CHEMICAL PROFILING OF EXHALED BREATH FROM CYSTIC FIBROSIS SUBJECTS USING COMPREHENSIVE TWO-DIMENSIONAL GAS CHROMATOGRAPHY

By
Mohammad Asif Iqbal
A thesis submitted for the
Degree of Doctor of Philosophy (Science)
University of Technology Sydney
January 2021

CERTIFICATE OF AUTHORSHIP AND ORIGINALITY

CERTIFICATE OF AUTHORSHIP AND ORIGINALITY

I certify that the work in this thesis has not previously been submitted for any other degree/s nor has it been submitted as part of the requirements for a degree/s except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all the information sources and literature used are indicated in the thesis.

This research is supported by an Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Signature of student:

DATE: 4/01/2021

DEDICATION

To my parents

And

My younger brother and his everyday struggle to survive through terminal illness

Love you and always with you until I am alive

ACKNOWLEDGEMENT

Firstly, thanks and regards to my primary supervisor, Dr. Shari L. Forbes for her continuous support and for sharing her knowledge. I am also grateful to Dr. Tapan Rai (my second primary supervisor), Dr. Katie Nizio (co-supervisor), Dr Maiken Ueland (co-supervisor), and Dr Peter G Middleton (external supervisor) for their great support all throughout my PhD study. I am especially grateful to all the participants of my study both from the outpatient Cystic Fibrosis Clinic, Westmead Hospital and from the Faculty of Science, University of Technology Sydney. Without their generous participation the project wouldn't be possible. I am also grateful to my co-workers in the Gas Chromatography Lab here in UTS. In particular, Amber Brown, Nicole Cattarossi, and Vitor Cesar Taranto. In addition, my special thanks for technical staff; Dr. Ronald Shimmon, Dr Regina V. Taudte, and Dr. Dayanne Bordin. I am also very fortunate to be a part of such a lovely research group and have so many fun memories.

Finally, I am thankful to my family, especially my lovely wife Ayesha and a beautiful boy (Zyad) we are blessed with. Her relentless support and numerous sacrifices made this PhD battle possible. I am also asking forgiveness for not giving enough time to my wife, my boy, my brother, and above all my parents. I do feel guilty of that.

Most importantly, my gratitude to the Australian Government and the Taxpayers, because it is their money allocated to me as tuition fee waivers and living allowances. Without that I can't even dream of pursuing this PhD degree.

TABLE OF CONTENTS

Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

(Listed alphabetically)

ACFDR	Australian Cystic Fibrosis Data Registry
BAL	Bronchoalveolar lavage
CAR	Carboxen
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
DMS	Dimethyldisulfide
DVB	Divinylbenzene
EBC	Exhaled breath condensate
GC×GC	Comprehensive two-dimensional gas chromatography
GC×GC-TOFMS	Comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry
GC–FID	Gas chromatography flame ionisation detection
GC–MS	Gas chromatography mass spectrometry
GC–TOFMS	Gas chromatography time-of-flight mass spectrometry
HCN	Hydrogen cyanide
HS-SPME	Headspace–solid-phase microextraction
IMR–MS	Ion-molecule reaction mass spectrometry
IMS	Ion mobility spectrometry
LDA	Linear discriminant analysis
LD	Linear discriminant
LOOCV	Leave-one-out cross-validation
NMR	Nuclear magnetic resonance
PDMS	Polydimethylsiloxane
ppb	Parts-per-billion
PTR–MS	Proton transfer reaction mass spectrometry

SIFT-MS	Selected ion flow tube mass spectrometry
SPME	Solid phase microextraction
STs	Sorbent tubes
TB	Tuberculosis
VAP	Ventilator-associated pneumonia
VOCs	Volatile organic compounds

PUBLICATIONS

1. Forensic decomposition odour profiling: A review of experimental designs and analytical techniques. MA Iqbal, KD Nizio, M Ueland, SL Forbes. ***TrAC Trends in Analytical Chemistry 91, 2017, 112-124.***
2. Recent advances in the estimation of post-mortem interval in forensic taphonomy. MA Iqbal, M Ueland, SL Forbes ***Australian Journal of Forensic Sciences, 52, 2020, 107-123.***

CONFERENCE PROCEEDINGS

1. Poster presentation in the 24th Australian Society of Medical Research (ASMR) NSW ANNUAL SCIENTIFIC MEETING held by Australian Society for Medical Research during 6th June 2016 at Powerhouse Museum, Sydney, Australia.
2. Co-presented in the Combined Health Science Conference Sydney New Horizons held during 21st November 2016 at the University of Technology Sydney, Australia.
3. Poster presentation in the 35th Combined Health Science Conference Sydney New Horizons 2018, 19-20 November 2018, Kolling Institute, St Leonards, Australia.

AWARDS

1. Winner of April 2017 Paper of the Month for SCHOOL OF MATHEMATICAL AND PHYSICAL SCIENCES (MAPS), University of Technology Sydney.

ABSTRACT

Chronic lung infections are the leading cause of death in subjects with cystic fibrosis (CF). To date, sputum culture is the most common technique for the diagnosis of lung infections in adult CF subjects. However, it requires several days or longer to obtain culture results. Therefore, a rapid diagnostic technique for lung infections would significantly improve CF healthcare. During recent decades, exhaled breath analysis has attracted interest as a rapid and non-invasive tool for the diagnosis of non-communicable diseases such as cancers and heart diseases. However, there is limited progress in the diagnosis of infectious diseases such as lung infections in CF subjects using volatile organic compounds (VOCs) as biomarkers of infection.

In this study, sputum and breath samples were collected from CF subjects and healthy controls (only breath) and profiled for VOCs using comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC–TOFMS). Multivariate analyses (e.g. principal component analysis and linear discriminant analysis (LDA)) were then performed to allow differentiation between: (i) CF subjects and healthy controls and (ii) CF subjects with/without *Pseudomonas aeruginosa* infections and those with no known lung infections as confirmed using their sputum culture results. This study identified a set of 16 VOCs which allowed differentiation between CF subjects and healthy controls. In particular, healthy controls were classified with 98% accuracy, while CF subjects were classified with 92% accuracy. It is important to note that all of the CF subjects that participated in this study are significantly different from control groups, not only in terms of their lung infection status but also in terms of numerous other factors (e.g. diet, lifestyle, medications, and other health complications). These factors can also impact the breath profiles obtained from the study group (CF subjects).

The analysis of matching sputum and breath samples collected from CF subjects provided a set of 24 core VOCs common between both sample types. LDA performed using these VOCs provided accurate classification of CF subjects according to their lung infection status (i.e. CF subjects with/without *Pseudomonas aeruginosa* infection). The outcome of LDA also showed that these common VOCs have better classification accuracy than the entire profile of the VOCs detected in sputum and breath samples. Finally, the comparison of breath profiles between CF subjects with/without *Pseudomonas aeruginosa* infection and those with no known lung infection showed that it is also

possible to allow differentiation between these contrasting groups using breath VOCs profiles.

Chapter 1:

INTRODUCTION

Chapter 1: INTRODUCTION

One in every 2,500 babies is born with cystic fibrosis (CF) in Australia, equating to one being born every four days (ACFDR, 2016). Individuals suffering from CF are prone to frequent lung infections and often undergo regular health check-ups (e.g. sputum culture, lung biopsy, and bronchoscopy) for the diagnosis of common bacterial species (e.g. *Pseudomonas aeruginosa and staphylococcus aureus*) in their lungs and airways (Bilton, 2008, Lyczak et al., 2002). These tests are mostly invasive and require a certain amount of time for the culture results (a couple of days to weeks). Invasive tests for the diagnosis of lung infections (e.g. lung biopsy and bronchoscopy) are also exhausting and stressful to subjects. For this reason, researchers have focused on developing non-invasive and faster methods for detecting infections to minimize the drawbacks of common invasive techniques.

Recently, the analysis of volatile disease biomarkers (i.e. VOCs: volatile organic compounds) in exhaled breath has attracted general interest as an important non-invasive diagnostic method that can be used to detect the signs of non-communicable diseases like cancers and metabolic disorders (Lourenço and Turner, 2014). However, the number of studies concerning the diagnosis of opportunistic infections (e.g. *Pseudomonas aeruginosa* infections in CF subjects) through the analysis of VOCs present in exhaled breath are very limited. There is, however, several studies which have reported VOCs from *in vitro* cultures of different bacterial species associated with CF lung infections (Bean et al., 2012, Nizio et al., 2016). Such *in vitro* approaches provide important insights into the discovery of potential biomarkers associated with lung infections in CF subjects, but also suffer from several challenges as recent studies reported diverse VOCs from growth medium and significant degradation of samples with prolonged storage time (Nizio et al., 2016). Considering this, it is important to characterise VOCs directly in the headspace of their natural growth media (i.e. sputum and bronchoalveolar lavage samples from CF subjects) as well as understanding their transition into breath samples collected from patients. However, there is a knowledge gap concerning the analysis of VOCs directly in the headspace of CF sputum samples. In particular, there is only two studies (Goeminne et al., 2012, Savelev et al., 2011), which has reported the VOCs in the headspace of sputum samples from CF subjects and none of these studies analysed corresponding breath samples.

1.1. Cystic fibrosis

1.1.1. Cause and mechanism of cystic fibrosis

Cystic Fibrosis (CF) is a recessive genetic disorder, caused by the presence of mutations in both copies of the cystic fibrosis transmembrane conductance regulator (CFTR) gene (Rommens et al., 1989, Zielenski et al., 1991). This gene provides instructions for the cells to make CFTR proteins, which function as a channel across the membrane of cells that produce mucus, sweat, saliva, tears, and digestive enzymes (Kälin et al., 1999). The channel transports salts in and out of cells and helps control the movement of water in tissues, which is necessary for the production of thin, freely flowing mucus that lubricates and protects the lining of the airways, digestive system, reproductive system, and other organs and tissues (Boucher, 2007, Matsui et al., 1998, NIH, 2016). Due to the defect of the CFTR gene, CF subjects develop an abnormal amount of excessively thick and sticky mucus within the lungs, airways and digestive system (Collins, 1992, Johansson et al., 2013) **(Figure 1.1)**. This leads to frequent and repeated lung infections causing irreversible damage (De Jong et al., 2004, Davidson et al., 1995).

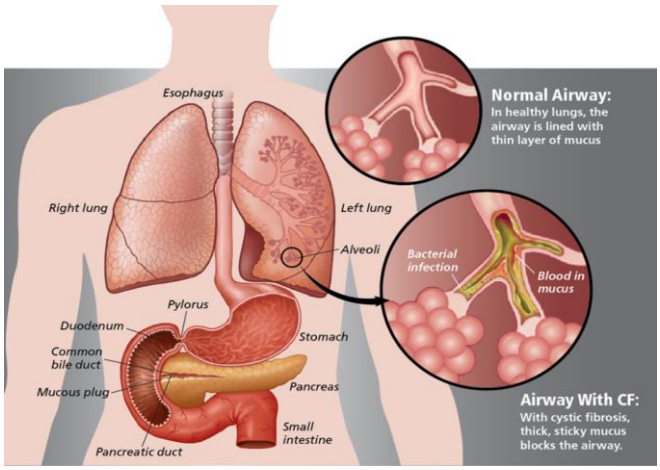


Figure 1.1: CF airways and digestive system [source: (Trivedi, 2013)]

At present, there is no cure for CF. Subjects need special care throughout their entire life span and are typically attached to a specialty centre with multidisciplinary care teams (i.e. cystic fibrosis centres). The management of CF is an ongoing and relentless process, which requires regular monitoring of diverse health conditions. Routine treatment generally involves: intensive daily physiotherapy to clear the lungs (Lannefors et al., 2004); routine exercise to help clear the airways and strengthen the body (Nixon et al.,

1992); long term use of antibiotics (e.g. azithromycin) to treat lung infections (Saiman et al., 2003, Equi et al., 2002, Clement et al., 2006); pancreatic enzyme replacement and fat-soluble vitamin supplementation, especially in children (Ramsey et al., 1992, Lancellotti et al., 1996, Borowitz et al., 2002); and lung transplantation if lung function continues to worsen (Yankaskas et al., 1998, Chaparro et al., 2001). With proper management and treatment, CF subjects can expect an average life span of 30~40 years; which has increased dramatically (almost doubled) during the last three decades (1980-2010) (Aris et al., 2005, MacKenzie et al., 2014, Bilton, 2008, ACFDR, 2016).

1.1.2. Prevalence and mortality in Australia

CF is the most common lethal genetic disorder in Caucasian populations (Ratjen and Döring, 2003, Bilton, 2008, Heijerman, 2008). It is also the most common life-limiting genetic condition in Australia; one in every 25 Australians carries a defective CF gene; while, the incidence is even higher in Tasmania, with one in every 20 people (ACFDR, 2012). Since 1986, all new-borns have been subject to a heel prick test for the screening of a number of conditions, including CF; which has helped with the declining prevalence of cystic fibrosis in Australia (Massie et al., 2010). However, heel prick tests are not conclusive for a positive diagnostic of CF, and may require other more specific tests like the Sweat Test (CFFA, 2016). In some cases, further diagnosis may also result after a baby is born with obvious CF symptoms (e.g., a bowel blockage or failure to thrive) (ACFDR, 2012). According to the 15th Annual Report by the Australian Cystic Fibrosis Data Registry (ACFDR), over 80% of infant diagnoses were completed by three months of age, assisted by neonatal screening programs that operate in Australia (ACFDR, 2012).

During 2016, a total of 19 CF-associated deaths were reported in Australia (17 adults and 2 children aged less than 18) (ACFDR, 2016). This number is lower than the numbers reported in 2012 (total 40, 34 adults and 6 children aged less than 18) (ACFDR, 2012). The 2016 figure was also lower than the number of total deaths reported in 2011 (total 27 deaths) and 2010 (total 21 deaths). During 2016, the median age at death for subjects was 32.6 years, up from a median of 31.6 in 2015. Pulmonary causes and post-transplant complications were responsible for 15 of the 19 deaths reported in 2016, while four causes of death were unknown or unstated (ACFDR, 2016).

1.1.3. Major health complications in cystic fibrosis

Chronic bacterial lung infections and contaminant airway inflammation are the leading causes of death in subjects with CF; these are responsible for at least 80% of CF-associated deaths (Lyczak et al., 2002, O'Sullivan and Freedman, 2009). Such chronic infections and inflammation cause permanent damage to the respiratory system of CF subjects, leading to fatal health conditions, like pulmonary insufficiency (Lyczak et al., 2002); major and massive haemoptysis (Barben et al., 2003, Flume et al., 2005b); and pneumothorax (Flume et al., 2005a). For this reason, CF subjects undergo frequent check-ups for the diagnosis of lung infections and are treated immediately following detection before any irreversible damage is caused (Lyczak et al., 2002).

CF subjects also suffer from other health complications, which include: (1) gastrointestinal complications such as gastro-oesophageal reflux (Blondeau et al., 2008), abnormal liver function (Colombo et al., 2002), cirrhosis or portal hypertension (McKeon et al., 2004, Efrati et al., 2003, Colombo et al., 2002), and pancreatitis (Colombo et al., 2002, Ooi et al., 2011); (2) endocrine disorders such as chronic insulin-dependent diabetes (Mackie et al., 2003, Rolon et al., 2001), intermittent insulin-dependent diabetes (Mackie et al., 2003), and other glucose abnormality (Hameed et al., 2010); (3) renal impairment such as renal calculi (Sidhu et al., 1998) and renal failure (Al-Aloul et al., 2005); (4) bone diseases such as osteoporosis (Döring and Conway, 2008) and osteopenia (Aris et al., 2005, Conway et al., 2000); and (5) other complications such as growth failure in children (Zemel et al., 2000), vitamin and essential fatty acid deficiency (Sinaasappel et al., 2002, Strandvik et al., 2001), and cancer (Maisonneuve et al., 2003, Maisonneuve et al., 2007).

1.1.4. Lung infections associated with cystic fibrosis

Typically, the respiratory system of CF subjects becomes colonised by *Haemophilus influenzae* or *Staphylococcus aureus* or both, during their infancy (at age 1) (Khan et al., 1995, Rosenfeld et al., 2001). Within a short period of time (2 to 3 years of age), *Pseudomonas aeruginosa* becomes the predominant organism in the respiratory tract of CF subjects and challenges the immune defences (Oliver et al., 2000). These bacterial populations are responsible for chronic biofilm lung infections in CF subjects; causing lung tissue damage and leading to lung failure, lung transplantation, and death (Høiby et al., 2010). In Australia, during 2012, 53% of CF subjects tested produced positive *P. aeruginosa* cultures. Its prevalence is greater in adult subjects with approx. 70% of adults

that tested positive which is two to three times higher than the proportion for adolescents and much higher than that for children (ACFDR, 2016).

The airways of CF subjects can also be colonised with *Burkholderia cepacia (B. cepacia)* complex (a complex of at least 18 different species, many of them having innate antibiotic resistance) (LiPuma, 2005, LiPuma, 2010). These species are highly virulent and transmissible in nature and can cause a rapid collapse in pulmonary function and increase mortality in CF subjects (LiPuma, 2005, Steinkamp et al., 2005, De Boeck et al., 2004, O'Sullivan and Freedman, 2009). Some other important pathogens include: *Stenotrophomonas maltophilia* (*S. maltophilia*), a multidrug-resistant organism that increases the risk of pulmonary exacerbations in CF subjects (Waters et al., 2013) and meticillin-resistant *S. aureus* that causes reduced lung function in CF subjects (Dasenbrook et al., 2008).

1.1.5. Common diagnostic techniques for lung infections

At present, sputum culture is the most common technique applied to detect and identify bacterial lung infections in adult CF subjects; although, bronchoalveolar lavage (BAL)/bronchoscopic culture is mainly applied in young children who have trouble producing sputum spontaneously (Stafler et al., 2011, ACFDR, 2016). Both of these techniques are routinely used to define airway microbiology and inflammation in CF subjects (Tunney et al., 2008). Some recent studies also reported the use of the sputum induction technique as a preclusion of BAL (Blau et al., 2014). However, most of these techniques are invasive and stressful for subjects, especially for children, and require a certain amount of time (typically 2 to 3 days or longer) to obtain results.

For this reason, there is a growing interest in the identification of volatile organic compounds (VOCs) released from sputum samples to facilitate rapid and non-invasive diagnosis of certain lung infections based on VOC profiles. A large number of studies have already reported the profiling of VOCs in the headspace of *in vitro* cultures from CF-associated microorganisms (Shestivska et al., 2011, Chippendale et al., 2014, Dryahina et al., 2016, Neerincx et al., 2016a, Labows et al., 1980, Carroll et al., 2005, Briard et al., 2016, Purcaro et al., 2018, Nizio et al., 2016, Baptista et al., 2019, Franchina et al., 2019, Hahn et al., 2020, Jenkins and Bean, 2019). However, there is limited progress in the detection of VOCs directly in the headspace of sputum samples. In particular, so far only two studies have reported the VOCs released from sputum samples collected from CF subjects and compared only *P. aeruginosa* positive with negative

cultures (Goeminne et al., 2012, Savelev et al., 2011). In addition, the number of studies related to the diagnosis of non-CF lung infections (e.g. ventilator-associated pneumonia (VAP) and pulmonary tuberculosis (TB)) using sputum VOCs are also limited (Lawal et al., 2018b, Kolk et al., 2010). Further studies are required in this area to understand the VOCs released in the headspace of sputum samples collected from CF subjects infected with other CF-associated microorganisms (e.g. *H. influenza*, *S. aureus, B. cepacia, and S. maltophilia*).

1.2. Exhaled breath analysis

1.2.1. Volatile organic compounds in exhaled human breath

Exhaled human breath is a complex mixture of several atmospheric inorganic gases (e.g., O_2, N_2, CO, NO, CO_2, H_2O, NH_3, H_2S), exhaled breath condensate (EBC), and diverse VOCs. The composition of VOCs in exhaled breath depends on human metabolism (endogenous) and is also influenced by exogenous sources (e.g., consumed air, water, and food) (Phillips et al., 1994). Endogenous VOCs can provide important information related to biological metabolism and can also reveal certain health complications (Amann et al., 2007). For example, diabetic subjects lack sufficient insulin to metabolize glucose, and instead decompose fat as an alternative source of energy. This process releases excessive amounts of acetone as a metabolite which is carried by blood to the lungs, where it partitions into exhaled breath, indicating uncontrolled diabetes (Owen et al., 1982).

During recent decades, the analysis of VOCs in exhaled human breath has attracted general interest as an important, non-invasive disease diagnosis technique (Kim et al., 2012, Lourenço and Turner, 2014, Pereira et al., 2015, Schleich et al., 2019). Interestingly, ancient Greek physicians like Hippocrates were already curious about the diagnostic potential of breath aroma caused by VOCs (e.g., the urine-like smell that accompanies failing kidneys and the sewer-like smell of a lung abscess) (Kim et al., 2012, Buszewski et al., 2007). However, the modern era of breath analysis commenced in 1971; when, Nobel Prize winner Linus Pauling and co-workers detected over 200 different VOCs in human breath and in the headspace of urine using gas chromatography (Pauling et al., 1971).

As reported by Phillips et al. (Phillips et al., 1999b), the breath of healthy humans (n=50) can contain an average of 204 VOCs, while the total number of different VOCs detected in all subjects was around 4000 (Phillips et al., 1999b). Beside the typical metabolites,

exhaled breath may also contain a wide range of biomarkers, reflecting the metabolic status and condition of body organs (e.g., lungs, liver, and kidneys) (Kim et al., 2012, Phillips et al., 2006, Phillips et al., 1999b). Recent advances in analytical techniques and numerous efforts from dedicated scientists make it possible to identify a number of biomarkers associated with certain diseases and health complications (Buszewski et al., 2007). For example, the biomarkers of some non-infectious diseases and health complications include: e.g., cancer: several alcohols, ketones, esters, heterocyclic compounds, and different hydrocarbons (Deng et al., 2004a, Deng et al., 2004b, Phillips et al., 2006, Phillips et al., 1999a, Phillips et al., 1999b, Poli et al., 2005, Smith et al., 2003, Wehinger et al., 2007, Xue et al., 2008, Yu et al., 2005); oxidative stress: ethane, pentane, and hydrogen peroxide (Amann and Smith, 2005, Horváth et al., 1998a); cholesterol metabolism: isoprene (Amann and Smith, 2005, Karl et al., 2001, Miekisch et al., 2004); uncontrolled diabetes mellitus: acetone; liver impairment: dimethylsulfide, methyl mercaptan, and ethyl mercaptan (Amann and Smith, 2005, Buszewski et al., 2007, Kim et al., 2012); uraemia and kidney impairment: ammonia, dimethylamine, trimethylamine (Amann and Smith, 2005, Miekisch et al., 2004, Narasimhan et al., 2001); and volatile compounds from subjects with asthma: nitric oxide, carbon monoxide, and hydrogen peroxide (Horváth et al., 1998a, Horváth et al., 1998b).

1.2.2. Exhaled breath analysis in lung infection

1.2.2.1 Exhaled breath analysis in different lung infections

The diagnostic potential of breath biomarkers is extensive and not limited to the diagnosis of non-communicable diseases (e.g. cancers and heart diseases). There is a large number of studies which has focused on the diagnosis of communicable diseases such as lung infections using breath biomarkers (Bos et al., 2013a, van der Schee et al., 2015, Chan et al., 2020). For instance, early studies by Phillips et al. (2007 and 2010) reported a set of VOCs which has allowed discrimination between subjects with/without active pulmonary tuberculosis (Phillips et al., 2012, Phillips et al., 2010, Phillips et al., 2007). In further studies, Beccaria et al. evaluated breath VOCs to diagnose active TB in subjects with confirmed Mycobacterium tuberculosis infections (Beccaria et al., 2018a, Beccaria et al., 2018c). In addition, several studies have reported the application of exhaled breath analysis for the diagnosis of lung infections in mechanically ventilated subjects at intensive care units (Bean et al., 2014, Filipiak et al., 2015). For instance, Filipiak et al. (2015) reported an exploratory study assessing the feasibility of breath VOC analyses for

the non-invasive diagnosis of infections in the lower respiratory tract of ventilated subjects (Filipiak et al., 2015). In another study, Purcaro et al. (2019) analysed VOCs in the exhaled breath of murine models to identify core VOCs differentiating between *Pseudomonas aeruginosa* infected and non-infected animals (Purcaro et al., 2019).

1.2.2.2 Exhaled breath analysis in cystic fibrosis

Several studies investigating exhaled breath of CF subjects have identified potential markers of oxidative stress and airway inflammation. For instance, Paredi et al. (2000) reported ethane as a potential non-invasive marker of oxidative stress in CF breath, which is correlated with airway obstruction in subjects (Paredi et al., 2000). Other studies analysed EBC from CF subjects and identified increased concentrations of 8-isoprostane and nitrotyrosine, associated with increased oxidative stress (Montuschi et al., 2000, Balint et al., 2001). Several studies also reported lower pH and nitric oxide values in the EBC of CF subjects, compared to control subjects, which may be indicative of CF airway inflammation (Carpagnano et al., 2004, Grasemann et al., 2000, Balint et al., 2001, Montuschi et al., 2000, Paredi et al., 2000, Grasemann et al., 1997, Tate et al., 2002, Robroeks et al., 2008, Rosias et al., 2010, Robroeks et al., 2010a).

Another major and recent aspect of CF breath analysis is to characterise bacterial lung infections in CF subjects (Barker et al., 2006). Carroll et al. (Carroll et al., 2005) analysed VOCs produced by different isolates of *P. aeruginosa* cultures in vitro from subjects with CF and reported significantly higher levels of hydrogen cyanide (HCN) in the headspace of *P. aeruginosa*-positive samples, compared to controls. A further study reported that HCN is elevated in the exhaled breath of children with CF (13.5 ppb), compared to children with asthma (2 ppb) (Enderby et al., 2009). Similarly, HCN has been reported as a potential biomarker in nose-exhaled breath from adult CF subjects, and described as a biomarker of chronic airway infection with *P. aeruginosa* (Gilchrist et al., 2013).

Recently reported potential biomarkers of CF lung infections include: (i) methyl thiocyanate, which was reported in the headspace of *P. aeruginosa* cultures and also in the breath of CF children (Shestivska et al., 2011); (ii) breath sulfides as potential non-invasive markers of respiratory colonisation in CF subjects (Kamboures et al., 2005); (iii) ethanol, acetate, 2-propanol, acetone, and methanol in EBC reported to allow discrimination between CF subjects and healthy controls and subjects with stable and unstable CF (Montuschi et al., 2011); (iv) acetic acid, which is reported to have elevated concentrations in breath samples from CF subjects (independent of their *P. aeruginosa*

infection status), compared to that of healthy controls (Smith et al., 2016); and 2-aminoacetophenone as a potential biomarker of *P. aeruginosa* in CF lung infections (Scott-Thomas et al., 2010).

Microorganisms like *P. aeruginosa* can produce diverse metabolic profiles during the early stages of lung infection in CF subjects (Jørgensen et al., 2015). Moreover, the number of VOCs released from a single strain of bacteria can also be very high (Nizio et al., 2016). For this reason, recent studies suggest that the assessment of a pattern or a profile of VOCs is likely to provide greater insight into infection status than a single compound (i.e. biomarker) (Thorn et al., 2011, Nizio et al., 2016, Robroeks et al., 2010b). For instance, Robroeks et al. (2010) analysed exhaled breath samples from 150 children (48 with CF + 57 controls), and reported a total of 1099 VOCs with a prevalence of at least 7%; according to that study, it was possible to identify CF subjects and controls by using 22 VOCs (with 100% certainty). Additionally, 92% of subjects were correctly classified with 10 VOCs; this technique was also reported to be able to discriminate between CF subjects with or without *Pseudomonas* colonisation (Robroeks et al., 2010b). Filipiak et al. (2012) investigated the release or consumption of VOCs by bacteria and reported that approximately 32 and 37 metabolites were released by *S. aureus* and *P. aeruginosa*, respectively. There were also differences in the bacteria-specific VOC profiles, especially with regard to aldehydes, which were taken up only by *P. aeruginosa,* but released by *S. aureus* (Filipiak et al., 2012).

