration in vivo samples [204]. While CRC and other gastrointestinal adenocarcinomas effect various parts of the intestinal tract, they share histological similarities. It is therefore reasonable to hypothesise that adenocarcinoma may produce hexanoic acid which can be either released to intestinal lumen or blood stream then exhaled via breath. Acetone and isoprene are the two common compounds from human breath related to energy and cholesterol metabolism [209, 210]. Here in the study, acetone showed significant difference from cancer or normal mucosa versus room air. Also, the trend for acetone from cancer mucosa and normal mucosa showed possible difference. If the number of sample size is bigger, the result may be that acetone is different from cancer versus normal mucosa, too. It can be related to the energy metabolism and normally cancer tissue may consume higher energy.

For selected VOCs there appeared to be an association between pathological T-stage and VOC concentration. Although sample size was limited, preventing formal statistical analysis, this is the first evidence that VOC production in CRC may be linked to tumour size/burden. This finding has important implications of future VOC bases tests. Many researchers working in this field are seeking to develop non-invasive methods for the detection of early CRC. Currently, ineffective referral pathways and diagnostic strategies mean that many patients with CRC are diagnosed with advance disease that is associated with worse overall survival. For many putative VOC biomarkers identified in the current study, lowest concentrations were observed in T1 (early stage) disease. It will therefore be necessary for future diagnostic tests to have suitable precision to detect those VOCs at low concentrations.

Despite this study demonstrating an association between specific VOCs and CRC, this underlying mechanisms for their local release from tumours is incompletely understood. Whilst the tumour itself is presumed to be a source of VOCs; the associated intestinal microbiome may also contribute. Previous studies have reported that bacteria within the colon can produce short chain fatty acids and other VOCs [211-213]. Further studies are now needed to examine the role of the onco-microbial axis and VOC production in CRC.

In the sample collection, there were two patients found to have no residual tumour after endoscopic resection. The pathological report for the resected malignant polyp showed resection margin involved by malignant tumour. Further operation for bowel resection was done and samples were collected at that time. They had polypectomy in within one month prior to the operation and the scar was still viable in the specimen. Besides, they had colonoscopy with scar tattooing to confirm the location of previous polyp. Current study revealed some possible risk factor for lymph node metastasis in T1 CRC. Left-sided CRC, submucosal invasion depth, poor histological grade, lymphatic invasion, vascular invasion, and tumour budding 2/3 were significant factors for the prediction of lymph node metastasis in T1 CRC cases [214]. Here in the study, they both had complete surgery because of the risk factor mentioned above, and one of them was found lymph node metastasis after bowel resection. while we did the sample collection, we could not know the lymph node status or if there were still residual malignant cell on the specimen. However, the results still showed significant different from tumour and normal tissue.

Table 23 *Remove patient 10 and 17*

	Room air		Cancer		Control			
	median	**IQR**	**Median**	**IQR**	**median**	**IQR**	**P1**	**P2**
Acetone	19.7	7.1	123.9	59.7	87.3	58.7	0.000	0.228
Phenol	6.0	3.6	8.2	6.9	7.7	3.6	0.045	0.079

Table 23 showed the comparison of two compounds, acetone, and phenol if we take out patient 10 and 17. Only minor changes were seen from the result. Although the two patients had no residual tumour in the specimen, the addition of the samples may not result in significant changes.

The other issue is about the bowel prepare. Bowels prepare was thought to be an important factor to reduce surgical site infection but this concept changes gradually with further research [215]. For most of the colorectal surgery, mechanical bowel prepare may not be a necessary step. In the study, we cannot totally ignore if mechanical bowel prepare may change the VOCs. While bowel prepare is to wash the bowel route with massive fluid, it is possible to change the microbes in the bowel. However, in the experiment, we did not record if these patients received any medicine for bowel prepare.

It may be interesting to collect some in situ CRC headspace samples. While we collect samples from the tumour headspace in the theatre, it is difficult to acquire samples from in situ (Tis) lesion. For Tis lesion, local excision is the proper treatment since there is no lymph node metastasis. However, understanding if mucosal lesion can produce these specific VOCs may be an interesting topic. If Tis lesion may not give these compounds as T1 or advanced lesion, it would be a great finding for future clinical practice how we can decide if polypectomy only or biopsy for uncertain colorectal lesion.

This study has several acknowledge limitations. Firstly, the method of assessing tumour headspace VOCs relied upon ex-vivo analysis of the colon specimen. The loss of blood supply to the specimen and the effect of ischemia may have influenced VOC release. In addition, the method of headspace isolation and sampling could have been influenced by contaminants within room air. Cross-platform analysis of samples with GC-MS would have allowed improved VOC identification. Finally, local VOC release from colorectal tumours were not corelated to VOC levels with exhaled breath or urine. Such analysis would help to provide a more robust association between local production and systemic release of VOC biomarkers in relation to CRC.

3.1.5 Conclusion

In conclusion this study supports the hypothesis that CRC is associated with aberrant VOC production which would be released into bowel lumen directly. Although it is still far away from using these compounds for disease detection, some similarity (2-butanone) was found from different study. The next phase of studies within this thesis will examine VOC within the exhaled breath of patients before, during and after treatment for CRC.

3.2 Colorectal neoplasm and breath analysis

3.2.1 Introduction

Colorectal cancer (CRC) is one of the leading cancers worldwide in its incidence and related death rate [216]. Early detection and diagnosis are still the key to treatment for both primary and relapsed disease [217-219]. While 5-year disease free survival rate can be over 90% for stage I CRC, it drops to less than 30% in stage IV disease. Inadequacies in disease detection and post-treatment surveillance contribute to the poor survival of this disease [220-222].

An ideal biomarker(s) for detection and monitoring of CRC would be accurate, affordable, and acceptable to patients. Existing CRC markers such as FOBT and CEA, do not meet all these criteria and as such there is room for further research in this area. Any new biomarker does not need to totally replace CEA or FOBT but can be a co-biomarker which improve both the sensitivity and specificity on detecting the early or asymptomatic tumour.

The field of metabolomics attracted considerable interest in recent years because of its potential to identify novel biomarkers of human disease, including cancer. Diseases, like cancer, are known to influence cell metabolism, leading to the aberrant production of individual metabolic by products [223-225]. Previous research has proved the potential of metabolites as biomarker [168, 226-228]. One well-known example is the accumulation of ammonia in liver cirrhosis. Such patients cannot detoxify ammonia in liver and the amount of ammonia in blood stream therefore raise.

Metabolites may be excreted via the lungs, kidney, or sweat after circulated in the blood stream. Some other route may include saliva, stool, and flatus. Considering the accessibility of each sample source, urine, stool, blood, or breath are more favoured. Breath sampling is particularly appealing as it is entirely non-invasive and therefore acceptable to patients [229]. Whilst disease

detection through breath 'odour' analysis is not a new idea, the field has undergone rapid growth in the last two decades because of technological advances. Application of mass spectrometry-based techniques has permitted research to detect and quantify volatile metabolites within breath at ultra-low concentrations. Numerous previous studies have reported the associated between breath metabolite and human disease [97, 109, 136, 196].

In the field of CRC, VOCs from breath, or headspace of blood, urine, or stool has been tested as possible biomarker [133, 136, 138, 166, 181, 196, 203, 230]. The limited scale of these studies coupled with their heterogenous methodologies has meant that VOC analysis for CRC has not made it to routing clinical practice. Some of the most compelling work in this field was conducted by Altomare and colleagues [167] who reported the VOC profile of the same patient before and after potentially curative therapy for CRC. This work critically showed how cancer specific VOCs may diminish after removal of the cancer. Due to the scale and duration of the research, there were no patients in this series with recurrent cancer mentioned. This would have been necessary to assess the hypothesis that VOCs can be used to detect disease recurrence. Therefore, further analysis with the recurrent disease from different institute is necessary to understand the potential of VOCs as biomarker.

The different sampling methods and analytical platforms used in CRC VC research has meant that it is not possible to draw firm conclusions from the literature [176]. Furthermore, most studies have lacked robust quality control processes. Since there is no single idea platform on sample analysis, the possible solution is to find the linkage between each platform and see if these data could be transferred in each MS.

Experimental work conducted thus far within this thesis acknowledges the importance of understanding methodological variability in the analysis of breath VOCs. Effort was therefore made to establish standard operating procedures for breath collection and analysis.

Utilising the knowledge and standardized methodology developed in the preceding section, the aim of the next chapter is to explore the exhaled VOC profile of CRC before and after treatment.

3.2.2 Material and methods:

3.2.2.1 Sample collection

3.2.2.1.1 Patients:

Cancer patients: patients with biopsy proven colorectal cancer in addition to colorectal cancer patients who had undergone potentially curative surgery were recruited for this study.

Preoperative patients were further subdivided into the following groups:

(i) treatment naïve (before neoadjuvant chemotherapy or surgery)

(ii) post neoadjuvant chemotherapy

(iii) post palliative chemotherapy.

Postoperative patients were classified as either:

(i) disease free post primary resection

(ii) disease free post salvage resection (of recurrent disease);

(iii) metastatic disease post resection of the primary tumour,

(iv) disease recurrence.

Where patients were sampled longitudinally, at different time points during their treatment (e.g., pre-, and post-operative). A minimum of six weeks was observed between any therapeutic intervention (e.g., surgery or chemotherapy) and breath sampling.

Healthy controls: patient undergoing routine colonoscopy, found to have a healthy colon and rectum were recruited as controls.

3.2.2.1.2 Patient identification

Patient with CRC were identified from the local multidisciplinary team meeting and surgical schedules for patients attending for elective resection of a CRC.

Healthy control subjects were initially sampled at clinic while they came for their appointment. These patients may receive stool occult blood test as in bowel cancer screening program with positive result. Further endoscopy would be arranged for these patients to rule out colonic lesions. We will follow the result of colonoscopy to confirm if they are healthy control. If cancer or high-grade dysplasia polyp found, the sample would be removed from healthy control.

Inclusion criteria

(i) Patients aged from 18 to 90 years old
(ii) Patients with histology proven CRC either before or after treatment (curative or palliative)
(iii) Control subjects with macroscopically normal colon on colonoscopy (with or without biopsy confirmation)
(iv) Provision of informed written consent

Exclusion criteria

(i) Patients with synchronous cancer(s) at another body site
(ii) Patients with active infection, or who had received antibiotic therapy within 6 weeks of recruitment
(iii) Patients who declined or who were unable to provide informed written consent

Ethical approval

This study was approved by the National Research Ethics Service (East of England - Essex Committee, REC Ref: 17/EE01/12) and the BCSP research advisory committee (ID: 189). The study was registered (NCT03699163) and reported according to the Standards for Reporting of Diagnostic Accuracy studies (STARD 2015) guidelines. This is part of the COlorectal cancer and BReath Analysis (COBRA) study. In the beginning, COBRA study was divided into two parts, to detect CRC, and to do the longitudinal follow up for CRC. More detailed regarding to the selection of patients would be describe later in the method part. All patients were required to provide

informed written consent prior to enrolment and provision of breath samples. For patients who were recruited at separate times points either before and after treatment, a new consent to participate was sought at each study time point. Although the COBRA study was divided into two parts, the samples taken were done together.

All the participants were free to withdraw from the research at any time point.

3.2.2.1.3 Sampling method and procedure:

Patients were recruited from Imperial NHS trust including St Mary's Hospital and Charing Cross Hospital. Patient recruitment and sample collections occurred between December 2016 and September 2019.

Breath samples were collected in one of three locations:

(i) Theatre admission unit
(ii) Endoscopy admission unit
(iii) Surgical outpatient clinic

All breath samples were collected by GPL or a trained NIHR Clinical Research Nurse.

Prior to breath sampling subjects were asked to sit in a rested state for 20 minutes. During this period informed written consent was obtain and the process for beath sample collection using the ReCIVA device was carefully explained to patients. In the study, we recorded if the fasten time and if they have bowel prepare prior to sample collection.

Patients provided a breath sample in a seated position using the optimised and standardized ReCIVA setting [187].

Prior to use, TD tubes were conditioned at 330°C for 40 minutes in a stream of nitrogen of 50 ml/min (TC-20 tube conditioner, Markes International Ltd). TD tubes were typically conditioned the day before sample collection. At each breath sampling event four TD tubes were collected simultaneously.

Subjects were asked to breathe tidally into the single use facemask connected to the ReCIVA device for a period of approximately five minutes. Breath (500ml) was collected on to each of the four TD tube at a rate of 125mls per minute (equivalent to 250 ml/min for each of the ReCIVA's two pumps that provide gas independently to two TD tubes).

3.2.2.1.4 Transport and storage of the sample

Immediately after completion of breath sampling TD tubes were removed from the ReCIVA device and sealed with brass caps.

Capped tubes were placed in a zip tie bag before transport to the central laboratory at St Mary's Hospital. Where possible, same day courier services or a 'hopper bus' (between Charing Cross and St Mary's Hospital's) was used to transport samples.

After arriving at the central laboratory at St Mary's Hospital TD tubes were logged prior to same day analysis. In situations where it was not possible to analyse samples on the day or receipt, they were stored at -80°C. Breath samples are known to be stable from up to 30 days when stored within TD tubes at -80°C (unpublished data Hanna group).

3.2.2.2 Analysis of the samples – GC-MS

Breath samples were analysed by GC–MS Agilent 7890B with 5977A MSD (Agilent Technologies, Cheshire, UK) coupled with a Markes TD-100 (Markes Ltd, Llantrisant UK) thermal desorption unit. For GC-MS analysis, TD tubes were desorbed using a constant flow of helium onto an electrically cooled, sorbent-packed focusing trap (U-T12ME-2S, Markes International Ltd, Llantrisant, UK),

secondary desorption was then conducted by ballistic heating, leading to quick sample introduction into the GC-MS. A detailed description of TD-GCMS analytical conditions is provided elsewhere [231]. Quality control was performed daily for the entire duration of the project on GC-MS. A detailed description is given below.

3.2.2.3 Quality control of the mass spectrometry

To ensure the quality of sample analysis by mass spectrometry, a daily quality control evaluation of the of the mass spectrometer was performed.

3.2.2.3.1 GC-MS quality control

Five TD tubes were loaded daily with VOCs standard mixture from the permeation unit as previously explained and then analysed. Retention time and peak area were evaluated to assure constancy of the instrument response. A relative standard deviation below 5% signified that the quality control was passed and that analysis of clinical samples could proceed. To quantify VOCs, calibration curves were performed with the use of the permeation unit, modulating the flow to load TD tube at different concentrations.

3.2.2.3.2 Calibration curve for acetone quantification with GC-MS

An 11-point calibration curve was prepared by serial dilution of acetone pure standard solution. Dilutions of the stock solution were prepared in methanol, obtaining concentrations ranging from 0 ppb (pure methanol) to 2 ppm. 1 ul of each stock solution was spiked on a TD tube using a Calibration Solution Loading Rig (CSLR, Markes International Ltd, Llantrisant, UK) and analysed by GC-MS. The curve was constructed plotting peak area and theoretical concentrations. All peak areas obtained from samples analysed in this study were plotted on the calibration curve, obtaining corresponding concentration for each signal.

3.2.2.4 Data processing – TD-GC-MS

Raw data obtained from the TD-GC-MS analysis were extracted using MassHunter software (Agilent Technologies, UK). Despite data analysis is possible using the same software, a new pipeline for GC data analysis was developed at Imperial College London to improve deconvolution of chromatographic peaks and annotation of VOC [232]. Following this data analysis path, data were analysed using MassHub software and the Global Natural Product Social Molecular Networking (GNPS). The software operates an intra/inter-sample mass drift correction, noise filtering and baseline correction, peak alignment, with detection, integration, and deconvolution. The output of this analysis was abundance of the single VOC (expressed in peak area) and mass spectra that were matched using on-line National Institute of Standards and Technology (NIST) library for potential identification.

3.2.2.5 Quality control of the breath sample data

Routine analysis of exhaled breath has yet to reach wider clinical practice. As such there remains no standardized guidelines for quality control. For this reason, a series of in-house quality control procedures were established for this clinical study. As TD tubes were used for breath collection, a quality control step to evaluate the presence of the breath collected inside the tube was established. Collection of breath in TD tubes lack the visual confirmation of breath presence that is possible when collecting breath in sample bags. The quality control method was based on the use of an abundance threshold for a reference compound. Unpublished data from our group obtained measuring breath VOCs in a large cohort of healthy subjects and cancer patients showed that compounds such as acetone, isoprene and ethanol are always present in breath. Acetone and isoprene were selected as reference compounds. Ethanol was not included as a reference compound due to its high abundance in ambient air within clinical environments (Please see section 2.3).

Acetone is produced by the metabolism of lipids and carbohydrates [233, 234]. The concentration of acetone in human breath varies especially following feeding status and presence of diabetes mellitus [234]. The prolonged fasting time increases the level of acetone, and patients with underlined diabetes also have higher acetone levels compared to healthy controls [235]. However,

even following this variability, acetone is always present in a measurable quantity in human breath, and for this reason was chosen as reference compound for GC-MS.

Isoprene is another abundant compound from human breath and the range of isoprene in healthy subjects is 12-580 ppb [236]. Levels of isoprene can vary usually following physical activity [237]. Although the concentrations are lower than acetone, isoprene is always present in human breath of healthy subjects and patients, as shown by our unpublished data. For this reason, isoprene has been chosen as second reference compounds, to confirm the presence of breath in TD tubes when they are analysed with PTR using NO^+ and O_2^+.

The established threshold for each reference compound were: 4,000,000 of area (corresponding to 100 ppb) of acetone for samples analysed with GC-MS; 50 ppb of acetone for samples analysed with PTR using H_3O+ ionisation; 2.5 ppb of isoprene for samples analysed using NO+ ionisation; and 5 ppb of isoprene using O_2+ ionisation. Samples in which these reference thresholds were not reached were discarded.

3.2.3 Results

Patients

In total 411 patients were recruited to this study including 363 CRC patients (cancer or disease-free patients) and 48 control patients. Patient characteristics are presented in Table 24.

Table 24 General characteristics of the patients recruited in the study

	Total number of samples collected: 411 cancer patients: 363 (including cancer and disease-free patients); Healthy control: 48	
Age	64.94 ± 13.94 (mean ± STD)	
Gender	male	225
	Female	163
	Unknown	23
Status	Newly diagnosed	86
	Under neoadjuvant therapy	44
	Recurrent or distant metastasis	69
	No residual tumour	151
	Healthy control	48
Location of tumour	right	105
	left	109
	Rectum	124
	Synchronous tumour	2
	Unknown	23
Stage	0 (complete response after CCRT)	2
	I	38
	II	74
	III	89
	IV	113
	Unknown	47
Hospital site	St Mary's	198
	Charing Cross	201
	Others	12

3.2.3.1 Selection of compounds

More than one thousand compounds were identified from an initial analysis using the MSHub platform. The presence or absence of these compounds and their peak area count however varied from sample to sample.

To ensure the quality of analysis and prevent bias, selection of compounds was performed prior to all the sample analysis. Compound selection was be based on two criteria: (i) the compound was previously reported to be associated with colorectal cancer in the published literature, and: (ii) visual inspection of chromatograms to assess for strong peaks. From the literature review, selected colorectal cancer associated compounds included: acetone, ethanol, aldehydes, and fatty acids.

Initial analysis therefore focused in thirteen compounds data for which is presented in Table 25.

Table 25 List of selected compounds

Compound	Formula	literature review	Potential biomarker
Propene	CH_2CHCH_3	[238]	yes
Acetone	C_3H_6O	[133] higher in CRC	yes
Dimethylsulphide	C_2H_6S	[137] stool samples	yes
2-Propenenitrile	C_3H_3N	[239] H.pylori infection	yes
Methylcyclopentane	C_6H_{12}	[240] direction of change not reported	yes
Cyclohexane	C_6H_{12}	[241] direction of change not reported	yes
3-methyl pentane	C_6H_{14}	[240] direction of change not reported	yes
Phenol	C_6H_6O	[109] possible gastroesophgeal cancer biomarker	yes
3-ethylhexane	C_8H_{18}	[138] from plasma	yes
4-methyloctane	C_9H_{20}	[242]	Yes
Nonanal	$C_9H_{18}O$	[243]	yes
Decanal	$C_{10}H_{20}O$	[111] [199]	yes
2-Phenoxyethanol	$C_8H_{10}O_2$	[244]	yes

3.2.3.2 Quality control of the samples

From previous studies (unpublished data), the concentration of acetone less than 4 x 10^6 (peak area count from GC-MS) may indicate inadequate breath sample collection on to TD Tube. Cancer cachexia may increase breakdown of energy stores leading to a potential rise in acetone [133, 197, 201].

A total 411 samples were collected from CRC patients and healthy controls. Of those breath samples 101 that were collected prior to September 2017 were excluded from further clinical analysis of disease status. Prior to September 2017 older (un-optimised) setting were routinely being used for the ReCIVA device that were subsequently found to be inadequate for breath collection and subsequent analysis.

A further 111 samples failed to pass acetone/isoprene threshold quality control. Incomplete data collection was found from fifteen samples requiring their exclusion.

In total 184 samples were therefore deemed suitable for further analysis of disease specific VOC biomarkers. The loss of the samples was shown in Figure 16.

Figure 16 Flow diagram of sample discarded

Total samples: 411

363 = active/treated colorectal cancer

48 = healthy controls

101 samples excluded

(Use of non-optimised sampling method)

Total samples:310

262 = active/treated colorectal cancer

48 = healthy controls

111 samples excluded

(Failed quality control)

Total samples: 199

175 = active/treated colorectal cancer

24 = healthy controls

15 samples excluded

(Incomplete patient level data)

Total samples: 184

160 = with active/treated colorectal cancer

24 = healthy controls

Potential factors influencing the exclusion of samples were:

Sampling location

Off the 111 samples that did not meet quality control criteria the majority (85/111; 77%) were collected from Charing Cross Hospital. These samples may have been affected by the reliance on research nurses for sample collection and managing as well as the requirement for transportation to the central laboratory at St Mary's Hospital. These factors may have influence sample quality leading to a high sample exclusion rate.

Storage of the tubes

Due to logistical challenges and instrument down-time the requirement to store TD tubes during large scale clinical breath research trials is unavoidable.

Experiments to confirm the feasibility of TD tubes storage at -80°C were conducted. Previous unpublished data from our group showed that tubes containing both breath and chemical standards were stable after being stored at -80°C for one month. These findings were confirmed by other authors [245]. During this clinical study, some samples were stored for more than one month for instrument down time. As some of these samples did not pass the quality control procedures prolonged storage may have contributed to their low quality and possible sample degradation.

3.2.3.3 Comparison of pre- and post- September 2017 samples

To assess the impact of modifying the ReCIVA settings, samples collect pre- (n=101) and post- (n=310) September 2017 were compared. Comparison was based on the number of samples meeting quality control criteria. Comparison suggested that use of the 'new' ReCIVA setting post September 2017 led to a lower rate of sample exclusion although this was not statistically significant (50.5% vs. 35.8% exclusion rate; *P*=0.157) (Table 26).

Table 26 Comparison on the failure rate of the two methods

ReCIVA method	Failed/total (number)	Failure rate
Old (pre-optimisation)	51/101	50.5%
New (post-optimisation)	111/310	35.8%

3.2.3.4 Comparison of datasets using old and new ReCIVA settings

A Comparison was made between breath that had passed quality controls collected using old or new ReCIVA settings.

Of the 13 VOCs that were assessed the levels of 12 were significantly different in breath samples collected using old and new ReCIVA setting (Table 27). Eleven of the 13 assessed VOC were detected at higher levels with the new ReCIVA setting.

Table 27 Comparison of the samples using different ReCIVA setting

Compound	Old ReCIVA settings n=101 Median	New ReCIVA settings n=310 Median	*p* value
Propene	10913	34696	0.000
Acetone	16455219	17274127	0.007
Dimethylsulphide	328	116750	0.000
2-Propenenitrile	562	26	0.000
Methylcyclopentane	105654	291054	0.005
Cyclohexane	18582	33446	0.147
3-Methylpentane	35277	25507	0.000
Phenol	1798	6434	0.001
3-Ethylhexane	20290	98195	0.002
4-Methyloctane	7898	14173	0.009
Nonanal	127852	1018869	0.000
Decanal	257829	1091751	0.005
2-Phenoxyethanol	0	83171	0.007

3.2.3.5 VOC signatures of colorectal cancer

3.2.3.5.1 Patient characteristics

The characteristics of included patients are reported in Table 28. All the sample collected were separated into five groups based on the status of cancer while they were collected

Group 1 **Colorectal cancer - pre treatment**
Sample were collected at clinic or in the theatre. These patients were pathologically proved adenocarcinoma.

Group 2 **Colorectal cancer - post neoadjuvant therapy**
In this group, patient may have resectable tumour under neoadjuvant (for better prognosis) or unresectable tumour under palliative therapy (for symptomatic control)

Group 3 **Post treatment - tumour recurrence**
In this group, patients were found to have recurrent tumour and would have further palliative therapy or currently under palliative therapy

Group 4 **Post treatment - disease free**
Patients post curative-intent surgery and pathologically and clinically reported with no residual tumour. Can be primary or salvage tumour resection or

Group 5 **Healthy control**
These patients were suggested to have colonoscopy at clinic because of the positive results on stool occult blood test. Sample would be collected at clinic and retrograded medically history analysis wound be done after colonoscopy to confirm their status. If any lesion found from colonoscopy, the collected samples would be moved another category

Table 28 characteristics of patients with qualified samples collected

		Group 1 Colorectal cancer - pre treatment N=40	Group 2 Colorectal cancer - post neoadjuvant therapy N=21	Group 3 Post treatment - tumour recurrence N=30	Group 4 Post treatment - disease free N=60	Group 5 Healthy controls N=24
Age	**(mean ± STD**	67.0 ± 15.4	58.4 ± 13.2	63.9 ± 13.5	63.9 ± 13.5	73.5 ± 13.8
Sex	**male**	24	11	17	29	10
	female	16	10	13	31	14
Tumour location	**right**	13	5	11	19	-
	left	12	5	9	17	-
	rectum	14	11	10	23	-
	Multiple	1			1	-
Stage	**I**	6	1	1	12	-
	II	11	1	3	19	-
	III	7	3	6	10	-
	IV	13	14	20	14	-
	Unknown	2	2		1	-
Hospital site	**SMH**	33	12	11	30	11
	CXH	7	9	9	30	13
Bowel preparation(+/-)		10/30	1/20	2/28	0/60	0/24
Duration of fasting (median hours)		2.5	3.1	3.3	2.9	2.7

Bowel prepare (+): had bowel preparation before sampling

Bowel prepare (-): no bowel preparation before sampling

Pre-operative patients were excluded from duration of fasting time due to medical advisement

Most of the newly diagnosed (pre-treatment) cancers were recruited from St. Mary's Hospital (SMH). Normally, colorectal surgery was performed in the SMH and hence this is where patients attended outpatient appointments.

Patients who were diagnosed stage IV CRC or locally advanced rectal cancer may have pre-operative chemotherapy or chemoradiotherapy in the purpose of increasing the possibility of radical resection [41, 246]. Therefore, in the group 2, the distribution of stage showed higher percentage of later stage.

The comparison of the VOC levels within the breath of included subjects are presented below in table 29.

Table 29 Peak area count of selected VOC from the five study groups

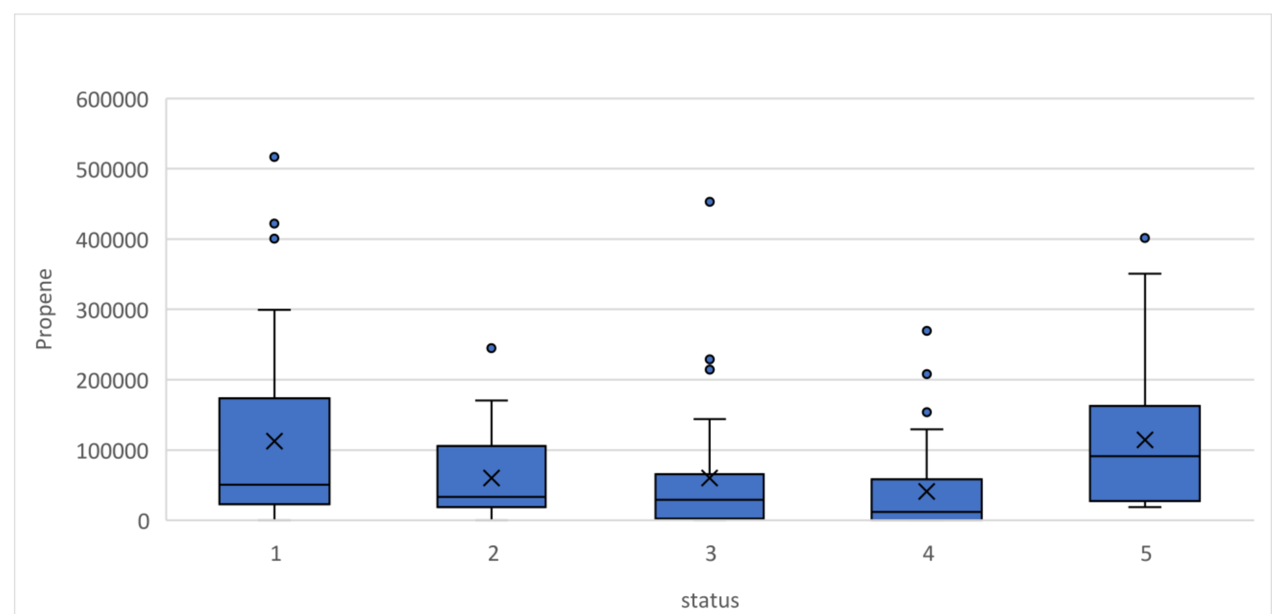
	Group 1 Newly Diagnosed cancer				Group 2 Post NA-therapy				Group 3 Patients with recurrent tumour				Group 4 Disease free				Group 5 Healthy control			
	medium	Q1		Q3	medium	Q1		Q3	medium	Q1		Q3	medium	Q1		Q3	medium	Q1		Q3
Propene	49903	22924	-	173720	31628	20351	-	101440	29244	3368	-	58813	12030	51	-	57556	103399	28222	-	160997
Acetone	22960973	8149758	-	39569530	16666095	9698023	-	29896840	14318717	7651063	-	22283797	19448761	9629620	-	30642650	20213087	13896917	-	34875151
Dimethylsulphide	98513	56187	-	189964	152811	81248	-	275123	89717	41348	-	268786	105098	26621	-	262264	225865	92150	-	316024
2-Propenenitrile	22	2	-	6181	18	7	-	15148	636	0	-	12860	8314	0	-	18055	12	0	-	30
Methyl-cyclopentane	378968	203154	-	528816	348771	211909	-	467103	218950	178657	-	342181	272445	171599	-	398441	370768	239050	-	608814
Cyclohexane	27820	11397	-	54773	43643	27592	-	55369	27661	16357	-	69315	31723	17435	-	64674	54725	27978	-	133945
3-methyl pentane	31630	20006	-	46785	18011	5765	-	30767	29433	17330	-	39515	23336	11678	-	38704	25966	16471	-	53384
Phenol	8242	3763	-	19729	6193	677	-	15868	3779	170	-	10008	2485	0	-	12246	17768	9755	-	40144
3-ethyl-hexane	137733	73759	-	237898	84328	51667	-	153561	73028	31984	-	149263	93714	39828	-	188915	127633	51807	-	259914
4-methyloctane	22708	3483	-	102233	17856	5190	-	59992	9877	508	-	35931	7771	0	-	29429	20670	1709	-	54623
Nonanal	1256305	804356	-	1802421	1124914	721432	-	1701029	936439	575513	-	1404125	878488	506943	-	1291738	1111141	781211	-	2103679
Decanal	1521143	882406	-	2510549	1127427	768190	-	2185467	928168	621247	-	1864416	1031549	535249	-	1396321	1102379	815358	-	2102680
2-Phenoxy-ethanol	54802	0	-	212982	243196	12854	-	369354	76435	33478	-	269603	159839	37308	-	369295	43480	1349	-	86194

Independent sample T-test was performed to compare cancer patients (Groups 1-3) and disease-free patients (Group 4). From the results, propene ($P<0.001$), Cyclohexane ($P=0.049$), 4-methyloctane ($P=0.004$), Decanal ($P=0.008$) and 2-Phenoxy-ethanol ($P=0.044$) were found to be different between patients with cancer compared to disease free status.

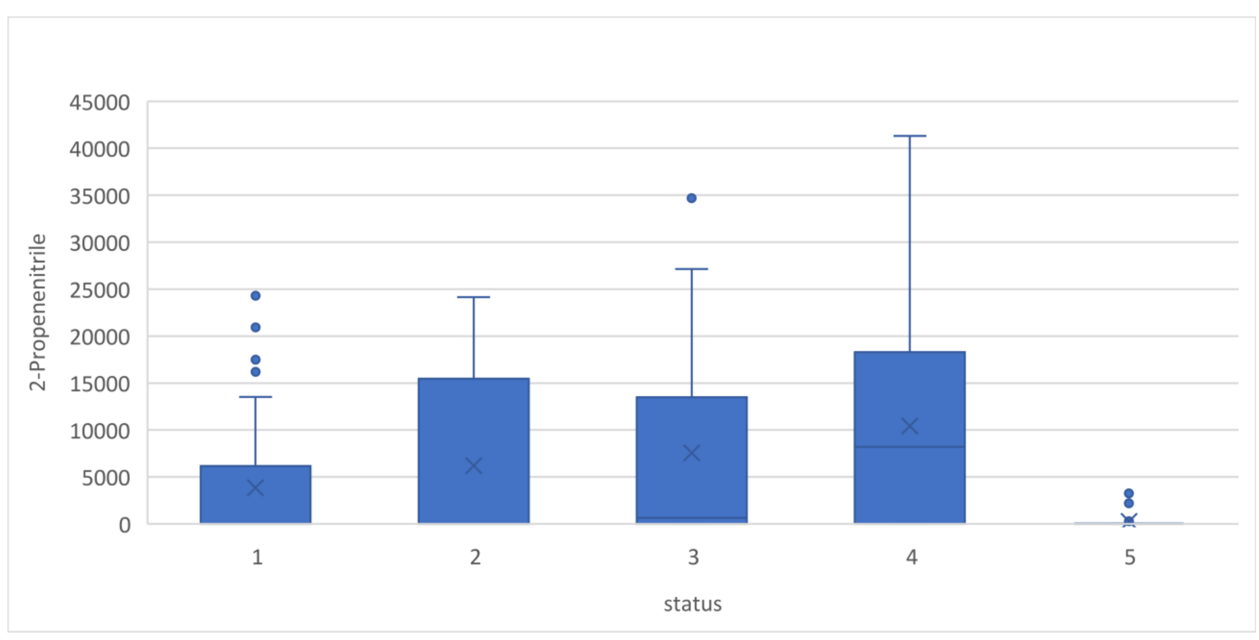
One-way ANOVA was performed to compare cancer patients (Groups 1-3), disease-free patients (Group 4) and healthy controls (Group 5). Propene ($P=0.044$), 2-Propenenitrile ($P<0.001$), and Cyclohexane ($P=0.010$) were found to be significantly different between the three groups.

No VOCs separated group 2 (patients receiving neoadjuvant therapy) from the other study groups. Likewise, patients in group 3 (recurrent CRC) showed poor association with discriminant VOCs.

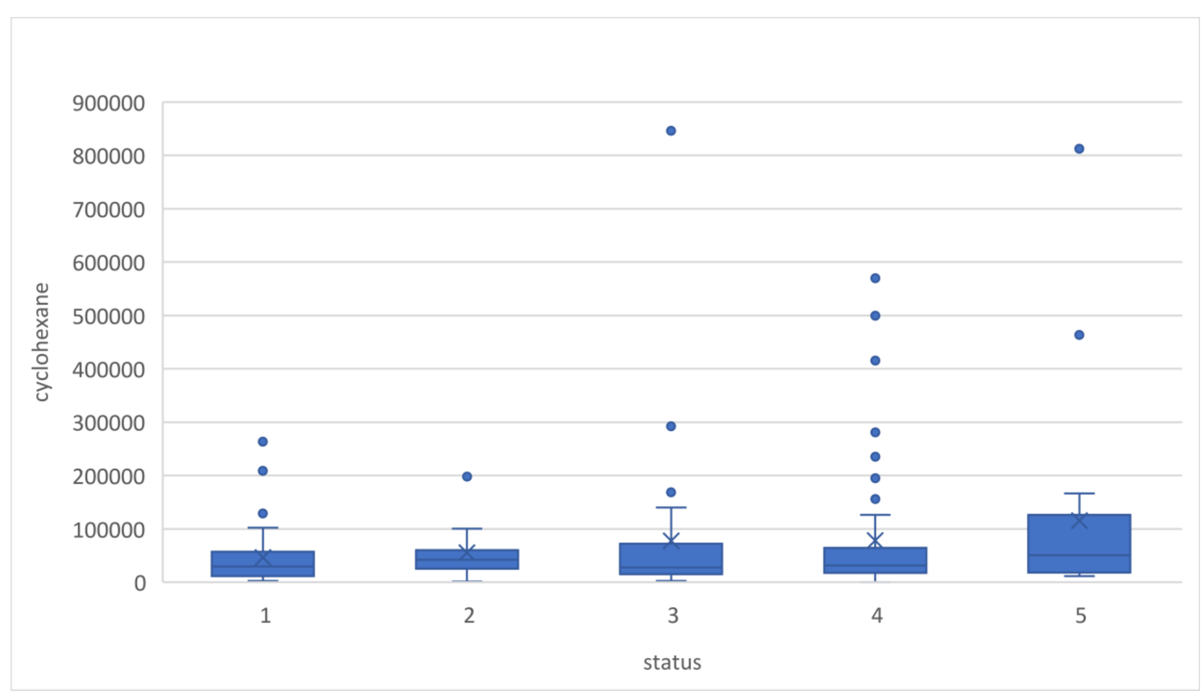
Figure 17 Significant difference from selected compounds



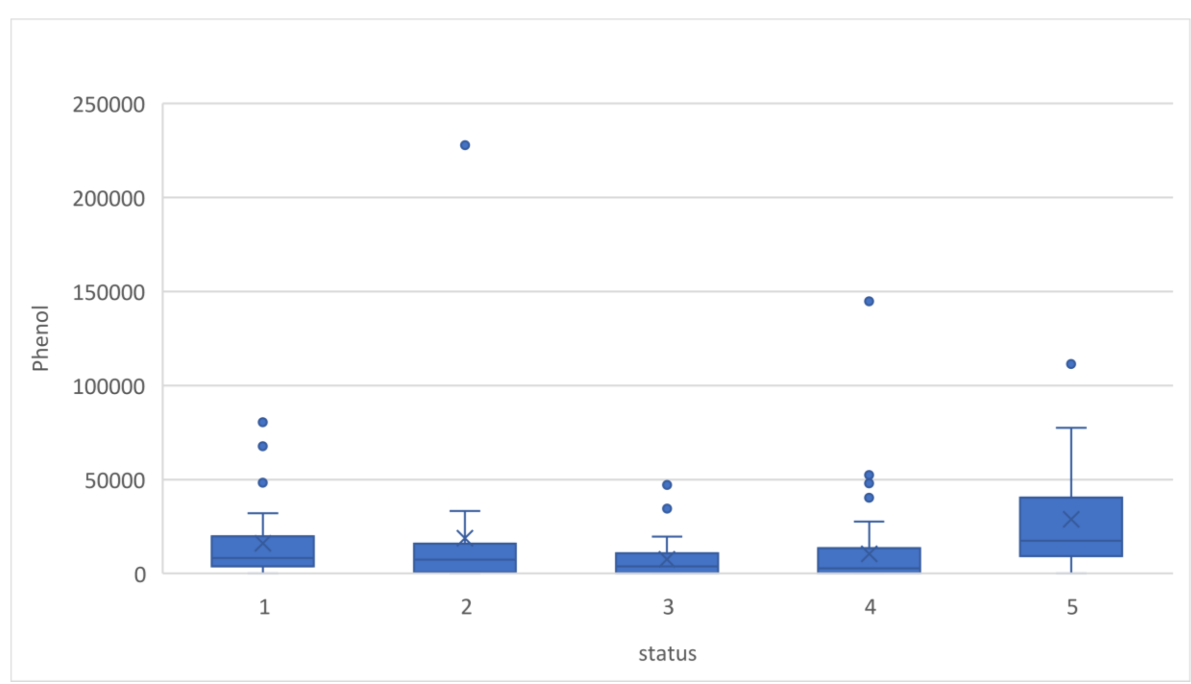
Group 1 and 4, and Group 4 and 5 showed significant different and the *p* value for each were 0.002 and 0.015.