1.3. Common analytical techniques for breath volatiles

1.3.1. Sampling devices

VOC sample collection is one of the major challenges involved in exhaled breath analysis and there are a number of parameters that require special attention in order to avoid misunderstanding about the origin of the identified VOCs (Pereira et al., 2015). These parameters include: the type and the number of breath collections (Amann et al., 2010, Pleil and Lindstrom, 1995); the portion of breath used (Amann et al., 2010, Schubert et al., 2004); duration and media of sample storage (Amann et al., 2010, Nizio et al., 2016) and; the interference of ambient VOCs (Amann et al., 2010, Pereira et al., 2015). Breath collection can be achieved through a single breath or multiple breaths. The collection of a single breath tends to be less time consuming and more appreciated by subjects, but on the other hand, multiple breaths may be required for the identification of a specific set of

biomarkers associated with certain diseases (Amann et al., 2010, Amann and Smith, 2005). Another approach is to collect a particular portion of breath, (e.g. alveolar air: air from the pulmonary alveoli), which is rich in volatile blood-borne compounds (Miekisch and Schubert, 2006).

Exhaled breath samples can be collected and stored in different polymer bags (e.g. Nalophan, Tedlar®, Teflon, Kynar and FlexFoil) (Mochalski et al., 2009) and canisters (Pleil and Lindstrom, 1995) **(Figure 1.2)**.

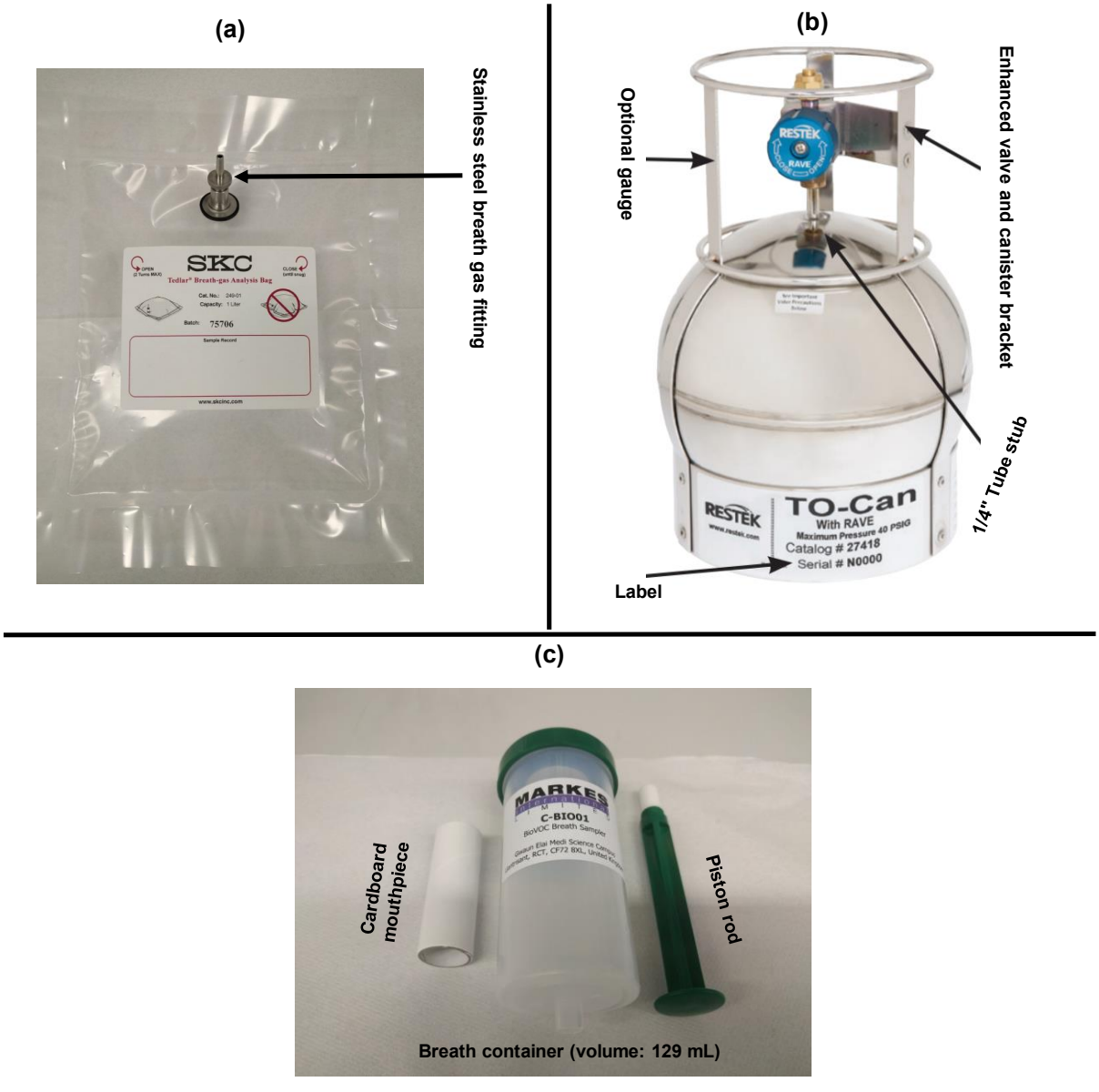


Figure 1.2: (a) 1 L Tedlar® breath sampling bag (SKC Inc., USA); (b) TO-Can Canister with RAVE Valve (Restek Corp., USA); and (c) Bio-VOC™ breath sampler (Markes International Limited, UK).

Previous studies reported the superior performance of Tedlar® bags over Nalophan, Teflon, Kynar and Flexfilm in terms of background emission and the stability of stored VOCs (Mochalski et al., 2013, Mochalski et al., 2009). In addition, Tedlar® bags were reported for their reusability (Mochalski et al., 2009). However, any polymer bag requires an appropriate cleaning protocol and proper handling to avoid any contamination from outgassing (of bag materials) and trace residual VOCs from previous samples (Beauchamp et al., 2008, Mochalski et al., 2009). Although bag samples need careful handling, the transportation of breath samples collected on bags are significantly easier than that of canisters considering the volume and weight of canisters **(Figure 1.2)**. That is why the use of canister sampling is not very common in exhaled breath analysis which is also limited by several other factors including sample loss, cost, space, and storage (Beauchamp et al., 2008, Pereira et al., 2015). A number of recent studies have reported the use of the Bio-VOC™ breath sampler (Markes International Limited, UK) for rapid sampling of VOCs in exhaled human breath samples. Bio-VOC™ consists of a small non-emitting plastic reservoir, a piston rod, and a cardboard mouthpiece **(Figure 1.2)** (Wilson and Monster, 1999, Lawal et al., 2017). This device is designed to collect the alveolar portion of breath (endogenous VOCs) and it is a method used more commonly for disease-specific VOC biomarkers found at high concentration. However, the limitation of Bio-VOC™ is that the sample cannot be stored and VOCs need to be extracted immediately after sampling. Several studies have also reported the use of gas-tight syringes to collect and store alveolar breath samples in pre-evacuated glass vials for further analysis (Miekisch et al., 2008, Wang et al., 2014). However, there is still a lack of consensus on the best sampling technique for the collection of VOCs in exhaled human breath, which hinders the standardisation process for the analysis of breath VOCs.

1.3.2. Extraction techniques

Sample extraction is an important consideration in the analysis of VOCs in exhaled breath samples, as most VOCs are present at trace levels and are highly volatile in nature (Amann et al., 2004, Pereira et al., 2015). There are several extraction techniques available, such as cryogenic trapping, solid phase microextraction (SPME), and adsorption into sorbent tubes (STs) (Lourenço and Turner, 2014, Pereira et al., 2015). STs are pen-sized glass or stainless steel tubes filled with different sorbent materials with differing capacities to retain VOCs, temperature limits, and hydrophobicity **(Figure 1.3a,b)** (Pankow et al., 2012, Woolfenden, 2010b, Woolfenden, 2010a, Iqbal and Kim,

2014, Iqbal et al., 2014a, Iqbal et al., 2014b). STs packed with Tenax® TA sorbent material are commonly used for breath analysis as they allow good performance in terms of extraction and transport of breath samples before processing (Reynolds et al., 2014, Van der Schee et al., 2012, Reynolds et al., 2010). In addition, breath samples collected on stainless steel STs packed with Tenax® TA are reported to stay stable for up to 14 days after sampling, while stored at cool temperatures (4 °C) (Harshman et al., 2016).

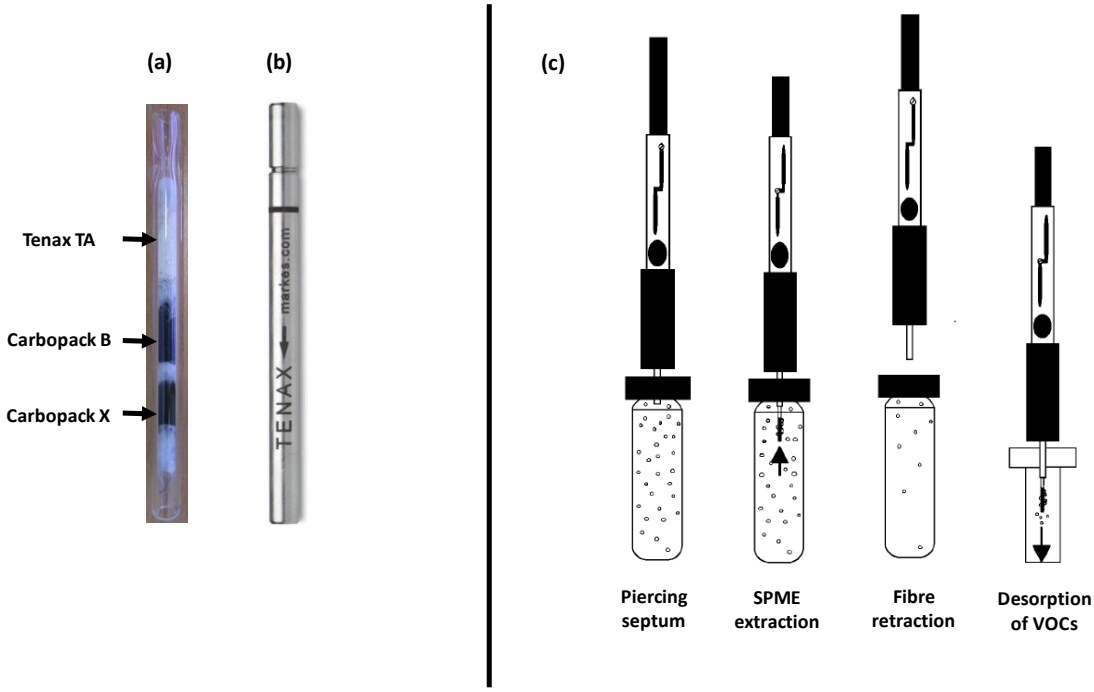


Figure 1.3: (a) A glass sorbent tube packed with three different sorbent materials (50 mg of each: Tenax® TA (60/80 mesh, Restek, USA), Carbopack™ B (60/80 mesh, Supelco, USA), and Carbopack™ X (40/60 mesh, Supelco, USA) (Iqbal and Kim, 2014, Iqbal et al., 2014a, Iqbal et al., 2014b); (b) Stainless steel sorbent tube packed with 100 mg Tenax® TA (Markes International Limited, UK); and (c) Extraction and thermal desorption of VOCs using SPME technique (adopted from (Rust, 2018)).

Headspace (HS)-SPME is another widely used extraction technique for VOCs in exhaled human breath (Poli et al., 2010, Tang et al., 2015, Raninen et al., 2016, Miekisch et al., 2008, Wang et al., 2014). This technique was first developed by Pawliszyn and colleagues as a means of rapid and solvent-free sample preparation for VOCs (Górecki et al., 1999). The advantage of SPME is the ease of sample desorption directly into the GC-inlet **(Figure 1.3c)** without any additional instrumentation (such as the thermal desorption unit which is required for ST samples). SPME works based on the adsorption of VOCs on a

coated silica fibre with a thin layer of suitable polymeric adsorbent like polydimethylsiloxane (PDMS), divinylbenzene (DVB), carboxen (CAR), combinations of these, and others (Arthur and Pawliszyn, 1990, Pawliszyn, 1999, Pawliszyn and Pedersen-Bjergaard, 2006). The type and thickness of fibres are selected based on the polarity and molecular weight of target VOCs. In breath analysis, the DVB/CAR/PDMS fibres are widely reported for best performance; however, the search for improved materials is still ongoing (Pawliszyn, 1999, Pawliszyn and Pedersen-Bjergaard, 2006, Pereira et al., 2015). Several recent studies have applied HS-SPME technique for in vitro analysis of metabolites from common CF-associated bacterial species (e.g. *P. aeruginosa, B. cenocepacia, H. influenzae, and S. maltophilia*) (Nizio et al., 2016, Shestivska et al., 2011, Shestivska et al., 2015). However, so far, there is no study concerning the analysis of VOCs in exhaled human breath which have reported the comparison of two major extraction techniques (ST-extraction vs SPME) based on GC×GC data. In this study, the performance of both techniques (ST-extraction vs SPME) was evaluated.

1.3.3. Chromatographic analysis

A number of analytical techniques have been reported throughout the literature for the detection and identification of VOCs in exhaled human breath including: gas chromatography flame ionisation detection (GC–FID) (Pauling et al., 1971, Xu et al., 2015); gas chromatography mass spectrometry (GC–MS) (Xu et al., 2015); gas chromatography time of flight mass spectrometry (GC–TOFMS) (Robroeks et al., 2010b); ion mobility spectrometry (IMS) (Westhoff et al., 2009, Jünger et al., 2012); and nuclear magnetic resonance (NMR) spectroscopy (Montuschi et al., 2011). Several on-line measurement techniques have been applied for real-time monitoring of VOCs present in exhaled breath *(in vivo)* and produced by pathogen cultures *(in vitro)*, including: selected ion flow tube mass spectrometry (SIFT-MS) (Shestivska et al., 2011, Carroll et al., 2005, Enderby et al., 2009, Smith et al., 2016, Chippendale et al., 2014); proton transfer reaction mass spectrometry (PTR–MS) (Bunge et al., 2008, King et al., 2010); and ion-molecule reaction mass spectrometry (IMR–MS) (Dolch et al., 2008, Millonig et al., 2010). Several studies also reported the use of electronic noses which do not identify specific VOCs but rely on pattern recognition (Boots et al., 2012, Röck et al., 2008).

GC–MS is considered the gold standard for the analysis of volatiles and semi-volatiles however, it has several limitations (e.g., insufficient peak capacity and restricted selectivity) (Filipiak et al., 2013, Nizio et al., 2016) which may result in co-eluting peaks,

chromatographic artefacts, and a dynamic range that is difficult to manage (Perrault et al., 2015). A GC-system with one capillary column (first dimension: ^{1}D) can separate 100–150 peaks in a single run, but is unable to facilitate proper separation of complex samples containing thousands of VOCs (e.g., petroleum). In past decades, comprehensive two-dimensional gas chromatography (GC×GC) has emerged as a powerful analytical technique in which an additional column (second dimension: ^{2}D) is employed for further separation of the effluent from a ^{1}D column. Typically, the first column is a longer (15 to 60 m) nonpolar or mid-polar column and the second column is short and relatively more polar than the first column, the combination of these two columns facilitates rapid and high-resolution separations (Dallüge et al., 2003). However, in order to pass an eluent into the second column, analytes must be focused using a modulator which is essential for the GC×GC system (Marriott and Shellie, 2002, Shellie et al., 2001). Another major challenge in GC×GC is the use of a detector capable to match the requirements (e.g., ability to acquire full range mass spectra using MS detectors) (Marriott and Shellie, 2002); TOFMS is frequently used because of its ability to acquire full range mass spectra and its faster acquisition capabilities (Dallüge et al., 2002, Edwards et al., 2011, Xu et al., 2015).

Figure 1.4 depicts the typical set-up of a GC×GC system which consists of a ^{1}D column in a ^{1}D oven, a modulator, a ^{2}D column placed in a ^{2}D oven, and a detector (TOFMS).

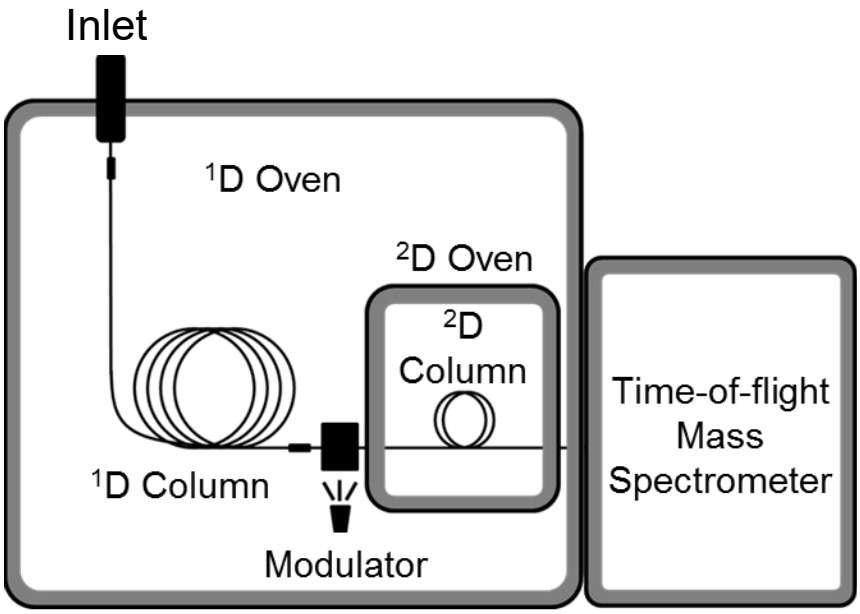


Figure 1.4: The typical set-up of a GC×GC system (image courtesy of Katelynn Perrault)

The modulator is used to repeatedly trap, concentrate, and release the effluent from the ^{1}D column for further separation on the ^{2}D column. The loss of the ^{1}D resolution is minimised by short modulation periods (Mondello et al., 2008). In addition, a high sampling rate is set to facilitate several modulations across each ^{1}D peak. As ^{1}D separations are carried out in a traditional capillary column of 15–60 m length, the peak widths are typically 10 – 30s. Therefore, ^{2}D separation needs to be completed within 2–8 s to obtain at least four modulation points across a ^{1}D peak (Mondello et al., 2008). Hence, a short capillary column of 1 – 2m is used in the ^{2}D. Based on this working principle, the result obtained from a GC×GC run is a collection of many short ^{2}D runs. This result can be visualised as a conventional one-dimensional (1D) chromatogram (Figure 1.5a). However, the interpretation of such data is complicated because it is difficult to identify the peaks that belong to the same compound. An additional processing is required called demodulation for better visualisation of the raw GC×GC chromatogram which is essentially the cutting of the sequential GC×GC data into sections the length of the modulation time and stacking them next to each other (Van Stee and Brinkman, 2016). As a result, 1D vector data are transformed into a matrix and visualised as two-dimensional (2D) contour (Figure 1.5b) or three-dimensional (3D) surface (Figure 1.5c) plots.

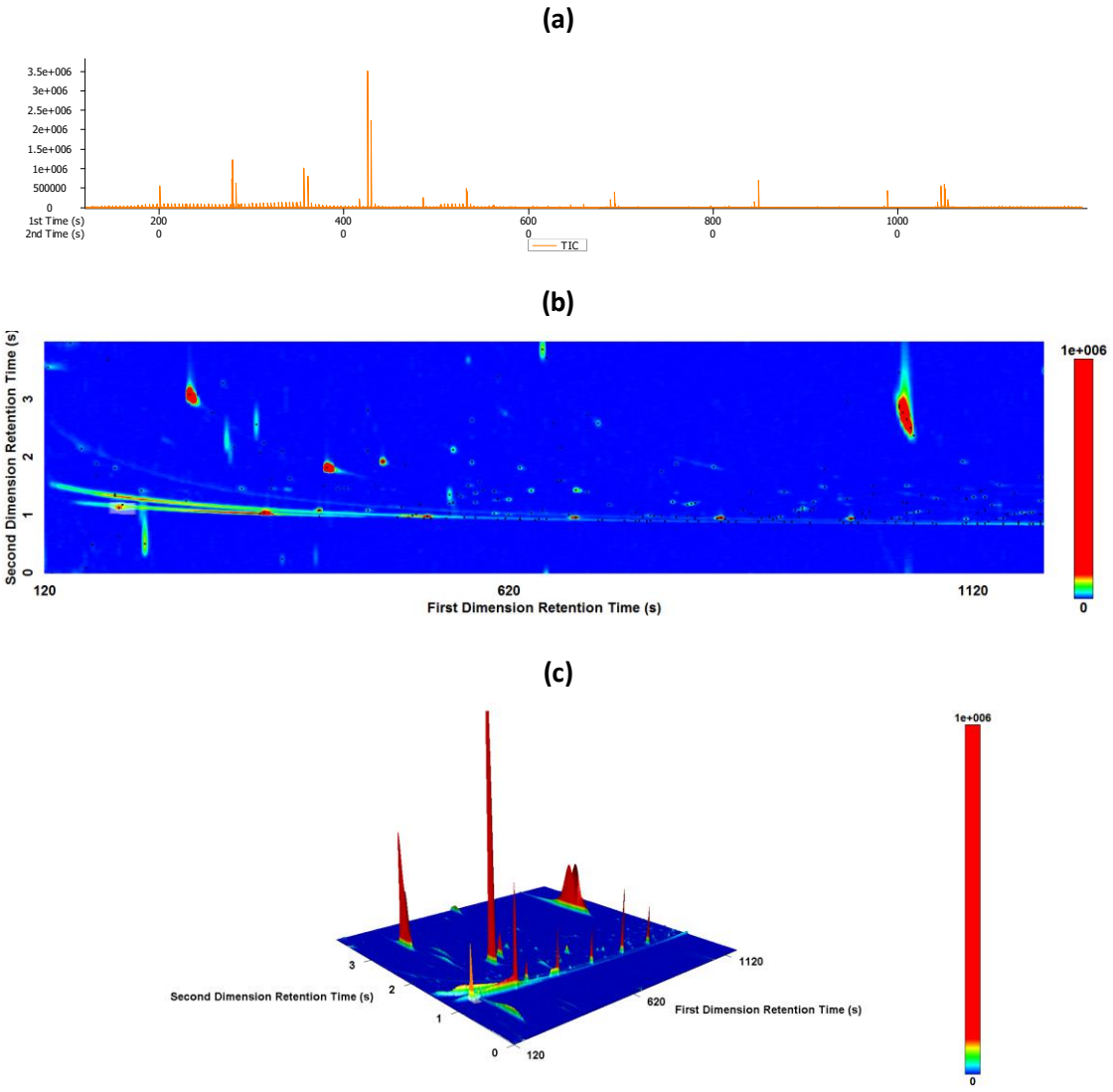


Figure 1.5: (a) 1D total ion current (TIC) plot of a sputum profile obtained through SPME-GC×GC-TOFMS technique; (b) 2D contour plot, where every dot represents an individual compound; and (c) 3D surface plot. The intensity of instrumental response is represented by the colour gradient (blue to red) in Figure b and c [This is the 3D view of the same sputum sample].

Different studies have reported the successful application of GC×GC for the analysis of highly complex samples, including: petroleum products (Marriott and Shellie, 2002, von Mühlen et al., 2006, Nizio et al., 2012), flavour compounds (Shellie et al., 2001, Tranchida et al., 2013), complex environmental and forensic samples (Frysinger and Gaines, 2002, Pani and Górecki, 2006, Sampat et al., 2016), and bacterial metabolites (Nizio et al., 2016, Phillips et al., 2013). A number of studies have also applied GC×GC techniques for the analysis of VOCs in exhaled breath samples (Sanchez and Sacks, 2006,

Libardoni et al., 2006, Caldeira et al., 2012, Phillips et al., 2013, Bean et al., 2012, Bean et al., 2015, Bean et al., 2016, Purcaro et al., 2019, Purcaro et al., 2018, Stefanuto et al., 2020). Recently, Bean et al. (Bean et al., 2012) employed a comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC×GC-TOFMS) system for the first time to analyse VOCs in the headspace of *P. aeruginosa* culture. The authors identified a total of 56 VOCs which almost doubled the list of previously published volatile metabolites for one of the most prevalent bacterial species associated with CF lung infections (Bean et al., 2012). More recently, Nizio et al. (Nizio et al., 2016) reported an average of 472 VOCs from the *in vitro* cultures of common CF-associated bacterial species (n=6), indicating an order-of-magnitude increase in the number of VOCs detected using GC×GC–TOFMS compared to the VOC profiles obtained using traditional one-dimensional GC-MS. So far, there are only two studies which have applied GC×GC for the analysis of VOCs either in the headspace of CF sputum (Hahn et al., 2020) or BAL (Nasir et al., 2018) samples. However, none of these studies reported VOCs in the corresponding exhaled breath samples collected from CF subjects.

1.4. Research aims and objectives

The aim of this study was to detect and identify certain lung infections in CF subjects through the profiling of VOCs in their exhaled breath using GC×GC–TOFMS. To facilitate this process, a set of matching sputum and breath samples were collected from adult CF subjects and profiled for VOCs. Additional breath samples were then collected from CF subjects to increase the validity of the breath sample set. In parallel, breath samples were also collected from healthy participants recruited as a control group and profiled for VOCs using the same technique. These profiles were used to perform multivariate analysis to interpret trends within the data and to identify potential biomarkers associated with lung infections in CF subjects. To achieve these aims, the following thesis objectives were conducted:

1. Optimisation of a sample collection and analytical technique for identification of VOCs in breath and sputum samples
2. Collection and analysis of sputum and breath samples from CF subjects and healthy controls using this optimized technique
3. Comparison of CF subjects with healthy controls using breath profiles to determine VOCs differentiating between CF and control population

4. Comparison among CF subjects with certain lung infections using a set of matching sputum and breath profiles to understand potential transition of VOCs from sputum to breath
5. Further comparison among CF subjects using an extended set of breath profiles

Chapter 2:
OPTIMISATION OF SPUTUM SAMPLING

Chapter 2: OPTIMISATION OF SPUTUM SAMPLING

2.1. Introduction

The lungs and airways of subjects with cystic fibrosis (CF) are characterised by chronic bacterial and/or fungal infections (Lyczak et al., 2002). To date, sputum (a thick mucus produced in the lungs and lower airways) culture is the primary method for the detection and identification of lung infections in CF subjects. However, culturing can take several days to weeks to produce results based on the type and complexity of the infections (Isles et al., 1984). For this reason, there is a growing interest in the identification of volatile organic compounds (VOCs) specific to certain infections to facilitate their early diagnosis based on VOC profiles. In particular, a large number of studies have reported the profiling of VOCs in the headspace of *in vitro* cultures from CF associated microorganisms (Shestivska et al., 2011, Chippendale et al., 2014, Dryahina et al., 2016, Neerincx et al., 2016a, Labows et al., 1980, Carroll et al., 2005, Briard et al., 2016, Purcaro et al., 2018, Nizio et al., 2016, Franchina et al., 2019, Veselova et al., 2019). A recent study by Nizio et al. (2016) reported diverse VOCs from culture media (e.g. LB-Lennox and BHI broth), while studying VOCs released by different species associated with lung infections in CF subjects (e.g. *Pseudomonad aeruginosa, Burkholderia cenocepacia, Haemophilus influenzae, Stenotrophomonas maltophilia, Streptococcus pneumoniae and Streptococcus milleri*) (Nizio et al., 2016). Hence, it is important to study the VOCs directly in the headspace of sputum samples from CF subjects to minimise the effect of the VOCs associated with the culture media. Only a few studies have reported VOCs directly in the headspace of sputum samples from subjects with lung infections (Goeminne et al., 2012, Savelev et al., 2011, Arslan et al., 2019, Hahn et al., 2020), and none of these studies reported the VOCs present in the corresponding breath samples from the same patient. This presents a gap in the literature as there is currently no information about the potential transition of VOCs from sputum to breath and whether the VOCs produced by bacterial infections in sputum can be readily detected in breath samples. In this study, both sputum and breath samples were collected from CF subjects and profiled for VOCs using comprehensive two-dimensional gas chromatography – time-of-flight mass spectrometry (GC×GC–TOFMS). This facilitated an extensive profiling of VOCs using an advanced analytical technique and allowed comparison between sputum and breath profiles.

This chapter will discuss the procedures involved in the collection of sputum samples from CF subjects, the extraction of VOCs from sputum samples using an optimised headspace solid-phase microextraction (HS-SPME) method, and their chemical characterisation using GC×GC–TOFMS. Discussion will also cover the evaluation of this HS-SPME-GC×GC–TOFMS method in terms of its ability to extract, detect and identify the VOCs associated with CF sputum and to allow differentiation between samples collected from different CF subjects. The following chapter will focus on the corresponding breath samples.

2.2. Materials and Method

In this chapter, a series of method optimisation experiments were performed to facilitate the suitable extraction and analysis of VOCs in the headspace of sputum samples collected from CF subjects. A total of nine individual sputum samples were used in the method optimisation experiments while the following HS-SPME parameters were tested: sample incubation time, duration of headspace extraction, and fibre desorption time. For all these parameters, the optimal sampling variable was determined in terms of the number of detected compounds for different tested durations. After the optimisation of SPME parameters, the stability of the optimised technique was then evaluated using another set of sputum samples collected from five individual CF subjects (CF01 to CF05). Triplicate HS-SPME samples were extracted from each of these five sputum samples and analysed for VOCs using GC×GC–TOFMS (this approach provided a total of 15 samples). In addition, a set of five SPME samples were also collected from blank SPME vials to determine any background VOCs (e.g. artefacts from septum or impurities from internal standards). This also allowed the identification of sputum related compounds without any influence of the background VOCs.

2.2.1. Sample collection

Sputum samples were collected from CF subjects during their regular visit to an Outpatient CF Clinic at Westmead Hospital, NSW, Australia during their routine visits. Signed consents (Appendix A) were collected from each participant before any sampling procedure was performed (under the approval # LNR/14/WMEAD/386, approved by the human research ethics committee of Western Sydney Local Health District). Participants were asked to sit and relax and provide a set of two sputum samples, one for routine microbial culture and another for the profiling of VOCs. The sample for the profiling of

VOCs was collected in a 20 mL SPME vial with airtight screw caps containing 1.3 mm thick polytetrafluoroethylene/silicone septum (Sigma-Aldrich, Castle Hill, NSW, Australia). The vial was sealed immediately after sampling and transported to the University of Technology Sydney (UTS) for VOCs analysis within 24h of collection to maintain sample integrity.

2.2.2. Sample extraction

Considering the biohazardous nature of the sputum samples, HS-SPME was used as the preferred technique for the extraction of VOCs in their headspace, which provided static extraction from a sealed system. The core HS-SPME procedure was adopted from a recent study carried out in the research group which optimised and reported the *in vitro* analysis of VOCs using GC×GC-TOFMS to differentiate the common bacteria associated with lung infections in CF subjects (Nizio et al., 2016). In brief, the VOCs were extracted manually using a divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) 24 Gauge Stableflex SPME fibre (length: 1 cm; diameter: 50 µm) attached to a manual fibre holder (Supelco, Bellefonte, PA, USA) **(Figure 2.1)**. The suitability of these fibres to collect bacterial volatiles is reported in recent studies concerning the diagnosis of lung infections in CF subjects (Bean et al., 2012, Nizio et al., 2016).

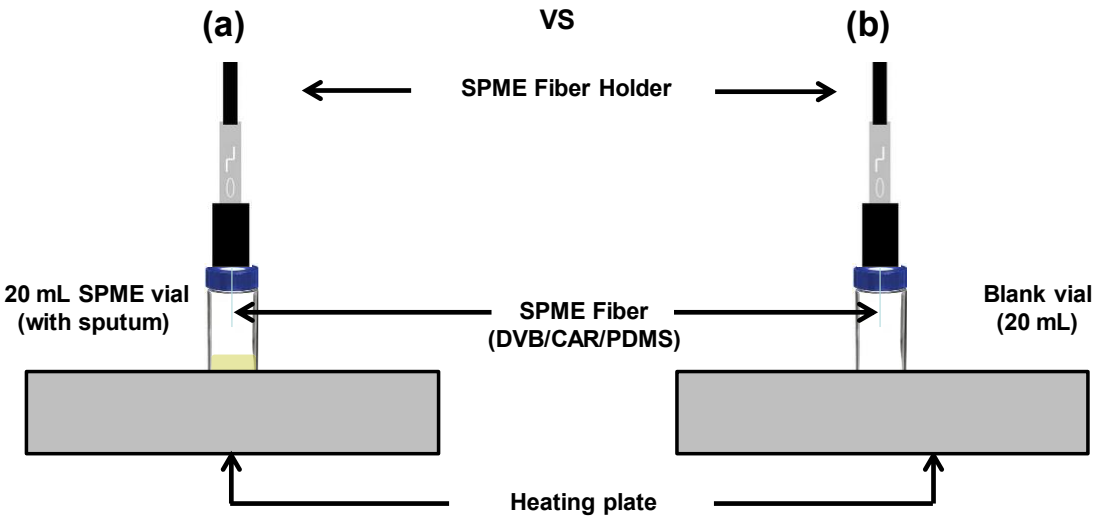


Figure 2.1: (a) Extraction of VOCs in the headspace of sputum and (b) blank vial samples

Before use, new fibres were conditioned for 60 min at 270 °C, according to the manufacturer's recommendations. Fibre reconditioning (5 min at 250 °C) was performed before the start of every sampling day and a fibre blank was also analysed. As the

concentration of VOCs was very low in the sputum headspace, no additional conditioning was performed between samples in a single sampling day. Prior to analysis, each SPME fibre was also preloaded with an internal standard to monitor the response of the instrument and to normalise samples relative to a fixed amount of standard. To facilitate this preloadings process, 200 µL of 100 ppm deuterated chlorobenzene prepared in methanol (HPLC grade, Sigma-Aldrich, Castle Hill, NSW) was placed inside a 20 mL sealed SPME vial. The fibre was manually exposed in the headspace of the internal standard for 15 s at room temperature for the preloading before removing it and placing it in a GC inlet for desorption of VOCs. Please note, the SPME extraction procedure and chromatographic parameters are identical for both sputum and vial blank samples. However, in the case of vial blanks, each vial was pre-cleaned by flushing with ultrapure N_2 (99.999%, BOC, Australia) for 2/3 minutes before use.

2.2.3. GC×GC-TOFMS analysis of sputum samples

All SPME samples were analysed using a Pegasus® 4D GC×GC–TOFMS (LECO, Castle Hill, NSW, Australia) equipped with a liquid nitrogen cryogenic quad jet modulator **(Figure 2.2)**.

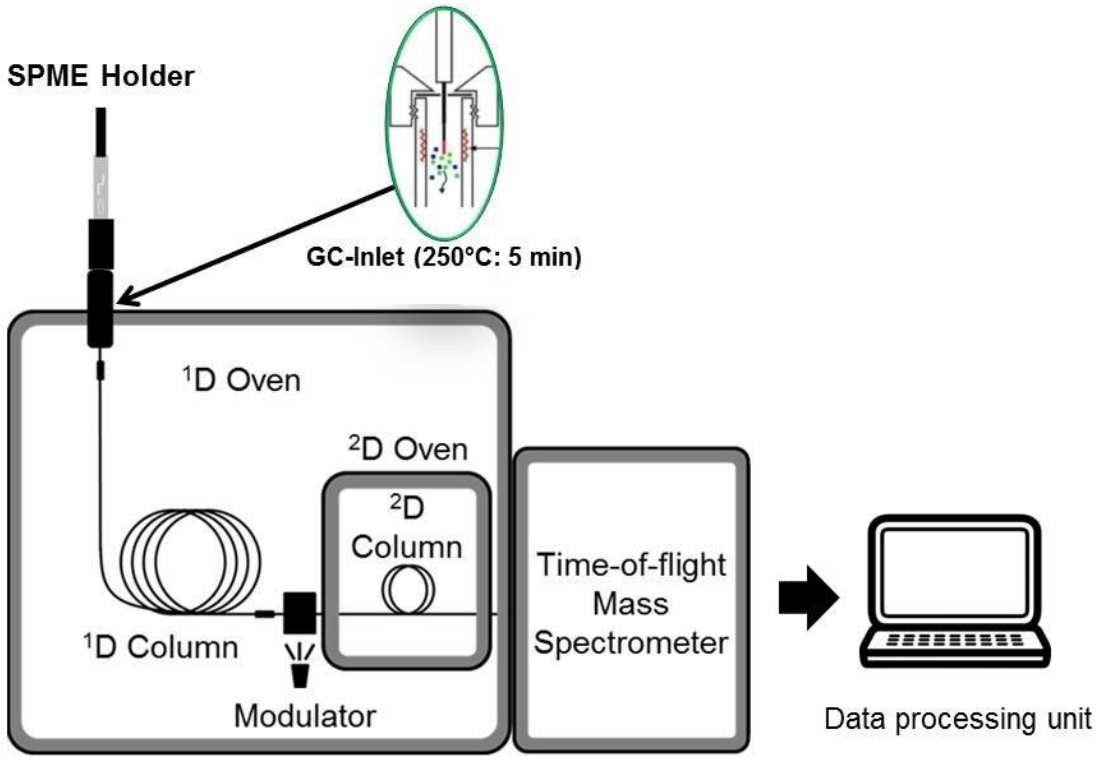


Figure 2.2: GC×GC–TOFMS system used for the analysis of sputum samples

The VOCs desorbed in the GC×GC inlet were first separated on a mid-polar first dimension (^{1}D) column (Rxi1-624Sil MS: 30 m *(length)* x 0.25 mm *(internal diameter)* x 1.40 µm *(film thickness)*; Restek Corporation, USA) **(Table 2.1)**. Further separation of VOCs was performed on a polar second dimension (^{2}D) column (Stabilwax (2 m *(length)* x 0.25 mm *(internal diameter)* x 0.50 µm *(film thickness)*; Restek Corporation, USA) which was connected to the ^{1}D column using a SilTite® µ-Union (SGE Analytical Science, Australia). High purity helium (BOC, Sydney, NSW, Australia) was used as the carrier gas at a constant flow rate of 2 mL/min. The GC oven setup was as follows: initial temperature 40 °C held for 0.2 min; ramp of 10 °C/min to 230 °C and held for 0.8 min; total run time of 20 min. The modulator offset was +30 °C relative to the ^{2}D oven and the ^{2}D oven temperature offset was +5 °C relative to the ^{1}D oven. The modulation period was 4 s with a hot pulse of 0.4 s and 1.6 s cooling time between stages. The MS transfer line was maintained at 250 °C and the mass acquisition was performed with a range between 25 and 500 atomic mass units at a rate of 200 Hz. The ion source temperature was set at 200 °C and the electron ionisation energy was -70 eV. All GC×GC-TOFMS parameters were controlled by ChromaTOF® version 4.51.6.0 (LECO).

Table 2.1: GC×GC-TOFMS parameters for the analysis of VOCs

i. Oven setup			**ii. ^{1}D column (Rxi-624Sil MS)**		
Initial temperature	40	°C	Length (*l*)	30	m
Ramping rate	10	°C/min	Internal diameter (*id*)	0.25	mm
Final temperature	230	°C	Film thickness:	1.4	µm
Initial hold time	0.2	min			
Final hold time	0.8	min	**iii. ^{2}D column (Stabilwax)**		
Total run time	20	min	Length (*l*)	2	m
Carrier gas	He	High purity	Internal diameter (*id*)	0.25	mm
Carrier gas flow	2	mL/min	Film thickness:	0.5	µm
iv. Modulator			**v. Detector**		
Temperature offset (^{1}D to ^{2}D)	+5	°C	MS transfer line temperature	250	°C
Temp offset (^{2}D to modulator)	+30	°C	Ionization mode/energy (EI)	-70	eV
Modulation period	4	s	Ion source temperature	200	°C
Hot pulse	0.4	s	TIC scan range	25-500	m/z
Cooling time	1.6	s	Scan rate	200	Hz

2.2.4. Processing of raw data

A computerised data acquisition and peak integration system (ChromaTOF® version 4.51.6.0) was used to process the GC×GC–TOFMS raw data. The baselines of the

chromatograms were automatically smoothed by the software with an 80% offset. The ^{1}D and ^{2}D peak width was set at 20 s and 0.1 s, respectively; while, the minimum signal-to-noise ratio (S/N) for the base peak and sub-peaks was set at 250 and 20, respectively. For acquisition, a minimum of two apexing masses had to have a SN > 250. The identification of compounds was performed using the mass spectral library database provided by the USA National Institute of Standards and Technology (NIST, 2011) and a match > 80% was required to assign a peak name. The alignment of peaks identified using the NIST database was performed using the Statistical Compare software feature in ChromaTOF®. Statistical Compare was performed with samples sorted into two classes (e.g. sputum samples or vial blanks). A signal-to-noise (S/N) of 20 was used to search for peaks not found during the initial peak finding step. In addition, a mass spectral match >60% was required for peaks to be identified as the same compound across chromatograms, while sub-peaks were combined when the second dimension retention time shift was ≤ 0.1 s. In addition, a peak had to appear in at least three samples to be retained in the peak table as all/most of the samples were analysed in triplicates. After alignment, the analyte peak areas were normalized using the internal standard peak area.

Further statistical comparison was made to identify the compounds with the highest between-class variance applying Fisher Ratio (FR) calculated using the same Statistical Compare software feature in ChromaTOF®. This was performed by comparing the calculated FR to a critical *F* value (F_{crit}) obtained through an *F*-test (Bean et al., 2015, Nizio et al., 2016). Analytes with higher FR values (or analytes absent from a class or only detected in a single sample in a class thus labelled as 'undefined') indicated compounds that statistically differed in abundance between the defined classes and were therefore retained. The application of FR filtering is reported to be very effective for identifying class-distinguishing compounds while processing complex GC×GC data (Beckstrom et al., 2011, Pierce et al., 2006, Mohler et al., 2007, Stefanuto et al., 2015). For all samples where an *F*-test was performed, compounds with FR above the F_{crit}, which includes those labelled as 'undefined', were exported as a *.csv file. The resulting *.csv file was imported into Microsoft Excel for the manual removal of chromatographic artefacts (i.e. column bleed and solvent peaks) and further multivariate analysis.

2.2.5. Principal component analysis

Principal component analysis (PCA) was performed using Unscrambler® X (version 10.3; CAMO Software, Oslo, Norway) to assess trends in the data and to group samples using

scores plots. Prior to PCA, the data was pre-processed using mean centering and scaling and unit vector normalisation. These pre-treatment steps have been reported in several recent studies analysing VOCs from complex biological samples (Goeminne et al., 2012, Nizio et al., 2016, Pesesse et al., 2019).

2.3. Results and Discussion

2.3.1. Optimisation of basic SPME parameters

To optimise the extraction of VOCs in the headspace of sputum samples, the following HS-SPME parameters were tested:

1. Duration of sample incubation (10, 30, and 60 min)
2. Duration of headspace extraction (10, 30, and 60 min)
3. Duration of fibre desorption in the GC -inlet (5, 10, and 15 min)

The first optimisation step was carried out to determine the suitable duration of sample incubation which is the first stage of an HS-SPME procedure. To facilitate this process, three different sputum samples were collected from three individual CF subjects at the CF clinic. Sample numbers were limited to ensure that all samples could be analysed on the day of collection without the requirement for cold storage. Samples were transported to the laboratory within four hours of collection and held at ambient temperature before any sampling procedure was performed. The quantity of each sample was measured using an analytical balance (model 220-4M, resolution 0.1 mg, KERN & SOHN GmbH, Germany) and recorded along with other parameters (e.g. the appearance of the samples in terms of colour and density). For HS-SPME, a vial containing the sputum sample was placed in a custom heating system designed to fit SPME vials (Thermoline Scientific, Wetherill Park, NSW, Australia). The incubation temperature was maintained at 37.5 °C to stimulate normal human body temperature. Three triplicates were then collected from the same sputum sample using different incubation times of 10, 30, and 60 min (the order of incubation was randomised so the extraction was not in order of increasing incubation time). **Figure 2.3** shows the diagram of incubation and headspace extraction to illustrate how multiple incubation times could be tested on a single sample. For all incubation times, the headspace extraction and fibre desorption time was maintained at 10 and 5 min, respectively. This allowed for one parameter (i.e. incubation time) to be tested at a time,

while the two other parameters remained constant. Desorption of SPME fibres was performed in the GC×GC inlet set at 250°C. The same procedure was applied to all triplicates. This provided a set of triplicates for each of the three tested durations (10, 30, and 60 min).

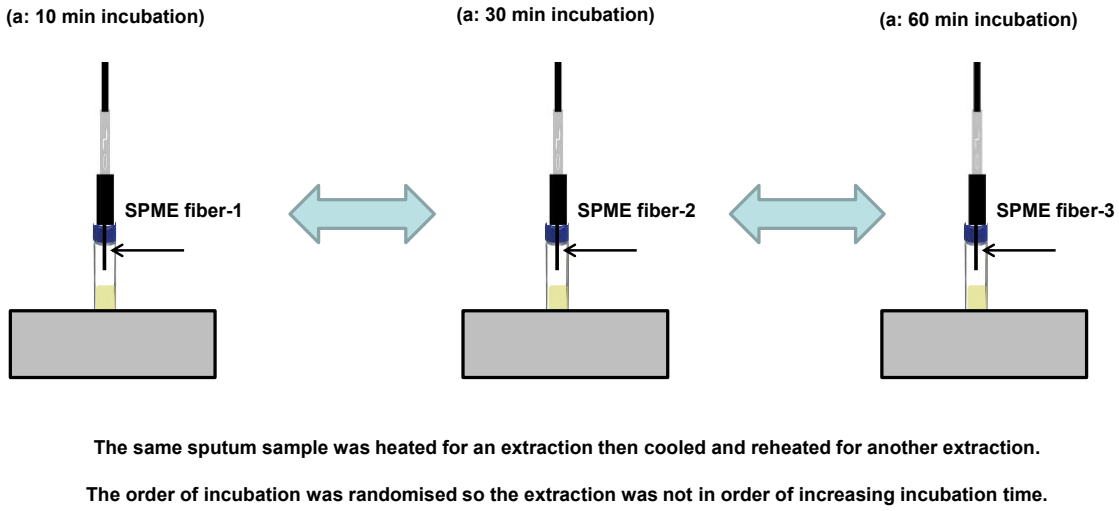


Figure 2.3: Diagram showing the incubation of the same sputum sample at different durations (10, 30, and 60 minutes) and the extraction of VOCs with different SPME fibers.

Table 2.2 lists the average number of VOCs detected for the different durations tested as well as the other two parameters optimised (and discussed below). There was no difference in the VOC profile between the different incubation times. Based on this observation, 10 min was selected as the optimal incubation time for sputum samples.

Table 2.2: Results of HS-SPME method optimisation experiments

Parameter	Average number of detected compounds for different durations					
	10 min	30 min	60 min	5 min	10 min	15 min
Incubation time for sputum samples	41	39	40			
Extraction time for VOCs	72	72	73			
Desorption time for SPME fibre				82	80	75

The second parameter optimised was the extraction time for VOCs in the headspace of sputum samples. Three durations were tested at 10, 30, and 60 min. To facilitate this process, another set of sputum samples was collected which contained three separate samples from three individual CF subjects. Each sample was incubated for 10 min at

37.5°C, while HS-SPME was performed in triplicate for the three different extraction durations (i.e. 10, 30, and 60 min) at the same temperature. The order of extraction was randomised, meaning, the extraction was not in order of increasing extraction time; therefore the first extraction could have occurred for 10 min, 30 min or 60 min. This procedure was maintained to assure that the number of detected VOCs was not biased by the order of extraction as it can be perceived that the first extraction would generate more VOCs. However, as shown in **Table 2.2**, the number of detected VOCs did not change with increasing extraction time. Based on this observation, the extraction time was set at 10 min for rapid analysis of samples.

Finally, three different desorption times (5, 10, and 15 min) were tested to determine the time required for complete desorption of VOCs extracted on SPME fibres. Like the other two parameters, this parameter was also optimised using three individual sputum samples in triplicate. Incubation and extraction times were both maintained at 10 min, while desorption times varied at 5, 10, and 15 min. Notably, the number of detected VOCs reduced with increasing desorption time (although the optimisation process was randomised). For instance, a 5 min desorption yielded a total of 82 VOCs, while only 75 VOCs were detected after 15 min of fibre desorption. The lower number of VOCs with the higher desorption time indicates possible degradation of VOCs with increasing desorption time at high temperatures (Emmons et al., 2019). Overall, the results of this experiment show that a 5 min desorption is optimal to retain a maximum number of breath VOCs as they are highly volatile in nature.

2.3.2. Evaluation of SPME extraction technique

After the optimisation of HS-SPME parameters, the stability of the technique was evaluated using a set of five sputum samples collected from five individual CF subjects (CF01 to CF05). Triplicate HS-SPME samples were extracted from each of these five sputum samples and analysed for VOCs using GC×GC–TOFMS, which provided a total of 15 different GC×GC chromatograms. In addition to the sputum samples, a set of five SPME samples were also collected from blank SPME vials to determine any background VOCs (e.g. artefacts from septum or impurities from internal standards) which allowed identification of sputum related compounds without any influence of the background VOCs. **Table 2.3** lists the detection frequencies of all 46 VOCs that were detected in the headspace of sputum samples and vial blanks. Among these 46 VOCs, 32 were detected only in sputum samples (order 1 to 32), 11 were common in both sample types (order 33

to 43), and 3 were only detected in vial blanks (order 44 to 46). The 32 VOCs detected only in sputum were denoted as characteristic VOCs related to sputum samples. The three VOCs detected only in the blanks (acetic acid ethenyl ester, 2-butoxy-ethanol, and 1-tridecanamine, N,N-dimethyl-) were probably from background/room air. In addition, the 11 VOCs presents in both sample types may have resulted from the sputum samples or the internal standard. However, among these 11 VOCs, nine have high (>1.5) relative abundance (calculated as the ratio of average peak area values: sputum/blanks) in sputum samples, thus were considered as sputum related VOCs.

Table 2.3: Volatile organic compounds detected in sputum samples and vial blanks

Order	***Compounds***	***Frequency of detection (in a total of 20 samples)***		***Relative abundance (sputum/blank)***
		Sputum (15)	***Blanks (5)***	
1	1,5-Hexadien-3-ol	11	0	-
2	1-Butanol, 3-methyl-	12	0	-
3	1-Hexanol	12	0	-
4	1-Octen-3-one	12	0	-
5	1-Pentadecanamine, N,N-dimethyl-	12	0	-
6	1-Pentanol	12	0	-
7	1-Propanol	12	0	-
8	2,3-Butanedione	12	0	-
9	2,3-Octanedione	11	0	-
10	2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	13	0	-
11	2-Nonanone	6	0	-
12	2-Piperidinone	4	0	-
13	3-Octanone	12	0	-
14	Acetic acid	8	0	-
15	Acetoin	15	0	-
16	Aziridine, 1-ethenyl-	11	0	-
17	Benzaldehyde	12	0	-
18	Butanal, 3-methyl-	3	0	-
19	Butylated Hydroxytoluene	3	0	-
20	Ethylbenzene	12	0	-
21	Hexanal	12	0	-
22	Indole	3	0	-
23	Methylene chloride	11	0	-
24	Morpholine, 4-octadecyl-	12	0	-
25	N-Morpholinomethyl-isopropyl-sulfide	12	0	-
26	Phosphonic acid, (p-hydroxyphenyl)-	8	0	-
27	Propane, 2,2-dimethoxy-	14	0	-
28	p-Xylene	12	0	-
29	Pyridine	8	0	-
30	Pyrrole	11	0	-
31	Tetrahydrofuran	15	0	-
32	Trichloromethane	12	0	-
33	**α-Terpineol**	**15**	**5**	**0.27**
34	1-Hexanol, 2-ethyl-	15	5	1.55
35	Toluene	15	5	1.55
36	Propanoic acid, 2-methyl-, 3-hydroxy-2,4,4-trimethylpentyl ester	15	5	3.76
37	2-Pentanone	15	5	5.02
38	Benzene	15	5	18.4
39	1-Octen-3-ol	15	3	192
40	2-Heptanone	15	1	67.9
41	Oxime-, methoxy-phenyl-_	12	5	1.66
42	**Ethyl Acetate**	**3**	**5**	**0.66**

43	Propanoic acid, 2-methyl-, 2,2-dimethyl-1-(2-hydroxy-1-methylethyl)propyl ester	3	5	3.26
44	Acetic acid ethenyl ester	0	5	-
45	Ethanol, 2-butoxy-	0	5	-
46	1-Tridecanamine, N,N-dimethyl-	0	5	-

In contrast, two of the VOCs with low (<1) relative abundance in sputum samples (α-terpineol and ethyl acetate; shown in bold in **Table 2.3**) were believed to be derived from the HPLC grade methanol used to prepared the internal standard (100 ppm deuterated chlorobenzene) and were excluded from further statistical analysis. The exclusion of the VOCs derived from the background (total three) and internal standards (total two) provided a list of 41 VOCs related to sputum samples, which indicates the effectiveness of the applied technique to extract VOCs in the headspace of sputum samples with little interference from the background environment and internal standard.

2.3.3. Evaluation of the multivariate analysis

After the evaluation of the SPME extraction technique, further evaluation was performed to assess the efficacy of the developed technique to discriminate between sample types. To facilitate this process, principal component analysis (PCA) was performed using the VOCs (total 41) detected in the triplicates collected in the headspce of sputum samples (CF 01 to CF 05). As shown in the PCA scores plot, triplicates from each of the five samples (highlighted in five different colors) clustered together **(Figure 2.4)**. For instance, triplicates collected in the headspce of sputum sample CF03 (green rectangles) clustered together at the third quadrant of the plot. As expected, the blanks also clustered together (red dots at the right-hand side of the plots). The clustering of samples in PCA indicates that the multivariate analysis performed in this study is able to group multiple extractions from the same sputum samples precisely. In addition, there is also some clustering in between different sputum samples probably based on their infection status or other factors. However, the number of samples (only five) is not sufficient to reach any conclusion. A further chapter of this thesis **(Chapter 5)** will discuss a large number of sputum samples with particular focus on the relationship between VOC profiles and microbial cluture results (infection status) of the sputum samples.

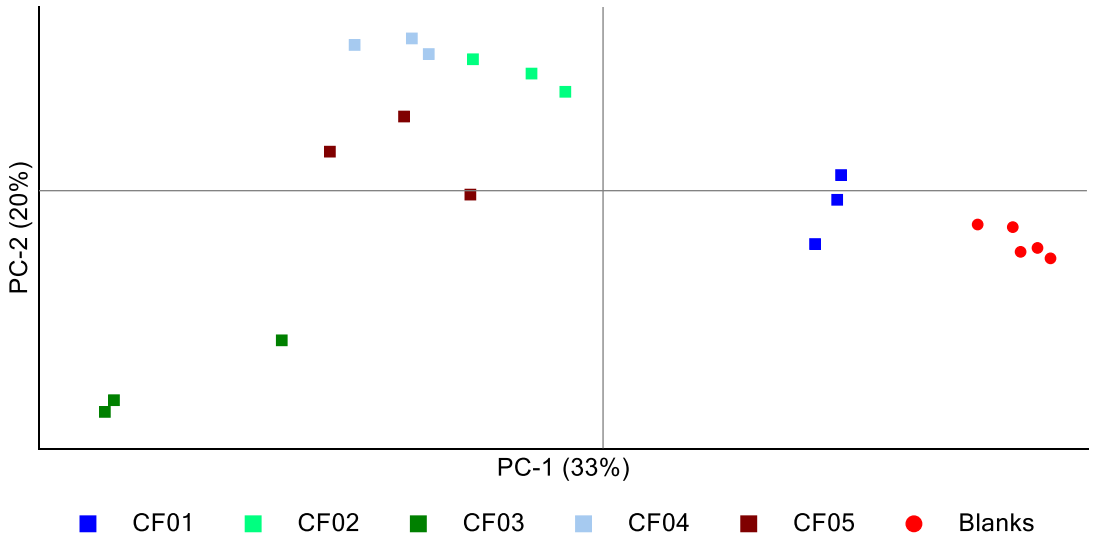


Figure 2.4: PCA scores plot prepared using pre-processed GC×GC-TOFMS peak area data obtained from sputum samples and blank SPME vials

2.4. Conclusion

This chapter reported the application of a SPME-based method for the optimal extraction of VOCs present in the headspace of sputum samples and their chemical characterisation using GC×GC-TOFMS. The analysis of background VOCs showed very little influence from the SPME vials and/or internal standards used in this study and also allowed elimination of any artefacts. Evaluation of the applied method indicates its capability to extract, detect and identify the VOCs associated with CF sputum. In addition, the accurate clustering of triplicate samples in the PCA plots indicates the suitability of the multivariate analysis performed in this study.