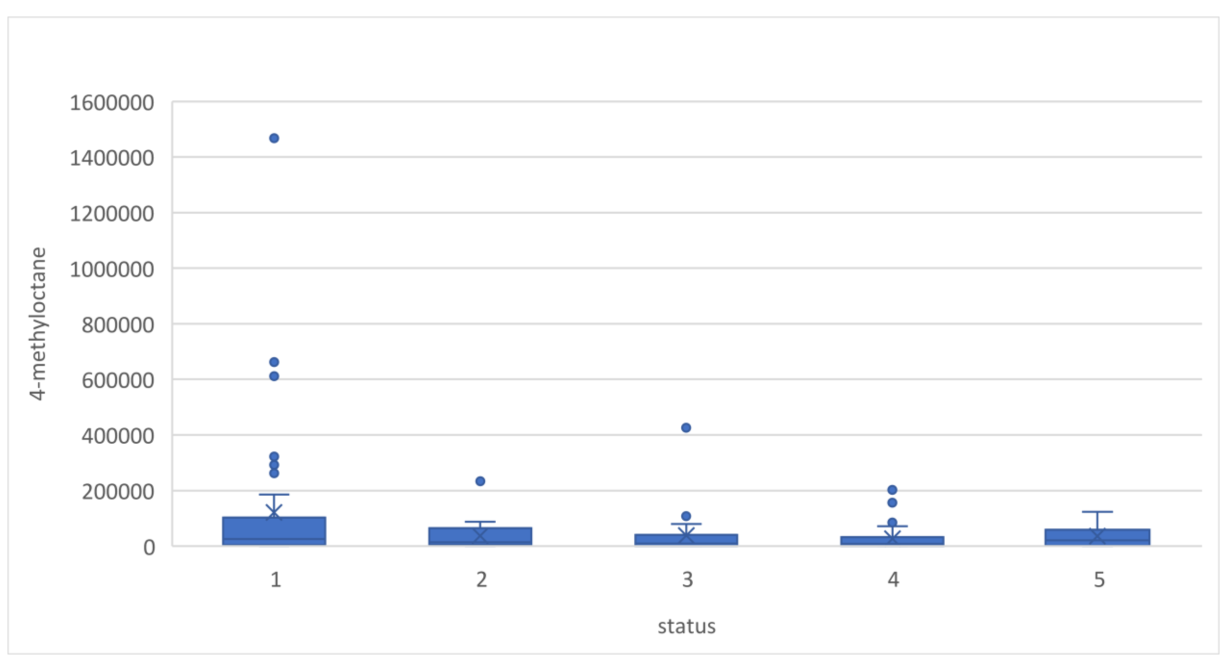
Group 1 and 4, Group 3 and group 5, and Group 4 and 5 showed significant different and the *p* value for each were 0.004, 0.039 and <0.001.



Group 1 and 5 showed significant different and the *p* value was 0.029. (One outlying value from group 5 was eliminated from this figure)



Group 3 and 5, Group 4 and group 5 showed significant different and the *p* value for each were 0.024, and 0.030.



Group 1 and 4 showed significant difference with the *p* value 0.006

3.2.3.5.2 Cancer lesion only

Further analysis was conducted to assess the influence of cancer stage on VOC levels. Breath samples acquired from CRC patients. Pathological cancer stage was based on the 8th AJCC colorectal cancer staging system [247]. The four stages were determined by the depth of tumour invasion, lymph node metastasis, and distant metastasis. Typically, CRC prognosis would be expected to correlate with disease stage.

Characteristics of patients included in the analysis are presented below in Table 30. The peak area count of selected compounds in breath from cancer patients were listed in the table 31.

Table 30 Characteristics of patients with CRC lesions

Characteristic		
Age	64.3 ± 14.6 (mean ± STD)	
Gender	male	49
	female	35
Location of tumour	right	28
	left	26
	rectum	30
Stage	I	7
	II	13
	III	10
	IV	31
	Recurrent cancer	23
Status	1 newly diagnosed	36
	2 under neoadjuvant therapy	19
	3 recurrent or distant metastasis	29

A single patient had synchronous tumours of the transverse colon and rectosigmoid colon junction. For this patient staging was based on the larger (more advance) tumour that was within the transverse colon.

As mentioned above, the stage was based on its pathological stage. For group 1 patients, the pathological stage was determined after tumour resection. For group 3 patients, whether they were diagnosed metastatic lesion or recurrent disease, they were stage IV disease. Staging for group 2 was obtained either radiologically or at the time of tumour resection. It would be difficult to answer if it would be better to use T stage only instead of pathological stage. Here, we tried to see if the severity of cancer itself can be correlated to the VOCs we found. Back to the general concept of these VOCs we found, it shall be produced with the altered metabolism. T stage itself may represent the invasion depth of tumour. While the VOCs were not produced by the tumour itself for us to detect them directly, there is not enough evidence to prove that larger tumour or deeper bowel wall invasion can be a strong factor to produce more VOCs. Therefore, we use the final pathological stage instead of T stage only.

From the results, it is difficult to find any trend how the stage of CRC changes the VOCs patterns. Considering staging was based on the AJCC staging system, it is noted that a more advanced stage does not necessarily correlate to greater tumour mass. If tumours not only change the metabolic pathway but also produce some specific compounds directly, it is impossible to neglect the size of the tumour as potentially relevant in VOC production. When the staging system of CRC was developed however, the size of the tumour showed little affect to the prognosis and was therefore not taken into consideration in the staging system.

Table 31 peak area counts according to cancer stage and cancer recurrence

	Stage I	Stage II	Stage III	Stage IV	Recurrent tumour
			Median		
Propene	45444	80397	47582	34597	27696
Acetone	38719957	28927866	20944240	11492772	12499092
Dimethylsulphide	87146	115084	108628	108160	107493
2-Propenenitrile	15	21	23	19	34
Methylcyclopentane	259058	263908	464372	354031	213612
Cyclohexane	31687	35482	26569	35469	27405
3-Methylpentane	30881	32433	29060	25504	26357
Phenol	7022	10546	13644	7139	3403
3-Ethylhexane	83203	101071	146334	110180	71996
4-Methyloctane	19841	34966	47090	14147	8205
Nonanal	1307841	1230451	1427406	1278052	928343
Decanal	1701218	1572273	1543088	1219050	930422
2-Phenoxyethanol	58899	42410	219603	48217	78981

3.2.3.7 Serial patient sampling

The ability to make serial measurement of exhaled VOC in patients with CRC may provide an opportunity to assess longitudinal variation and therapeutic response.

Current assessment of treatment response is through clinical history, radiological and endoscopic assessment, and blood markers such as CEA. Whilst radiological and endoscopic investigations are potentially helpful in assessing therapeutic response and/or tumour recurrence, they are expensive and invasive, as such are used sparingly within defined surveillance pathways. The blood marker CEA is widely used for post-operative follow-up of CRC patients, but its sensitivity for disease recurrent is low and may not be useful for those patients whose initial CEA was normal before operation.

During this clinical study, a small cohort patients provided breath samples on more than one occasion offering the opportunity for longitudinal assessment of VOC levels.

Characteristic of the patients who underwent repeat sampling are presented in Table 32. Whilst longitudinal samples were initially acquired from 54 patients after quality control, data for 15 patients was suitable for analysis. The principal reasons for data/patient exclusion were sampling with the earlier (pre-September 2017) ReCIVA method and failed analysis of samples by GC-MS.

Table 32 Characteristics of patients who provided serial breath samples

Patient number	Sex	Age	Initial stage	Location	Sample numbers	Status	Tumour change between sampling
1	female	74	ypT0N0	rectum	1	Post-op (no recurrence)	
					2	Post-op (no recurrence)	Stable
2	male	72	M1	rectum	1	Post-op (with distant metastasis or recurrence)	
			Stage IV		2	Post-op (with distant metastasis or recurrence)	Regressed
3	male	36	ypT2N0M0	rectum	1	Post-op (no recurrence)	
			Stage I		2	Post-op (no recurrence)	Stable
4	female	84	cT3	rectum	1	Newly diagnosed	
			Stage II		2	Post-systemic therapy (origin tumour)	Progressed
5	male	60	M1	rectum	1	Post-op (no recurrence)	
			Stage IV		2	Post-op (with distant metastasis or recurrence)	Progressed
					3	Post-op (no recurrence)	Regressed
					4	Post-op (with distant metastasis or recurrence)	Progressed
6	female	68	ypT4N0	A-colon	1	Post-systemic therapy (origin tumour)	
			Stage II		2	Post-op (no recurrence)	Regressed
					3	Post-op (no recurrence)	Stable
7	male	75	ypT4N2	sigmoid	1	Post-systemic therapy (origin tumour)	
			Stage III		2	Post-op (no recurrence)	Regressed
8	female	51	M1	A-colon	1	Post-op (with distant metastasis or recurrence)	
			Stage IV		2	Post-op (with distant metastasis or recurrence)	Progressed
9	female	32	M1	rectum	1	Post-systemic therapy (origin tumour)	
			Stage IV		2	Post-op (no recurrence)	Regressed
					3	Post-op (with distant metastasis or recurrence)	Progressed
10	male	52	M1	sigmoid	1	Post-systemic therapy (origin tumour)	
			Stage IV		2	Post-op (no recurrence)	Regressed
11	female	68	M1	appendix	1	Post-op (with distant metastasis or recurrence)	
			Stage IV		2	Post-op (with distant metastasis or recurrence)	Progressed
					3	Post-op (with distant metastasis or recurrence)	Progressed
12	female	68	pT3N0	caecum	1	Newly diagnosed	
			Stage II		2	Post-op (no recurrence)	Regressed
13	male	75	pT3N1	A-colon	1	Post-op (no recurrence)	
			Stage III		2	Post-op (no recurrence)	Stable
14	male	55	pT4N0	T-colon	1	Post-op (no recurrence)	
			Stage II		2	Post-op (no recurrence)	Stable
15	female	50	pT1N0	sigmoid	1	Newly diagnosed	
			Stage I		2	Post-op (no recurrence)	Regressed

The small size of the patient cohort and different time points meant that it was not possible to make robust statistical comparisons within the dataset of longitudinal samples. However, it was possible to gain some insight into the effects of disease status in patients who provided multiple samples throughout their treatment.

For patients with progressive disease, cyclohexane, nonanal, and decanal showed possible increased in peak area count. Although the size of collected samples was small, almost a two-fold increase in peak area count could be observed with nonanal and decanal.

For patients with disease regression or surgical resection, ethanol, acetone, 3-methylpentane, and decanal showed higher variation. Lower acetone level could be measured after tumour resection. Normally, acetone comes from fat burning which often occurs with cancer patients to gain more energy. The lower acetone may be an index that no extra energy necessary. 3-methylpentane were found to be possible cancer biomarker before. The result of decanal may require further investigation because of the unmatched results in the progressed and regressed disease.

3.2.4 Discussion

Based on the result, an appropriate power calculation is difficult to achieve, especially the method was under modification during the study. The principal outcome of this clinical study was recognition of the importance of a standardized and quality-controlled methodology for breath collection. Selected VOCs were found to be associated with CRCs compared to successfully treated cancers and healthy controls.

As mentioned in previous chapter, there was no well-established method for sample collection and analysis. It was difficult to do the power calculation. If we find any VOCs, we cannot make sure if there is any interaction between the VOCs we found. Therefore, it was less likely that we can decide the statistical method before we saw the data. Here in the thesis, we can give the indicative results only.

This study has emphasised the vulnerability for volatile breath biomarker research to sources of methodological variability. High rates of data exclusion reflected a change in the study protocol (after optimizations of ReCIVA settings) and the strict application of robust quality control measures. Changing ReCIVA setting in the current study was associated with a significant change in all VOCs levels, except for cyclohexane. A change in the ReCIVA settings followed work other work within the volatile biomarker research indicating that the collection of higher sample volume resulted in more reliable VOC detection.

I noted that breath sampling in general was well tolerated and acceptable to patients. All patients were able to complete the test and there were no reports of discomfort or distress during testing. These findings are in keeping with a previous report that

showed high the acceptability of breath testing within patients attending the GP for non-specific gastrointestinal symptoms [229]

It is also noteworthy that samples the failed to pass quality control were more often collected from Charing Cross Hospital. Unlike at St Mary's Hospital where all samples were collected by the primary research (GPL), a considerable proportion of samples from Charing Cross Hospital were collected by a team of NIHR research nurses. Where a team of individuals is conducting sampling, there is the potential for variation in execution of the sampling technique. Even small deviations from the standard procedure for breath sampling can have a significant effect on the obtained VOCs. A previous studies did find that multi-centre breath sampling is feasible with a majority of collected samples meeting quality control criteria [229]. For large-scale clinical trial, it is not possible to recruit all the patients and sample by one researcher. Although we trained each research nurse for taking samples and left manuscription for them, some changes of the member may result in some unqualified sample taking. A better solution is to communicate with NIHR team every time they recruit new member and make sure they can fully understand the whole sampling procedure. It shall be part of the standardization in the research. it will be important whether the device for future breath sampling will be.

Most of the sample collected at early stage failed to pass the quality control. Comparing to other simple and cheap test such as stool occult blood test, breath test in the study showed its weak side that we cannot use a standardized method to collect sample and analysis. However, after we modified the method, we found that the rate for sampled tube to pass the quality control did increase. It can an unpreventable

process to the destination of an even easier test. Also, another advantage of breath test is its potential as a point-of-care test because the accessibility of the samples.

In chapter 2, we discuss about the repeatability of ReCIVA and found that the four tubes could collect similar compounds. Also, we could conclude that TD-PTR-MS can give some similar result as TD-GC-MS. However, the results here was only a single tube from each patient to be analysed on GC-MS. It is possible to analyse other tubes with PTR-MS or GC x GC-MS while more than one tube could be obtained with ReCIVA with a single sample collection. Besides, the sorbent in the TD tube were kept the same since the beginning of the study. The reason only on tube was analysed is that we tried to decrease variable, therefore we choose GC-MS only for breath sample analysis. The other reason is the setting of ReCIVA. While we change the setting of ReCIVA in the middle of sample collection, we were testing the so-called optimized method with clinical trial. With the proceeding of the exam, we started to do quality control before sample analysis. The method we used was different from the beginning to the end of the study. It would be reasonable to analyse one tube only at the beginning of large-scale clinical trial.

From the data presented herein, it does appear that selected VOCs may be associated with the presence of CRC. Specifically, Propene and Decanal tended to be higher in the exhaled breath of patients with CRC compared to patients who had been successful treated for this disease. Propene however was higher in the breath of healthy controls compared to patients with CRC. Altomare has previously reported finding higher Decanal levels in the exhaled breath of CRC patients [111]. Decanal was also associated with a number of other human cancers [248].

In the current analysis exhaled Decanal, and other VOCs, were not obviously linked to cancer stage or tumour recurrence. The significance of this finding remains uncertain but could indicated that the volatile signature of CRC is a consequence of more than the tumour mass itself. Currently the role of the tumour associated microbiome and systemic metabolic changes in VOC production remain unknown. As different analytical platforms and different VOCs markers in the tumour headspace study presented in section 3.1 above, it is not possible to draw clear comparisons between the local (tumour headspace) and breath VOC levels in cancer. Certainly, it would be interested to examine whether exhaled CRC VOCs markers were also enriched in the headspace above the tumour.

There was tentative evidence that decanal cyclohexane and nonanal were associated with disease progression in patients who were provided serial breath samples as part of a longitudinal study. A lack of uniformity in the timing of sampling and patient status at the time of sampling mean that it is not possible to make a definitive conclusion based on these findings.

This study suffers from a number of acknowledged limitations, some of which have already been commented on. High sample exclusion meant that a considerable proportion of breath samples were not used in the analysis. This underscores the importance of high quality and reliable sample collection, storage, and analysis. It may have also meant that the current study was underpowered to determined difference between patient groups. The number of patients who provided longitudinal samples was also small due to necessary exclusions and there was variability in the timing of sample collection. Future studies should aim to complete a more comprehensive longitudinal study of CRC at defined timepoint during their treatment pathway.

The use of bowel preparation and at the time of sampling was common in patients who were recruited at the time of colonoscopy an on the day of surgery. Bowel preparation has the potentially to affect the volatolomics through it actions within the bowel and of systemic fluid balance. In a another (unpublished) study conducted by our group bowel preparation did not emerge as a diagnostic feature in a machine learning model for the prediction of CRC. Notwithstanding a proposed future breath test for CRC would be intended for use in primary care and is unlikely to require pre-test conditions such as bowel preparation.

Owing to technical issues and instrument downtime different analytical platforms were use in the clinical studies reported in sections 3.1. and 3.2. The meant that it was not possible to directly compare the VOC biomarkers of tumour headspace and breath studies. In particular the GC-MS instrument in the configuration that it was used in the current study was not able to reliably detect polar compounds such as short chain fatty acids.

For serial breath sampling, the limited number of samples could be analysed made it difficult to do any further statistical analysis. Here, we simplified the status of disease into only three categories, disease progression, stable disease, and disease regression. Progression can be the initial diagnosis of tumour recurrence or the increase of tumour burden by CT or tumour marker (CEA). Stable disease can be patients without recurrence of tumour, or no progression found during therapy. These patients would be those after operation without relapsed tumour. However, the limited number made it difficult to show if there were any clue that adjuvant chemotherapy may change the VOCs pattern. Patients 3, 13, and 14 had adjuvant chemotherapy but the regiment was

different. Patient 9 is under both systemic palliative chemotherapy and pressurized intraperitoneal aerosolized chemotherapy (PIPAC). Since it is not possible to find similarities with these patients, we can only investigate the result and see if any changes of the disease may change the VOCs. We could find some little trend that some VOCs can be changed with the status of CRC. Ideally, we may need a long term follow up and a more detailed program to collect samples. besides, we need a clearer definition of the status of disease to give a better interpret the results.

3.2.4.1 Conclusions

In conclusions this study has provided tentative evidence that selected exhaled VOCs are linked to CRC and that the compounds may change as a response to therapeutic intervention and disease status. Whilst the breath test was acceptable to patients, further work is needed to refine the methodology for sample collection and handling to ensure reliability of results.

FINAL CONCLUSIONS

The overall objective of this thesis was to develop a VOC based model for the assessment and monitoring of colorectal cancer. Secondary aims were to: explore the influence of off-line sampling methodology on the detection of exhaled VOCs; perform multiplatform correlation of VOCs in human breath, and; determine the relationship between VOCs detected in colonic gas and exhaled breath. This thesis broadly met these stated aims.

An initial systematic literature review demonstrated a body of evidence supporting the premiss that colorectal cancers are associated with a specific VOC signature. Significant methodological variation between studies however meant that it was not possible to directly compare studies and derive a unifying panel of colorectal cancer specific VOCs. The association between cancer and VOCs is well established and has been the subject of numerous papers and review articles [248].

A major challenge that faces the field of volatolomics and breath analysis is the necessity for robust and standardized protocols for sample collection, managing and analysis. The absence of an agreed consensus for the standardized VOC analysis has meant that researchers have adopted a wide range of strategies. It has generally not been possible to replicate of validate finding between different studies.

Initial studies within this thesis focuses on understanding specific aspect of the methodology for breath collection, storage, and analysis.

The utility of different devices for collection of breath on to thermal desorption tubes was assessed (Section 2.1). It was found that no device has all the ideal characteristic for the proposed task. Whilst the breath concentrator devices yielded the most reliable profile of captured VOCs, with low contamination, it was limited by its status an experimental device that is not commercially available. Unsurprisingly it therefore did not score the highest in human factor analysis of relating to usability. By comparison, the ReCIVA device scored highly on the human factor analysis, had acceptable VOC capture, but was limited by contamination from exogenous sources presumed to be associated with the device itself. Another important limitation of the ReCIVA device is it high initial cost and the requirement for a single use face masks for each sampling event. Breath collection using ReCIVA was however shown be both feasible and acceptable in a large study of patients in primary care [229].