Chapter 3:
OPTIMISATION OF BREATH SAMPLING

Chapter 3: OPTIMISATION OF BREATH SAMPLING

3.1. Introduction

The selection of a sampling device and extraction technique are two important parameters for the optimal analysis of volatile organic compounds (VOCs) present in exhaled human breath samples (Wilde et al., 2019). Tedlar® bags (Mochalski et al., 2009) and canisters (Pleil and Lindstrom, 1995) are two reported choices for the sampling and storage of VOCs; however, the handling and transportation of bag samples are significantly easier than that of canisters considering their volume and weight. A number of recent studies have also reported the use of the Bio-VOC™ breath sampler for rapid sampling of VOCs in exhaled human breath (Phillips et al., 2014, Kramer et al., 2015, Marco and Grimalt, 2015). However, Bio-VOC™ is not designed for sample storage and VOCs need to be extracted immediately onto sorbent tubes (STs) for further analysis using a GC equipped with a thermal desorption (TD) unit (for the desorption of VOCs collected on STs). In contrast, breath samples collected in Tedlar® bags can either be extracted using a solid-phase microextraction (SPME) fibre or an ST packed with sorbent material.

In this study, optimisation of breath sampling was performed in a stepwise manner. Initially, the performance of two breath sampling devices (Bio-VOC™ breath sampler vs Tedlar® breath sampling bag) was evaluated to determine the most suitable container for VOC collection. Based on the results which indicated that Tedlar® breath sampling bags were the optimal container, two different extraction techniques (SPME vs ST-extraction) were evaluated using breath samples to determine the most appropriate combination of sampling device-extraction technique.

3.2. Materials and Method

To evaluate the performance of the two sampling devices, breath samples were collected from five individual healthy participants at the same time using both Bio-VOC™ breath sampler and Tedlar® bags. The samples were extracted immediately onto sorbent tubes (STs). In parallel to breath samples, an equal number of corresponding background air samples were also collected at the same sampling location. The performance of these sampling devices were then evaluated based on their capability to distinguish between healthy human breath and background air collected simultaneously at the same location.

After the determination of the most suitable sampling device, further experiments were performed to determine the optimal combination of sampling device-extraction technique. To facilitate this process, breath samples were collected in Tedlar® bags from two contrasting groups: CF subjects with different lung infections (n=15) vs healthy participants without any known lung infections (n=15). The VOCs present in these samples were then extracted using both SPME fibres and STs packed with sorbent material. The performance of these extraction techniques will be evaluated in terms of their ability to differentiate between the two contrasting study groups.

3.2.1. Recruitment of subjects

Healthy participants were recruited from the Faculty of Science, University of Technology Sydney (UTS). Signed consents (Appendix B) were collected from each participant before any sampling procedures were performed (under the approval # ETH17-1806, approved by the human research ethics committee of UTS). CF subjects were recruited from the Outpatient CF Clinic at Westmead Hospital, NSW, Australia and breath samples were collected along with signed consent forms (under the approval # LNR/14/WMEAD/386, approved by the human research ethics committee of Western Sydney Local Health District). **Table 3.1** lists the anthropometric characteristics of the study population for the study concerning the optimal combination of sampling device-extraction technique. All parameters are in close range for both groups as only adults were recruited in this study.

Table 3.1: Anthropometric characteristics of the study populations

Parameters	CF subjects (n=15)	Control subjects (n=15)
Age (y: average ± SD (Min/Max))	26.2 ± 7.49 (18/45)	31.1 ± 9.71 (23/56)
Sex, M/F (ratio)	8/7 (1.14)	7/8 (0.86)
Height (cm: average ± SD (Min/Max))	169 ± 7.12 (156/182)	171 ± 8.24 (158/186)
Weight (kg: average ± SD (Min/Max))	63.7 ± 5.77 (50/72)	70.0 ± 16.5 (50/105)

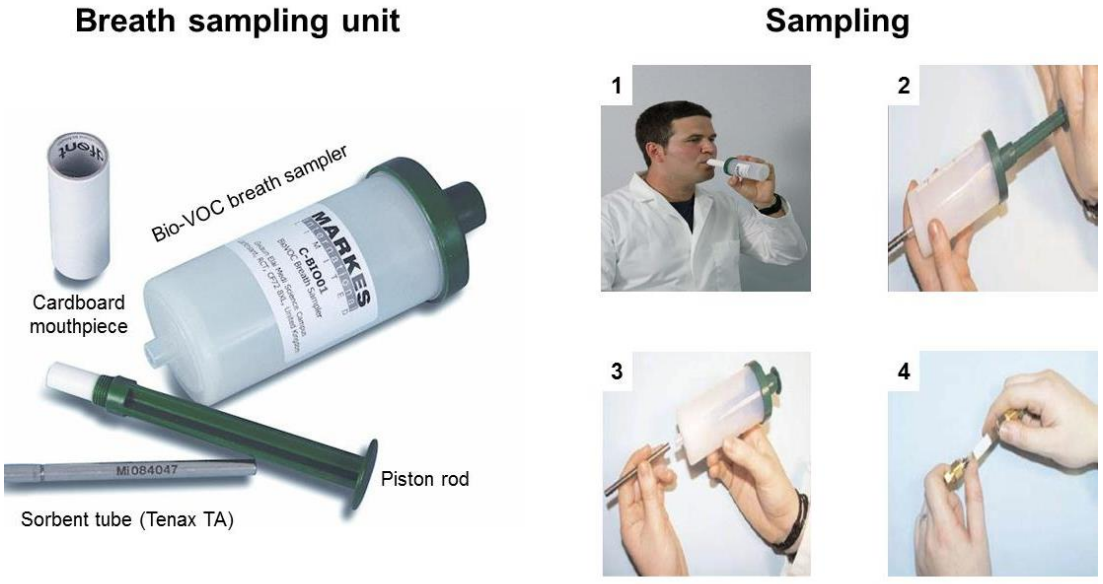
To note, breath samples were only collected from subjects over 18 years old, either with CF (for CF group), or with no CF and any known lung infection/s (for control group). Exclusion criteria for both groups included: regular smoker, clinical history of halitosis, poor dentition/gingivitis/gum disease, eaten a meal/had a drink/s within 1 hour, and inability to provide breath sample. Any subjects with cardiovascular disease, chronic anaemia, panic disorders, and any digestive disease/s were excluded from the healthy control group. In addition, for the CF group all details of medication and lung infection

status data was collected along with anthropometric data. However, some other parameters that could potentially impact breath VOC profiles, such as food habits and/or teeth brushing pattern of participants, were not recorded or accounted for in this study.

3.2.2. Sampling devices

3.2.2.1. Bio-VOC™ samples

Bio-VOC™ was a commercially available breath sampling device (now replaced by an updated version Bio-VOC-2™) marketed by Markes International Limited, UK (Wilson and Monster, 1999). For the testing of this device, breath sampling was performed in the UTS analytical chemistry laboratory where there is no storage of solvents or chemicals and adequate ventilation. Participants were asked to sit in a chair and rest until they were fully relaxed; then they were asked to take a deep breath and slowly exhale a complete single breath (expel until the full capacity of the lungs was reached) into the sampler **(Figure 3.1)**.



[Source: https://www.markes.com/Products/Sampling-accessories/Sampling/BIO-VOC-breath-sampler.aspx]

Figure 3.1: Breath sample collection using the Bio-VOC™ breath sampler

The sampler consisted of a PTFE syringe (capacity 129 mL of air) and an orifice (dia. 3 mm) at the opposite end to where the breath is introduced. The function of the orifice is to allow flow through, so that the initial dead-space air is displaced progressively by air from the alveolar region of the lungs as exhalation proceeds (Phillips et al., 2014). After complete exhalation, a hand-operated plunger was used to slowly force the breath through the orifice into a clean stainless-steel ST packed with 100 mg of Tenax® TA sorbent

material (35/60 mesh; Markes International Limited, UK). The manoeuvre was repeated 3 times for extraction of the VOCs (Jareño-Esteban et al., 2013).

Historically, Tenax® TA has been reported as the preferred adsorbent for trace VOCs due to its hydrophobic nature and the capability to retain a wide range of compounds (McCaffrey et al., 1994, Maria, 1996, Helmig and Vierling, 1995). Tenax® TA is also reported as a suitable sorbent material to collect and store VOCs in exhaled breath samples for up to 14 days (Harshman et al., 2016). Three samples were collected from each of five participants using three different STs packed with Tenax® TA which provided a total of 15 Bio-VOC™-ST breath samples. Corresponding room air samples (blank Bio-VOC™ samples filled with room air) were also collected simultaneously for comparison purposes. The same sampling procedure was used for all subjects; however, a different Bio-VOC™ sampler was used for each participant and each sampler was flushed with ultrapure N_2 (99.999%, BOC, Australia) for 2 to 3 minutes before use.

After sampling, each ST was immediately sealed with brass long-term storage caps, wrapped in aluminium foil, placed in a glass air-tight container, and stored at 4 °C in this condition, until GC×GC-TOFMS analysis was carried out in accordance with the U.S. Environmental Protection Agency Method TO-17 (Woolfenden and McClenny, 1999). All samples were analysed within 14 days of extraction as recommended by Hrashman et al. (2012) (Harshman et al., 2016). Before analysis, each sample was also loaded with 1 µL of 100 ppm deuterated chlorobenzene prepared in methanol (HPLC grade, Sigma-Aldrich, Castle Hill, NSW) as an internal standard to facilitate peak area normalisation (Knobel et al., 2018). For the loading of the internal standard, 1 µL of the standard was taken from a stock internal standard solution (prepared at 10 mL volume and stored in a sealed SPME vial at 4 °C) using a micro-syringe (capacity 5 µL, SGE Analytical Science, Wetherill Park, NSW, Australia). This 1 µL of internal standard was then loaded on the top end of the ST.

3.2.2.2. Tedlar® bag samples

Tedlar® bags are widely used for the sampling and storage of VOCs. In this study, 1 L transparent Tedlar® bags (SKC Inc., USA) were used which are specially designed for breath sampling and fitted with a stainless steel breath-gas fitting **(Figure 3.2a)**. The superior performance of these Tedlar® bags over Kynar and Flexfilm bags in terms of background emission and the stability of 41 VOC species, was reported in a previous study (Mochalski et al., 2013). Before collecting breath samples, all bags were thoroughly

cleaned to remove residual contaminants by flushing with ultrapure N_2 (99.999%, BOC, Australia) 10-14 times; this also allowed the re-use of the bags (Mochalski et al., 2009). In addition, no bag was used more than five times for the collection of breath samples.

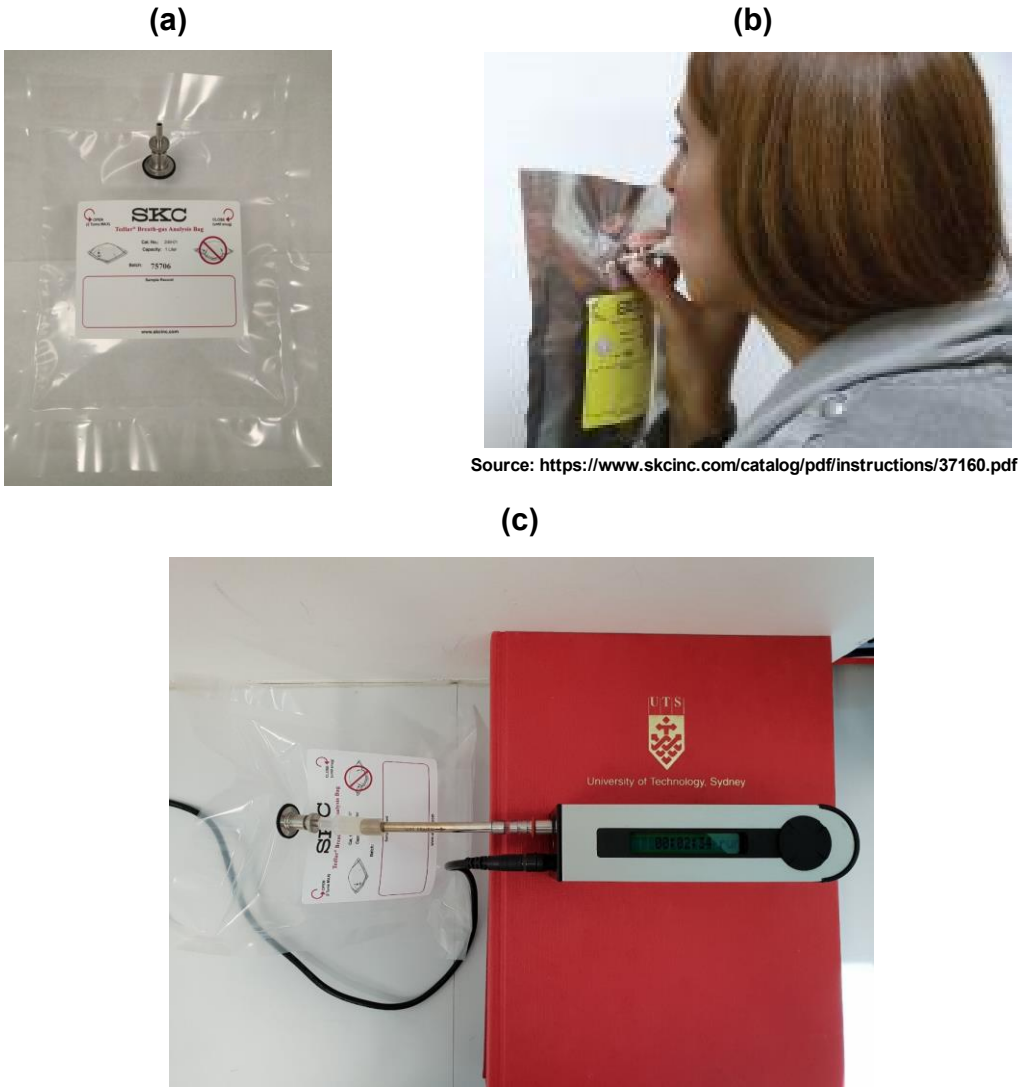


Figure 3.2: (a) 1 L Tedlar® breath sampling bag; (b) Exhalation in to bag; (c) Sorbent tube extraction of breath samples

The bag sampling procedure was performed at the same laboratory location and at the same time as the Bio-VOC™ samples. Participants were asked to sit and relax; then take a deep breath and exhale out the initial part of the breath (mouth air) and were then asked to fill the 1 L Tedlar® bag, expiring the remaining part of the breath as the sample using a disposable mouthpiece **(Figure 3.2b)**. This approach allowed the removal of any VOCs potentially from exogenous sources such as foods/beverages and to collect the remaining part of the breath (Harshman et al., 2016). To note, samples were only collected from

participants who had not ingested any foods/drinks at least in the past one hour. Three samples were collected from each of five participants in three different bags which provided a total of 15 bag samples. Corresponding room air samples (1 L Tedlar® bags filled with room air) were also collected simultaneously for comparison purposes. To collect room air samples, a pre-cleaned empty Tedlar® bag was attached to the output of a mini-pump interfaced with a mass flow controller (CHEMATEC FL-1001, Markes International Limited, UK) using clear Tygon tubing. The pump flow rate was set at 250 mL for four minutes to fill the bag with 1 L of room air sample.

The breath and room air samples collected on Tedlar® bags were extracted on stainless steel STs packed with Tenax® TA (100 mg; 35/60 mesh; Markes International Limited, UK). To facilitate this process, the bag filled with breath sample was connected to one end of the ST; the other end of the ST was connected to a mini-pump interfaced with a mass flow controller (CHEMATEC FL-1001, Markes International Limited, UK) to dynamically draw the VOCs onto the sorbent tube **(Figure 3.2c)**. The dynamic sampling was performed at 250 mL/min for four minutes to completely empty the bag onto the ST and collect a total of 1 L of breath sample. After sampling, STs were stored according to the guideline provided by the U.S. Environmental Protection Agency (Method TO-17) and analysed within 14 days of extraction (Harshman et al., 2016, Woolfenden and McClenny, 1999). Before GC×GC-TOFMS analysis, each ST was also loaded with 1 µL of internal standard (100 ppm deuterated chlorobenzene) to facilitate peak area normalisation (Knobel et al., 2018).

3.2.3. Extraction techniques

3.2.3.1. SPME extraction of breath VOCs

To allow comparison between extraction techniques, breath samples were collected in the same 1 L Tedlar® bags (SKC Inc., USA) used for the sampling device study. For SPME extraction of breath VOCs, three phase (DVB/CAR/PDMS) fibres were used. These are the same fibres used for the suitable extraction of sputum samples collected in the prior study. The suitability of these fibres to collect bacterial volatiles are also reported in several recent studies (Bean et al., 2012, Nizio et al., 2016). HS-SPME extraction from bag samples was performed within 48 h of sample collection to keep the potential loss of any VOCs minimal (due to adsorption to and/or diffusion through the bag wall) (Beauchamp et al., 2008). Until extraction, bag samples were stored at room temperature (~22 °C). The extraction of VOCs was also performed at room temperature by exposing

the SPME fibre to the bag for 10 min **(Figure 3.3)**. After the extraction of VOCs, the SPME fibre was preloaded with an internal standard by exposing the fibre for 15 s in the headspace of 200 µL of a 100 ppm deuterated chlorobenzene prepared in methanol (HPLC grade, Sigma-Aldrich, Castle Hill, NSW). For VOC analysis, SPME fibres were thermally desorbed in a GC×GC inlet for 5 min at 250 °C.

Figure 3.3: SPME extraction of breath samples collected in to 1 L Tedlar® bag

3.2.3.2. ST extraction of breath VOCs

After SPME collection, the Tedlar® bag sample was then extracted onto an ST packed with Tenax® TA. These are the same STs used for the evaluation of the sampling devices. The ST sampling and storage conditions were previously described in **section 3.2.2.2**. It is important to note that the collection of a SPME sample prior to ST sampling might have an impact on the profile of VOCs obtained for the ST samples. However, considering the large volume of sample (1 L) and the typical size and capacity of a SPME fibre (length: 1 cm; diameter: 50 µm), the effect would be minimal.

3.2.4. GC×GC-TOFMS analysis of breath VOCs

3.2.5.1. Sorbent tube samples

The thermal desorption of STs was performed using a Markes Unity 2 thermal desorber (TD) equipped with a Series 2 ULTRA™ multi-tube auto-sampler (Markes International Limited, UK) **(Figure 3.4)**. This auto-sampler TD unit was controlled by Maverick Version 4.1.29 software (Markes International Limited, UK). Prior to first use, all freshly packed sorbent tubes were initially conditioned under a 70 mL/min flow of ultrapure N_2

(BOC, Sydney, NSW, Australia) for 2 h at 320 °C, followed by 30 min at 335 °C, according to the manufacturer's recommendations. In addition, the regular conditioning of tubes was performed at 300 °C for 1 h at a flow rate of 70 mL/min (ultrapure N_2; BOC, Sydney, NSW, Australia) using a TC-20™ tube conditioner (Markes International Limited, UK).

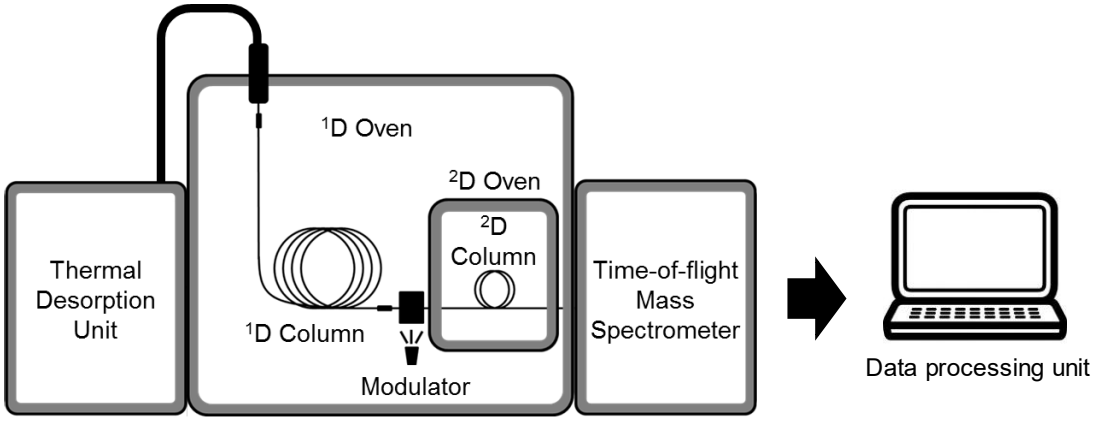


Figure 3.4: TD–GC×GC–TOFMS-unit for the analysis of breath samples collected by STs

The thermal desorption parameters adopted in this study are described in previous research concerning the analysis of VOCs from biological samples (e.g. decomposition odour (Forbes et al., 2016) and burned remains (Nizio and Forbes, 2018)). In brief, STs were placed on the multi-tube auto-sampler and transferred to the TD-unit for thermal desorption; thermal desorption was performed at 300 °C for 4 min and the VOCs were adsorbed onto a general purpose cold trap (Tenax® TA/Cerograph 1TD) at -10 °C followed by trap desorption at 300 °C for 3 min. A 5:1 split ratio was applied while the split flow was maintained at 20 mL/min. The TD unit was connected to a Pegasus® 4D GC×GC–TOFMS (LECO, Castle Hill, NSW, Australia) via a heated transfer line (1 m long uncoated fused silica transfer line; Markes International Limited, UK) maintained at 140 °C. The heated transfer line connected the TD unit to the first dimension (^{1}D) GC column inside the ^{1}D GC oven using an Ultimate Union Kit (Agilent Technologies, Mulgrave, NSW, Australia). The column configuration and GC×GC–TOFMS parameters were the same as the HS-SPME samples. In brief, the ^{1}D oven was set at: initial temperature 40 °C held for 0.2 min; ramp of 10 °C/min to 230 °C and held for 0.8 min; total run time of 20 min. The modulator offset was +30 °C relative to the ^{2}D oven and the ^{2}D oven temperature offset was +5 °C relative to the ^{1}D oven. The modulation period was 4 s with a hot pulse of 0.4 s and 1.6 s cooling time between stages. The MS transfer line

was maintained at 250 °C and the mass acquisition was performed with a range between 25 and 500 atomic mass units at a rate of 200 Hz. The ion source temperature was set at 200 °C and the electron ionisation energy was -70 eV.

3.2.5.2. SPME samples

The GC×GC-TOFMS operating conditions for the analysis of breath VOCs extracted using SPME fibres were previously described in **section 2.1.3**.

3.2.5. Processing of raw data

The processing of raw data was previously described in **section 2.2.4**.

3.2.6. Principal component analysis

Principal component analysis (PCA) was previously described in **section 2.2.5**.

3.3. Results and Discussion

3.3.1. Evaluation of sampling devices

To allow comparison between sampling devices, breath samples were collected in triplicates (5~10 min gaps between the collection of each sample) from a total of five participants using both Bio-VOC™ and Tedlar® bag. In parallel, an equal number of corresponding background air samples were also collected at the same sampling location. Overall, this approach provided a total of 60 samples **(Table 3.2)**

Table 3.2: List of breath samples collected for the evaluation of sampling devices

Order	Sampling device-extraction technique	Sample type	Sample count
1	Bio-VOC™-ST	Breath	15
2	Bio-VOC™-ST	Background air	15
3	Tedlar® bag-ST	Breath	15
4	Tedlar® bag-ST	Background air	15 (5 were lost)

However, 5 of the Tedlar® bag-ST background samples were lost due to a leak in the TD instrument. The raw chromatograms obtained from Bio-VOC™ and Tedlar® bag samples were processed separately with their corresponding background samples. For instance, Tedlar® bag breath samples were grouped into two classes (breath vs background) and processed according to the procedure described in **section 2.2.4**. The number of VOCs remaining (after processing) in Tedlar® bag samples (breath + background) was 60, while the number of VOCs retained from Bio-VOC™ samples (breath + background) was 43. Among the VOCs detected (total 84), 19 were common between both Bio-VOC™ and

Tedlar® bag samples (1,2-dimethyl cyclopropene; 1,3-pentadiene, (Z)-; 1-butanol; 1-hexanol, 2-ethyl-; 1pPropene, 2-methoxy-; acetaldehyde; á-pinene; benzene; benzonitrile; butanal; butane, 2,2,3,3-tetramethyl-; cyclohexanone; cyclopentane, methyl-; hexanal; limonene; methacrolein; methyl isobutyl ketone; pentane, 2-methyl-; and sulfide, allyl methyl). To allow quantitative comparison between Bio-VOC™ and Tedlar® bag samples, the average concentration of a common VOC, benzene, was compared between the sampling devices. The average concentration was calculated from the triplicate samples collected from the same control subject.

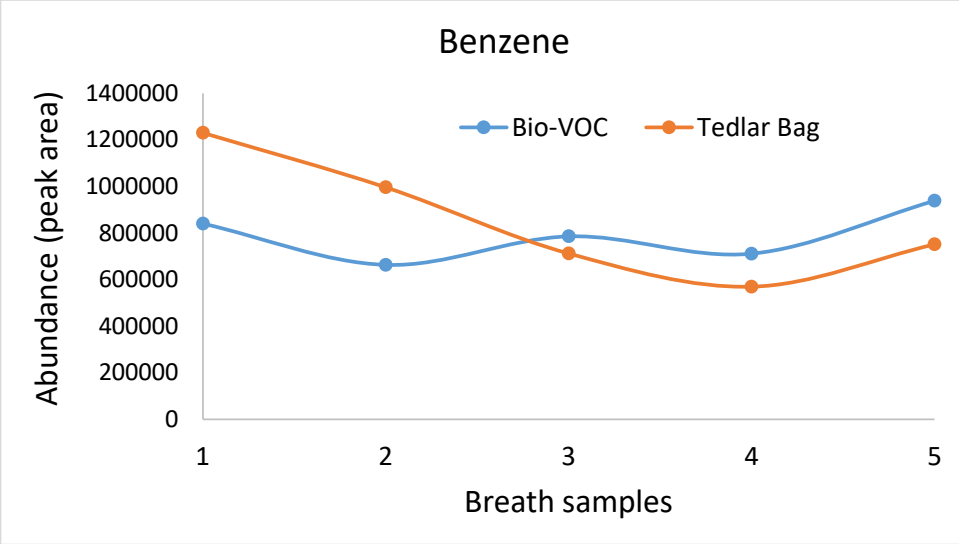


Figure 3.5: The average concentration of benzene in the breath samples collected from five healthy control subjects using two different breath sampling devices (Bio-VOC™ breath sampler vs Tedlar® bag).

As depicted in **Figure 3.5**, the two sampling devices were comparable in terms of their ability to retain benzene. However, the concentration of benzene was relatively more stable between subjects when sampling was performed using Bio-VOC™ breath sampler. In contrast, bag samples exhibited relatively higher benzene concentrations in the first two subjects while in the case of the three other subjects, the concentration of benzene was slightly lower in Tedlar® bag samples. However, these differences were not significant and overall, both techniques were able to retain benzene effectively.

For further investigation of the performance of these two sampling devices, PCA was performed. **Figure 3.6** shows the PCA scores plot for both Bio-VOC™ and Tedlar® bag samples prepared using 43 and 60 VOCs, respectively. PCA of the Bio-VOC™ samples contained 14 breath data points as one of the data points was an outlier (the third of the

triplicates from the fifth subject: CB-5_3). The correction for outliers was performed with the Hotelling's T^2 distribution Test with 95% confidence limit (Nizio et al., 2016). There were no outliers found in Tedlar® bag samples.