These findings would seem to indicate that further work is needed to improve the method for breath collection and storage. At the present time thermal desorption tubes do appear to be the best solution to long term storage of breath assuming that they are managed in accordance with protocols to ensure that they are clean and uncontaminated. A new device is therefore needed to initially capture the breath and pass it to the thermal desorption tube. Within the Hanna lab a bespoke breath collection device has been build which requires breath to be first collected into a single use PTFE sample bag before being transferred to thermal desorption tubes with the aid of a precision pump (Figure 18). The relative simplicity of this process makes breath sampling simple for patients and staff.

A subsequent study examined the comparability of different platforms for VOC analysis (section 2.2). Importantly this study demonstrated acceptable correlation between

direct injection and TD-PTR-ToF-MS as well as TD-PTR-ToF-MS and TD-GC-MS. Correlation between TD-PTR-ToF-MS and TD-GC-MS was mainly for the abundant compounds that were studied. Thermal desorption tubes as already mentioned offer a reliable method for breath storage. The analysis of VOC trapped within thermal desorption tubes can be achieved by a number of analytical techniques that can be selected partly based on the specific analyte(s) of interest. It is however important to understand correlation of results acquired by these instruments.

A study of ambient VOC levels withing different clinical environments using TD PTR-ToF-MS was performed to examine the potential impact of sampling location on breath analysis (Section 2.3). Finding broadly supported previous studies that identified selected VOC are found in high levels in certain environments, presumably due to local contaminants [249]. A typical example of this is ethanol that is widely used in all areas of the hospital for the purpose of hand sanitation. This study helped to provide a broad understanding of potential contaminants. It should however be noted that all potential biomarkers identified at the discovery phase (and in subsequent studies) show be carefully evaluated both in terms of their relative change with respect to disease/physiological status, presumed biological origin and potential for contamination from exogenous sources, such as ambient air. It is an acknowledge limitation that the majority of VOCs studied in this section were not ultimately examined in the clinical studies, principally due to changes in the analytical procedure.

Two clinical studies explored the VOC profiles of colorectal cancers in-situ and within exhaled breath. Whilst it was possible to identify selected VOC that were associated with colorectal cancers it was not possible to draw direct comparisons between tumour head space and exhaled breath experiments. This was primarily due to the

different analytical platforms used for these two studies. From breath studies the aldehyde Decanal did emerge as a potential marker of colorectal cancer. This would appear to corroborate the findings for other studies [111, 248].

The breath analysis study did look at the change in VOCs in response to treatment and treatment response. Whilst a small number of VOCs did appear to vary in response to therapeutic intervention (including surgery), it was not possible to derive a clear pattern of reliable change. This was most likely due to differences in sampling conditions and small patient numbers. A similar outcome was seen in a limited longitudinal study and a study of the effect of tumour T stage.

The clinical breath study was principally limited by the substantial number of sample exclusions that were required in respect to strict quality control measures. Whilst this can be considered a strength of the study it also resulted in significant depletions of sample numbers, affecting both cross sectional and longitudinal studies. It is recognised that with improvement in sample collection and handling methods the number of exclusions can be reduced.

FUTURE WORK

A consequence of this thesis there have been major developments in the method used for breath collection, handling, and analysis. As previously mentioned, the Hanna lab has developed a custom breath sampling device comprised of a single use sample bag and precision pump that transfers breath to thermal desorption tubes (Figure 18).

In addition, there has been a significant upgrade in the GC-MS instruments available within the Hanna lap to allow multiplatform polar and non-polar GC-MS for biomarker discovery and two-dimensional GC-MS for confirmation of biomarker identify. These instrument and the analytical pathway have been optimised for biomarker discovery and confirmation. The pathway also retained the rigorous quality control applied within this thesis.

A future large scale clinical trial is now planned to examine the VOC profile of colorectal cancer. this trial will utilise the new breath sampling device and multiplatform GC-MS analytical pathway. Breath sampling will be performed in patients who have been referred from primary care with suspected colorectal cancer. Within this population colorectal cancer rates are approximately 4.5%. Compared to this patient population used in the current thesis the proposed population for the future trial is more in keeping with the intended target population for the breath, with is symptomatic patients attending primary care.

Figure 18 Custom breath sampling device

The patient blows into a single use PTFE sample bag (right of picture) with the breath subsequently being passed on thermal desorption tubes (centre) using a precision pump (left)

The planned clinical trial in colorectal cancer will be run in parallel to similar trials for oesophagogastric and pancreatic adenocarcinoma. As the biomarkers for each of these cancers are expected to be different [132, 180, 250] it may be possible in the future to make a single 'pan-cancer' test.

Concurrent work is also planned to examine the biological origin of VOC in colorectal and other intestinal cancer. This work will involve in vitro cell culture studies that will examine cancer cell specific and external (microbiome, immune cell related) pathways of VOC production.

Further optimisation of the sampling protocol and mitigation of external contaminants will further support wider adoption of VOC analysis within future clinical practice.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries**. *CA: A Cancer Journal for Clinicians* 2018, **68**(6):394-424.
2. Soriano LC, Soriano-Gabarró M, García Rodríguez LA: **Trends in the contemporary incidence of colorectal cancer and patient characteristics in the United Kingdom: a population-based cohort study using The Health Improvement Network**. *BMC Cancer* 2018, **18**(1):402.
3. Chan AT, Giovannucci EL: **Primary prevention of colorectal cancer**. *Gastroenterology* 2010, **138**(6):2029-2043.e2010.
4. Watson AJM, Collins PD: **Colon Cancer: A Civilization Disorder**. *Digestive Diseases* 2011, **29**(2):222-228.
5. Cunningham D, Atkin W, Lenz H-J, *et al*: **Colorectal cancer**. *The Lancet* 2010, **375**(9719):1030-1047.
6. Ekbom A, Helmick C, Zack M, Adami HO: **Ulcerative colitis and colorectal cancer. A population-based study**. *N Engl J Med* 1990, **323**(18):1228-1233.
7. Savoca PE, Ballantyne GH, Cahow CE: **Gastrointestinal malignancies in Crohn's disease. A 20-year experience**. *Dis Colon Rectum* 1990, **33**(1):7-11.
8. **Cancer Stat Facts: Colorectal Cancer**
9. Del Vecchio Blanco G, Cretella M, Paoluzi OA, *et al*: **Adenoma, advanced adenoma and colorectal cancer prevalence in asymptomatic 40- to 49-year-old subjects with a first-degree family history of colorectal cancer**. *Colorectal Dis* 2013, **15**(9):1093-1099.
10. Shida H, Ban K, Matsumoto M, *et al*: **Asymptomatic colorectal cancer detected by screening**. *Diseases of the Colon & Rectum* 1996, **39**(10):1130-1135.
11. Stewart SL, Wike JM, Kato I, Lewis DR, Michaud F: **A population-based study of colorectal cancer histology in the United States, 1998–2001**. *Cancer* 2006, **107**(S5):1128-1141.
12. Kijima S, Sasaki T, Nagata K, Utano K, Lefor AT, Sugimoto H: **Preoperative evaluation of colorectal cancer using CT colonography, MRI, and PET/CT**. *World J Gastroenterol* 2014, **20**(45):16964-16975.

13. Petrillo A, Fusco R, Petrillo M, *et al*: **Standardized Index of Shape (DCE-MRI) and Standardized Uptake Value (PET/CT): Two quantitative approaches to discriminate chemo-radiotherapy locally advanced rectal cancer responders under a functional profile**. *Oncotarget* 2017, **8**(5):8143-8153.
14. Wagner F, Hakami YA, Warnock G, Fischer G, Huellner MW, Veit-Haibach P: **Comparison of Contrast-Enhanced CT and [(18)F]FDG PET/CT Analysis Using Kurtosis and Skewness in Patients with Primary Colorectal Cancer**. *Mol Imaging Biol* 2017, **19**(5):795-803.
15. Catalano OA, Coutinho AM, Sahani DV, *et al*: **Colorectal cancer staging: comparison of whole-body PET/CT and PET/MR**. *Abdom Radiol (NY)* 2017, **42**(4):1141-1151.
16. Duffy MJ, van Dalen A, Haglund C, *et al*: **Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines**. *European Journal of Cancer* 2003, **39**(6):718-727.
17. Fakih MG, Padmanabhan A: **CEA monitoring in colorectal cancer. What you should know**. *Oncology (Williston Park)* 2006, **20**(6):579-587; discussion 588, 594, 596 passim.
18. Wolpin BM, Mayer RJ: **Systemic Treatment of Colorectal Cancer**. *Gastroenterology* 2008, **134**(5):1296-1310.e1291.
19. Rodel C, Bruch HP, Hofheinz R, Lang H, Arnold D: **[Treatment of rectal cancer]**. *Onkologie* 2010, **33 Suppl 4**:19-23.
20. Meyerhardt JA, Mangu PB, Flynn PJ, *et al* 3rd: **Follow-up care, surveillance protocol, and secondary prevention measures for survivors of colorectal cancer: American Society of Clinical Oncology clinical practice guideline endorsement**. *J Clin Oncol* 2013, **31**(35):4465-4470.
21. Labianca R, Nordlinger B, Beretta GD, Brouquet A, Cervantes A, Group ObotEGW: **Primary colon cancer: ESMO Clinical Practice Guidelines for diagnosis, adjuvant treatment and follow-up**. *Annals of Oncology* 2010, **21**(suppl_5):v70-v77.
22. Ong MLH, Schofield JB: **Assessment of lymph node involvement in colorectal cancer**. *World journal of gastrointestinal surgery* 2016, **8**(3):179-192.

23. Tang C-L, Eu K-W, Tai B-C, Soh JGS, Machin D, Seow-Choen F: **Randomized clinical trial of the effect of open versus laparoscopically assisted colectomy on systemic immunity in patients with colorectal cancer**. *BJS* 2001, **88**(6):801-807.
24. Kitano S, Kitajima M, Konishi F, Kondo H, Satomi S, Shimizu N: **A multicenter study on laparoscopic surgery for colorectal cancer in Japan**. *Surgical Endoscopy And Other Interventional Techniques* 2006, **20**(9):1348-1352.
25. Park JJ, Cheon JH, Kwon JE, *et al*: **Clinical outcomes and factors related to resectability and curability of EMR for early colorectal cancer**. *Gastrointestinal Endoscopy* 2011, **74**(6):1337-1346.
26. Bories E, Pesenti C, Monges G, *et al*: **Endoscopic Mucosal Resection for Advanced Sessile Adenoma and Early-Stage Colorectal Carcinoma**. *Endoscopy* 2006, **38**(03):231-235.
27. Ha RK, Han KS, Sohn DK, *et al*: **Histopathologic risk factors for lymph node metastasis in patients with T1 colorectal cancer**. *Ann Surg Treat Res* 2017, **93**(5):266-271.
28. Bianco F, De Franciscis S, Belli A, *et al*: **T1 colon cancer in the era of screening: risk factors and treatment**. *Tech Coloproctol* 2017, **21**(2):139-147.
29. Simmonds PC, Primrose JN, Colquitt JL, Garden OJ, Poston GJ, Rees M: **Surgical resection of hepatic metastases from colorectal cancer: A systematic review of published studies**. *British Journal of Cancer* 2006, **94**(7):982-999.
30. Regnard J-F, Grunenwald D, Spaggiari L, *et al*: **Surgical treatment of hepatic and pulmonary metastases from colorectal cancers**. *The Annals of Thoracic Surgery* 1998, **66**(1):214-218.
31. Selvaggi F, Cuocolo A, Sciaudone G, Maurea S, Giuliani A, Mainolfi C: **FGD-PET in the follow-up of recurrent colorectal cancer**. *Colorectal Disease* 2003, **5**(5):496-500.
32. Fahy BN: **Follow-up after curative resection of colorectal cancer**. *Ann Surg Oncol* 2014, **21**(3):738-746.
33. Rodrigues RV, Pereira da Silva J, Rosa I, *et al*: **Intensive Follow-Up After Curative Surgery for Colorectal Cancer**. *Acta Med Port* 2017, **30**(9):633-641.

34. Scott NA, Jeacock J, Kingston RD: **Risk factors in patients presenting as an emergency with colorectal cancer**. *BJS (British Journal of Surgery)* 1995, **82**(3):321-323.

35. McArdle CS, Hole DJ: **Emergency presentation of colorectal cancer is associated with poor 5-year survival**. *BJS (British Journal of Surgery)* 2004, **91**(5):605-609.

36. Cuffy M, Abir F, Audisio RA, Longo WE: **Colorectal cancer presenting as surgical emergencies**. *Surgical Oncology* 2004, **13**(2):149-157.

37. Zhang Y, Shi J, Shi B, Song C-Y, Xie W-F, Chen Y-X: **Self-expanding metallic stent as a bridge to surgery versus emergency surgery for obstructive colorectal cancer: a meta-analysis**. *Surgical Endoscopy* 2012, **26**(1):110-119.

38. Atukorale YN, Church JL, Hoggan BL, *et al*: **Self-Expanding Metallic Stents for the Management of Emergency Malignant Large Bowel Obstruction: a Systematic Review**. *Journal of Gastrointestinal Surgery* 2016, **20**(2):455-462.

39. Pozzo C, Barone C, Kemeny NE: **Advances in neoadjuvant therapy for colorectal cancer with liver metastases**. *Cancer Treatment Reviews* 2008, **34**(4):293-301.

40. Provenzale D, Gupta S, Ahnen DJ, *et al*: **NCCN Guidelines Insights: Colorectal Cancer Screening, Version 1.2018**. 2018, **16**(8):939.

41. Glynne-Jones R, Wyrwicz L, Tiret E, *et al*: **Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up[†]**. *Annals of Oncology* 2017, **28**:iv22-iv40.

42. Brandi G, De Lorenzo S, Nannini M, *et al*: **Adjuvant chemotherapy for resected colorectal cancer metastases: Literature review and meta-analysis**. *World journal of gastroenterology* 2016, **22**(2):519-533.

43. Midgley R, Kerr DJ: **Adjuvant chemotherapy for stage II colorectal cancer: the time is right!** *Nat Clin Pract Oncol* 2005, **2**(7):364-369.

44. Gervaz P, Rubbia-Brandt L, Andres A, *et al*: **Neoadjuvant chemotherapy in patients with stage IV colorectal cancer: a comparison of histological response in liver metastases, primary tumors, and regional lymph nodes**. *Ann Surg Oncol* 2010, **17**(10):2714-2719.

45. Cunningham D, Humblet Y, Siena S, *et al*: **Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer**. *N Engl J Med* 2004, **351**(4):337-345.
46. Horisberger K, Treschl A, Mai S, *et al*: **Cetuximab in combination with capecitabine, irinotecan, and radiotherapy for patients with locally advanced rectal cancer: results of a Phase II MARGIT trial**. *Int J Radiat Oncol Biol Phys* 2009, **74**(5):1487-1493.
47. Tol J, Koopman M, Cats A, *et al*: **Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer**. *N Engl J Med* 2009, **360**(6):563-572.
48. Ezzeldin HH, Acosta EP, Mattison LK, Fourie J, Modak A, Diasio RB: **13C-5-FU breath test current status and future directions: a comprehensive review**. *Journal of Breath Research* 2009, **3**(4):047002.
49. Ngan SY, Burmeister B, Fisher RJ, *et al*: **Randomized trial of short-course radiotherapy versus long-course chemoradiation comparing rates of local recurrence in patients with T3 rectal cancer: Trans-Tasman Radiation Oncology Group trial 01.04**. *J Clin Oncol* 2012, **30**(31):3827-3833.
50. Latkauskas T, Pauzas H, Gineikiene I, *et al*: **Initial results of a randomized controlled trial comparing clinical and pathological downstaging of rectal cancer after preoperative short-course radiotherapy or long-term chemoradiotherapy, both with delayed surgery**. *Colorectal Disease* 2012, **14**(3):294-298.
51. Li J, Liu H, Yin J, *et al*: **Wait-and-see or radical surgery for rectal cancer patients with a clinical complete response after neoadjuvant chemoradiotherapy: a cohort study**. *Oncotarget* 2015, **6**(39):42354-42361.
52. Habr-Gama A, Sabbaga J, Gama-Rodrigues J, *et al*: **Watch and Wait Approach Following Extended Neoadjuvant Chemoradiation for Distal Rectal Cancer: Are We Getting Closer to Anal Cancer Management?** *Diseases of the Colon & Rectum* 2013, **56**(10):1109-1117.
53. Probst CP, Becerra AZ, Aquina CT, *et al*: **Watch and Wait?--Elevated Pretreatment CEA Is Associated with Decreased Pathological Complete Response in Rectal Cancer**. *J Gastrointest Surg* 2016, **20**(1):43-52; discussion 52.

54. Dossa F, Chesney TR, Acuna SA, Baxter NN: **A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis**. *Lancet Gastroenterol Hepatol* 2017, **2**(7):501-513.
55. Townsend AR, Chong LC, Karapetis C, Price TJ: **Selective internal radiation therapy for liver metastases from colorectal cancer**. *Cancer Treatment Reviews* 2016, **50**:148-154.
56. Henke L, Kashani R, Robinson C, *et al*: **Phase I trial of stereotactic MR-guided online adaptive radiation therapy (SMART) for the treatment of oligometastatic or unresectable primary malignancies of the abdomen**. *Radiotherapy and Oncology* 2018, **126**(3):519-526.
57. Simmonds PC: **Palliative chemotherapy for advanced colorectal cancer: systematic review and meta-analysis. Colorectal Cancer Collaborative Group**. *BMJ (Clinical research ed)* 2000, **321**(7260):531-535.
58. Costi R, Di Mauro D, Giordano P, *et al*: **Impact of Palliative Chemotherapy and Surgery on Management of Stage IV Incurable Colorectal Cancer**. *Annals of Surgical Oncology* 2010, **17**(2):432-440.
59. Renehan AG, Egger M, Saunders MP, O'Dwyer ST: **Impact on survival of intensive follow up after curative resection for colorectal cancer: systematic review and meta-analysis of randomised trials**. *BMJ* 2002, **324**(7341):813.
60. Renehan AG, Dwyer O' ST, Whynes DK: **Cost effectiveness analysis of intensive versus conventional follow up after curative resection for colorectal cancer**. *BMJ* 2004, **328**(7431):81.
61. Mokhles S, Macbeth F, Farewell V, *et al*: **Meta-analysis of colorectal cancer follow-up after potentially curative resection**. *BJS (British Journal of Surgery)* 2016, **103**(10):1259-1268.
62. Koo S, Neilson LJ, Von Wagner C, Rees CJ: **The NHS Bowel Cancer Screening Program: current perspectives on strategies for improvement**. *Risk Manag Healthc Policy* 2017, **10**:177-187.
63. MEKLIN J, SYRJÄNEN K, ESKELINEN M: **Fecal Occult Blood Tests in Colorectal Cancer Screening: Systematic Review and Meta-analysis of Traditional and**

New-generation Fecal Immunochemical Tests. *Anticancer Research* 2020, **40**(7):3591-3604.

64. Shapiro JA, Bobo JK, Church TR, *et al*: **A Comparison of Fecal Immunochemical and High-Sensitivity Guaiac Tests for Colorectal Cancer Screening**. *The American journal of gastroenterology* 2017, **112**(11):1728-1735.
65. Wolf AMD, Fontham ETH, Church TR, *et al*: **Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society**. *CA: A Cancer Journal for Clinicians* 2018, **68**(4):250-281.
66. Navarro M, Nicolas A, Ferrandez A, Lanas A: **Colorectal cancer population screening programs worldwide in 2016: An update**. *World journal of gastroenterology* 2017, **23**(20):3632-3642.
67. Denters MJ, Deutekom M, Fockens P, Bossuyt PMM, Dekker E: **Implementation of population screening for colorectal cancer by repeated fecal occult blood test in the Netherlands**. *BMC gastroenterology* 2009, **9**:28-28.
68. Katsoula A, Paschos P, Haidich A-B, Tsapas A, Giouleme O: **Diagnostic Accuracy of Fecal Immunochemical Test in Patients at Increased Risk for Colorectal Cancer: A Meta-analysis**. *JAMA Internal Medicine* 2017, **177**(8):1110-1118.
69. Elsafi SH, Alqahtani NI, Zakary NY, Al Zahrani EM: **The sensitivity, specificity, predictive values, and likelihood ratios of fecal occult blood test for the detection of colorectal cancer in hospital settings**. *Clin Exp Gastroenterol* 2015, **8**:279-284.
70. Lurie JD, Welch HG: **Diagnostic Testing Following Fecal Occult Blood Screening in the Elderly**. *JNCI: Journal of the National Cancer Institute* 1999, **91**(19):1641-1646.
71. Guo F, Chen C, Schöttker B, Holleczek B, Hoffmeister M, Brenner H: **Changes in colorectal cancer screening use after introduction of alternative screening offer in Germany: Prospective cohort study**. *International Journal of Cancer*, **n/a**(n/a).
72. Imperiale TF, Ransohoff DF, Itzkowitz SH, *et al*: **Multitarget Stool DNA Testing for Colorectal-Cancer Screening**. *New England Journal of Medicine* 2014, **370**(14):1287-1297.

73. Ahlquist DA, Skoletsky JE, Boynton KA, *et al*: **Colorectal cancer screening by detection of altered human DNA in stool: Feasibility of a multitarget assay panel**. *Gastroenterology* 2000, **119**(5):1219-1227.

74. Mohammadi A, Mansoori B, Baradaran B: **The role of microRNAs in colorectal cancer**. *Biomedicine & Pharmacotherapy* 2016, **84**:705-713.

75. Sabry D, El-Deek SEM, Maher M, *et al*: **Role of miRNA-210, miRNA-21 and miRNA-126 as diagnostic biomarkers in colorectal carcinoma: impact of HIF-1alpha-VEGF signaling pathway**. *Mol Cell Biochem* 2018.

76. Kanth P, Hazel MW, Boucher KM, et al: **Small RNA sequencing of sessile serrated polyps identifies microRNA profile associated with colon cancer**. *Genes Chromosomes Cancer* 2018.

77. Yörüker EE, Holdenrieder S, Gezer U: **Blood-based biomarkers for diagnosis, prognosis and treatment of colorectal cancer**. *Clinica Chimica Acta* 2016, **455**:26-32.

78. Reinert T, Schøler LV, Thomsen R, *et al*: **Analysis of circulating tumour DNA to monitor disease burden following colorectal cancer surgery**. *Gut* 2016, **65**(4):625-634.

79. Lecomte T, Berger A, Zinzindohoué F, *et al*: **Detection of free-circulating tumor-associated DNA in plasma of colorectal cancer patients and its association with prognosis**. *International Journal of Cancer* 2002, **100**(5):542-548.

80. Butler TM, Spellman PT, Gray J: **Circulating-tumor DNA as an early detection and diagnostic tool**. *Current Opinion in Genetics & Development* 2017, **42**:14-21.

81. Zhang SY, Lin M, Zhang HB: **Diagnostic value of carcinoembryonic antigen and carcinoma antigen 19-9 for colorectal carcinoma**. *Int J Clin Exp Pathol* 2015, **8**(8):9404-9409.

82. Yakabe T, Nakafusa Y, Sumi K, *et al*: **Clinical significance of CEA and CA19-9 in postoperative follow-up of colorectal cancer**. *Ann Surg Oncol* 2010, **17**(9):2349-2356.

83. Lee JH, Park SH, Lee SS, *et al*: **CT colonography in patients who have undergone sigmoid colostomy: a feasibility study**. *AJR Am J Roentgenol* 2011, **197**(4):W653-657.

84. Yucel C, Lev-Toaff AS, Moussa N, Durrani H: **CT colonography for incomplete or contraindicated optical colonoscopy in older patients**. *AJR Am J Roentgenol* 2008, **190**(1):145-150.
85. Rex DK, Rahmani EY, Haseman JH, Lemmel GT, Kaster S, Buckley JS: **Relative sensitivity of colonoscopy and barium enema for detection of colorectal cancer in clinical practice**. *Gastroenterology* 1997, **112**(1):17-23.
86. McCarthy PA, Rubesin SE, Levine MS, *et al*: **Colon cancer: morphology detected with barium enema examination versus histopathologic stage**. *Radiology* 1995, **197**(3):683-687.
87. Hetta W, Niazi G, Abdelbary MH: **Accuracy of 18F-FDG PET/CT in monitoring therapeutic response and detection of loco-regional recurrence and metastatic deposits of colorectal cancer in comparison to CT**. *Egyptian Journal of Radiology and Nuclear Medicine* 2020, **51**(1):37.
88. O'Connor OJ, McDermott S, Slattery J, Sahani D, Blake MA: **The Use of PET-CT in the Assessment of Patients with Colorectal Carcinoma**. *International Journal of Surgical Oncology* 2011, **2011**:846512.
89. Carethers JM: **Fecal DNA Testing for Colorectal Cancer Screening**. *Annu Rev Med* 2020, **71**:59-69.
90. Wang S, Ang HM, Tade MO: **Volatile organic compounds in indoor environment and photocatalytic oxidation: State of the art**. *Environment International* 2007, **33**(5):694-705.
91. Kansal A: **Sources and reactivity of NMHCs and VOCs in the atmosphere: A review**. *Journal of Hazardous Materials* 2009, **166**(1):17-26.
92. Camarasu CC: **Headspace SPME method development for the analysis of volatile polar residual solvents by GC-MS**. *Journal of Pharmaceutical and Biomedical Analysis* 2000, **23**(1):197-210.
93. San Juan PM, Carrillo JD, Tena MT: **Fibre selection based on an overall analytical feature comparison for the solid-phase microextraction of trihalomethanes from drinking water**. *Journal of Chromatography A* 2007, **1139**(1):27-35.
94. Linehan WM, Srinivasan R, Schmidt LS: **The genetic basis of kidney cancer: a metabolic disease**. *Nature Reviews Urology* 2010, **7**(5):277-285.

95. Coller HA: **Is Cancer a Metabolic Disease?** *The American Journal of Pathology* 2014, **184**(1):4-17.

96. Seyfried TN, Shelton LM: **Cancer as a metabolic disease**. *Nutrition & Metabolism* 2010, **7**(1):7.

97. Krilaviciute A, Heiss JA, Leja M, Kupcinskas J, Haick H, Brenner H: **Detection of cancer through exhaled breath: a systematic review**. *Oncotarget* 2015, **6**(36):38643-38657.

98. Peng G, Hakim M, Broza YY, *et al*: **Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors**. *British Journal of Cancer* 2010, **103**(4):542-551.

99. Amann A, Costello BdL, Miekisch W, *et al*: **The human volatilome: volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva**. *Journal of Breath Research* 2014, **8**(3):034001.

100. Abaffy T, Möller MG, Riemer DD, Milikowski C, DeFazio RA: **Comparative analysis of volatile metabolomics signals from melanoma and benign skin: a pilot study**. *Metabolomics* 2013, **9**(5):998-1008.

101. de Boer NKH, de Meij TGJ, Oort FA, *et al*: **The Scent of Colorectal Cancer: Detection by Volatile Organic Compound Analysis**. *Clinical Gastroenterology and Hepatology* 2014, **12**(7):1085-1089.

102. Wolfsdorf J, Glaser N, Sperling MA: **Diabetic Ketoacidosis in Infants, Children, and Adolescents**. *A consensus statement from the American Diabetes Association* 2006, **29**(5):1150-1159.

103. Lockwood AH, Yap EWH, Wong W-H: **Cerebral Ammonia Metabolism in Patients with Severe Liver Disease and Minimal Hepatic Encephalopathy**. *Journal of Cerebral Blood Flow & Metabolism* 1991, **11**(2):337-341.

104. Pauling L, Robinson AB, Teranishi R, Cary P: **Quantitative Analysis of Urine Vapor and Breath by Gas-Liquid Partition Chromatography**. *Proceedings of the National Academy of Sciences* 1971, **68**(10):2374-2376.

105. de Lacy Costello B, Amann A, Al-Kateb H, *et al*: **A review of the volatiles from the healthy human body**. *Journal of Breath Research* 2014, **8**(1):014001.

106. Phillips M, Cataneo RN, Cummin ARC, *et al*: **Detection of Lung Cancer With Volatile Markers in the Breatha**. *Chest* 2003, **123**(6):2115-2123.

107. Phillips M, Cataneo RN, Ditkoff BA, *et al*: **Volatile Markers of Breast Cancer in the Breath**. *The Breast Journal* 2003, **9**(3):184-191.
108. Schallschmidt K, Becker R, Jung C, *et al*: **Comparison of volatile organic compounds from lung cancer patients and healthy controls-challenges and limitations of an observational study**. *J Breath Res* 2016, **10**(4):046007.
109. Kumar S, Huang J, Abbassi-Ghadi N, *et al*: **Mass Spectrometric Analysis of Exhaled Breath for the Identification of Volatile Organic Compound Biomarkers in Esophageal and Gastric Adenocarcinoma**. *Annals of Surgery* 2015, **262**(6):981-990.
110. Krilaviciute A, Heiss JA, Leja M, Kupcinskas J, Haick H, Brenner H: **Detection of cancer through exhaled breath: a systematic review**. *Oncotarget* 2015, **6**(36):38643-38657.
111. Altomare DF, Di Lena M, Porcelli F, *et al*: **Exhaled volatile organic compounds identify patients with colorectal cancer**. *Br J Surg* 2013, **100**(1):144-150.
112. Filipiak W, Sponring A, Filipiak A, *et al*: **TD-GC-MS Analysis of Volatile Metabolites of Human Lung Cancer and Normal Cells In vitro**. *Cancer Epidemiology Biomarkers & Prevention* 2010, **19**(1):182-195.
113. Phillips M, Cataneo RN, Cheema T, Greenberg J: **Increased breath biomarkers of oxidative stress in diabetes mellitus**. *Clinica Chimica Acta* 2004, **344**(1):189-194.
114. Pijls KE, Smolinska A, Jonkers DMAE, *et al* **A profile of volatile organic compounds in exhaled air as a potential non-invasive biomarker for liver cirrhosis**. *Scientific Reports* 2016, **6**(1):19903.
115. Shirasu M, Touhara K: **The scent of disease: volatile organic compounds of the human body related to disease and disorder**. *The Journal of Biochemistry* 2011, **150**(3):257-266.
116. Wilson AD, Baietto M: **Advances in electronic-nose technologies developed for biomedical applications**. *Sensors (Basel)* 2011, **11**(1):1105-1176.
117. Niccolai E, Baldi S, Ricci F, *et al*: **Evaluation and comparison of short chain fatty acids composition in gut diseases**. *World Journal of Gastroenterology* 2019, **25**(36):5543-5558.

118. Vockley J, Ensenauer R: **Isovaleric acidemia: New aspects of genetic and phenotypic heterogeneity**. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 2006, **142C**(2):95-103.

119. Syhre M, Chambers ST: **The scent of Mycobacterium tuberculosis**. *Tuberculosis* 2008, **88**(4):317-323.

120. Garner CE, Smith S, Costello BdL, *et al*: **Volatile organic compounds from feces and their potential for diagnosis of gastrointestinal disease**. *The FASEB Journal* 2007, **21**(8):1675-1688.

121. Yao S-S, Guo W-F, Lu Y, Jiang Y-X: **Flavor Characteristics of Lapsang Souchong and Smoked Lapsang Souchong, a Special Chinese Black Tea with Pine Smoking Process**. *Journal of Agricultural and Food Chemistry* 2005, **53**(22):8688-8693.

122. Fens N, Zwinderman AH, van der Schee MP, *et al*: **Exhaled Breath Profiling Enables Discrimination of Chronic Obstructive Pulmonary Disease and Asthma**. *American Journal of Respiratory and Critical Care Medicine* 2009, **180**(11):1076-1082.

123. Delfino RJ, Gong H, Linn WS, Hu Y, Pellizzari ED: **Respiratory symptoms and peak expiratory flow in children with asthma in relation to volatile organic compounds in exhaled breath and ambient air**. *Journal of Exposure Science & Environmental Epidemiology* 2003, **13**(5):348-363.

124. Cikach FS, Dweik RA: **Cardiovascular Biomarkers in Exhaled Breath**. *Progress in Cardiovascular Diseases* 2012, **55**(1):34-43.

125. Arai M, Yuzawa H, Nohara I, *et al*: **Enhanced carbonyl stress in a subpopulation of schizophrenia**. *Arch Gen Psychiatry* 2010, **67**(6):589-597.

126. Mitchell KJ, Porteous DJ: **Rethinking the genetic architecture of schizophrenia**. *Psychol Med* 2011, **41**(1):19-32.

127. Brooke BS, Schwartz TA, Pawlik TM: **MOOSE Reporting Guidelines for Meta-analyses of Observational Studies**. *JAMA Surgery* 2021, **156**(8):787-788.

128. Mezmale L, Leja M, Lescinska AM, et al: **Identification of Volatile Markers of Colorectal Cancer from Tumor Tissues Using Volatilomic Approach**. *Molecules* 2023, **28**(16).

129. Smielowska M, Ligor T, Kupczyk W, Szeliga J, Jackowski M, Buszewski B: **Screening for volatile biomarkers of colorectal cancer by analyzing breath and fecal samples using thermal desorption combined with GC-MS (TD-GC-MS)**. *J Breath Res* 2023, **17**(4).

130. Boulind CE, Gould O, de Lacy Costello B, *et al*: **Urinary Volatile Organic Compound Testing in Fast-Track Patients with Suspected Colorectal Cancer**. *Cancers (Basel)* 2022, **14**(9).

131. De Vietro N, Aresta A, Rotelli MT, *et al*: **Relationship between cancer tissue derived and exhaled volatile organic compound from colorectal cancer patients. Preliminary results**. *J Pharm Biomed Anal* 2020, **180**:113055.

132. Markar SR, Chin ST, Romano A, *et al*: **Breath Volatile Organic Compound Profiling of Colorectal Cancer Using Selected Ion Flow-tube Mass Spectrometry**. *Ann Surg* 2019, **269**(5):903-910.

133. Amal H, Leja M, Funka K, *et al*: **Breath testing as potential colorectal cancer screening tool**. *International Journal of Cancer* 2016, **138**(1):229-236.

134. Altomare DF, Di Lena M, Porcelli F, *et al*: **Effects of Curative Colorectal Cancer Surgery on Exhaled Volatile Organic Compounds and Potential Implications in Clinical Follow-up**. *Annals of Surgery* 2015, **262**(5):862-866; discussion 866-867.

135. Wang C, Ke C, Wang X, *et al*: **Noninvasive detection of colorectal cancer by analysis of exhaled breath**. *Analytical & Bioanalytical Chemistry* 2014, **406**(19):4757-4763.

136. Bond A, Greenwood R, Lewis S, *et al*: **Volatile organic compounds emitted from faeces as a biomarker for colorectal cancer**. *Alimentary Pharmacology & Therapeutics* 2019, **49**(8):1005-1012.

137. Batty CA, Cauchi M, Lourenco C, Hunter JO, Turner C: **Use of the Analysis of the Volatile Faecal Metabolome in Screening for Colorectal Cancer**. *PLoS ONE [Electronic Resource]* 2015, **10**(6):e0130301.

138. Kim S, Yin X, Prodhan MAI, Zhang X, Zhong Z, Kato I: **Global Plasma Profiling for Colorectal Cancer-Associated Volatile Organic Compounds: a Proof-of-Principle Study**. *Journal of Chromatographic Science* 2019, **57**(5):385-396.

139. Wang C, Li P, Lian A, *et al*: **Blood volatile compounds as biomarkers for colorectal cancer**. *Cancer Biology & Therapy* 2014, **15**(2):200-206.

140. Rozhentsov AA, Koptina AV, Mitrakov AA, *et al*: **A new method to diagnose cancerbased on image analysis of mass chromatograms of volatile organic compounds in urine**. *Sovremennye Tehnologii v Medicine* 2014, **6**:151-157.

141. Gohlke RS: **Time-of-Flight Mass Spectrometry and Gas-Liquid Partition Chromatography**. *Analytical Chemistry* 1959, **31**(4):535-541.

142. Tsivou M, Kioukia-Fougia N, Lyris E, *et al*: **An overview of the doping control analysis during the Olympic Games of 2004 in Athens, Greece**. *Analytica Chimica Acta* 2006, **555**:1-13.

143. Smith PA, Lepage CJ, Lukacs M, Martin N, Shufutinsky A, Savage PB: **Field-portable gas chromatography with transmission quadrupole and cylindrical ion trap mass spectrometric detection: Chromatographic retention index data and ion/molecule interactions for chemical warfare agent identification**. *International Journal of Mass Spectrometry* 2010, **295**(3):113-118.

144. Adams NG, Smith D: **The selected ion flow tube (SIFT); A technique for studying ion-neutral reactions**. *International Journal of Mass Spectrometry and Ion Physics* 1976, **21**(3):349-359.

145. Seiyama T, Kagawa S: **Study on a Detector for Gaseous Components Using Semiconductive Thin Films**. *Analytical Chemistry* 1966, **38**(8):1069-1073.

146. Schüler M, Helwig N, Schütze A, Sauerwald T, Ventura G: **Detecting trace-level concentrations of volatile organic compounds with metal oxide gas sensors**. In: *SENSORS, 2013 IEEE: 3-6 Nov. 2013 2013*; 2013: 1-4.

147. Bouvet M, Parra V, Locatelli C, Xiong H: **Electrical transduction in phthalocyanine-based gas sensors: from classical chemiresistors to new functional structures**. *Journal of Porphyrins and Phthalocyanines* 2009, **13**(01):84-91.

148. Ghoorchian A, Alizadeh N: **Chemiresistor gas sensor based on sulfonated dye-doped modified conducting polypyrrole film for high sensitive detection of 2,4,6-trinitrotoluene in air**. *Sensors and Actuators B: Chemical* 2018, **255**:826-835.

149. Persaud K, Dodd G: **Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose**. *Nature* 1982, **299**(5881):352-355.

150. Röck F, Barsan N, Weimar U: **Electronic Nose: Current Status and Future Trends**. *Chemical Reviews* 2008, **108**(2):705-725.

151. Westenbrink E, Arasaradnam RP, O'Connell N, *et al*: **Development and application of a new electronic nose instrument for the detection of colorectal cancer**. *Biosensors and Bioelectronics* 2015, **67**:733-738.

152. Covington JA, van. der Schee MP, Edge ASL, Boyle B, Savage RS, Arasaradnam RP: **The application of FAIMS gas analysis in medical diagnostics**. *Analyst* 2015, **140**(20):6775-6781.

153. Saba J, Bonneil E, Pomiès C, Eng K, Thibault P: **Enhanced Sensitivity in Proteomics Experiments Using FAIMS Coupled with a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer**. *Journal of Proteome Research* 2009, **8**(7):3355-3366.

154. Bonneil E, Pfammatter S, Thibault P: **Enhancement of mass spectrometry performance for proteomic analyses using high-field asymmetric waveform ion mobility spectrometry (FAIMS)**. *Journal of Mass Spectrometry* 2015, **50**(11):1181-1195.

155. Costanzo MT, Boock JJ, Kemperman RHJ, Wei MS, Beekman CR, Yost RA: **Portable FAIMS: Applications and future perspectives**. *International Journal of Mass Spectrometry* 2017, **422**:188-196.

156. Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, *et al*: **Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis**. *PLoS One* 2014, **9**(9):e108750.

157. Arasaradnam RP, Ouaret N, Thomas MG, *et al*: **A Novel Tool for Noninvasive Diagnosis and Tracking of Patients with Inflammatory Bowel Disease**. *Inflammatory Bowel Diseases* 2013, **19**(5):999-1003.

158. Ghimenti S, Lomonaco T, Bellagambi FG, *et al*: **Comparison of sampling bags for the analysis of volatile organic compounds in breath**. *Journal of Breath Research* 2015, **9**(4):047110.

159. Steeghs MML, Cristescu SM, Harren FJM: **The suitability of Tedlar bags for breath sampling in medical diagnostic research**. *Physiological Measurement* 2006, **28**(1):73-84.

160. Beauchamp J, Herbig J, Gutmann R, Hansel A: **On the use of Tedlar® bags for breath-gas sampling and analysis**. *Journal of Breath Research* 2008, **2**(4):046001.