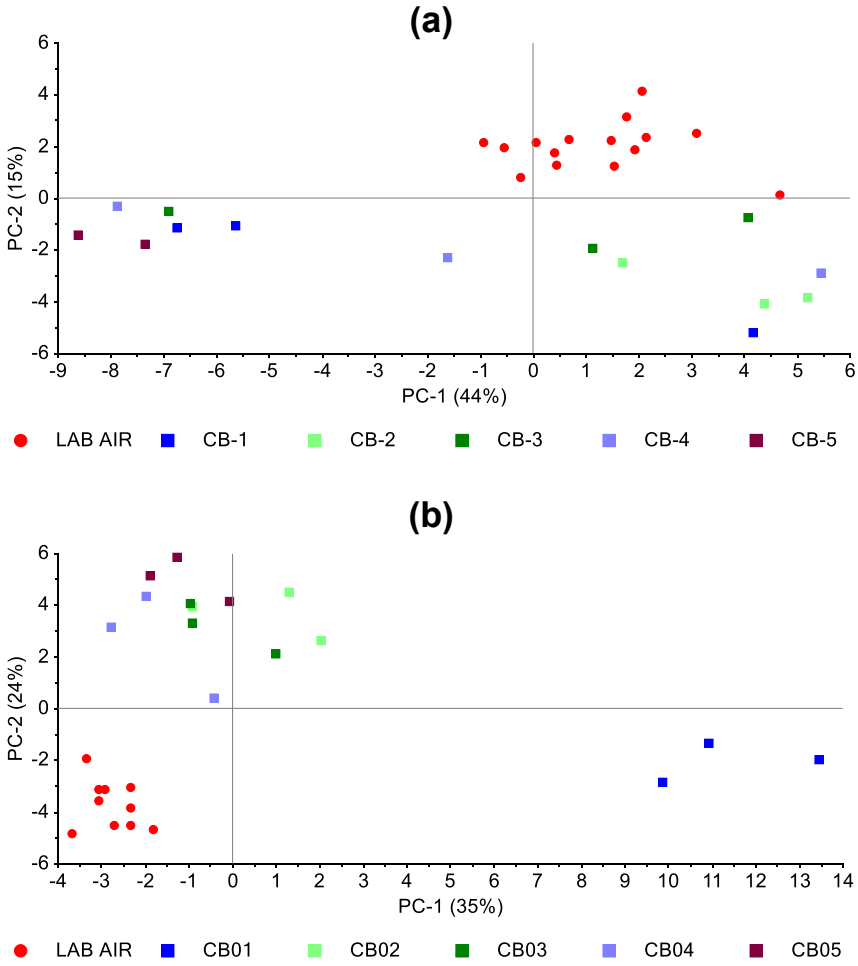


Figure 3.6: PCA scores plot prepared using pre-processed TD-GC×GC-TOFMS peak area data: (a) Bio-VOC™ breath samples vs background air and (b) Tedlar® bag breath samples vs background air

As depicted in the PCA, both sampling devices were able to retain VOCs that allowed differentiation between sample types (breath samples vs background/lab air). However, the samples collected using Tedlar® bags provided relatively better precision as triplicates collected from the same subject clustered in a systematic manner **(Figure 3.6b)**, as did the background samples of laboratory air. In particular, the Euclidean distance between breath triplicates is very low in the Tedlar® bag samples. For instance, in the Tedlar® bag samples, the three triplicates of CB01 have a distance of 4 (10 ~ 14) toward PC-1 **(Figure 3.6b)**. In contrast, the same triplicates (CB01) collected using Bio-VOC™ breath sampler

have a distance of 11 (-7 ~ +4) toward PC-1 **(Figure 3.6a)**. In addition, the background air samples collected with Tedlar® bags also exhibited close clustering with low Euclidean distance between samples. In contrast, Bio-VOC™ samples demonstrated high scattering between triplicates collected from the same subject at the same time (e.g. triplicates from CB-4 in **Figure 3.6a**). Moreover, the Euclidean distance between the breath and background air cluster is also lower in the PCA score plots of Bio-VOC™ samples compare to the PCA score plot of Tedlar® bag samples. Based on this observation Tedlar® bags were selected as the preferred device for breath sampling, which also provided some additional benefits over the Bio-VOC™ breath sampler. For instance, bag sampling is relatively easier than Bio-VOC™ as it takes at least three exhalations to collect a single Bio-VOC sample, while a single exhalation is sufficient to collect a 1 L Tedlar® bag breath sample. Furthermore, the procedures involved in Bio-VOC™ sampling were relatively time consuming and not feasible for application in a busy clinic (e.g. the outpatient CF clinic in the Westmead Hospital which is open once every week for two hours only). In this short time frame, each patient has to see several practitioners (e.g. doctor/specialist, pharmacists, physiotherapist, dietician, pathologist, and psychiatrist). For this reason, the collection of samples using Bio-VOC™ would be very challenging as it consumes a considerable amount of the time available for the patient. Tedlar® bags, in contrast, provided easy handling and storage of samples and were thus selected as suitable devices for breath sampling in this study.

3.3.2. Evaluation of extraction techniques

3.3.2.1. Optimisation of SPME parameters for breath samples

To determine the optimal extraction technique for breath VOCs, breath samples collected in Tedlar® bags were extracted using two common VOC-extraction techniques (SPME vs ST-extraction). Initially, method optimisation was performed for two basic SPME parameters: extraction time and fibre desorption time, while the metric for optimal extraction/desorption was the number of detected compounds for different tested durations. To facilitate this process, three different breath samples were collected in Tedlar® bags from three individual heathy participants. VOCs present in each bag sample were extracted for three tested durations (10, 30, and 60 min) in a random manner. The samples were desorbed in the GC×GC inlet for 5 min at 250°C and analysed using GC×GC-TOFMS. The same procedure was applied to all three breath samples. This provided a set of three samples for each of the three tested durations (10, 30, and 60 min).

The average number of VOCs detected for the three extraction durations of 10, 30, and 60 min was 46, 46, and 47, respectively **(Table 3.3)**.

Table 3.3: Results of Tedlar bag®-SPME method optimization experiments

Parameters	Average number of detected compounds for different durations					
	10 min	30 min	60 min	5 min	10 min	15 min
Extraction time for VOCs	46	46	47			
Desorption time for fibre				81	80	81

The results indicate no significant impact of extraction time on the number of detected VOCs, when characterisation was performed using GC×GC-TOFMS. Based on this result, the extraction time was selected as 10 min, which allowed rapid sampling of VOCs from breath samples collected in Tedlar® bags. In addition to extraction time, fibre desorption time for breath VOCs collected on SPME fibres was also tested for 5, 10, and 15 min using the same approach. The extraction of VOCs was performed at a fixed duration of 10 min, while desorption times varied. However, the number of VOCs were stable with increasing desorption time and a 5 min desorption was sufficient to allow complete desorption of all the VOCs extracted on the SPME fibre. Based on these optimisation parameters, further SPME extraction of breath samples collected on Tedlar® bags was performed at room temperature (~22 °C) by exposing the fibre to the headspace in the bag for 10 min.

3.3.2.2. Comparison between extraction techniques

To allow comparison between extraction techniques, a set of 30 breath samples were collected in Tedlar® bags from two contrasting groups: CF subjects with different lung infections (n=15) vs healthy participants as a control group without any known lung infections (n=15). Each of the 30 samples were extracted using both SPME fibres and STs and analysed using SPME-GC×GC-TOFMS and TD-GC×GC-TOFMS, respectively. This approach provided a total of 60 breath profiles. The raw chromatograms obtained from this experiment were processed separately according to their extraction technique. The 30 SPME samples were grouped into two classes (CF breath vs healthy breath) and processed according to the procedure described in **Section 2.2.4**. In parallel, the ST samples were also separated into the same two classes and processed according to the same procedure. In brief, the data processing allowed the identification of class-distinguishing compounds and retained them for further statistical analysis (i.e. PCA).

The total number of VOCs remaining (after processing) from the SPME samples was 16, which is lower than the number remaining for the ST samples which was 40. PCA performed using these VOCs exhibit that both techniques demonstrated some distinction between groups (**Figure 3.7)**.

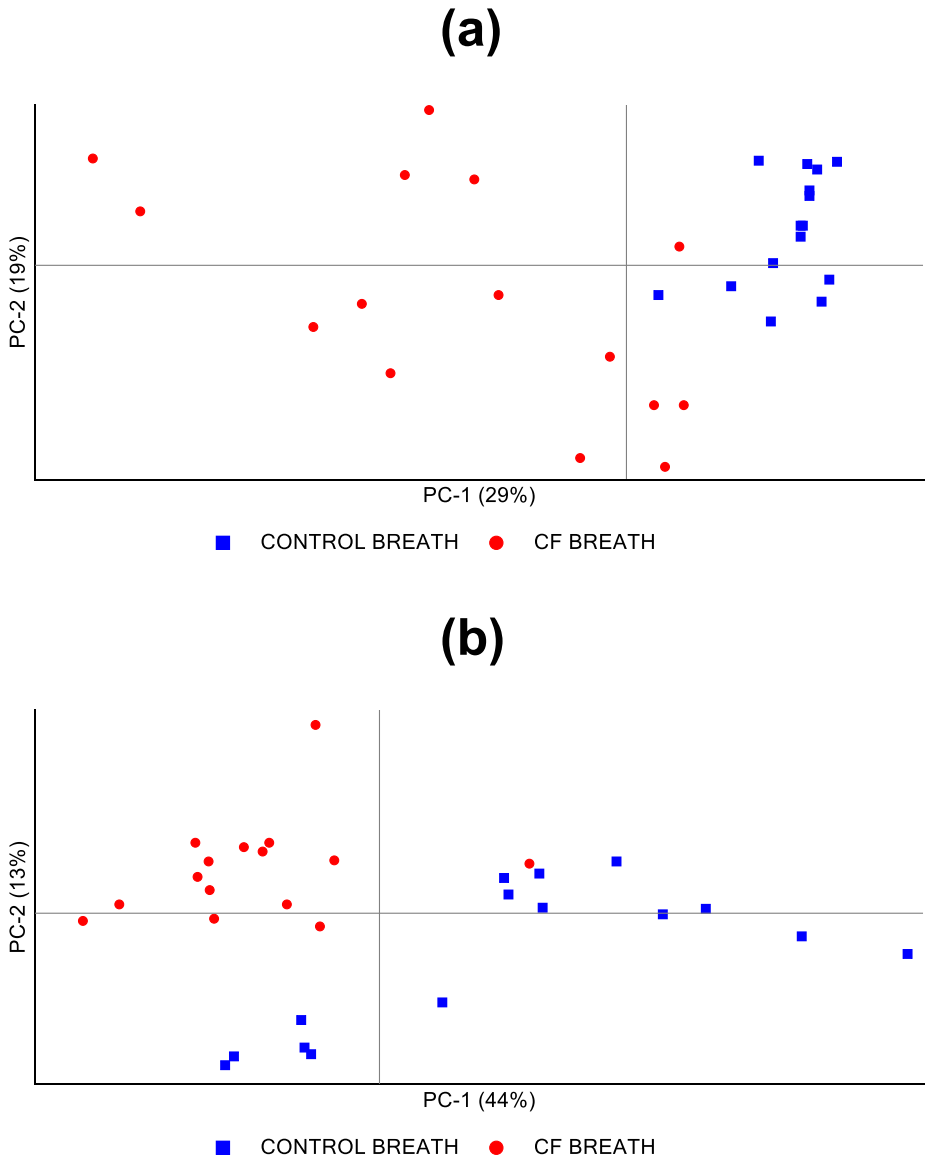


Figure 3.7: PCA scores plot prepared using pre-processed GC×GC-TOFMS peak area data obtained from breath samples collected in Tedlar® bags and extracted using: (a) SPME fibres and (b) Sorbent tubes.

However, the pattern of clustering was different bewteen the two extraction techniques. For instance, in the case of SPME samples, the breath profiles of healthy participants clustered together in the right-hand side of the plot, while CF profiles formed several

clusters **(Figure 3.7a)**. In contrast, the PCA performed using ST samples shows a different pattern as CF samples clustered together, while the breath profiles of healthy participants formed several clusters **(Figure 3.7b)**. From the pattern of samples observed in the PCA, it is not posssible to determine which extraction technique is better as both techniques allowed separation between contrasting groups. However, the pattern of ST samples shows good reproducibility for the CF subjects but less potential to differentiate these samples based on other factors such as infection status. Based on this observation, SPME was selected as the preferred preconcentartion technique which exhibited potential to allow discrimination between CF subjects. Another supplementary reason was that the use of SPME allowed direct comparison between sputum and breath profile as SPME extraction was the method most suitable for the sputum samples.

3.4. Conclusion

This chapter evaluated the performance of two widely reported breath sampling devices (Bio-VOC™ vs Tedlar® bag) and two common extraction techniques available for the extraction of VOCs present in breath samples (SPME vs ST-extraction). The comparison between sampling devices showed better performance of the Tedlar® bags compared to the Bio-VOC™ breath sampler. However, breath samples collected using both devices allowed differentiation between exhaled human breath and background/room air samples, while samples collected in Tedlar® bag exhibited relatively better precision in terms of triplicate analysis. Based on this result, Tedlar® bags were selected as the sampling device for further breath sampling.

Further comparison between extraction techniques (SPME vs ST-extraction) using Tedlar® bag samples proved the strength of both techniques. In particular, both SPME and ST-extraction was able to allow separation between breath samples collected from CF subjects and healthy participants. However, SPME allowed relatively better separation among CF subjects in terms of their breath VOC profiles. For this reason, the combination of Tedlar® bag sampling and SPME was selected as the suitable technique for further breath sampling from a larger study population. The next chapter will discuss the collection of this sample set and the characteristics of the study population.

Chapter 4: COMPARISON OF BREATH PROFILES BETWEEN CF SUBJECTS WITH LUNG INFECTIONS AND HEALTHY CONTROLS

Chapter 4: COMPARISON OF BREATH PROFILES BETWEEN CF SUBJECTS WITH LUNG INFECTIONS AND HEALTHY CONTROLS

4.1. Introduction

During recent decades, a large number of studies explored the potential of breath analysis for the diagnosis of non-communicable diseases like cancers and heart diseases (Amann et al., 2007, Boots et al., 2012, Fu et al., 2014, Phillips et al., 2006, Phillips et al., 1999a, Poli et al., 2005, Poli et al., 2010, Wang et al., 2014, Finamore et al., 2019). There is also a growing interest in the diagnosis of communicable diseases like lung infections using breath biomarkers (Bos et al., 2013a, van der Schee et al., 2015, Beccaria et al., 2018b, Beccaria et al., 2018c). For instance, a number of studies compared breath VOCs between CF subjects with lung infections and healthy controls (Barker et al., 2006, Gaisl et al., 2018, Kamboures et al., 2005, Robroeks et al., 2010b, White et al., 2013). In an early study, Kamboures et al. (2005) determined the concentration of carbonyl sulfide, dimethylsulfide (DMS), and carbon disulphide in the breath of CF subjects (n=20) and healthy controls (n=23) using gas chromatography-mass spectrometry (GC-MS). In another GC-MS-based study, Baker et al. (2006) studied breath samples from adult CF (n=20) and adult control (n=20) subjects and concluded that adult subjects with CF exhibited higher pentane and lower DMS output in their exhaled breath. Studies by Kamboures et al. (2005) and Baker et al. (2006) showed that the concentration of some VOCs are different between CF subjects and healthy control subjects. However, none of these studies collected sputum samples from the subjects during breath sampling and thus have no culture results to compare. In addition, the studies only measured certain targeted VOCs and did not attempt to perform non-targeted analysis of VOCs to screen for potential biomarker/s of lung infections in adult CF subjects.

In a relatively recent study, Robroeks et al. (2010) reported a combination of 10 VOCs as the most discriminating between children with CF (n=48) and healthy control children (n=57). Most importantly, the authors performed non-targeted analysis of VOCs using GC-MS and also collected matching sputum samples, while performing breath sampling. In another study, White et al. (2013) applied proton-transfer-reaction mass spectrometry

(PTR-TOFMS) for the comparison of exhaled breath from children with CF (n=10) and healthy control children (n=4). The sputum culture of 10 CF subjects found five infected and five non-infected CF subjects. However, as reported by the authors, it was not possible to obtain complete separation between groups. Finally, in a recent study, Gaisl et al. (2018) reported that a total of 49 exhaled breath features were different between adult CF subjects (n=30) in comparison to healthy controls (n=30), when secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS) analysis was performed. Due to the type of analytical technique used, the authors were only able to provide a set of tentative chemical formula rather than a set of specific biomarkers. Based on the available literature, to date there is no study which has characterised VOCs from adult CF subjects with lung infections and adult healthy controls without lung infections using a GC-based analytical technique which is capable of providing a set of potential biomarkers rather that features/tentative chemical formula. The only such study was conducted by Robroeks et al. (2010) but the participants were children.

In this study, we carried out a direct comparison between CF subjects with different lung infections and healthy controls without lung infections using their breath VOC profiles. It is important to note that differentiation between CF and non-CF was not a goal of this study. Since 1986, all newborns in Australia have been subject to a heel prick test for the screening of a number of conditions, including CF (Massie et al., 2010). The goal of this study was to determine whether it was possible to allow differentiation between CF subjects with confirmed lung infections and healthy controls without any known lung infections to ensure optimisation of the method for subsequent analyses. To facilitate this process, breath samples were collected in Tedlar® bags, extracted using an optimised solid-phase microextraction (SPME) technique, and characterised for VOCs using GC×GC–TOFMS. Obtained breath profiles were then assessed through principal component analysis (PCA) to compare the two contrasting groups in terms of their lung infection status (CF subjects with lung infections vs healthy participants without any lung infection). Finally, linear discriminant analysis (LDA) was performed using RStudio (version 3.5.3) to determine a set of VOCs that allowed differentiation between the two groups with/without lung infections and demonstrated the capability of the analytical method for subsequent analyses.

4.2. Materials and Method

4.2.1. Collection of breath samples

Breath samples (total 82) were collected from CF subjects during their regular visit to an outpatient CF Clinic at Westmead Hospital (NSW, Australia). For recruitment, each participant was provided with a brief description of the study along with inclusion/exclusion criteria **(Table 4.1)**.

Table 4.1: Inclusion/exclusion criteria for sampling

Criteria	CF subjects	Control subjects
Inclusion criteria		
Individuals over 18 years of age	√	√
Exclusion criteria		
Have CF and/or any known pulmonary disease		√
Regular smoker	√	√
Clinical history of halitosis	√	√
Poor dentition/gingivitis/gum disease	√	√
Inability to provide sputum and/or breath sample	√	√
Eaten a meal/had a drink/s within 1hr	√	√
Cardiovascular disease		√
Chronic anemia		√
Panic disorders		√
Any digestive disease (e.g. irritable bowel syndrome)		√

Written consents were collected from every participant prior to sampling. For breath sampling, participants were asked to sit and relax; then take a deep breath in and exhale out the initial part of their breath (the mouth air), and then to fill the 1 L Tedlar® bag (SKC Inc., USA), expiring the remaining part of the breath as the sample using a disposable mouthpiece (SKC Inc., USA). Samples collected in the CF clinic were transported securely to the University of Technology Sydney (UTS) for the profiling of VOCs using GC×GC-TOFMS. In parallel to the CF breath samples, a total of 77 control breath samples were also collected from healthy participants recruited from the Faculty of Science, UTS. The breath sampling procedure was the same for both CF subjects and healthy controls. However, a portion of the samples (17 CF and 26 control breath samples) was lost due to instrumental issues and unavoidable situations (e.g. sudden power-shutdowns and instrumental malfunctions) (**Table 4.2**). To note, the samples collected in this study were analysed in triplicates where possible to achieve optimal outcome and were averaged to obtain a single breath profile from each subject.

Table 4.2: Sample count for breath samples collected from CF subjects and controls

Sample type	Samples collected	Lost samples*	Total samples remaining
CF breath	82 (5/09/2017 ~ 27/11/2018)	17	65
Control breath	77 (9/05/2018 ~ 21/12/2018)	26	51

*Samples lost due to instrumental issues

4.2.2. Extraction of VOCs present in breath samples

VOCs present in breath samples were extracted using a three-phase (divinylbenzene/carboxen/polydimethylsiloxane) SPME fibre attached to a manual fibre holder (Supelco, Bellefonte, PA, USA). The details of the VOC extraction is described in **section 3.2.3.1)**. In brief, SPME collection from Tedlar® bag samples was performed within 48 h of sample collection; until extraction, samples were stored at room temperature (~22 °C). The extraction of VOCs was also performed at room temperature by exposing a SPME fibre to the bag for 10 min. The fibre was then preloaded with internal standard by exposing it for 15 s in the headspace of 200 µL of a 100 ppm deuterated chlorobenzene prepared in methanol standard (HPLC grade, Sigma-Aldrich, Castle Hill, NSW). The thermal desorption of VOCs was performed in a GC×GC (Pegasus® 4D GC×GC–TOFMS: LECO, Castle Hill, NSW, Australia) inlet for 5 min at 250 °C. In addition to breath samples, a set of bag blanks (10 different Tedlar® bags filled with room air) were also extracted using the same procedure for the purpose of background correction.

4.2.3. GC×GC-TOFMS analysis of samples extracted using SPME

The detailed operating conditions of the GC×GC-TOFMS used for the characterisation of breath VOCs are discussed in **section 3.2.4**. In brief, the VOCs desorbed in the GC×GC inlet were first separated on a 30 m mid-polar GC-column. Further separation was performed on a short (2 m) polar column. High purity helium (BOC, Sydney, NSW, Australia) was used as the carrier gas at a constant flow rate of 2 mL/min. The GC oven setup was as follows: initial temperature 40 °C held for 0.2 min; ramp of 10 °C/min to 230 °C and held for 0.8 min; total run time of 20 min.

4.2.4. Processing of raw data

As described earlier, breath samples were collected from a total of 82 CF and 77 control subjects. However, a number of samples (CF = 17 and Control = 26) were lost due to unavoidable situations. The remaining samples (CF = 65 CF and Control = 51) were

analysed for VOCs using GC×GC–TOFMS. This provided a total of 181 different chromatograms for CF samples, as most of the samples were analysed in triplicates (51×3=153), while some were analysed in duplicates (14×2=28). In addition, a total of 133 different chromatograms were obtained from control samples, as most of the samples were analysed in triplicates (31×3=93) and some in duplicates (20×2=40). The processing of raw chromatograms was performed using the Statistical Compare software feature in ChromaTOF® (additional details are provided in the first paragraph of **section 2.2.4**).

After the processing of raw chromatograms, filtering of VOCs was performed with samples sorted into three classes: (i) CF breath (n=181); (ii) Control breath (n=133); and (iii) Bag blanks (pre-cleaned Tedlar® bags filled with room air: n=10). This process allowed the filtering of the compounds with the highest between-class variance applying Fisher Ratio (FR) calculated using the Statistical Compare software feature in ChromaTOF®. This was performed by comparing the calculated FR to a critical *F* value (F_{crit}) obtained through an *F*-test (Bean et al., 2015, Nizio et al., 2016). Analytes with higher FR values (or analytes absent from a class or only detected in a single sample in a class thus labelled as 'undefined') indicated compounds that statistically differed in abundance between the defined classes and were therefore retained. For all samples where an *F*-test was performed, compounds with FR above the F_{crit}, which includes those labelled as 'undefined', were exported as a *.csv file. The resulting *.csv file was imported into Microsoft Excel for the manual removal of chromatographic artefacts (i.e. column bleed, solvent peaks, and bag artefacts) and further statistical analysis. To note, the identification of compounds in this study were performed using the mass spectral library database provided by the USA National Institute of Standards and Technology (NIST, 2011) and a match > 80% was required to assign a peak name. Any compound named in this study is thus recognised as level two in metabolite identifications/putatively annotated compounds (i.e. compounds identified without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries) (Sumner et al., 2007). The CAS number of all named compounds are compiled in APPENDIX C.

4.2.5. Multivariate analysis

4.2.5.1. Principal component analysis

After the processing of raw data, the triplicates from each breath sample were averaged to obtain a single breath profile for each subject, which provided a total of 116 unique

breath profiles (65 from CF subjects and 51 from controls). Principal component analysis (PCA) was performed using the VOCs present in these breath profiles (in Unscrambler® X version 10.3, CAMO Software, Oslo, Norway) to assess the pattern of the data and to understand the clustering between two contrasting study groups. Prior to PCA, the data was pre-processed using mean centering, scaling and unit vector normalisation as reported in several recent studies analysing VOCs from complex biological samples (Goeminne et al., 2012, Nizio et al., 2016, Pesesse et al., 2019).

4.2.5.2. Liner discriminant analysis

In parallel to PCA, linear discriminant analysis (LDA) was also performed using the VOCs present in the breath profiles obtained from CF subjects and healthy controls. The goal was to determine a set of VOCs which allowed distinction between two contrasting groups (CF subjects with different lung infections vs healthy controls without lung infections). LDA was chosen for its ability to determine a linear combination of features that allow distinction between two or more groups of objects or events (Aleix and Avinash, 2001). LDA is also closely related to PCA as both analyses look for linear combinations of variables to reduce the dimensionality of the data and to explain the data in the best possible manner (Aleix and Avinash, 2001). However, unlike PCA, LDA explicitly attempts to model the difference between the groups. In this study, LDA was performed using RStudio (version R 3.5.3, 2019).

4.3. Results and Discussion

4.3.1. Population characteristics

Table 4.3 highlights the anthropometric characteristics of the study population (Details in Appendix D and E). All parameters are in close range for both groups as only adults were recruited in this study. A student's *t*-test was also performed (Owen, 1965) to check any difference between groups (CF vs. control) in terms of age, height, and weight which provided *p*-values of 0.15, 0.17, and 0.10, respectively for these three parameters.

Table 4.3: Anthropometric characteristics of the study populations

Parameters	CF subjects (n=65)	Control subjects (n=51)
Age (y: mean ± SD (Min/Max))	27.6 ± 8.71 (18/59)	29.7 ± 6.99 (21/56)
Sex, M/F (ratio)	39/26 (1.50)	28/23 (1.22)
Height (cm: mean ± SD (Min/Max))	169 ± 8.57 (146/187)	172 ± 8.23 (158/198)
Weight (kg: mean ± SD (Min/Max))	65.6 ± 11.4 (39/95)	69.6 ± 13.7 (50/105)

However, CF subjects were significantly different from controls in terms of lung infection status since most had a range of different infections, while none of the controls had any known health complications including lung infections. In particular, 43 out 65 CF subjects (66.2%) had at least a single species, while 22 (33.8%) had multiple species. *Pseudomonas aeruginosa* was the dominating species in CF subjects infecting 67.7% of subjects followed by *Staphylococcus aureus* (26.2%), and other infections (26.2%). Six CF subjects had no infection detected in their sputum culture **(Table 4.4:** details of culture results for individual CF subjects are recorded in APPENDIX F**)**. In addition, 14% of CF subjects also had asthma. 80% of the CF subjects were treated with antibiotics such as Azithromycin, Flucloxacillin, Bactrim, Colomycin, Tobramycin, Minocycline, Ciprofloxacin, Rifaximin, and Ceftazidime (Details of treatment for individual CF subjects are recorded in APPENDIX G). Almost all the CF subjects were treated with bronchodilators (short and/or long acting), while some (12.3%) were also treated with inhaled corticosteroids.

Table 4.4: Additional characteristics of the CF populations (n=65)

Microorganisms, n (%)	
Single species	43 (66.2%)
Multiple species	22 (33.8%)
Pseudomonas aeruginosa	44 (67.7%)
Staphylococcus aureus	17 (26.2%)
Others	17 (26.2%)
None	6 (9.23)
Asthma, n (%)	9 (14%)
Treatment, n (%)	
Antibiotics	52 (80%)
Bronchodilators (short acting)	57 (87.7%)
Bronchodilators (long acting)	62 (95.4%)
Inhaled corticosteroids	8 (12.3%)

4.3.2. Volatile organic compounds detected in breath samples

In this study, breath samples were collected from both CF subjects (total 65) and healthy controls (n=51) and profiled for VOCs using GC×GC–TOFMS. In addition, VOCs were also analysed from a set of bag blanks (n=10) for background correction and to retain all

breath-related compounds. **Table 4.5** listed all the VOCs (total 81) which were statistically different between these three sample classes (CF vs control vs blanks) as determined using Fisher Ratio (FR) values. 65 out of these 81 VOCs were only detected in breath samples, and thus have undefined FR. Among these 65 VOCs, 64 were common in both CF and control breath, while 3-hexanone, 2,5-dimethyl-4-nitro- was detected only in five CF breath profiles (not present in any control breath or bag blanks), thus they also have undefined FR values (highlighted bold in **Table 4.5**). The remaining VOCs (total 16, shaded in **Table 4.5**) were detected simultaneously in all three sample classes (CF breath, control breath, and bag blanks). However, these 16 VOCs have the highest between-class variance with FR values above critical *F* value obtained through an *F*-test, which indicates that the pattern of these VOCs are significantly different between classes and thus also regarded as breath related VOCs (Bean et al., 2015, Nizio et al., 2016). Overall, the processing of raw chromatograms and initial filtering of VOCs performed in this study identified a set of 81 breath-borne VOCs.