161. Kasper PL, Oxbøl A, Hansen MJ, Feilberg A: **Mechanisms of Loss of Agricultural Odorous Compounds in Sample Bags of Nalophan, Tedlar, and PTFE**. *Journal of Environmental Quality* 2018, **47**(2):246-253.

162. Boshier PR, Marczin N, Hanna GB: **Repeatability of the measurement of exhaled volatile metabolites using selected ion flow tube mass spectrometry**. *Journal of the American Society for Mass Spectrometry* 2010, **21**(6):1070-1074.

163. Volden J, Thomassen Y, Greibrokk T, Thorud S, Molander P: **Stability of workroom air volatile organic compounds on solid adsorbents for thermal desorption gas chromatography**. *Analytica Chimica Acta* 2005, **530**(2):263-271.

164. **Sample Preparation Techniques in Analytical Chemistry**; 2003.

165. 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**. *British Journal of Cancer* 2011, **105**(12):1894-1904.

166. Di Lena M, Porcelli F, Altomare DF: **Volatile organic compounds as new biomarkers for colorectal cancer: a review**. *Colorectal Disease* 2016, **18**(7):654-663.

167. Altomare DF, Di Lena M, Porcelli F, *et al*: **Effects of Curative Colorectal Cancer Surgery on Exhaled Volatile Organic Compounds and Potential Implications in Clinical Follow-up**. *Annals of Surgery* 2015, **262**(5):862-867.

168. Spratlin JL, Serkova NJ, Eckhardt SG: **Clinical Applications of Metabolomics in Oncology: A Review**. *Clinical Cancer Research* 2009, **15**(2):431-440.

169. Herbig J, Beauchamp J: **Towards standardization in the analysis of breath gas volatiles**. *Journal of Breath Research* 2014, **8**(3):037101.

170. Beauchamp JD, Pleil JD: **Simply breath-taking? Developing a strategy for consistent breath sampling**. *Journal of Breath Research* 2013, **7**(4):042001.

171. Khalid TY, Costello BDL, Ewen R, *et al*: **Breath volatile analysis from patients diagnosed with harmful drinking, cirrhosis and hepatic encephalopathy: a pilot study**. *Metabolomics* 2013, **9**(5):938-948.

172. Lewis JR, Utesch BS, Maher DE: **UMUX-LITE: when there's no time for the SUS**. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Paris, France: Association for Computing Machinery; 2013: 2099–2102.

173. Berkman MI, Karahoca D: **Re-assessing the usability metric for user experience (UMUX) scale**. *J Usability Studies* 2016, **11**(3):89–109.

174. Borsci S, Federici S, Bacci S, Gnaldi M, Bartolucci F: **Assessing User Satisfaction in the Era of User Experience: Comparison of the SUS, UMUX, and UMUX-LITE as a Function of Product Experience**. *International Journal of Human–Computer Interaction* 2015, **31**(8):484-495.

175. Project Management Institute (2021). A guide to the project management body of knowledge (PMBOK guide). Project Management Institute (7th ed.). Newtown Square, PA. ISBN 978-1-62825-664-2.

176. Chin S-T, Romano A, Doran SLF, Hanna GB: **Cross-platform mass spectrometry annotation in breathomics of oesophageal-gastric cancer**. *Scientific Reports* 2018, **8**(1):5139.

177. Moretto H-H, Schulze M, Wagner G: **Silicones**. In: *Ullmann's Encyclopedia of Industrial Chemistry.* edn.

178. ***Quantifying the User Experience.*** **edn**. Edited by Sauro J, Lewis JR. Boston: Morgan Kaufmann; 2012: xv.

179. Kind T, Wohlgemuth G, Lee DY, *et al*: **FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry**. *Analytical Chemistry* 2009, **81**(24):10038-10048.

180. Markar SR, Wiggins T, Antonowicz S, *et al*: **Assessment of a Noninvasive Exhaled Breath Test for the Diagnosis of Oesophagogastric Cancer**. *JAMA Oncology* 2018, **4**(7):970-976.

181. Markar SR, Chin ST, Romano A, *et al*: **Breath Volatile Organic Compound Profiling of Colorectal Cancer Using Selected Ion Flow-tube Mass Spectrometry**. *Annals of Surgery* 2019, **269**(5):903-910.

182. Romano A, Doran S, Belluomo I, Hanna GB: **High-Throughput Breath Volatile Organic Compound Analysis Using Thermal Desorption Proton Transfer**

Reaction Time-of-Flight Mass Spectrometry. *Analytical Chemistry* 2018, **90**(17):10204-10210.

183. Smith D, Španel P: **Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis**. *Mass Spectrom Rev* 2005, **24**:661.
184. Smith D, Španěl P, Herbig J, Beauchamp J: **Mass spectrometry for real-time quantitative breath analysis**. *J Breath Res* 2014, **8**(2):027101.
185. Spaněl P, Smith D: **Selected ion flow tube mass spectrometry for on-line trace gas analysis in biology and medicine**. *Eur J Mass Spectrom (Chichester)* 2007, **13**(1):77-82.
186. Romano A, Hanna GB: **Identification and quantification of VOCs by proton transfer reaction time of flight mass spectrometry: An experimental workflow for the optimization of specificity, sensitivity, and accuracy**. *Journal of Mass Spectrometry* 2018, **53**(4):287-295.
187. Doran SLF, Romano A, Hanna GB: **Optimisation of sampling parameters for standardised exhaled breath sampling**. *Journal of Breath Research* 2017, **12**(1):016007.
188. Kumar S, Huang J, Abbassi-Ghadi N, Španěl P, Smith D, Hanna GB: **Selected Ion Flow Tube Mass Spectrometry Analysis of Exhaled Breath for Volatile Organic Compound Profiling of Esophago-Gastric Cancer**. *Analytical Chemistry* 2013, **85**(12):6121-6128.
189. Španěl P, Smith D: **Quantification of trace levels of the potential cancer biomarkers formaldehyde, acetaldehyde and propanol in breath by SIFT-MS**. *Journal of Breath Research* 2008, **2**(4):046003.
190. Lehnert A-S, Behrendt T, Ruecker A, Pohnert G, Trumbore S: **Performance of SIFT-MS and PTR-MS in the measurement of volatile organic compounds at different humidities**; 2019.
191. Mochalski P, Wzorek B, Sliwka I, Amann A: **Suitability of different polymer bags for storage of volatile sulphur compounds relevant to breath analysis**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, **877**(3):189-196.
192. Brůhová Michalčíková R, Dryahina K, Smith D, Španěl P: **Volatile compounds released by Nalophan; implications for selected ion flow tube mass**

spectrometry and other chemical ionisation mass spectrometry analytical methods. *Rapid Communications in Mass Spectrometry* 2020, **34**(5):e8602.

193. Anderson JC: **Measuring breath acetone for monitoring fat loss: Review**. *Obesity (Silver Spring)* 2015, **23**(12):2327-2334.
194. Miekisch W, Kischkel S, Sawacki A, Liebau T, Mieth M, Schubert JK: **Impact of sampling procedures on the results of breath analysis**. *Journal of Breath Research* 2008, **2**(2):026007.
195. Filipiak W, Ruzsanyi V, Mochalski P, *et al*: **Dependence of exhaled breath composition on exogenous factors, smoking habits and exposure to air pollutants**. *Journal of breath research* 2012, **6**(3):036008-036008.
196. van Keulen KE, Jansen ME, Schrauwen RWM, Kolkman JJ, Siersema PD: **Volatile organic compounds in breath can serve as a non-invasive diagnostic biomarker for the detection of advanced adenomas and colorectal cancer**. *Alimentary Pharmacology & Therapeutics* 2020, **51**(3):334-346.
197. Woodfield G, Belluomo I, Lin G, *et al* **PTU-146 Breath testing for GI cancers-scaling up for clinical practice**. *Gut* 2018, **67**(Suppl 1):A268-A269.
198. Oakley-Girvan I, Davis SW: **Breath based volatile organic compounds in the detection of breast, lung, and colorectal cancers: A systematic review**. *Cancer Biomarkers: Section A of Disease Markers* 2017, **21**(1):29-39.
199. Bhattacharyya D, Kumar P, Mohanty SK, Smith YR, Misra M: **Detection of Four Distinct Volatile Indicators of Colorectal Cancer using Functionalized Titania Nanotubular Arrays**. *Sensors* 2017, **17**(8):04.
200. 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**. *Journal of breath research* 2010, **4**(3):031001.
201. Wang G, Li Y, Liu M, *et al*: **Determination of volatile organic compounds in SW620 colorectal cancer cells and tumor-bearing mice**. *Journal of Pharmaceutical & Biomedical Analysis* 2019, **167**:30-37.
202. Liu M, Li Y, Wang G, *et al*: **Release of volatile organic compounds (VOCs) from colorectal cancer cell line LS174T**. *Analytical Biochemistry* 2019, **581**:113340.
203. Mozdiak E, Wicaksono AN, Covington JA, Arasaradnam RP: **Colorectal cancer and adenoma screening using urinary volatile organic compound (VOC)**

detection: early results from a single-centre bowel screening population (UK BCSP). *Techniques in Coloproctology* 2019, **23**(4):343-351.

204. Adam ME, Fehervari M, Boshier PR, *et al*: **Mass-Spectrometry Analysis of Mixed-Breath, Isolated-Bronchial-Breath, and Gastric-Endoluminal-Air Volatile Fatty Acids in Esophagogastric Cancer**. *Anal Chem* 2019, **91**(5):3740-3746.

205. Romano A, Doran S, Belluomo I, Hanna GB: **High-Throughput Breath Volatile Organic Compound Analysis Using Thermal Desorption Proton Transfer Reaction Time-of-Flight Mass Spectrometry**. *Anal Chem* 2018, **90**(17):10204-10210.

206. Romano A, Hanna GB: **Identification and quantification of VOCs by proton transfer reaction time of flight mass spectrometry: An experimental workflow for the optimization of specificity, sensitivity, and accuracy**. *J Mass Spectrom* 2018, **53**(4):287-295.

207. **R: a language and environment for statistical computing** [https://www.R-project.org/]

208. Condello I, Santarpino G, Fiore F, *et al*: **Propofol pharmacokinetics and pharmacodynamics-a perspective in minimally invasive extracorporeal circulation**. *Interact Cardiovasc Thorac Surg* 2021, **33**(4):625-627.

209. Saasa V, Malwela T, Beukes M, Mokgotho M, Liu C-P, Mwakikunga B: **Sensing Technologies for Detection of Acetone in Human Breath for Diabetes Diagnosis and Monitoring**. *Diagnostics* 2018, **8**(1):12.

210. Salerno-Kennedy R, Cashman KD: **Potential applications of breath isoprene as a biomarker in modern medicine: a concise overview**. *Wiener klinische Wochenschrift* 2005, **117**(5):180-186.

211. McBride D: **Colonic biofilms are associated with right-sided colorectal cancer**. *ONS Connect* 2015, **30**(1):59.

212. Louis P, Hold GL, Flint HJ: **The gut microbiota, bacterial metabolites and colorectal cancer**. *Nature Reviews Microbiology* 2014, **12**(10):661-672.

213. Nyangale EP, Mottram DS, Gibson GR: **Gut microbial activity, implications for health and disease: the potential role of metabolite analysis**. *Journal of Proteome Research* 2012, **11**(12):5573-5585.

214. Fujino S, Miyoshi N, Kitakaze M, *et al*: **Lymph node metastasis in T1 colorectal cancer: Risk factors and prediction model**. *Oncol Lett* 2023, **25**(5):191.
215. Kumar AS, Kelleher DC, Sigle GW: **Bowel Preparation before Elective Surgery**. *Clin Colon Rectal Surg* 2013, **26**(3):146-152.
216. **Cancer** [https://www.who.int/health-topics/cancer#tab=tab_1]
217. Ishihara S, Nishikawa T, Tanaka T, *et al*: **Benefit of primary tumor resection in stage IV colorectal cancer with unresectable metastasis: a multicenter retrospective study using a propensity score analysis**. *Int J Colorectal Dis* 2015, **30**(6):807-812.
218. Eker B, Ozaslan E, Karaca H, *et al*: **Factors affecting prognosis in metastatic colorectal cancer patients**. *Asian Pac J Cancer Prev* 2015, **16**(7):3015-3021.
219. Deliu IC, Georgescu EF, Bezna MC: **Analysis of prognostic factors in colorectal carcinoma**. *Rev Med Chir Soc Med Nat Iasi* 2014, **118**(3):808-816.
220. Manceau G, Karoui M, Breton S, *et al*: **Right colon to rectal anastomosis (Deloyers procedure) as a salvage technique for low colorectal or coloanal anastomosis: postoperative and long-term outcomes**. *Dis Colon Rectum* 2012, **55**(3):363-368.
221. Coburn MC, Pricolo VE, Soderberg CH: **Factors affecting prognosis and management of carcinoma of the colon and rectum in patients more than eighty years of age**. *J Am Coll Surg* 1994, **179**(1):65-69.
222. Wanebo JH, Stearns M, Schwartz MK: **Use of CEA as an indicator of early recurrence and as a guide to a selected second-look procedure in patients with colorectal cancer**. *Ann Surg* 1978, **188**(4):481-493.
223. Coburn LA, Horst SN, Allaman MM, *et al*: **L-Arginine Availability and Metabolism Is Altered in Ulcerative Colitis**. *Inflamm Bowel Dis* 2016, **22**(8):1847-1858.
224. Graziano F, Ruzzo A, Giacomini E, *et al*: **Glycolysis gene expression analysis and selective metabolic advantage in the clinical progression of colorectal cancer**. *Pharmacogenomics J* 2016.
225. Tian Y, Xu T, Huang J, *et al*: **Tissue Metabonomic Phenotyping for Diagnosis and Prognosis of Human Colorectal Cancer**. *Scientific Reports* 2016, **6**:20790.

226. Kibi M, Nishiumi S, Kobayashi T, Kodama Y, Yoshida M: **GC/MS and LC/MS-based Tissue Metabolomic Analysis Detected Increased Levels of Antioxidant Metabolites in Colorectal Cancer**. *Kobe Journal of Medical Sciences* 2019, **65**(1):E19-E27.
227. Huang J, Mondul AM, Weinstein SJ, *et al*: **Serum metabolomic profiling of prostate cancer risk in the prostate, lung, colorectal, and ovarian cancer screening trial**. *British Journal of Cancer* 2016, **115**(9):1087-1095.
228. Tan B, Qiu Y, Zou X, *et al*: **Metabonomics identifies serum metabolite markers of colorectal cancer**. *Journal of Proteome Research* 2013, **12**(6):3000-3009.
229. Woodfield G, Belluomo I, Boshier PR, *et al*: **Feasibility and acceptability of breath research in primary care: a prospective, cross-sectional, observational study**. *BMJ Open* 2021, **11**(4):e044691.
230. Bosch S, Berkhout DJ, Ben Larbi I, de Meij TG, de Boer NK: **Fecal volatile organic compounds for early detection of colorectal cancer: where are we now?** *Journal of Cancer Research & Clinical Oncology* 2019, **145**(1):223-234.
231. Chin ST, Romano A, Doran SLF, Hanna GB: **Cross-platform mass spectrometry annotation in breathomics of oesophageal-gastric cancer**. *Sci Rep* 2018, **8**(1):5139.
232. Aksenov AA, Laponogov I, Zhang Z, *et al*: **Auto-deconvolution and molecular networking of gas chromatography-mass spectrometry data**. *Nat Biotechnol* 2021, **39**(2):169-173.
233. Kalapos MP: **On the mammalian acetone metabolism: from chemistry to clinical implications**. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2003, **1621**(2):122-139.
234. Kalapos MP: **Possible physiological roles of acetone metabolism in humans**. *Medical Hypotheses* 1999, **53**(3):236-242.
235. Reichard GA, Jr., Haff AC, Skutches CL, Paul P, Holroyde CP, Owen OE: **Plasma Acetone Metabolism in the Fasting Human**. *The Journal of Clinical Investigation* 1979, **63**(4):619-626.
236. Fenske JD, Paulson SE: **Human Breath Emissions of VOCs**. *Journal of the Air & Waste Management Association* 1999, **49**(5):594-598.

237. King J, Koc H, Unterkofler K, *et al*: **Physiological modeling of isoprene dynamics in exhaled breath**. *Journal of Theoretical Biology* 2010, **267**(4):626-637.

238. Vernia F, Valvano M, Fabiani S, *et al*: **Are Volatile Organic Compounds Accurate Markers in the Assessment of Colorectal Cancer and Inflammatory Bowel Diseases? A Review**. *Cancers* 2021, **13**(10):2361.

239. Mochalski P, Leja M, Gasenko E, *et al*: **Ex vivo emission of volatile organic compounds from gastric cancer and non-cancerous tissue**. *Journal of Breath Research* 2018, **12**(4):046005.

240. Altomare D F DLMPFTLTETMDSMV, de Gennaro G: **Exhaled volatile organic compounds identify patients with colorectal cancer**. *Br J Surg* 2013, **100**:144.

241. Bhattacharyya D, Kumar P, Mohanty SK, Smith YR, Misra M: **Detection of Four Distinct Volatile Indicators of Colorectal Cancer using Functionalized Titania Nanotubular Arrays**. *Sensors (Basel)* 2017, **17**(8):1795.

242. Amal H, Leja M, Funka K, *et al*: **Breath testing as potential colorectal cancer screening tool**. *International Journal of Cancer* 2016, **138**(1):229-236.

243. Altomare DF, Picciariello A, Rotelli MT, *et al*: **Chemical signature of colorectal cancer: case–control study for profiling the breath print**. *BJS Open* 2020, **4**(6):1189-1199.

244. Xu Zq, Broza YY, Ionsecu R, *et al*: **A nanomaterial-based breath test for distinguishing gastric cancer from benign gastric conditions**. *British journal of cancer* 2013, **108**(4):941-950.

245. Kang S, Paul Thomas CL: **How long may a breath sample be stored for at -80 °C? A study of the stability of volatile organic compounds trapped onto a mixed Tenax:Carbograph trap adsorbent bed from exhaled breath**. *J Breath Res* 2016, **10**(2):026011.

246. **NCCN guideline** [https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1428]

247. Cancer AJCo: **AJCC Cancer Staging Manual, 8th Edition** Springer; 2016.

248. Hanna GB, Boshier PR, Markar SR, Romano A: **Accuracy and Methodologic Challenges of Volatile Organic Compound-Based Exhaled Breath Tests for Cancer Diagnosis: A Systematic Review and Meta-analysis**. *JAMA Oncol* 2019, **5**(1):e182815.

249. 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* 2010, **4**(3):031001.

250. Markar SR, Brodie B, Chin ST, Romano A, Spalding D, Hanna GB: **Profile of exhaled-breath volatile organic compounds to diagnose pancreatic cancer**. *Br J Surg* 2018, **105**(11):1493-1500.

APPENDICES

1. breath concentrator -- **UWE sorbent tube breath sampling device**

Introduction

A device to permit the sampling of the breath of a user during exhalation was designed and constructed. The device was connected to a computer via a USB interface. A computer program was written to control the device, and to log data from it.

Design considerations

The device consists of a tube, into which a user breathes, with a tube branching from the main tube through which a sample of the exhaled breath may be pumped. An absorption tube packed with a suitable absorbent material may be incorporated into this branch to accumulate a sample of the exhaled air.

Although the concept of the device is simple, the implementation of it required consideration of the following aspects:

1. The device needed to be portable, so all components needed to be housed in a small enclosure.
2. Breath has a high level of humidity, therefore a primary consideration was to reduce as far as possible the chance of water condensing inside the unit. A combination of (a) heating the tubes, and (b) maintaining a gentle flow through the main tube at all times, except when sampling, was adopted.

3. A replaceable mouthpiece fitted with a non-return valve to prevent the user drawing air out of the unit was essential. A similar non-return valve was attached to the outlet of the device.
4. Sensors to measure the temperature and humidity inside the main tube were required.

5. In order to measure the volume of breath exhaled and the volume of the breath sample pumped through the absorption tube mass flow sensors were required in both branches.
6. As the absorbent material is packed tightly in the absorption tube, a compact pump capable of pumping air through such a tube was required.
7. The absorption tubes should be easily accessible and easily replaceable from the outside of the unit;
8. A USB link to a laptop computer was required so that software could be used to:
   a. control the operation of the sampling pump;
   b. control the operation of the fan generating the gentle air flow through the main tube;
   c. log data from the sensors;
   d. pass instructions to the user;
   e. display the current status and results.

Construction

The breath sampling device was designed to fit into an ABS case with aluminium end panels of dimensions 76 x 160 x 250 mm [Hammond; 1598GBK; F; 722686]1. All components were mounted on a base-plate cut from 4 mm polystyrene sheet to fit the outer case. A diagram of the device is shown in Figure 1a. Two photographs of the device are shown in Figure 1b.

A standard domestic plumbing plastic push-in T-piece for 15 mm diameter tubing was used as the main chamber into which the user would exhale. The stem of the T was used for the user's input, whilst the other two branches were used to connect to the high-flow (0 to 200 l/min) mass flow sensor [Sensirion; EM1; F; 1207226] and to the absorption tube. Holes were drilled in the bar of the T to accommodate a temperature sensor [National Semiconductor; LM35DZ ; F; 9488200] and a humidity sensor [Honeywell; HIH-4000-001; F; 1187547] each mounted on 4-way connectors [Binder; Series 719; RS; 464-404]. The outlet from the EM1 mass flow sensor was connected via a 40 mm diameter, 12 V fan [EBM Papst; 412FH; F; 9601279] mounted on end panel

2 of the outer case. Heating elements [DBK; HP03-1/04-24; F; 4408263], mounted on an aluminium plate to distribute the heat, were fitted to the base plate below the T-piece and the EM1 mass flow sensor. A thermal switch [Honeywell S&C; 2455R-100-73; F; 1082191] (cut-out at 50oC; reset at 35oC) was connected in series with the power connection to the heaters to protect the unit against overheating. A second LM35DZ temperature sensor was used to monitor the temperature within the outer case outside the flow tube.

The absorption tubes were of diameter 6 mm and length 89 mm [Perkin-Elmer (Supelco); Y37071]. Each end could be sealed using a Swagelok fitting. A Swagelok fitting with a tapered thread on one end and a standard tubing fitting on the other was modified by incorporating an O-ring so that an absorption tube inserted through it would form an airtight seal. The tapered thread was screwed into a length of 15mm

outside diameter domestic plumbing plastic tubing. The plain end of the plastic tube was inserted through a 15mm hole drilled in end panel 1 of the outer case into the push-in T-piece. An absorption tube could then be inserted through the modified Swagelok fitting from the outside of the case into the T-piece. The outer end of the absorption tube was fitted with a Swagelok connector and connected via two right-angle connectors to another Swagelok connector mounted on end plate 1 adjacent to the absorption tube connector, which was connected to the second mass flow sensor (0 to 400 ml/min) [Sensirion; ASF1400; F; 1207217].

The outlet from the ASF1400 mass flow sensor was connected through a flow restrictor to a miniature 12 V pump [NMP830 KNDC 12V; KNF Neuberger UK Ltd., Avenue Two, Station Lane Industrial Estate, Witney, Oxon OX28 4FA, UK] fitted with a filter/silencer on the exhaust port. The pump was bolted to the base plate using O-rings as washers to minimise noise and vibration.

The output signals from both Sensirion mass flow sensors were via RS-232 interfaces. Modules to convert the RS-232 signal to USB were used [FTDI; EVAL232R; F; 1146041]. The RS-232 connector (9-pin D-connector) and the USB connector were de-soldered

and removed from each of the EVAL232R modules to reduce the size. The 9-pin D-connectors were replaced with 4-pin right-angle Molex KK square pin connectors [Molex; 22-05-7048; F; 9731628] and the USB connections were wired directly to the printed circuit board.

An 11-channel voltage to USB board [ADC11/12 USB data-logger; Pico Technology, James House, Marlborough Road, Colmworth Business Park, Eaton Socon, St Neots, Cambridgeshire PE19 8YP, UK] was used to convert voltage signals and the outputs

from the temperature and humidity sensors into a USB signal. As well as 11 analogue voltage inputs, the Pico ADC11/12 board possessed two digital outputs (DO1 & DO2).

A control circuit was designed and constructed on stripboard. A schematic of this circuit is shown in Figure 1c. This board performed a number of functions, as follows:

1. supplying power to, and conditioning the signals from, the temperature sensors;
2. supplying power to, and conditioning the signals from, the humidity sensor;
3. supplying power to the heaters;
4. supplying power to the fan, switchable via a relay [Meder; SIL05-1A72-71D; F; 1079459] controlled by Pico DO1;
5. supplying power to the pump, switchable via a relay [TE Connectivity/Schrack; RYA31005; F; 9659730]controlled by Pico DO2;
6. supplying power to the mass flow sensors EM1 and ASF1400;
7. creating voltage signals to permit the software to determine the (on or off) power state of the whole unit, the fan and the pump;
8. supplying power to an on/off indicator LED mounted on end panel 2.

The three USB outputs were routed through a 4-port USB hub [Newlink; F; 8704341] mounted inside the case so that a single USB cable could be used to connect the device to the computer.

The power inlet socket to the device was mounted on end panel 2. Power was supplied by a 12 V, 2.4 A medically approved regulated power supply [F; 1183913; no longer

available]. A 5 V rail was generated by using a 12 V to 5 V dc/dc convertor [Murata Power Solutions; LME1205SC; F; 1021405] mounted on the stripboard.

A block schematic diagram of the interconnections between the various component parts is shown in Figure 1d.

Software

Software was written to:

1. Control the operation of the breath sampling device.
2. Decode the data from the mass flow sensors.
3. Log and display information from the sensors, for example:
    a. Temperatures;
    b. Humidity;
    c. Rate of exhalation;
    d. Volume of air exhaled;
    e. Volume of air pumped through the sample tube.
4. Display instructions to the user.

The software is available under licence from UWE2.

Operation of the breath sampling device

A screenshot of the main software window is shown in Figure 1e.

To initiate a session the user selects ‘Run’ from the main menu bar. If the unit is not powered when the user selects ‘Run’, then he/she is prompted to switch the unit on.

On power-up the software measures the system temperature. If the temperature is less than the minimum operating temperature (user selectable; default 35 oC) a message is displayed requesting the user to wait until the unit has reached this

temperature, and the controls are locked to prevent the user from proceeding further. Once the minimum operating temperature has been attained the controls are unlocked. Also on power-up the fan is switched on to provide a constant low flow of air (< 0.2 l/min)through the main tube and EM1 mass flow sensor to purge the system of previously exhaled air and to prevent the build-up of water vapour on internal surfaces. The fan runs continuously, except when a sample is being extracted.

When the user is ready to begin he/she clicks the 'Start' button. The user is prompted to check that an absorption tube has been inserted, and is also given a choice of options for the saving of the data (full, partial or none). If the user chooses to save some or all of the data then he/she is prompted for a file name and path. The user is then prompted to exhale into the tube. The bargraph display on the left side of the screen shows the rate of exhalation. The lower part of the bargraph is colour-coded yellow, the mid-range green and the upper range red; the user is encouraged to attempt to maintain a steady exhalation in the mid (green) range. When the exhalation rate exceeds a user-defined threshold the fan is switched off and the pump switched on to extract a sample of breath through the absorption tube. The pump may be programmed to continue operating once the exhalation has ceased by setting user-defined parameter called the 'Pump Run-on Time'. At the end of the exhalation, and after the Pump Run-on Time (if selected), all on-screen data fields are updated and a window showing a graphical representation of the exhalation profile is displayed. The user is prompted to repeat the exhalation until the required volume of breath has been pumped through the absorption tube.

At the end of a session the absorption tube may be removed from the apparatus, sealed with Swagelok fittings and returned to the laboratory for analysis.

User-selectable parameters and manual overrides may be accessed via the 'Utilities' window displayed by clicking 'Utilities' from the main menu bar.

Figures 1a-1e

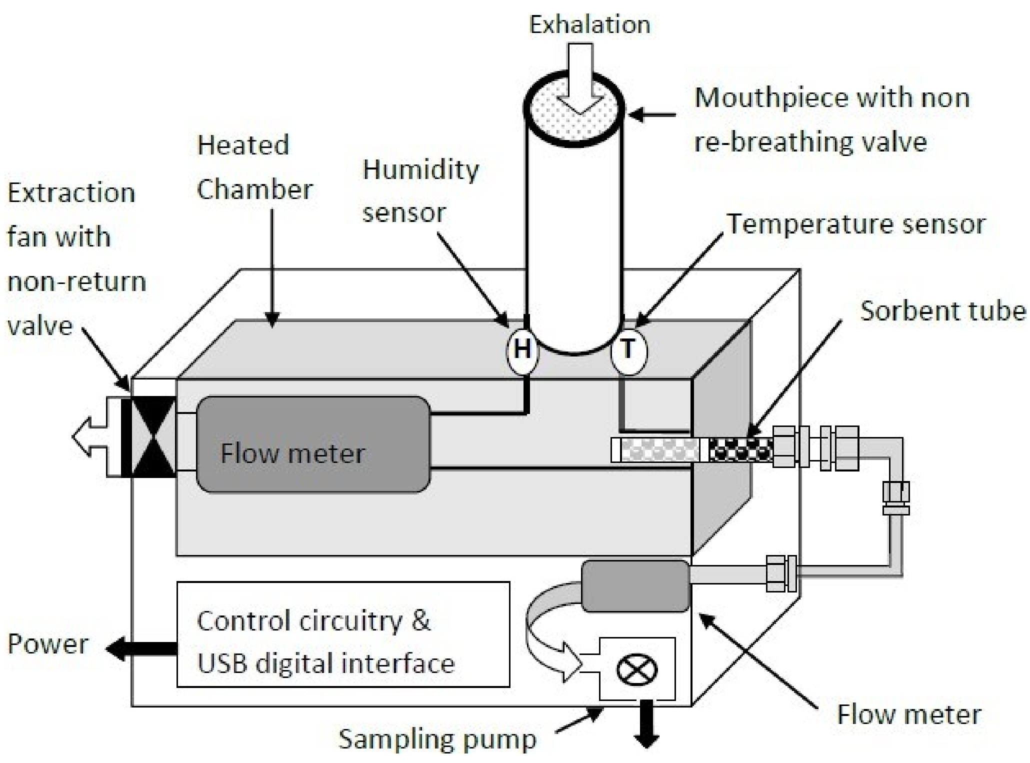
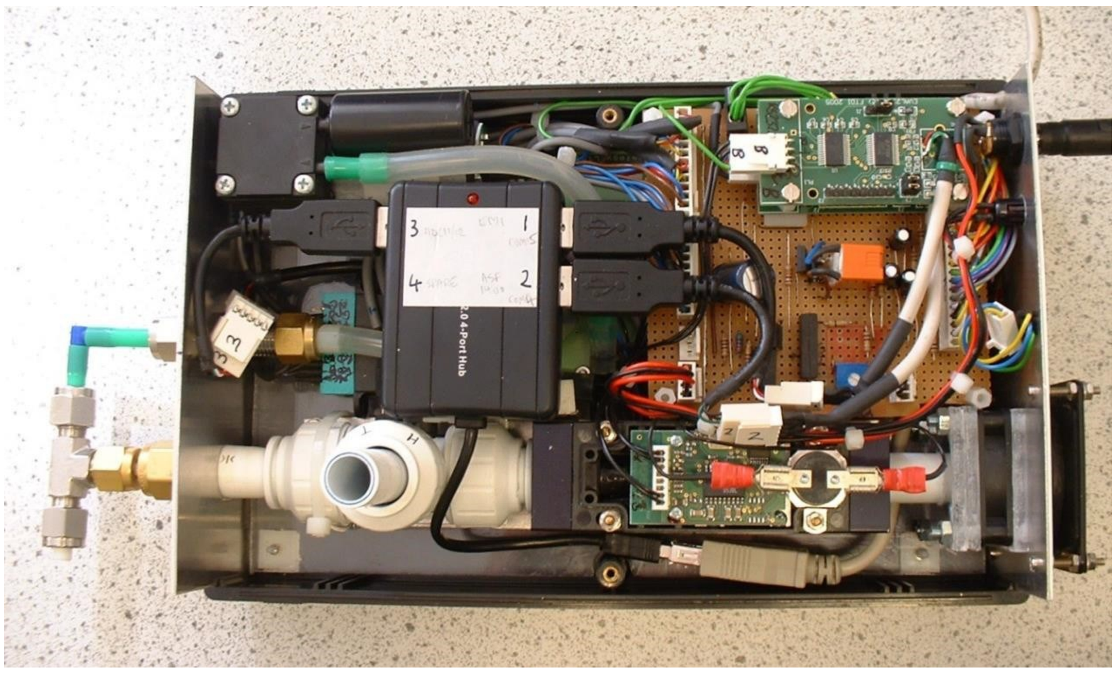


Figure 1a. Schematic diagram of sorbent tube breath sampling device.



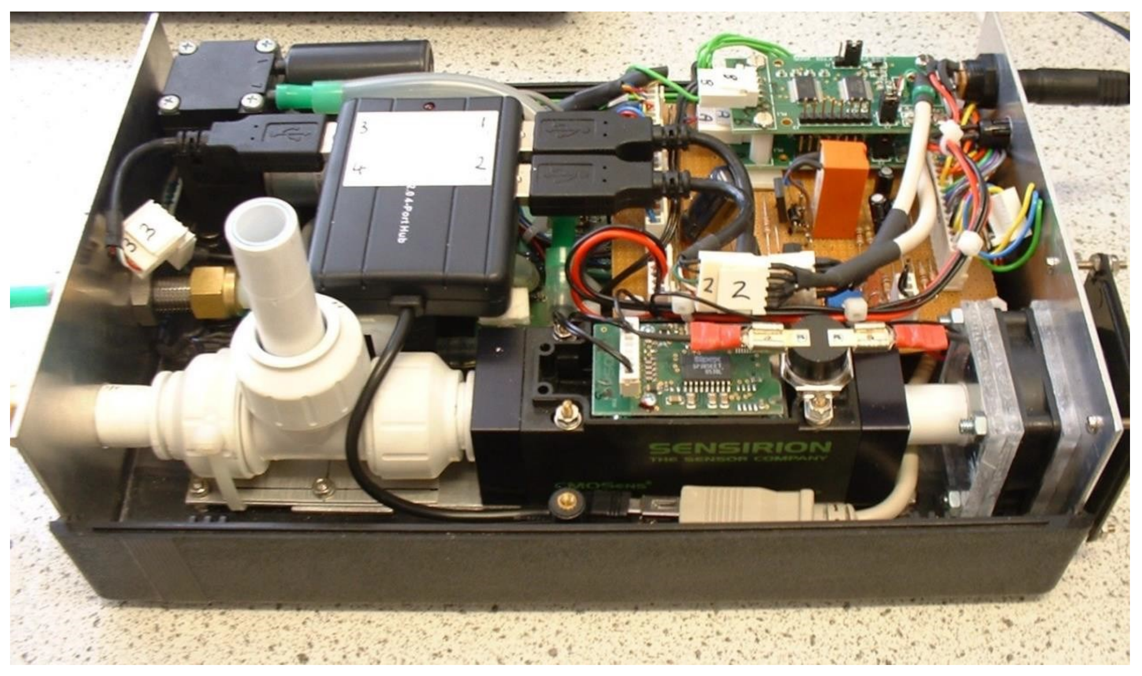


Figure 1b. Photographs of the breath sampling device with the top cover removed.

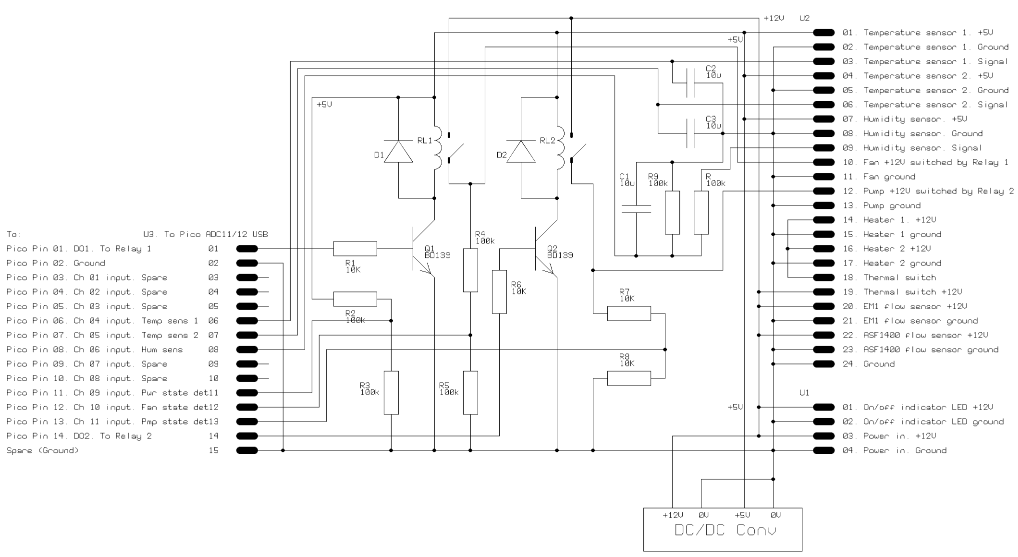


Figure 1c. Schematic of the control circuit.

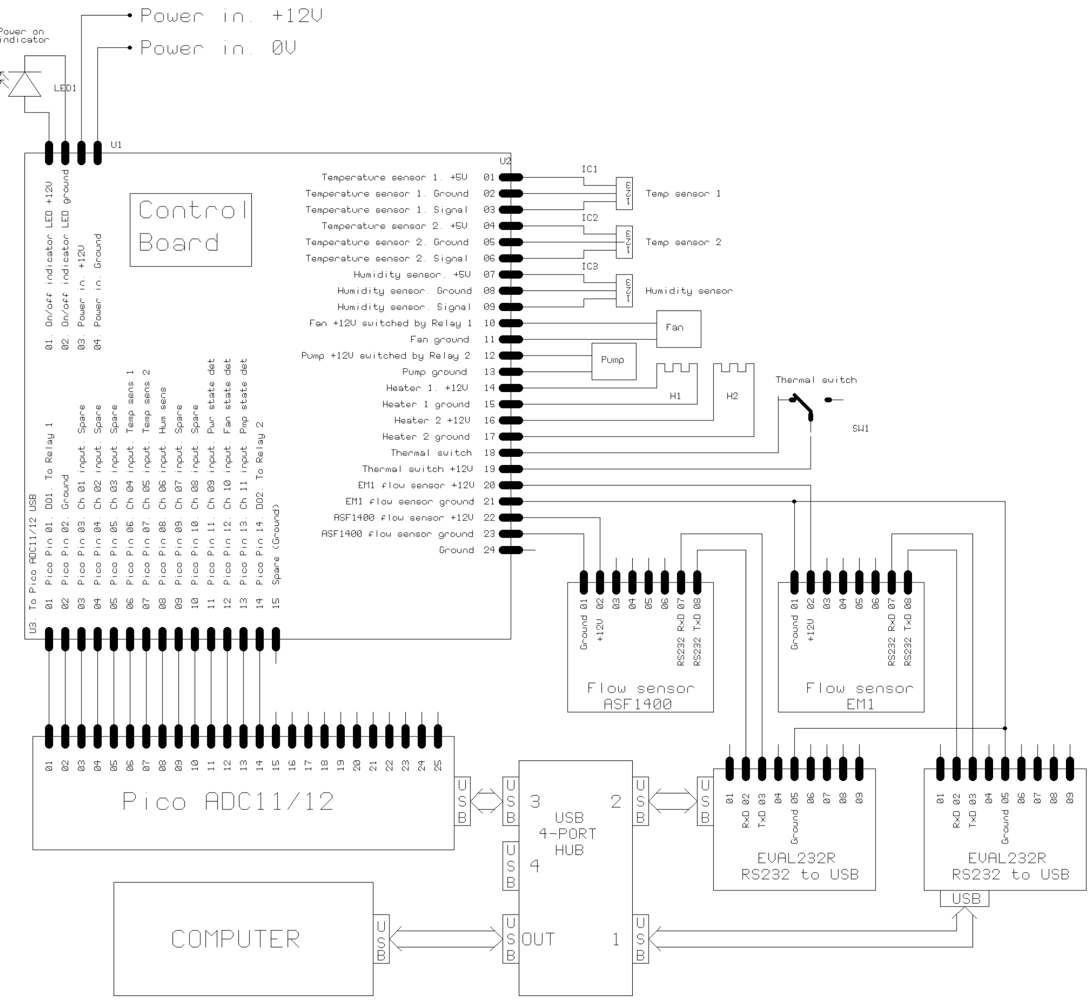


Figure 1d. Schematic block diagram of circuit interconnections.