Table 4.5: List of VOCs detected in breath samples

Order	*Compounds (name according to NIST mass spectral library database)*	*Compound class*	*Fisher ratio*
1	1-Propanol	Alcohols	Undefined
2	4-Methoxy-4-methyl-2-pentanol	Alcohols	Undefined
3	Cyclohexanol, 3,3,5-trimethyl-	Alcohols	Undefined
4	Ethanol	Alcohols	Undefined
5	Ethanol, 2-(hexyloxy)-	Alcohols	Undefined
6	1-Propanol, 2-(2-hydroxypropoxy)-	Alcohols	Undefined
7	DL-3-Methyl-2-butanol	Alcohols	Undefined
8	Ethanol, 2-(2-ethoxyethoxy)-	Alcohols	Undefined
9	2',4'-Dihydroxy-3'-methylacetophenone	Aromatics	Undefined
10	Benzene, 1,2,3,5-tetramethyl-	Aromatics	Undefined
11	Benzene, 1,3-dichloro-	Aromatics	Undefined
12	Benzene, 1-ethyl-4-methyl-	Aromatics	Undefined
13	Benzene, 1-methyl-3-(1-methylethyl)-	Aromatics	Undefined
14	Benzenemethanol, α,α-dimethyl-	Aromatics	Undefined
15	Benzofuran, 4,5,6,7-tetrahydro-3,6-dimethyl-	Aromatics	Undefined
16	Butylated Hydroxytoluene	Aromatics	Undefined
17	Levomenthol	Aromatics	Undefined
18	Oxime-, methoxy-phenyl-_	Aromatics	Undefined
19	Tetrahydrofuran	Aromatics	Undefined
20	Butanoic acid	Carboxylic acids	Undefined
21	Propanoic acid	Carboxylic acids	Undefined
22	Acetic acid	Carboxylic acids	Undefined
23	Acetic acid, phenyl ester	Esters	Undefined
24	Benzoic acid, methyl ester	Esters	Undefined
25	Carbamic acid, N-[1,1-bis(trifluoromethyl)ethyl]-, 4-(1,1,3,3-tetramethylbutyl)phenyl ester	Esters	Undefined
26	Methyl propionate	Esters	Undefined
27	Pentane, 3-ethyl-2,2-dimethyl-	Esters	Undefined
28	Propane, 2,2-dimethoxy-	Esters	Undefined
29	Propanoic acid, 2-oxo-, methyl ester	Esters	Undefined
30	Phthalic acid, isobutyl tridec-2-yn-1-yl ester	Esters	Undefined
31	Butyrolactone	Esters	Undefined
32	1,2-Benzenedicarboxylic acid, butyl 2-ethylhexyl ester	Esters	Undefined
33	Tetrachloroethylene	Halogenated	Undefined
34	Trichloromethane	Halogenated	Undefined
35	2,2,4,4-Tetramethyloctane	Hydrocarbons	Undefined
36	Butane, 2,2,3,3-tetramethyl-	Hydrocarbons	Undefined
37	Heptane, 2,2,4,6,6-pentamethyl-	Hydrocarbons	Undefined
38	Heptane, 5-ethyl-2,2,3-trimethyl-	Hydrocarbons	Undefined
39	Octane	Hydrocarbons	Undefined

40	Pentane, 2,2,3,3-tetramethyl-	Hydrocarbons	Undefined
41	Decane	Hydrocarbons	Undefined
42	Tridecane	Hydrocarbons	Undefined
43	2-Cyclohexen-1-one	Ketones	Undefined
44	3-Methylheptyl acetate	Ketones	Undefined
45	3-Octanone	Ketones	Undefined
46	Acetoin	Ketones	Undefined
47	Isophorone	Ketones	Undefined
48	Tricyclo[5.2.2.0(2,6)]undecan-11-one-8,9-dicarboxylic anhydride, 3-[(2-methoxyethoxy)methoxy]-2-methyl-	Ketones	Undefined
49	**3-Hexanone, 2,5-dimethyl-4-nitro-**	Ketones	Undefined
50	Methyl Isobutyl Ketone	Ketones	Undefined
51	Acetamide, N-methyl-	Nitrogen containing	Undefined
52	Acetonitrile	Nitrogen containing	Undefined
53	Urea, tetramethyl-	Nitrogen containing	Undefined
54	2-Oxazolidinone, 3-amino-5-(4-morpholinylmethyl)-	Nitrogen containing	Undefined
55	1-Dodecanamine, N,N-dimethyl-	Nitrogen containing	Undefined
56	N-(tert-Butoxycarbonyl)glycine	Nitrogen containing	Undefined
57	1-Propene, 1-(methylthio)-	Sulfur containing	Undefined
58	Propane, 1-(methylthio)-	Sulfur containing	Undefined
59	Sulfide, allyl methyl	Sulfur containing	Undefined
60	α-Bourbonene	Terpenoids	Undefined
61	α-Pinene	Terpenoids	Undefined
62	Camphene	Terpenoids	Undefined
63	γ-Terpinene	Terpenoids	Undefined
64	Limonene	Terpenoids	Undefined
65	Eucalyptol	Terpenoids	Undefined
66	1-Hexanol, 2-ethyl-	Alcohols	4.21
67	Hexanal	Aldehydes	6.34
68	Nonanal	Aldehydes	6.92
69	2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	Aromatics	12.17
70	Benzene	Aromatics	5.23
71	Benzene, 1,3-dimethyl-	Aromatics	4.63
72	Ethylbenzene	Aromatics	3.96
73	Toluene	Aromatics	46.3
74	Propanoic acid, 2-methyl-, 3-hydroxy-2,4,4-trimethylpentyl ester	Esters	4.68
75	Methylene chloride	Halogenated	6.95
76	Decane, 2,5-dimethyl-	Hydrocarbons	52.1
77	Undecane	Hydrocarbons	44.7
78	2-Butanone	Ketones	9.18
79	Cyclohexanone	Ketones	8.29
80	Cyclohexane, isothiocyanato-	Nitrogen containing	4.93
81	Morpholine, 4-octadecyl-	Nitrogen containing	40.2

As shown in **Figure 4.1**, a large number of these VOCs are aromatics (16) followed by esters (11), hydrocarbons (10), ketones (10), alcohols (9), nitrogen containing compounds (8), terpenoids (6), carboxylic acids (3), halogenated compounds (3), sulfur containing compounds (3), and aldehydes (2)**.**

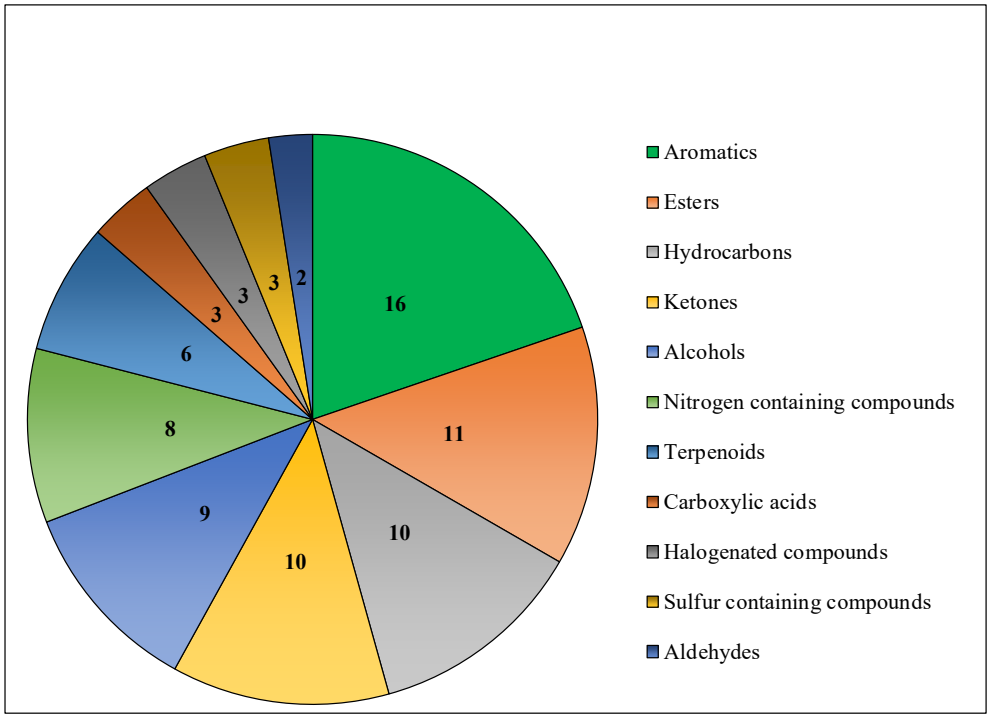


Figure 4.1: Number distribution of VOCs detected in breath samples categorised into different compound classes according to their chemical composition.

4.3.3. Univariate comparison of breath profiles

In this study, a total of 116 breath profiles were obtained from individual CF (65) and control subjects (51). In total, 81 different VOCs were detected in these profiles. The average number of VOCs detected in each profile was 55±9, while the average number of VOCs detected in CF breath profiles (58±9) was higher than their control counterpart (51±9). In addition, there was a significant difference in the pattern of individual VOCs between the two study groups. For instance, only 16 out of 81 VOCs have lower relative abundance (calculated as the ratio of average peak area values between CF and control breath profiles) in CF subjects (0.02 to 0.89: order 1 to 16 in **Table 4.6** also highlighted in light blue shade) than in control breath samples. Among the remaining VOCs, 23 have moderately higher abundance (1.02-1.53 times: order 17 to 39) in CF breath than controls, while more than half of the VOCs (42 out of 81) have significantly higher abundance

(1.53-26.4 times: order 40 to 81 in **Table 4.6** also highlighted in light olive shade) in breath samples from CF subjects than in control breath. In addition, the detection frequency (calculated as detected in % of samples from a study group) of VOCs were also different between study groups. For instance, 27 out of 81 VOCs have higher detection frequency in control samples (marked in bold in **Table 4.6**), while the remaining VOCs (total 54) were detected more frequently in CF subjects. Overall, the pattern of individual VOCs shows that the breath profiles from CF subjects have relatively higher average number of VOCs with more abundance than their control counterpart. This univariate comparison of breath volatiles between CF subjects with different lung infections and healthy control subjects with no known health complications including lung infections show clear distinction between these two contrasting group. However, it is tough to be conclusive on the reason behind such difference of the composition of VOCs between these two contrasting groups considering probable influence from different exogenous (e.g. CF medication, diet, and lifestyle) and endogenous sources (e.g. lung infection status of CF subjects).

Table 4.6: Relative abundance and detection frequencies of VOCs between CF and control breath. Order 1 to 16 present VOCs with higher abundance in control breath, while the remaining VOCs have higher abundance in CF breath. The VOCs with higher detection frequencies in control breath are in bold.

Order	VOCs	Relative abundance	Detection frequency (%)	
		(CF/Control)	Control	CF
1	**Cyclohexanol, 3,3,5-trimethyl-**	**0.02**	**5.88**	**3.08**
2	**α-Bourbonene**	**0.03**	**15.7**	**4.62**
3	**Propane, 1-(methylthio)-**	**0.22**	**88.2**	**64.6**
4	**Sulfide, allyl methyl**	**0.25**	**74.5**	**49.2**
5	**1-Propene, 1-(methylthio)-**	**0.31**	**92.2**	**75.4**
6	**Benzofuran, 4,5,6,7-tetrahydro-3,6-dimethyl-**	**0.38**	**23.5**	**6.15**
7	2-Oxazolidinone, 3-amino-5-(4-morpholinylmethyl)-	0.54	47.1	66.2
8	Trichloromethane	0.57	9.80	29.2
9	Levomenthol	0.58	80.4	96.9
10	**4-Methoxy-4-methyl-2-pentanol**	**0.61**	**70.6**	**60.0**
11	**Acetamide, N-methyl-**	**0.62**	**68.6**	**60.0**
12	Methylene chloride	0.74	96.1	98.5
13	γ-Terpinene	0.77	56.9	75.4
14	**Propane, 2,2-dimethoxy-**	**0.82**	**96.1**	**95.4**
15	Carbamic acid, N-[1,1-bis(trifluoromethyl)ethyl]-, 4-(1,1,3,3-tetramethylbutyl)phenyl ester	0.83	54.9	66.2
16	**Acetonitrile**	**0.89**	**82.4**	**80.0**
17	2-Cyclohexen-1-one	1.02	82.4	89.2
18	Acetic acid, phenyl ester	1.04	100	100
19	**Benzene, 1-methyl-3-(1-methylethyl)-**	**1.09**	**100**	**98.5**
20	**Benzene, 1,3-dichloro-**	**1.12**	**84.3**	**80.0**
21	**1-Propanol**	**1.13**	**88.2**	**80.0**
22	Propanoic acid, 2-methyl-, 3-hydroxy-2,4,4-trimethylpentyl ester	1.19	39.2	72.3
23	DL-3-Methyl-2-butanol	1.22	29.4	60.0
24	**Ethylbenzene**	**1.23**	**98.0**	**96.9**
25	**Benzenemethanol, α,α-dimethyl-**	**1.24**	**90.2**	**83.1**
26	Benzoic acid, methyl ester	1.26	88.2	98.5
27	**Benzene, 1,3-dimethyl-**	**1.28**	**100**	**98.5**
28	Urea, tetramethyl-	1.28	96.1	100
29	**Oxime-, methoxy-phenyl-_**	**1.29**	**94.1**	**93.8**
30	**Ethanol, 2-(hexyloxy)-**	**1.30**	**98.0**	**90.8**
31	Tetrachloroethylene	1.39	39.2	41.5
32	**Eucalyptol**	**1.39**	**100**	**98.5**
33	Benzene, 1-ethyl-4-methyl-	1.40	90.2	93.8
34	Nonanal	1.41	100	100
35	Methyl propionate	1.44	56.9	76.9
36	Cyclohexanone	1.44	100	100
37	Hexanal	1.45	78.4	93.8
38	Limonene	1.45	100	100
39	1-Hexanol, 2-ethyl-	1.50	100	100

40	Butylated Hydroxytoluene	1.53	76.5	93.8
41	**Cyclohexane, isothiocyanato-**	**1.53**	**100**	**96.9**
42	α-Pinene	1.54	66.7	75.4
43	**Tridecane**	**1.57**	**100**	**98.5**
44	Acetic acid	1.62	23.5	24.6
45	**Ethanol, 2-(2-ethoxyethoxy)-**	**1.62**	**100**	**90.8**
46	Benzene	1.64	98.0	98.5
47	2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	1.65	78.4	86.2
48	Propanoic acid, 2-oxo-, methyl ester	1.68	66.7	69.2
49	**Butyrolactone**	**1.68**	**92.2**	**89.2**
50	Tetrahydrofuran	1.70	98.0	98.5
51	1-Propanol, 2-(2-hydroxypropoxy)-	1.88	100	100
52	Phthalic acid, isobutyl tridec-2-yn-1-yl ester	1.92	5.88	10.8
53	1-Dodecanamine, N,N-dimethyl-	1.95	35.3	60.0
54	Methyl Isobutyl Ketone	1.97	96.1	100
55	Benzene, 1,2,3,5-tetramethyl-	1.98	41.2	87.7
56	Isophorone	1.98	27.5	60.0
57	2',4'-Dihydroxy-3'-methylacetophenone	2.32	41.2	75.4
58	3-Methylheptyl acetate	2.33	56.9	86.2
59	Butane, 2,2,3,3-tetramethyl-	2.35	43.1	52.3
60	**Butanoic acid**	**2.35**	**13.7**	**4.62**
61	**Undecane**	**2.41**	**100**	**96.9**
62	N-(tert-Butoxycarbonyl)glycine	2.59	21.6	46.2
63	**Decane**	**2.82**	**100**	**98.5**
64	2-Butanone	2.90	25.5	49.2
65	3-Octanone	2.98	47.1	80.0
66	Morpholine, 4-octadecyl-	3.03	33.3	46.2
67	Pentane, 2,2,3,3-tetramethyl-	3.22	52.9	92.3
68	Heptane, 5-ethyl-2,2,3-trimethyl-	3.40	51.0	86.2
69	**Heptane, 2,2,4,6,6-pentamethyl-**	**3.43**	**98.0**	**96.9**
70	Toluene	3.55	100	100
71	Camphene	3.62	51.0	55.4
72	Tricyclo[5.2.2.0(2,6)]undecan-11-one-8,9-dicarboxylic anhydride, 3-[(2-methoxyethoxy)methoxy]-2-methyl-	3.62	7.84	20.0
73	1,2-Benzenedicarboxylic acid, butyl 2-ethylhexyl ester	5.16	3.92	21.5
74	Octane	5.60	5.88	13.8
75	Decane, 2,5-dimethyl-	6.32	90.2	96.9
76	Propanoic acid	12.6	9.80	10.8
77	Acetoin	14.9	31.4	66.2
78	2,2,4,4-Tetramethyloctane	15.0	25.5	58.5
79	Ethanol	22.8	5.88	41.5
80	Pentane, 3-ethyl-2,2-dimethyl-	26.4	11.8	84.6
81	3-Hexanone, 2,5-dimethyl-4-nitro-	-	0.00	7.69

4.3.4. Multivariate comparison of breath profiles

4.3.4.1. Principal component analysis

To allow further comparison between CF subjects and healthy controls based on their breath VOC profiles, principal component analysis (PCA) was performed using the peak area values of all 81 VOCs detected in breath profiles. The benefit of PCA over univariate comparison is that it provides better visualization of the multivariate structure of the data (Nizio et al., 2016). All 116 breath profiles were used to perform PCA, while 7 of the profiles (6 CF and 1 control: CF02, CF10, CF11, CF15, CF16, CF17 and CB12) were later excluded from the PCA plot, classed as outliers by means of the Hotelling's T^2 with 95% confidence limit (Nizio et al., 2016). As shown in the PCA scores plot **(Figure 4.2)**, there is moderate separation between study groups based on their breath VOC profiles with principal component (PC) scores of 7% (PC-2) and 13% (PC-1). In particular, breath profiles from two contrasting study groups are minimally separated along PC-2 with most of the control samples located in the first and second quadrants of PCA while most of the CF samples were located in the third and fourth quadrants of the PCA score plot. To note, in addition to the two-dimensional (2D) PCA score plots which includes PC-1 and PC-2, three-dimensional (3D) score plots were also prepared for all of the PCA performed in this thesis. This allowed us to check whether there is better separation between groups if PC-3 is included on the score plot. However, as depicted in APPENDIX H which includes all of the 3D score plots for Chapter 4, the separation between groups were not improved by the addition of PC-3. In addition to these 3D plots (containing PC-1, PC-2, and PC-3), every PCA performed in this study was also checked for higher separation in any other PC (up to PC-7) using scree plots. However, PC-1 and PC-2 consistently accounted for the optimal amount of possible separation between groups and thus is preferentially discussed in the PCA sections throughout the thesis.

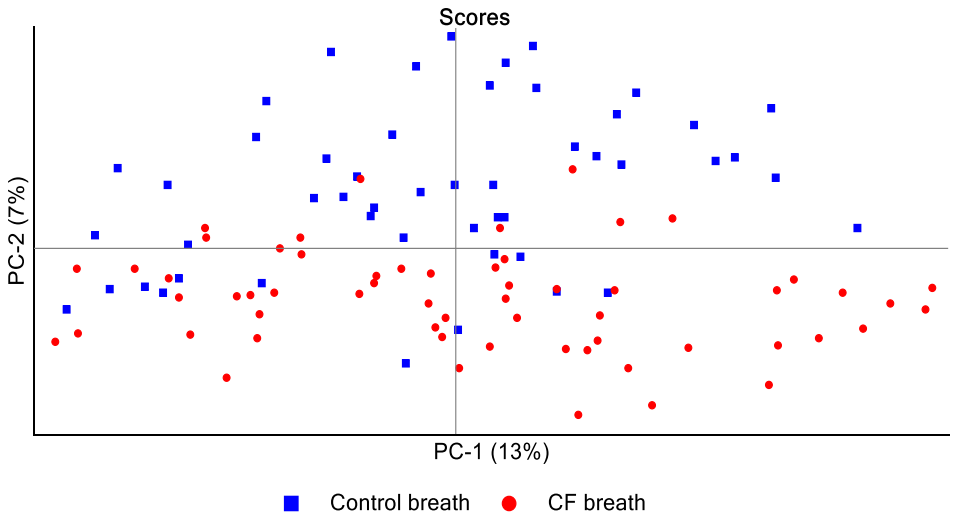


Figure 4.2: PCA scores plot of breath profiles obtained from CF subjects and healthy controls. The plot was prepared using 81 VOCs detected in breath profiles.

For further investigation of the pattern of VOCs and to determine the VOCs which are statistically different between study groups, a Student's *t*-test was performed (Owen, 1965). To facilitate this process, breath profiles were classified into two groups, CF and control, and Student's *t*-test was performed for all 81 VOCs detected in breath profiles using their peak area values. Among all 81 VOCs, only 26 had a *p-value* lower than 0.05 (difference is significant at 95% confidence level). **Table 4.7** lists these 26 VOCs which were statistically different between study groups as determined using Student's *t*-test. This approach significantly reduced the number of VOCs which are different between study groups.

Table 4.7: List of VOCs statistically different between study groups as determined using Student's *t-test*

Order (correspond to Table 4.6)	**VOCs**	***p*-value**	**Confidence level**	**Relative abundance (CF/Control)**
75	Decane, 2,5-dimethyl-	5.78E-10	99%	6.32
70	Toluene	6.30E-08	99%	3.55
68	Heptane, 5-ethyl-2,2,3-trimethyl-	2.91E-07	99%	3.40
61	Undecane	8.11E-07	99%	2.41
65	3-Octanone	1.71E-05	99%	2.98
69	Heptane, 2,2,4,6,6-pentamethyl-	4.96E-05	99%	3.43
63	Decane	5.77E-05	99%	2.82
80	Pentane, 3-ethyl-2,2-dimethyl-	3.41E-04	99%	26.4
78	2,2,4,4-Tetramethyloctane	1.01E-03	99%	15.0
51	2-(2-hydroxypropoxy)-1-Propanol	1.82E-03	99%	1.88
3	Propane, 1-(methylthio)-	1.98E-03	99%	0.22
64	2-Butanone	3.72E-03	99%	2.90
58	3-Methylheptyl acetate	4.31E-03	99%	2.33
54	Methyl Isobutyl Ketone	4.99E-03	99%	1.97
66	4-octadecyl- Morpholine	6.40E-03	99%	3.03
5	1-Propene, 1-(methylthio)-	6.47E-03	99%	0.31
67	Pentane, 2,2,3,3-tetramethyl-	7.18E-03	99%	3.22
43	Tridecane	1.35E-02	95%	1.57
57	2',4'-Dihydroxy-3'-methylacetophenone	2.20E-02	95%	2.32
49	Butyrolactone	2.36E-02	95%	1.68
55	1,2,3,5-tetramethyl- Benzene,	2.42E-02	95%	1.98
62	N-(tert-Butoxycarbonyl)glycine	2.98E-02	95%	2.59
34	Nonanal	3.20E-02	95%	1.41
50	Tetrahydrofuran	3.37E-02	95%	1.70
2	α-Bourbonene	4.01E-02	95%	0.03
4	Sulfide, allyl methyl	4.39E-02	95%	0.25

Further PCA performed using these 26 VOCs shows that it is possible to allow relatively better separation between study groups (in terms of PC values), while excluding the VOCs not statistically significant (at least at 95% confidence level) between CF and control breath profiles **(Figure 4.3)**. The findings of this study highlighted that a Student's *t*-test can effectively reduce the number of potential VOC biomarkers differentiating between CF subjects and healthy controls, and multivariate analysis (i.e. PCA) performed using the refined data (through Student's *t*-test) would provide better separation/clustering between the contrasting study population.

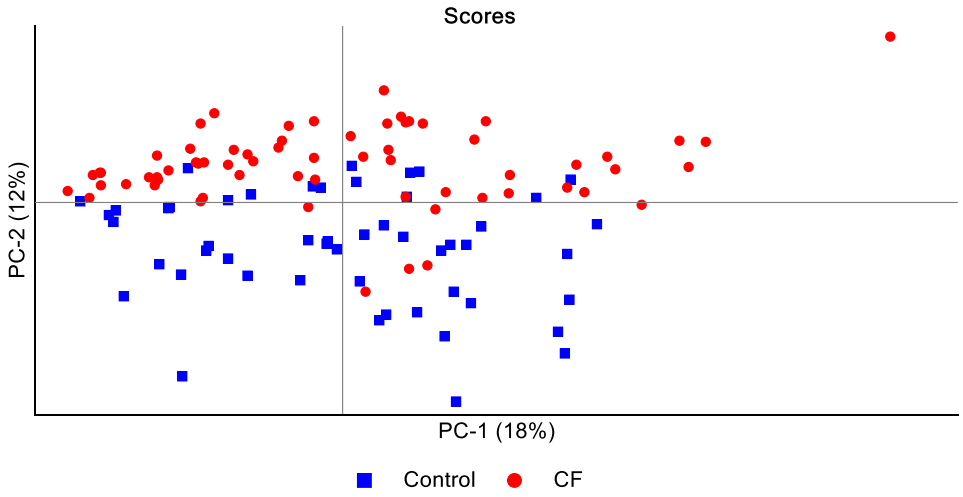


Figure 4.3: PCA scores plot of breath profiles obtained from CF subjects and healthy controls. The plot was prepared using 26 VOCs which were statistically significant between study groups as determined using a Student's *t*-test.

4.3.4.2. Linear discriminant analysis

In this study, a total of 81 VOCs were detected in breath samples collected from CF subjects and healthy controls. Student's *t*-test performed in this chapter reduced the number of VOCs to 26 (**Table 4.7** listed all VOCs). PCA performed using these 26 VOCs exhibited moderate separation between study groups. For further refinement of data and to determine whether it is possible to find a linear combination of VOCs to allow differentiation between CF subjects and healthy controls, linear discriminant analysis (LDA) was performed. In this process, samples (total 116: 65 CF + 51 Control) were classified in two groups (CF and Control). Each group was assigned with a score on a group measure; all control samples were assigned a value of 0, while all CF samples were assigned a value of 1. All 26 VOCs obtained through the Student's *t*-test were used as a

set of unknown variables to classify the samples into two groups (CF vs Control) based on their linear discriminant (LD) scores calculated through LDA. However, optimal differentiation (in terms of the number of correct classification) was obtained when a set of 16 VOCs was used to perform the LDA. To note, while performing LDA, multiple combinations of VOCs were checked (from the 26 VOCs obtained) and the combination of a set of 16 VOCs allowed the optimal differentiation between the two groups. In particular, 50 of the control samples were classified correctly (1 misclassification) with 57 correct classifications for CF samples (8 misclassifications). The performance of LDA was evaluated through leave-one-out cross-validation which resulted in 4 misclassifications for control subjects and 10 misclassifications for CF subjects.

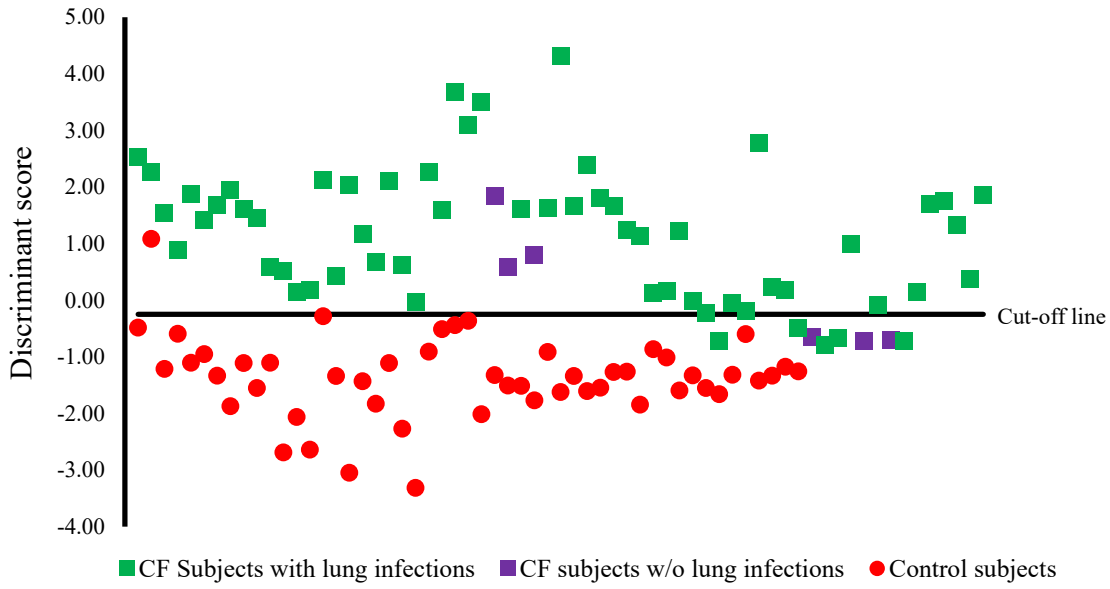


Figure 4.4: Linear discriminant scores of CF subjects and healthy controls

Figure 4.4 shows the LD scores of subjects and the separation between groups. The only red dot above the cut-off line highlighted the one misclassified control subject, while the squares below the cut-off line are the misclassified CF subjects. However, as recorded in Table 4.3, all 51 control subjects have no known lung infections. In contrast, 59 out 65 CF subjects have lung-infections, while 6 CF subjects have no lung infection during the time of breath sampling as confirmed using their sputum culture results. To note, the CF subjects with no lung infections were split between CF group and Control group in the LD score plot (the purple squares in **Figure 4.4**). If the LDA performed in this study is considered as a discriminatory test between subjects with/without lung infections, the accuracy of classification will be as follows: [54 (true positive)+50 (true negative Control

subjects)+3 (true negative CF subjects)]/[All samples total 116] = 92.2% (Baratloo et al., 2015).

4.3.5. Most discriminating breath VOCs between CF subjects and healthy controls

4.3.5.1. General characteristics

Table 4.8 lists 16 VOCs which allowed optimal discrimination between CF and control subjects. Among these VOCs, eight (VOC 3 to VOC 5 and VOC 8 to VOC 12) were detected more frequently in CF subjects (shaded in **Table 4.8**). Three out of the eight VOCs (VOC 3 to VOC 5) were ketones (1 aromatic + two aliphatic), while five were hydrocarbons (VOC 8, 9, 11 and 12) and the last one was a nitrogen containing compound (VOC 10). Among the remaining VOCs, two were detected more frequently in control samples (VOC 2 and VOC 13: highlighted bold in **Table 4.8**). For the rest of the VOCs (total six: VOC 1, VOC 6, VOC 7, VOC 14, VOC 15, and VOC 16), the difference of detection frequency between the two study groups was less than 5%.

Notably, the two VOCs with higher detection frequency in control subjects are the only sulfur containing VOCs in the list, 1-(methylthio)-1-propene and 1-(methylthio)-propane). However, both VOCs were reported previously to be associated with healthy human breath. For instance, Phillips et al. (1999) reported 1-(methylthio)-1-propene, while studying the variation of VOCs in the breath of 50 normal humans (Phillips et al., 1999c). In another study, Dadamio et al. (2012) reported that 1-(methylthio)-propane, was detected in 57% of the breath samples from healthy controls, but was not detected in any of the target breath samples from subjects with liver cirrhosis (Dadamio et al., 2012). Based on the finding of this study and available literature, these two sulfur containing VOCs can be considered as associated with control breath. It was also interesting to note that these two sulfur containing VOCs are the only VOCs with higher relative abundance in control breath than their CF counterpart. For instant, the relative abundance of 1-(methylthio)-propane was almost five times higher in control breath than that of CF breath. However, the remaining VOCs have higher relative abundance in breath samples from CF subjects with values ranging between 1.57 (tridecane) to 26.4 (3-ethyl-2,2-dimethyl-pentane) (**Table 4.8**).

Table 4.8: List of VOCs most discriminating between CF and control breath. VOCs detected more frequently in CF breath are highlighted in grey shades.

Sample code	VOCs	Detection frequency (DF: %)			Average of peak area values		Standard deviation of peak area values		Relative abundance
		CF	Control	CF/Control	CF	Control	CF	Control	(CF/control)
VOC 1	1-Propanol, 2-(2-hydroxypropoxy)-	100	100	1.00	116271	61873	95688	85837	1.88
VOC 2	**1-Propene, 1-(methylthio)-**	**75**	**92**	**0.82**	**6456**	**20940**	**39131**	**12845**	**0.31**
VOC 3	2',4'-Dihydroxy-3'-methylacetophenone	75	41	1.83	4033	1735	6134	4418	2.32
VOC 4	2-Butanone	49	25	1.96	8604	2965	6575	12153	2.90
VOC 5	3-Octanone	80	47	1.70	4628	1555	3064	4014	2.98
VOC 6	Butyrolactone	89	92	0.97	12080	7178	12995	9822	1.68
VOC 7	Decane	98	100	0.98	58083	20616	16480	61772	2.82
VOC 8	Decane, 2,5-dimethyl-	97	90	1.08	32176	5094	7657	27481	6.32
VOC 9	Heptane, 5-ethyl-2,2,3-trimethyl-	86	51	1.69	7517	2212	3562	6123	3.40
VOC 10	Morpholine, 4-octadecyl-	46	33	1.39	44949	14857	30447	71795	3.03
VOC 11	Pentane, 2,2,3,3-tetramethyl-	92	53	1.74	4267	1326	2205	7349	3.22
VOC 12	Pentane, 3-ethyl-2,2-dimethyl-	85	12	7.08	4671	177	564	8600	26.4
VOC 13	**Propane, 1-(methylthio)-**	**65**	**88**	**0.74**	**4172**	**19023**	**37188**	**4152**	**0.22**
VOC 14	Toluene	100	100	1.00	300986	84813	145885	230541	3.55
VOC 15	Tridecane	98	100	0.98	11249	7148	8660	8666	1.57
VOC 16	Undecane	97	100	0.97	52144	21674	20429	37149	2.41

4.3.5.2. Important VOCs between CF and control breath as reported in literature

A search of the available literature produced several studies which have compared VOCs in exhaled breath samples from CF subjects and healthy controls. In an early study, Kamboures et al. (2005) reported elevated level of DMS in exhaled breath of both CF subjects (n=20) and healthy controls (n=23), while no significant difference was found between the study groups in term of DMS concentration. In another study, Baker et al. (2006) analysed a set of VOCs from 20 CF subjects and 20 healthy controls and reported higher concentrations of n-pentane and lower concentrations of DMS in CF breath. The finding of Baker et al. (2006) is similar to our observation as we also detected lower concentrations of sulfur compounds (i.e. 1-(methylthio)-1-propene and 1-(methylthio)-propane) in CF breath compared to the control samples. In fact, sulfur containing VOCs were the only class of discriminating compounds (between CF and control) which had a higher abundance in control subjects.

Early studies by Kamboures et al. (2005) and Baker et al. (2006) provided important clues that the concentrations of some VOCs are different between CF subjects and healthy controls. However, none of these studies attempted to classify two contrasting groups (CF vs control) using a set of VOCs. There is, however, one study by Robroeks et al. (2010) which reported a combination of 10 VOCs as the most discriminating between CF and control subjects. The authors analysed breath samples from 105 children (48 with CF, 57 controls) and were able to differentiate between two study groups with 92% accuracy using the following 10 VOCs: 3,3-dimethylhex-1-ene; 2-buten-1-ol; N-methyl-2-methylpropylamine; C_8H_{16} hydrocarbon; tolualdehyde (o-, m-, or p- isomers); C_{16} poly-unsaturated hydrocarbon; C_{12} saturated hydrocarbon; C_{13} saturated hydrocarbon; benzothiazole; long chain alkylbenzene. Several of these 10 discriminating VOCs (C_{12} and C_{13} saturated hydrocarbons) were also determined as important VOCs in our study. In particular, VOC 8 (2,5-dimethyl-Decane) and VOC 9 (5-ethyl-2,2,3-trimethyl-Heptane) in **Table 4.8** are both C_{12} saturated hydrocarbons. In addition, VOC 15 (tridecane) is a C_{13} saturated hydrocarbon. However, none of the other VOCs were common between our study and the study by Robroeks et al. (2010).

In recent studies, researchers have also reported the application of real-time measurement techniques like PTR-TOFMS (White et al., 2013) and SESI-HRMS (Gaisl et al., 2018) for the comparison of exhaled breath from CF subjects and healthy controls. For instance, White et al. (2013) analysed breath samples from 10 CF subjects and 4 healthy controls

using PTR-TOFMS and concluded that PTR-TOF-MS analysis of tidal breath can distinguish between infected and non-infected CF subjects. Gaisl et al. (2018) reported that a total of 49 exhaled breath features were altered in CF subjects (n=30) in comparison to healthy controls (n=30), when SESI-HRMS analysis was performed. In addition, authors reported significant difference of oxidative stress metabolites such as fatty acids between subjects with CF and healthy controls (Gaisl et al., 2018).

4.3.5.3. Probable origin of important VOCs detected in this study

In this study, a set of 16 VOCs were determined which allowed differentiation between CF subjects with lung infections and healthy control without any known health complications including lung infections. Among these 16 VOCs, 8 were detected more frequently in CF subjects. The concentration of these VOCs were also higher in CF subjects in comparison to the control samples. Based on the pattern of these 8 VOCs between CF and control breath, they were considered as associated with lung infections in CF subjects. To find out whether these VOCs were reported previously in literature as related to bacterial/fungal metabolites a customised search was performed in Google Scholar. The customised search facilitated the filtering of literature which must contain the specific VOC name while the other keywords (i.e. bacterial metabolite CF breath) were set to create links between the VOC name and their origin (i.e. bacterial metabolite/CF breath). **Table 4.9** lists the results of the literature search. The first VOC, 2',4'-Dihydroxy-3'-methylacetophenone, was reported in only two articles (Carvalho, 2014, Prompanya, 2018). Carvalho et al. (2014) reported 2',4'-Dihydroxy-3'-methylacetophenone isolated from the extracts of fungi *Neosartorya siamensis*, which is further cited by Prompanya (2018). The second compound, 2-butanone, is reported in several studies as an important metabolite released by bacterial species associated with lung infections in CF subjects. Shestivska et al. (2015) reported VOCs in the headspace of *Stenotrophomonas strains cultures* (total 20 strains, both clinical and environmental isolates) and detected 2-butanone in the headspace of all the strains (Shestivska et al., 2015). 2-butanone is also reported as an important metabolite released by *P. aeruginosa* (Bos et al., 2013b, Zechman and Labows Jr, 1985, Shestivska et al., 2012). The third VOC, 3-octanone is reported to be associated with both fungal and bacterial infection in human (Perl et al., 2011, Küntzel et al., 2016). Perl et al. (2011) reported 3-octanone as a characteristic VOC of fungus *A. fumigatus* (Perl et al., 2011). Additionally, Küntzel et al. (2016) reported very high concentrations of 3-octanone in the headspace of

Mycobacterium avium ssp. Paratuberculosis (Küntzel et al., 2016). The 4th VOC, 2,5-dimethyl-decane is reported in only one study as associated with *M. tuberculosis* infection in human (Mellors, 2018). The author reported 2,5-dimethyl-decane along with several other VOCs while using breath to diagnose and monitor *M. tuberculosis* infection in human (Mellors, 2018). As discussed in section 4.3.5.2, both 2,5-dimethyl-decane and 5-ethyl-2,2,3-trimethyl- heptane are C_{12} saturated hydrocarbons reported in the literature as important VOCs which allowed differentiation between children with CF (n=48) and healthy control children (n=57) (Robroeks et al., 2010b). Among the remaining VOCs, 4-octadecyl- morpholine and 2,2,3,3-tetramethyl- pentane were not reported in any literature to be associated with breath/bacterial metabolite. However, 3-ethyl-2,2-dimethyl- pentane was reported by Phillips (2008) as a breath biomarker of patients with halitosis/oral malodour (Phillips, 2008).

Table 4.9: List of VOCs associated CF subjects and the results of literature search

Order	VOCs	Author	Metabolite	
1	2',4'-Dihydroxy-3'-methylacetophenone	Carvalho et al (2014)	Fungus	*Neosartorya siamensis*
2	2-Butanone	Shestivska et al (2015)	Bacteria	*Stenotrophomonas strains*
		Shestivska et al (2012)		*Pseudomonas aeruginosa*
3	3-Octanone	Perl et al (2011)	Fungi	*Aspergillus fumigatus*
		Küntzel et al (2016)	Bacteria	*Mycobacterium avium ssp. Paratuberculosis*
4	Decane, 2,5-dimethyl-	Mellors (2018)	Bacteria	*Mycobacterium tuberculosis*
5	Heptane, 5-ethyl-2,2,3-trimethyl-	Robroeks et al (2010)		
6	Morpholine, 4-octadecyl-	NR*		
7	Pentane, 2,2,3,3-tetramethyl-	NR		
8	Pentane, 3-ethyl-2,2-dimethyl-	Phillips (2008)	*Halitosis/oral malodor*	

*Not reported

4.4. Conclusion

In this study, comparisons were carried out between CF subjects with different lung infections and healthy controls without any lung infections using their breath VOC profiles. The chemical characterisation of VOCs was performed using GC×GC–TOFMS, while PCA and LDA was performed to allow comparison between study groups in terms of their lung infection status. A total of 81 VOCs were detected in all the breath samples, while 80 of them were common in both CF and control breath. However, there was a significant difference in the pattern of VOCs between infected CF and non-infected control subjects. In particular, breath profiles from CF subjects have relatively higher average numbers of VOCs with more abundance than their control counterpart. Such difference of breath VOC profiles was also evident in multivariate analysis. For instance, both groups formed their own cluster separate from each other in the PCA score plot. Moreover, LDA performed in this study allowed clear classification of CF subjects with lung infections and healthy controls without lung infections using a combination of 16 VOCs. In particular, it was possible to identify controls without lung infections with 98% accuracy while the accuracy of classification for lung-infected CF subjects was slightly lower at 92%. Overall, the combination of an advanced analytical technique and multivariate analysis performed in this study shows that it is possible to allow differentiation between subjects with/without lung infections.

However, it is important to note that all of the CF subjects participated in this study have significant differences from the control groups, not only in terms of lung infection status but also in terms of numerous other factors (i.e. 14% of CF subjects had asthma compared to none in the control group; 80% of the CF subjects were treated with different antibiotics during the study period; almost all the CF subjects were treated with bronchodilators, while some were also treated with inhaled corticosteroids). All these factors can significantly impact the breath profiles obtained from these CF subjects. Yet, the 16 VOCs discriminating between groups were common between both groups (infected vs non-infected). Most of these VOCs were detected more frequently with higher concentration in infected CF subjects. An increased concentration and prevalence of discriminating VOCs in CF breath is possibly due to their lung infection status. However, the probability of any influence from other sources such as CF-related medications, special CF diet, and lifestyle cannot be ignored. Hence, it is important to perform further

research which will consider these factors in greater detail. However, for the purposes of this study, the differentiation between CF subjects with lung infections and control subjects without lung infections provided a confirmation of the capability of the analytical method. This method was subsequently used for further analysis to investigate differences in CF subjects with a variety of lung infections and to compare their sputum VOCs with breath VOCs.

Chapter 5:
COMPARISON OF VOCS BETWEEN SPUTUM & BREATH SAMPLES FROM CF SUBJECTS

Chapter 5: COMPARISON OF VOCS BETWEEN SPUTUM & BREATH SAMPLES FROM CF SUBJECTS

5.1. Introduction

The lungs and airways of subjects with cystic fibrosis (CF) are characterised by the chronic infection of different bacterial (e.g. *Pseudomonas aeruginosa, Staphylococcus aureus, Stenotrophomonas maltophilia, and Haemophilus influenza*) and fungal species (e.g. *Aspergillus fumigatus and Scedosporium aurantiacum*) (Lyczak et al., 2002, Baptista et al., 2019). Sputum culture has been considered the primary method for the detection and identification of these species in CF lungs and airways. However, it can take several days to weeks to receive culture results based on the type and complexity of infections. For this reason, there is a growing interest in the determination of volatile organic compounds (VOCs) specific to certain species to facilitate their detection and identification in exhaled breath (Hahn et al., 2020, Koehler et al., 2020). In particular, a large number of studies have reported the profiling of VOCs in the headspace of in vitro cultures from CF associated species. Labows et al. (1980) reported a GC-MS-based analysis of VOCs released by different strains of *P. aeruginosa* (Labows et al., 1980). In another study, Carroll et al. (2005) reported a SIFT-MS-based analysis of VOCs from the cultures of bacterial isolates from cough and swab samples from CF subjects (Carroll et al., 2005). The authors reported significantly higher concentrations of hydrogen cyanide (HCN) from *P. aeruginosa* positive cultures (Carroll et al., 2005). Several recent SIFT-MS-based studies also reported HCN as an important biomarker of *P. aeruginosa* (Dryahina et al., 2016, Neerincx et al., 2015).

However, the first reported application of GC×GC-TOFMS for the analysis of VOCs in the culture of *P. aeruginosa* detected a total of 56 different compounds, 28 of which had not previously been reported as *P. aeruginosa*-derived volatiles (Bean et al., 2012). Further study from the same research group reported a total of 391 different VOCs associated with the growth and metabolism of 24 different isolates of *P. aeruginosa* (Bean et al., 2016). The diversity of VOCs released by bacteria is also highlighted in a recent study, which reported a total of 472 different VOCs form a total of six bacterial species associated with lung infections in CF (e.g. *P. aeruginosa, B. cenocepacia, H. influenzae,*

S. maltophilia, S. pneumoniae and S. milleri) (Nizio et al., 2016). In addition, significantly larger numbers (total 397) of VOCs were also reported only from growth media in which in vitro cultures were grown. Such impact of growth media is also reported in a study concerning the differentiation of *Staphylococcus* species by VOCs (Jenkins and Bean, 2019). Based on their observations, the authors recommended identifying a profile/s of VOCs to differentiate between bacterial species (Nizio et al., 2016, Jenkins and Bean, 2019).

The aim of such culture-based studies is to identify VOC biomarkers associated with different bacterial/fungal species that can be detected in breath to provide a rapid and non-invasive diagnosis of certain lung infections in CF subjects. Considering this aim, it is also important to study VOCs directly in the headspace of sputum samples from CF subjects to minimise the effect of VOCs associated with culture media. To date, only a few studies have reported VOCs directly in the headspace of sputum samples (Goeminne et al., 2012, Savelev et al., 2011), but none have reported the VOCs present in corresponding breath samples from the same subject. This represents a gap in the literature as there is limited understanding of the potential transition of VOCs from sputum to breath. In this study, both sputum and breath samples were collected from CF subjects and profiled for VOCs using GC×GC-TOFMS. This facilitated an extensive profiling of VOCs using an advanced analytical technique and allowed comparison between sputum and breath VOC profiles from the same subject.

5.2. Materials and Method

5.2.1. Sample collection

A set of 32 matching sputum and breath samples were collected from CF subjects during their regular visit to an outpatient CF clinic in Westmead hospital. Signed consent forms were collected from each participant before any sampling procedure was performed (under the approval # LNR/14/WMEAD/386, approved by the human research ethics committee of Western Sydney Local Health District). The details of sputum and breath sampling is described in **Section 2.2.1 and Section 3.2**, respectively. In brief, participants were asked to sit and relax and provide a set of two sputum samples, one for routine microbial culture and another for the profiling of VOCs. The sample for the profiling of VOCs was collected in a 20 mL SPME vial with airtight screw caps containing 1.3 mm thick polytetrafluoroethylene/silicone septum (Sigma-Aldrich, Castle Hill, NSW,

Australia). A corresponding breath sample was then collected in a 1 L transparent Tedlar® bag (SKC Inc., USA). The sealed vials containing sputum samples and the Tedlar® bags containing breath samples were transported to the University of Technology Sydney (UTS) for the extraction and analysis of VOCs within 24 h (sputum) ~ 72 h (breath) of collection to maintain sample integrity.

5.2.2. Extraction of VOCs

The VOCs in the headspace of sputum samples or in the breath samples were extracted using a three-phase (divinylbenzene/carboxen/polydimethylsiloxane) SPME fibre attached to a manual fibre holder (Supelco, Bellefonte, PA, USA). The optimisation of sputum and breath sampling is described in **Chapter 2 and Chapter 3**, respectively.

5.2.3. GC×GC-TOFMS analysis of samples

All SPME samples were analysed using a Pegasus® 4D GC×GC–TOFMS (LECO, Castle Hill, NSW, Australia) equipped with a liquid nitrogen cryogenic quad jet modulator. Refer to **section 2.2.3 and section 3.2.4** for details of the GC×GC–TOFMS analysis of sputum and breath samples, respectively.

5.2.4. Data processing and statistical analysis

In this study, a set of 32 matching sputum and breath samples were collected from CF subjects and were analysed for VOCs in triplicates. This provided a total of 192 different chromatograms (32×3 = 96 for sputum and 32×3 = 96 for breath samples). The processing of raw chromatograms obtained from "sputum" and "breath" samples were performed separately using the Statistical Compare software feature in ChromaTOF® (additional details of chromatographic data processing are provided in the first paragraph of **section 2.2.4**). However, the procedure was the same for both sample types. For example, sputum samples were processed with samples sorted into two classes: (i) sputum (n=96) and (ii) vial blank (n=5). Similarly, breath samples were processed with samples sorted into similar classes: (i) breath (n=96) and (ii) bag blank (pre-cleaned Tedlar® bags filled with room air: n=5). Technically, this approach allowed the extraction of any VOCs which were not from the vial materials or bag artefacts.

After the processing of raw chromatograms, filtering of VOCs was carried out to identify the compounds with the highest between-class (i.e. "sputum vs vial blank" and "breath vs vial blank") variance by applying a Fisher Ratio (FR) calculated using the same

Statistical Compare software feature in ChromaTOF®. This was performed by comparing the calculated FR to a critical F value (F_{crit}) obtained through an F-test (Bean et al., 2015, Blanca et al., 2017, Nizio et al., 2016). Analytes with higher FR values (or analytes absent from a class or only detected in a single sample in a class thus labelled as 'undefined') indicated compounds that statistically differed in abundance between the defined classes and were therefore retained. For all samples where an *F*-test was performed, compounds with FR above the F_{crit}, which includes those labelled as 'undefined', were exported as a *.csv file. The resulting *.csv file was imported into Microsoft Excel for the manual removal of chromatographic artefacts (i.e. column bleed and solvent peaks) and further statistical analysis. The final list of compounds of this study can be designated as level two in metabolite identifications/as putatively annotated compounds (i.e. compounds identified without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries) (Sumner et al., 2007).

5.2.5. Principal component analysis

After the processing of raw data, the triplicates from each sputum and breath sample were averaged to obtain a single sputum/breath profile for each subject, which provided a total of 64 unique sputum (n=32) and breath (n=32) profiles. Principal component analysis (PCA) was performed using the VOCs present in these profiles (in Unscrambler® X version 10.3, CAMO Software, Oslo, Norway) to assess the patterns of the data and to understand the clustering between CF subjects with different infection status (i.e. *Pseudomonas aeruginosa positive vs Pseudomonas aeruginosa negative CF subjects*). Prior to PCA, the data was pre-processed using mean centering, scaling and unit vector normalisation as reported in several recent studies analysing VOCs from complex biological samples (Goeminne et al., 2012, Nizio et al., 2016, Pesesse et al., 2019).

5.2.6. Linear discriminant analysis

In parallel to PCA, linear discriminant analysis (LDA) was also performed using the VOCs present in breath/sputum samples. The goal was to allow differentiation between CF subjects with/without *Pseudomonas aeruginosa* infection using their sputum and/or breath VOC profiles. LDA was chosen as a supervised analysis and for its ability to determine a linear combination of features that allow distinction between two or more groups of objects or events (Aleix and Avinash, 2001). In this study, LDA was performed using RStudio (version R 3.5.3, 2019).

5.3. Results and Discussion

5.3.1. Anthropometric data and sputum culture results of CF subjects

Table 5.1 highlights the anthropometric characteristics of the 32 CF subjects recruited in this study. All parameters are in close range for both groups as only adults were recruited in this study. To note, these CF subjects are the subset of CF subjects from Chapter 4.

Table 5.1: Anthropometric characteristics of the study populations

Parameters	CF subjects (n=32)
Age (y: mean ± SD (Min/Max))	27.4 ± 9.03 (18/59)
Sex, M/F (ratio)	22/10 (2.2)
Height (cm: average ± SD (Min/Max))	173 ± 8.03 (151/187)
Weight (kg: average ± SD (Min/Max))	68.3 ± 12.7 (39/93)

Of the 32 CF subjects, 18 (56.3%) had only *P. aeruginosa* as determined through sputum culture **(Table 5.2)**. Of the 18 CF subjects with *P. aeruginosa*, 12 had only mucoid *P. aeruginosa*, three had non-mucoid *P. aeruginosa*, while three had both mucoid and non-mucoid *P. aeruginosa*. Seven of the CF subjects had other infections (except one CF subject without any species detected in culture). The rest of the CF subjects (total seven) had a combination of *P. aeruginosa* with other infections as shown in **Table 5.2**.

Table 5.2: Sputum culture results of individual CF subjects

Order	Subject	Bacteria	Fungus
[1] Only P. aeruginosa			
1	CF03	*Mucoid P. aeruginosa*	-
2	CF04	*Mucoid P. aeruginosa*	-
3	CF06	*Mucoid P. aeruginosa*	-
4	CF10	*Mucoid P. aeruginosa*	-
5	CF12	*Mucoid P. aeruginosa*	-
6	CF19	*Mucoid P. aeruginosa*	-
7	CF22	*Mucoid P. aeruginosa*	-
8	CF23	*Mucoid P. aeruginosa*	-
9	CF25	*Mucoid P. aeruginosa*	-
10	CF26	*Mucoid P. aeruginosa*	-
11	CF27	*Mucoid P. aeruginosa*	-
12	CF29	*Mucoid P. aeruginosa*	-
13	CF02	*Non-mucoid P. aeruginosa*	-
14	CF05	*Non-mucoid P. aeruginosa*	-
15	CF40	*Non-mucoid P. aeruginosa*	-
16	CF15	*Mucoid P. aeruginosa; Non-mucoid P.*	-
17	CF28	*Mucoid P. aeruginosa; Non-mucoid P.*	-
18	CF30	*Mucoid P. aeruginosa; Non-mucoid P.*	-
[2] Combined (P. aeruginosa + Others)			

19	CF24	*Mucoid P. aeruginosa*	*A. fumigatus*
20	CF18	*Non-mucoid P. aeruginosa*	*S. apiospermum*
21	CF14	*Mucoid P. aeruginosa; Non-mucoid P.*	*S. aurantiacum*
22	CF07	*Mucoid P. aeruginosa; Non-mucoid P.*	-
23	CF11	*Non-mucoid P. aeruginosa; S. aureus*	-
24	CF20	*Mucoid P. aeruginosa; S. aureus*	-
25	CF39	*Non-mucoid P. aeruginosa; S. aureus*	-
[3] Others			
26	CF31	None	-
27	CF16	*S. aureus*	-
28	CF17	*S. maltophilia*	-
29	CF21	*A. xylosoxidans*	
30	CF33	*A. fumigatus; N. farcinica*	
31	CF13	*S. aureus*	*T. mycotoxinivorans*
32	CF38	*S. aureus; S. maltophilia; H. influenza*	*S. aurantiacum*

5.3.2. VOCs detected in sputum and breath samples

A total of 32 corresponding sputum and breath samples were collected from CF subjects and profiled for VOCs using SPME-GC×GC–TOFMS. One hundred (100) different VOCs were detected in the headspace of the 32 individual sputum samples. However, the number of VOCs detected in the corresponding breath samples (total 56 VOCs) was significantly lower than the sputum samples. **Table 5.3** lists all the VOCs (total 132) detected in the sputum and breath samples. Among these VOCs, only 24 were common between both sputum and breath samples (order 1 to 24 in **Table 5.3**), 77 were detected only in sputum samples (order 25 to 100 in **Table 5.3**), and 32 only in breath samples (order 101 to 132 in **Table 5.3**).

Table 5.3: List of compounds detected in sputum and breath samples collected from CF subjects. Compounds shaded in grey are common in both sample types, while tan and blue highlights indicate compounds detected only in sputum and breath samples, respectively.