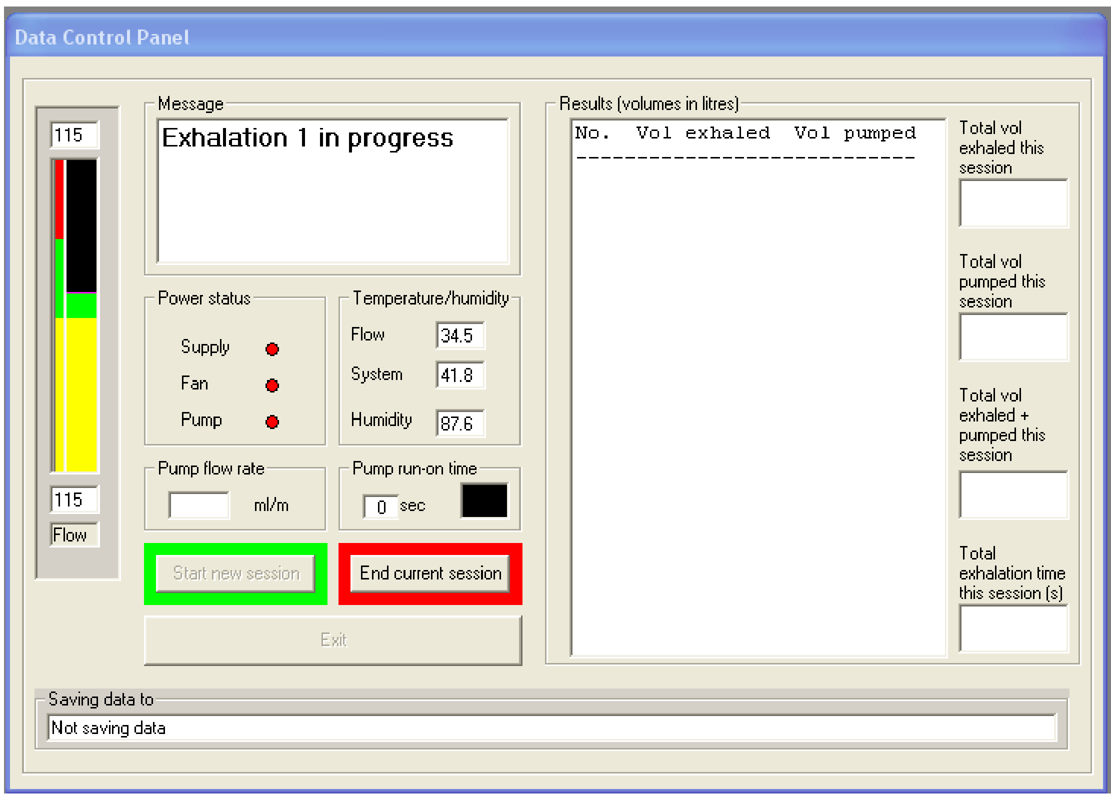


Figure 1e. Screenshot of the main screen of the breath sampling device software.

References

[1] Where the source of the component was RS Components (Birchington Road, Corby, Northants, NN17 9RS, UK), or Farnell, (150 Armley Road, Leeds, LS12 2QQ, UK), the source information will be in the following format:

[Manufacturer; Manufacturer's code (if known); Source; Order code], where the source will be abbreviated to RS or F respectively.

[2] Dr R. J. Ewen, University of the West of England, Frenchay Campus, Coldharbour Lane, Bristol BS16 1QY, UK. Email: Richard.Ewen@uwe.ac.uk

2. Questionnaire for the device comparison study

1. Information About You

We would like to collect some information about you.

1. Your occupation …

____

2. Time since last consumed food (hours)

____

3. Are you a smoker? *(please, mark only one box)*

- ☐ No
- ☐ Yes
- ☐ Yes, but only e-cigarette

3.1. (Only if you are a smoker) When did you last have a cigarette/e-cigarette? *(please, mark only one box)*

- ☐ Less than an hour ago
- ☐ One hour ago
- ☐ Two hours ago
- ☐ Three hours ago

4. Do you drink alcohol?

- ☐ No/Prefer not to say
- ☐ Yes, occasionally.
- ☐ Yes, quite regularly

4.1.(Only if you drink alcohol) When did you last drink alcohol?

(please, mark only one box)

- ☐ Less than an hour ago
- ☐ One hour ago
- ☐ Two hours ago
- ☐ Three hours ago
- ☐ Yesterday

5. Are you currently under any form of medication? *(please, mark only one box)*

- ☐ Yes. Please explain ___________________________
- ☐ No

6. Have you got any health problems or medical condition you feel we need to know? *(please, mark only one box)*

- ☐ Yes. Please explain ___________________________
- ☐ No

RECIVA

The following statements have been used by others to assess the perceived experience in the use of a tools.

In this case, we will ask to you to answer the questions in relation to the use of **RECIVA**. Please follow the instructions and fill in the questionnaire

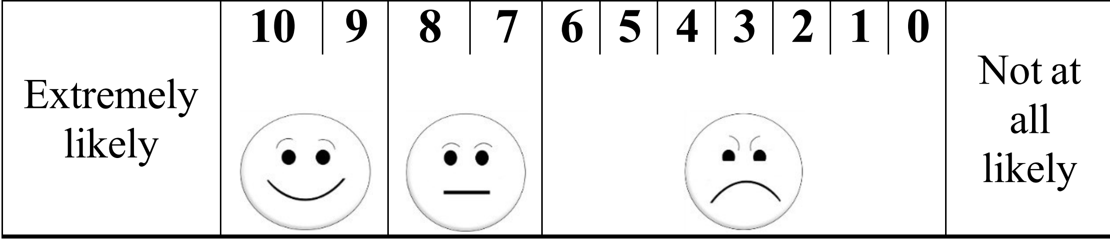
Instructions: **For each of the following statements, mark one box that best describes your overall experience with RECIVA:**

	Strongly Disagree						Strongly Agree
RECIVA capabilities meet my requirements	1	2	3	4	5	6	7
RECIVA is easy to use	1	2	3	4	5	6	7

* See for reference – Finstad, K. (2010) The usability metric for user experience.

On the basis of your experience with RECIVA, please answer the following question:

How likely is it that you would recommend the use of this RECIVA to collect breath samples at this hospital?
<u>Please</u>: Mark one box from 10 (Extremely like) to 0 (Not at all likely)



Extremely likely	10	9	8	7	6	5	4	3	2	1	0	Not at all likely

BIO VOC

The following statements have been used by others to assess the perceived experience in the use of a tools.

In this case, we will ask to you to answer the questions in relation to the use of BIO VOC. Please follow the instructions and fill in the questionnaire

Instructions: **For each of the following statements, mark one box that best describes your overall experience with BIO VOC:**

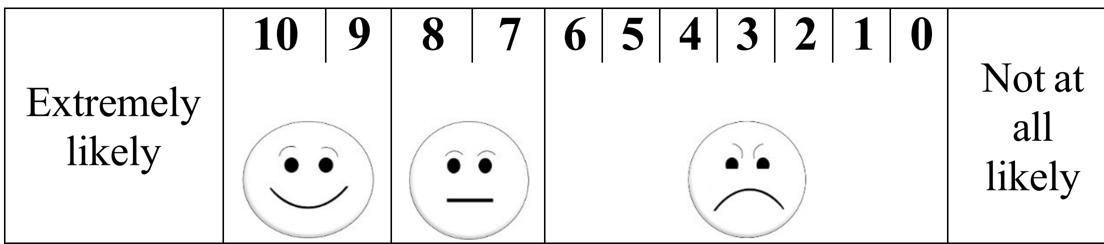
	Strongly Disagree						Strongly Agree
BIO VOC capabilities meet my requirements	1	2	3	4	5	6	7
BIO VOC is easy to use	1	2	3	4	5	6	7

* See for reference – Finstad, K. (2010) The usability metric for user experience

On the basis of your experience with the BIO VOC, please answer the following question:

How likely is it that you would recommend the use of this BIOVOC to collect breath samples at this hospital?

Please: Mark one box from 10 (Extremely like) to 0 (Not at all likely)



BREATH CONCENTRATOR

The following statements have been used by others to assess the perceived experience in the use of a tools.

In this case, we will ask to you to answer the questions in relation to the use of BREATH CONCENTRATOR. Please follow the instructions and fill in the questionnaire

Instructions: **For each of the following statements, mark one box that best describes your overall experience with BREATH CONCENTRATOR:**

	Strongly Disagree						Strongly Agree
BREATH CONCENTRATOR capabilities meet my requirements	1	2	3	4	5	6	7
BREATH CONCENTRATOR is easy to use	1	2	3	4	5	6	7

* See for reference – Finstad, K. (2010) The usability metric for user experience

On the basis of your experience with the **BREATH CONCENTRATOR**, please answer the following question:

How likely is it that you would recommend the use of this BREATH CONCENTRATOR to collect breath samples at this hospital?

<u>Please</u>: Mark one box from 10 (Extremely like) to 0 (Not at all likely)

Extremely likely	10	9	8	7	6	5	4	3	2	1	0	Not at all likely

On the basis of your experience today, please compare the easiness of use of DEVICE 1 and DEVICE 2 (mark one value).

DEVICE 1				DEVICE 2
1	2	**3**	4	5
DEVICE 1 is easier to use **than DEVICE 2**	DEVICE 1 **is quite easy to use compared** to DEVICE 2	**Equally**	DEVICE 2 **is quite easy to use compared** to DEVICE 1	DEVICE 2 **is easier to use** than DEVICE 1

On the basis of your experience today, please compare the easiness of use of DEVICE 1 and DEVICE 3 (mark one value).

DEVICE 1				DEVICE 3
1	2	**3**	4	5
DEVICE 1 is easier to use **than DEVICE 3**	DEVICE 1 **is quite easy to use compared** to DEVICE 3	**Equally**	DEVICE 3 **is quite easy to use compared** to DEVICE 1	DEVICE 3 **is easier to use** than DEVICE 1

On the basis of your experience today, please compare the easiness of use of DEVICE 2 and DEVICE 3 (mark one value).

DEVICE 2				DEVICE 3
1	2	3	4	5
DEVICE 2 is easier to use **than DEVICE 3**	DEVICE 2 **is quite easy to use compared** to DEVICE 3	**Equally**	DEVICE 3 **is quite easy to use compared** to DEVICE 2	DEVICE 3 **is easier to use** than DEVICE 2

THANK YOU FOR YOUR HELP!

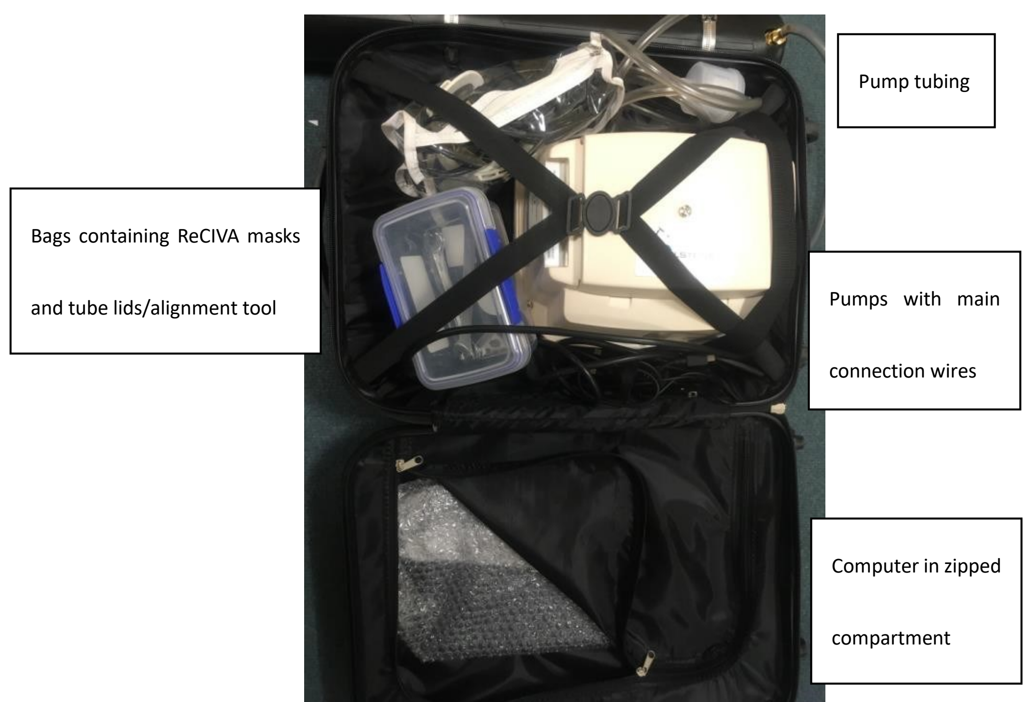
Your gender is ☐ Male ☐ Female ☐ I prefer not to say
Your Age: ________________

3. ReCIVA® List of tasks

ReCIVA Sequential Instructions, for COBRA study- May 2017 <u>Contents of the ReCIVA suitcase</u>

Open Suitcase

Scrubber



Phase 1- Preparation before the patient arrives

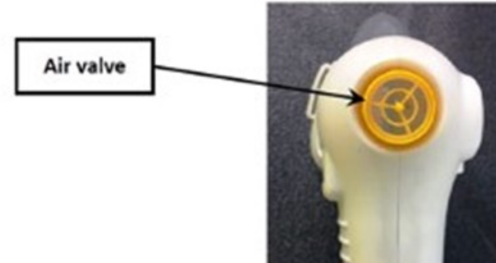
1. Ensure that the Study ID (written on the sample bag containing the tubes) is copied across to the top of the consent form and patient details forms.

   The study ID is written like this on each sample bag: CO_ _0001.

   Please fill in the second letter of the first name and second letter of the surname to make the full anonymous study ID.

2. Wear gloves

3. Inspect the air valve (yellow part in picture below) to check the membrane in the valve isn't damaged and is sitting correctly.



4. Collect the unopened Mask and 4 Sorbent tubes

5. **Prepare the tubes**

   a. Using the spanner tool, remove the brass caps from the ends of the sorbent tubes. Insert the tubes with the grooves facing upwards into the alignment tool.
   NB. T=Top

b. Seal the 4 tubes together with the metal lids. Tubes are safe to be kept like this until a patient comes to be sampled.

6. Prepare the computer

a) Plug the computer into the mains. The password is written on the keyboard. Click on "BSC" on the screen and the ReCIVA programme will open. It will say "please connect mask" on screen.

b) Connect the handheld ReCIVA to the computer using the black connector cord via the USB port.

The computer screen will show that it has detected the device and a moving display will appear. You will see the option to "start breath collection" at the bottom.

7. Prepare the pump.

a) The pump is strapped into the suitcase and does not need to be removed. The scrubber is the long cylindrical device encased in a plastic tube connected to the suitcase.
Connect the pump tubing to the scrubber using the pump connector. You will need to unscrew the lid of the scrubber to expose the connector. It simply pushes in to connect.

b) Connect the pump lead to the mains. You will hear it start. Turn it off at the plug until you need to use it.
Once the rest of the equipment is removed from the suitcase, you can zip the suitcase back up and keep the pump hidden, as you will not need to adjust this again.

c) Connect the end of the pump tubing to the hand held ReCIVA.

Phase 2: You are now ready to sample a patient

8. Firstly, explain the study and the test, and offer a patient information leaflet. Ask the patient to sign a consent form and give them the opportunity to ask questions.
9. If they agree to participate, open a new rubber mask, so that the patient can see that it is clean.
    a. Remove the top lid from the 4 tubes. Insert the tubes into the 4 holes in the bottom of the mask from below. The grooves in the tubes should be facing upwards so that they are inside the mask.

    b. You then need to remove the bottom lid, so that the tubes are dangling freely.

Sorbent tubes in mask

10. Insert the mask first into the ReCIVA sampler, then push the tubes into the holes

Mask inserted –
NOTE - tubes not inserted yet.

11. If you are finding it difficult to push the tubes down into the holes by hand, you can use the square metal tube lid to push the tubes down into position.

12. Connect the air hose from the pump to the attachment on the ReCIVA device, if it not already connected (see step 7c). Flow is around 30 Litres/min. Remember to turn the pump on at the plug.

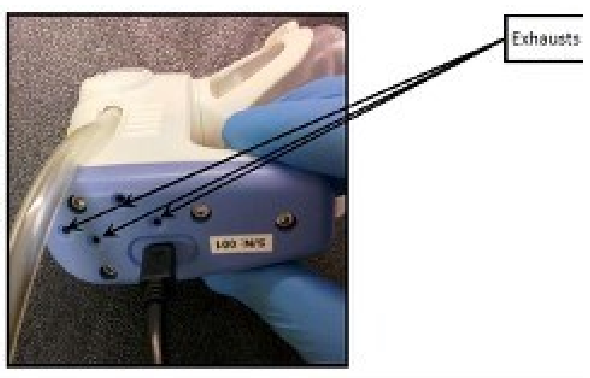
13. If not already done, connect ReCIVA sampler to the computer through the USB port place at the bottom part of ReCIVA sampler (see step 6b).
If the ReCIVA is already connected, I suggest closing down the programme and reopening it (BSC) if the programme has been running idle for a long time.
This will ensure that it is ready to go.

14. Final check-
Ensure pump is on at the plug
Ensure BCS Programme is loaded on the computer
Ensure mask is set up

15, Press" start breath collection• on thecomputer screen. An error message will appear. This alwayshappens. Click "continue". The computer will then wait until the patient starts breathing, and you will see a moving trace appear.

16, Ask patient to put the mask in place, forming a seal with the skin, and start to breathe normally. Either nasal or mouth breathing isacceptable. It will say "learning" for a few secondsin the bottom of the screen, and then "collecting".

NOTE: make sure the patient does not block Exhausts.



17, Check and correct (if needed) the mask position if there are any leaks.

18, Explain to patients to breathe normally:

- <<Please breathe normally until the end of the process, I will let you know when the sample collection is over, it will take approximately 2 or 3 minutes,>>
- Encourage the patient to keep a good seal between their face and the mask.

IMPORTANT NOTE:Patientscan talk during the collection, but this will slow thecollection.
If the patient removes the mask, thisisnot a problem; the sampler will stop sampling automatically and will restart when breathing commences again. However, encourage the patient not to do this, asthecollection will take longer.
If the patient is able, this can be a good time to fill in the patient details (CRF) form.

If the patient is distressed - stop the test

	19.	When software indicates that the sampling is completed, the test is finished
End	20.	Thank patient/discharge/debrief

Phase 3: Labelling samples and Disassembling the equipment

Disassemble

21. Pull out the mask and the Sorbent tubes from the ReCIVA. Remove and dispose of the mask.

22. Pull all the tubes out and seal them by screwing the brass nuts onto the ends of the tubes, using the spanner device if needed. Ensure these are tight and cannot be pulled off. The bottom of the brass cap should reach the groove if it is on correctly.

23. Insert the 4 sealed TD tubes into the sample packs that they came in. You can ensure that it is the correct bag by checking the barcode numbers that are written on the label of the bag.

24. Make sure the unique patient ID is complete by adding the second letter of the first name then the second letter of the second name.

25. Write in the practice name, date and time of sampling onto the label that is affixed onto the plastic sample bag.

26. Put the whole bag into the Tupperware box, and seal it so that it is airtight. All sample bags can go in the same Tupperware box back to the lab for processing.

27. Ensure that the patient information form and consent form are placed in the opaque folder. This folder should be transported with the samples, as it will need to be stored in a locked room at St Mary's.

28. To pack away the equipment, it is relatively straight-forward to unplug everything and put it all back in the suitcase.
One tip: TO DISCONNECT PUMP TUBING FROM SCRUBBER- this involves pushing down on the cylindrical connector so that the fitting is released. Do not pull it out! It will not release until you push the fitting down on the scrubber.