Order	*VOC # corresponding to PCA*		*Compounds*	*Compound class*
	Sputum	*Breath*		
1	10	2	1-Hexanol, 2-ethyl-	Alcohol
2	16	5	1-Propanol	Alcohol
3	22	7	2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	Aromatic
4	28	9	2-Oxazolidinone, 3-amino-5-(4-morpholinylmethyl)-	Nitrogen containing
5	34	12	Acetoin	Ketone
6	35	13	Acetonitrile	Nitrogen containing
7	37	15	α-Pinene	Terpenoid
8	40	16	Benzene	Aromatic
9	42	19	Benzene, 1,3-dimethyl-	Aromatic
10	51	24	Camphene	Terpenoid
11	53	26	Cyclohexanone	Ketone
12	61	33	Eucalyptol	Terpenoid
13	64	34	Heptane, 2,2,4,6,6-pentamethyl-	Hydrocarbon
14	71	35	Limonene	Terpenoid
15	73	37	Methyl propionate	Ester
16	75	38	Morpholine, 4-octadecyl-	Nitrogen containing
17	78	39	Octanal	Aldehyde
18	79	40	Octane	Hydrocarbon
19	87	45	Propane, 2,2-dimethoxy-	Ester
20	88	46	Propane, 2-methoxy-2-methyl-	Ester
21	89	47	Propanoic acid	Carboxylic acid
22	91	48	Propanoic acid, 2-oxo-, methyl ester	Ester
23	99	51	Toluene	Aromatic
24	100	52	Trichloromethane	Halogenated
25	1		α-Bourbonene	Terpenoid
26	2		Carvone	Terpenoid
27	3		1,2-Benzenedicarboxylic acid, bis(2-methylpropyl) ester	Ester
28	4		1,2-Benzenedicarboxylic acid, butyl 2-ethylhexyl ester	Ester
29	5		1,3-Cyclohexadiene, 5-(1,5-dimethyl-4-hexenyl)-2-methyl-	Terpenoid
30	6		1-Butanol	Alcohol

31	7	1-Butanol, 3-methyl-	Alcohol
32	8	1H,1H,2H,2H-Perfluorooctan-1-ol	Alcohol
33	9	1-Hexanol	Alcohol
34	11	1-Methyl-2-pyrrolidinone	Ketone
35	12	1-Nonen-4-ol	Alcohol
36	13	1-Pentadecanamine, N,N-dimethyl-	Nitrogen containing
37	14	1-Pentanol	Alcohol
38	15	1-Penten-3-ol	Alcohol
39	17	1-Undecene	Hydrocarbon
40	18	2,3-Butanedione	Ketone
41	19	2,3-Hexanedione	Ketone
42	20	2,3-Pentanedione	Ketone
43	21	2,4,7,9-Tetramethyl-5-decyn-4,7-diol	Alcohol
44	23	2-Heptanone	Ketone
45	24	2-Heptanone, 6-methyl-	Ketone
46	25	2-Nonanone	Ketone
47	26	2-Octanone	Ketone
48	27	2-Octenal	Alcohol
49	29	2-Piperidinone	Nitrogen containing
50	30	3-Octanone	Ketone
51	31	7-Octen-2-ol, 2,6-dimethyl-	Alcohol
52	32	Acetic acid	Carboxylic acid
53	33	Acetic acid, methyl ester	Ester
54	36	Acetyl valeryl	Ketone
55	38	Aziridine, 1-ethenyl-	Nitrogen containing
56	39	Benzaldehyde	Aldehyde
57	41	Benzene, 1-(1,5-dimethyl-4-hexenyl)-4-methyl-	Aromatic
58	43	Benzeneacetaldehyde	Aldehyde
59	44	Butanal, 3-methyl-	Aldehyde
60	45	Butane, 1,1,3,4-tetrachloro-1,2,2,3,4,4-hexafluoro-	Halogenated
61	46	Butanenitrile, 3-methyl-	Nitrogen containing
62	47	Butanoic acid	Carboxylic acid
63	48	Butanoic acid, 2-methyl-	Carboxylic acid
64	49	Butanoic acid, 3-methyl-	Carboxylic acid
65	50	Butylated Hydroxytoluene	Aromatic
66	52	γ-Terpinene	Terpenoid

67	54		Cyclohexanone, 2-methyl-5-(1-methylethenyl)-, trans-	Ketone
68	55		Dimethyl sulfone	Sulfar containing
69	56		Dimethyl trisulfide	Sulfar containing
70	57		Disulfide, dimethyl	Sulfar containing
71	58		DL-Glyceraldehyde	Aldehyde
72	59		Ethanol, 2-(2-ethoxyethoxy)-	Alcohol
73	60		Ethylbenzene	Aromatic
74	62		Furan, 2,5-dihydro-	Aromatic
75	63		Furan, 2-pentyl-	Aromatic
76	65		Hexadecane	Hydrocarbon
77	66		Hexanal	Aldehyde
78	67		Hydrazine	Nitrogen containing
79	68		Hydroxyurea	Nitrogen containing
80	69		Indole	Aromatic
81	70		Levomenthol	Aromatic
82	72		Methyl isovalerate	Ester
83	74		Methylene chloride	Halogenated
84	76		n-Hexane	Hydrocarbon
85	77		Nonanal	Aldehyde
86	80		o-Menthan-8-ol	Alcohol
87	81		p-Cymene	Terpenoid
88	82		Pentadecanoic acid, 14-methyl-, methyl ester	Ester
89	83		Pentanal, 2-methyl-	Alcohol
90	84		Phenylethyl Alcohol	Alcohol
91	85		Propanal, 2-methyl-	Alcohol
92	86		Propane, 1-(methylthio)-	Sulfar containing
93	90		Propanoic acid, 2-methyl-, 1-(1,1-dimethylethyl)-2-methyl-1,3-propanediyl ester	Ester
94	92		Pyrazine, 2,5-dimethyl-	Aromatic
95	93		Pyrazine, methyl-	Aromatic
96	94		Pyridine	Aromatic
97	95		Pyrrole	Aromatic
98	96		Tetradecane	Hydrocarbon
99	97		Tetrahydrofuran	Aromatic
100	98		Thiocyanic acid, methyl ester	Sulfar containing
101		1	1-Dodecanamine, N,N-dimethyl-	Nitrogen containing
102		3	1H-Pyrazole, 4,5-dihydro-5,5-dimethyl-4-isopropylidene-	Nitrogen containing

103	4	1-Nonadecanamine, N,N-dimethyl-	Nitrogen containing
104	6	1-Propanol, 2,2'-oxybis-	Alcohol
105	8	2,6-Dimethyldecane	Hydrocarbon
106	10	2-Propanone, 1-(acetyloxy)-	Ketone
107	11	Acetic acid, phenyl ester	Ester
108	14	α-Phellandrene	Terpenoid
109	17	Benzene, 1,2,3-trimethyl-	Aromatic
110	18	Benzene, 1,3-dichloro-	Aromatic
111	20	Benzene, 1-ethenyl-4-ethyl-	Aromatic
112	21	Benzene, 1-ethyl-3-methyl-	Aromatic
113	22	Benzene, 1-methyl-4-(1-methylethenyl)-	Aromatic
114	23	Butane, 2,2,3,3-tetramethyl-	Hydrocarbon
115	25	Cyclohexane, isocyanato-	Nitrogen containing
116	27	Decane	Hydrocarbon
117	28	Decane, 2,5-dimethyl-	Hydrocarbon
118	29	Dodecane	Hydrocarbon
119	30	Ethanol	Alcohol
120	31	Ethanol, 2-(hexyloxy)-	Alcohol
121	32	Ethanone, 1-[1-(4-amino-1,2,5-oxadiazol-3-yl)-5-methyl-1H-1,2,3-triazol-4-yl]-2-morpholino-	Nitrogen containing
122	36	Methyl Isobutyl Ketone	Ketones
123	41	Octane, 2,4,6-trimethyl-	Hydrocarbon
124	42	o-Cymene	Terpenoid
125	43	Pentane, 3-ethyl-2,2-dimethyl-	Ester
126	44	Phthalic acid, isobutyl tridec-2-yn-1-yl ester	Ester
127	49	p-Xylene	Aromatic
128	50	Styrene	Aromatic
129	53	Tridecane	Hydrocarbon
130	54	Undecane	Hydrocarbon
131	55	Undecane, 2-methyl-	Hydrocarbon
132	56	Urea, tetramethyl-	Nitrogen containing

To investigate the composition of sputum and breath samples, all detected VOCs (total 132) were grouped into 11 different compound classes (Alcohols, Aldehydes, Aromatics, Carboxylic acids, Esters, Halogenated compounds, Hydrocarbons, Ketones, Nitrogen containing compounds, Sulfur containing compounds, and Terpenoids).

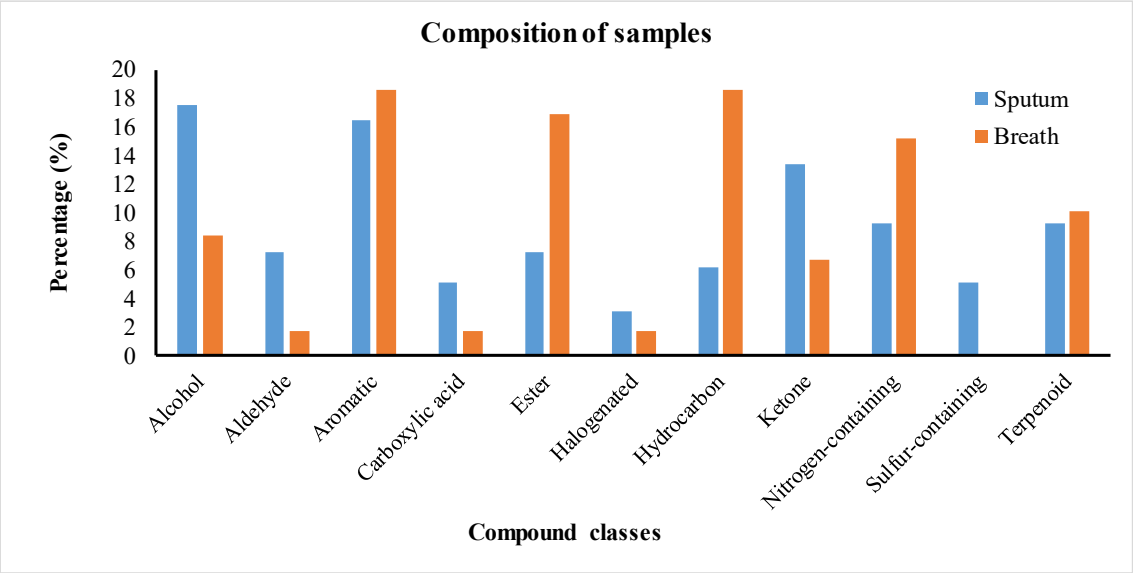


Figure 5.1: Comparison between sputum and breath profiles based on their chemical composition with compound classes in X-axis and the percentage (%) value of each class between sputum and breath samples in Y-axis.

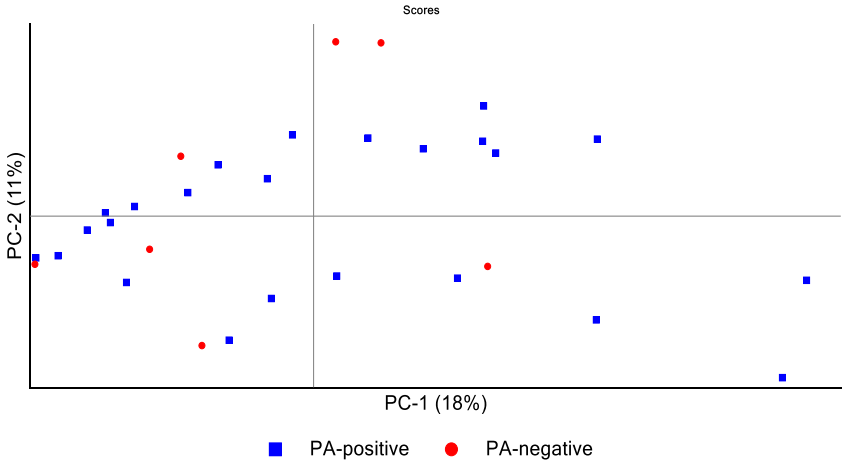
As depicted in Figure 5.1, the percentage of compounds from classes such as alcohols, aldehydes, carboxylic acids, halogenated compounds, and ketones was higher in sputum samples compared to breath samples. In contrast, compound classes such as aromatics, esters, hydrocarbons, nitrogen containing compounds, and terpenoids demonstrated higher percentage values in breath samples. Of all the compound classes, hydrocarbons exhibited the greatest percentage difference between the two sample types. In particular, only 6% of the VOCs detected in sputum samples were hydrocarbons compared to 20% in breath samples. However, compound classes such as alcohols and aldehydes had a significantly higher composition in the sputum samples. In particular, only 2% of the VOCs detected in breath were aldehydes compared to 7% in sputum samples. In addition, sulfur containing compounds were detected only in sputum samples and were not present in any of the breath samples.

5.3.3. Principal component analysis

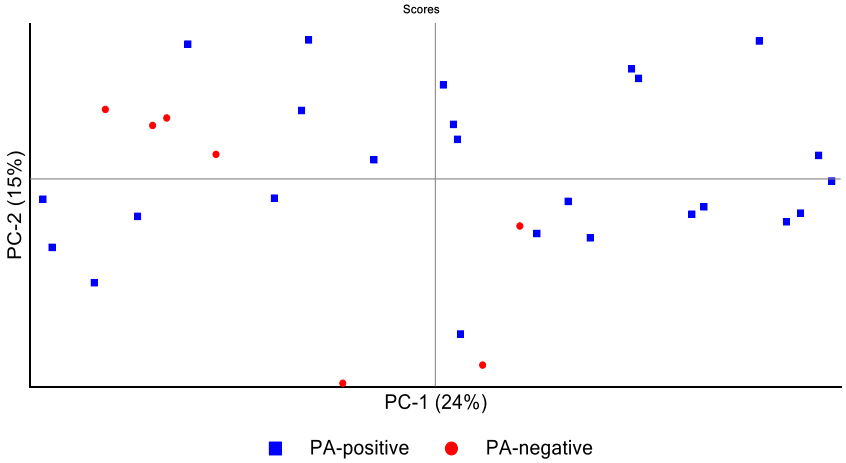
Among the 132 VOCs detected, only 24 were common between the two different sample types. Based on this observation, it was important to further investigate the structure of the data obtained from these samples through multivariate analysis (e.g. PCA). Such comparison allowed for the evaluation of performance between sputum and breath profiling in terms of the capability to differentiate between CF subjects with different lung infections. To facilitate this process, CF subjects were divided into two different groups according to their *P. aeruginosa* (PA) infection status: (i) PA-positive (n=25) and (ii) PA-negative (n=7). PCAs were performed to differentiate between these two groups using the following approaches; (a) PCA using all detected VOCs (total 132 VOCs, both sputum and breath profiles were used together), (b) PCA using only breath VOCs (total 56 VOCs), (c) PCA using only sputum VOCs (total 100 VOCs). In addition, PCAs were also prepared using only the VOCs common between both sample type. The following three approaches were used: (a) PCA using the VOCs common between both sputum and breath profiles (total 24 VOCs, both sputum and breath profiles were used together), (b) PCA using the common VOCs (total 24 VOCs, only breath profiles), and (c) PCA using the common VOCs (total 24, only sputum profiles).

Figure 5.2a shows the PCA scores plot prepared using all the VOCs detected in sputum and breath samples. This plot consists of a total of 31 CF subjects, as one of the subjects (from PA-positive group) was determined as an outlier by means of Hotelling's T^2 with a 95% confidence limit and was excluded from the plot. There were no outliers in the PCA scores plot prepared using only the breath VOCs **(Figure 5.2b)**, while the PCA scores plot prepared using sputum VOCs consists of two outliers (both from the PA-positive group) and these were excluded from the plot **(Figure 5.2c)**. However, as evident from the three PCA score plots of **Figure 5.2**, there was no clear clustering between CF subjects with/without PA infection. There is only a small cluster of PA-negative breath profiles in the 2nd quadrant of **Figure 5.2 b**. Based on the pattern of profiles in **Figure 5.2**, it was not possible to allow clear distinction between CF subjects with/without PA infection, either for sputum or breath profiles, or both profiles together while using all detected VOCs.

(a) PCA using all detected VOCs



(b) PCA using only breath VOCs



(c) PCA using only sputum VOCs

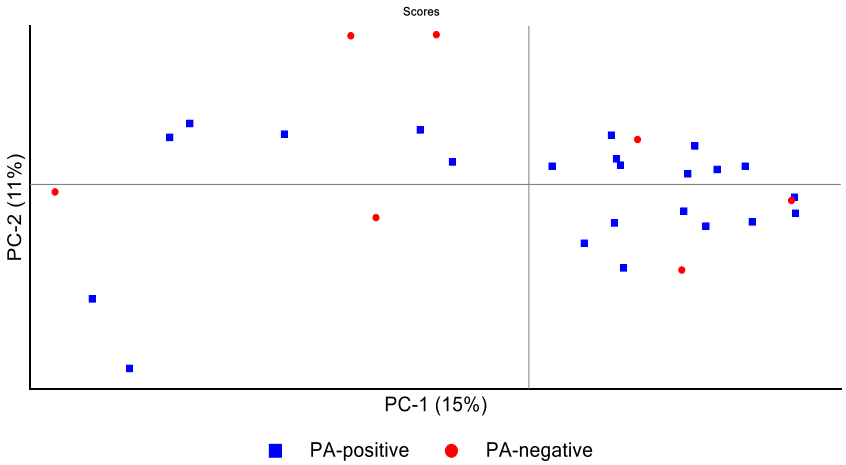
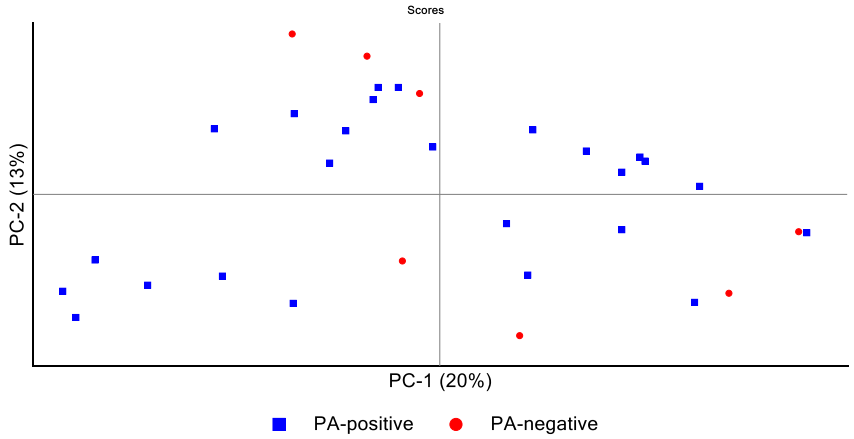


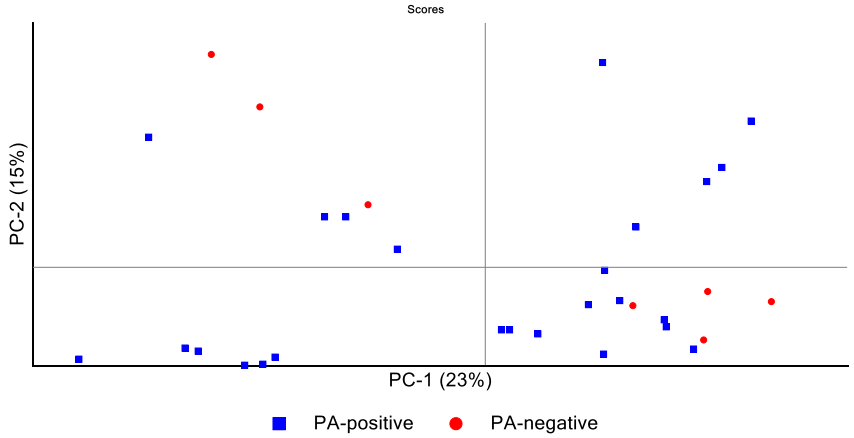
Figure 5.2: PCA scores plots prepared: (i) using all detected VOCs (total 132, both sputum and breath profiles were used), (ii) using only breath VOCs (total 56), (iii) using only sputum VOCs (total 100).

For this reason, further PCAs were performed to determine whether it was possible to allow differentiation between PA-positive and PA-negative CF subjects using only the VOCs common between sputum and breath samples **(Figure 5.3)**. However, similar to **Figure 5.2**, there was no clear clustering between CF subjects with/without PA infections when PCAs were performed with only common VOCs. However, the PC values are slightly higher in **Figure 5.3a** and **Figure 5.3c** compared to **Figure 5.2a** and **Figure 5.2c**. However, neither approach allowed any clear differentiation between the contrasting groups.

(a) PCA using common VOCs (sputum and breath profiles together)



(b) PCA using common VOCs (only breath profiles)



(c) PCA using common VOCs (only sputum profiles)

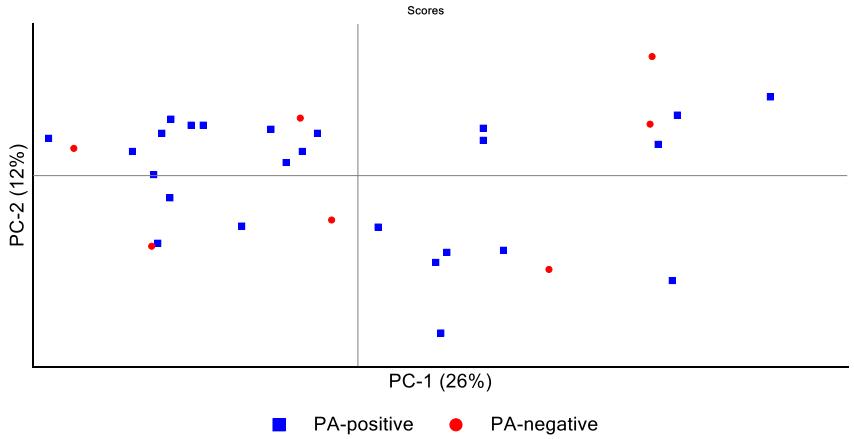


Figure 5.3: PCA scores plots prepared using the VOCs common between both sputum and breath samples (total 24 VOCs): (a) both profiles together, (b) only breath profiles, and (c) only sputum profiles.

5.3.4. Linear discriminant analysis

Based on the findings of PCAs performed to allow differentiation between PA-positive and PA-negative CF subjects, linear discriminant analysis (LDA) was performed as a supervised approach to allow any possible differentiation between these two groups. In this process, each group was assigned with a score on a group measure; all PA-positive samples were assigned a value of 0 (total 25), while all PA-negative samples were assigned a value of 1 (total 7). The first set of LDAs were then performed using all the VOCs detected in sputum and breath profiles (100 VOCs for sputum samples and 56 VOCs for breath samples). **Figure 5.4** shows the linear discriminant (LD) scores of subjects from two contrasting groups in two approaches: (a) LDA performed using all breath VOCs and (b) LDA performed using all sputum VOCs.

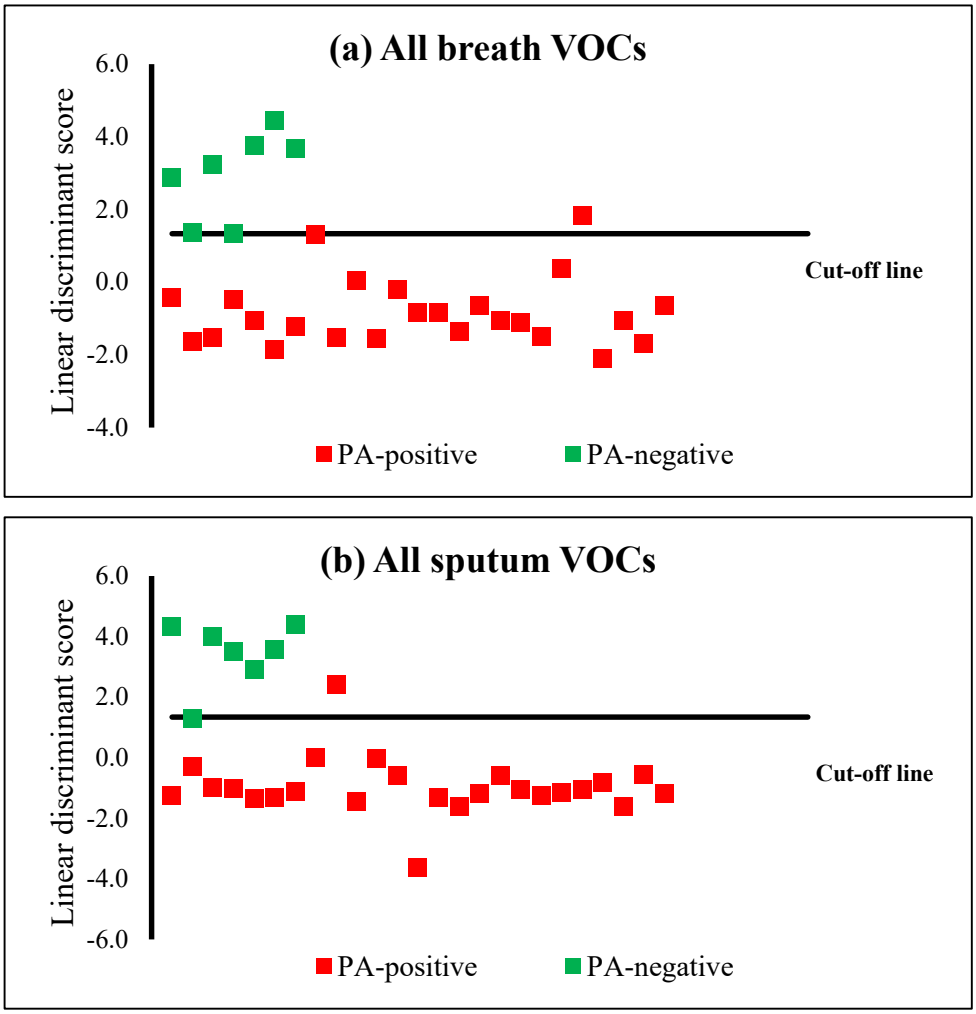


Figure 5.4: LD scores of CF subjects from two contrasting groups: (a) LDA performed using all breath VOCs and (b) LDA performed using all sputum VOCs.

The LD-score plot was capable of allowing classification between CF subjects with different lung infection status. However, sputum profiles performed relatively better with two misclassifications (one from both groups) while breath profiles produced three misclassifications (two from PA-negative group and one from PA-positive group).

In the next stage, the goal was set to evaluate the performance of the VOCs common between sample types in terms of their capability to classify PA-positive and PA-negative CF subjects. To facilitate this process, LDA was performed using the 24 common VOCs and through the following approaches: (a) sputum and breath profiles applied together, (b) only breath profiles used, and (c) only sputum profiles were used. The first approach (breath and sputum profiles together) classified 24 PA-positive and six PA-negative subjects correctly (with one misclassification from each group). Left-one-out cross-validation (LOOCV) performed on this approach produced nine misclassifications (out of 25) for the PA-positive group and two misclassifications (out of 7) for the PA-negative group. Relatively better results were obtained when only breath profiles were used. All the subjects were classified correctly according to their *Pseudomonas aeruginosa* infection status (PA-positive vs PA-negative). The outcome of LOOCV on this approach also provided a better outcome than the first approach with six misclassifications for the PA-positive group (out of 25 samples). However, two of the subjects from the PA-negative group were again misclassified as PA-positive in the LOOCV. The third approach (LDA using only sputum profiles) produced the same outcome as the first approach with one misclassification for both the PA-positive and PA-negative groups. However, LOOCV performed on this approach provided the worst outcome with 9 misclassifications for the PA-positive (out of 25 samples) and 5 misclassifications (out of 7 samples) for the PA-negative group.

Figure 5.5 shows the linear discriminant (LD) scores of CF subjects with/without *Pseudomonas aeruginosa* infection (PA-positive vs PA-negative CF subjects) while LDAs were performed using the common VOCs (total 24) and the following LDA approaches were tested: (a) LDA performed using both sputum and breath profiles together; (b) LDA performed using only breath profiles; and (c) LDA using only sputum profiles.

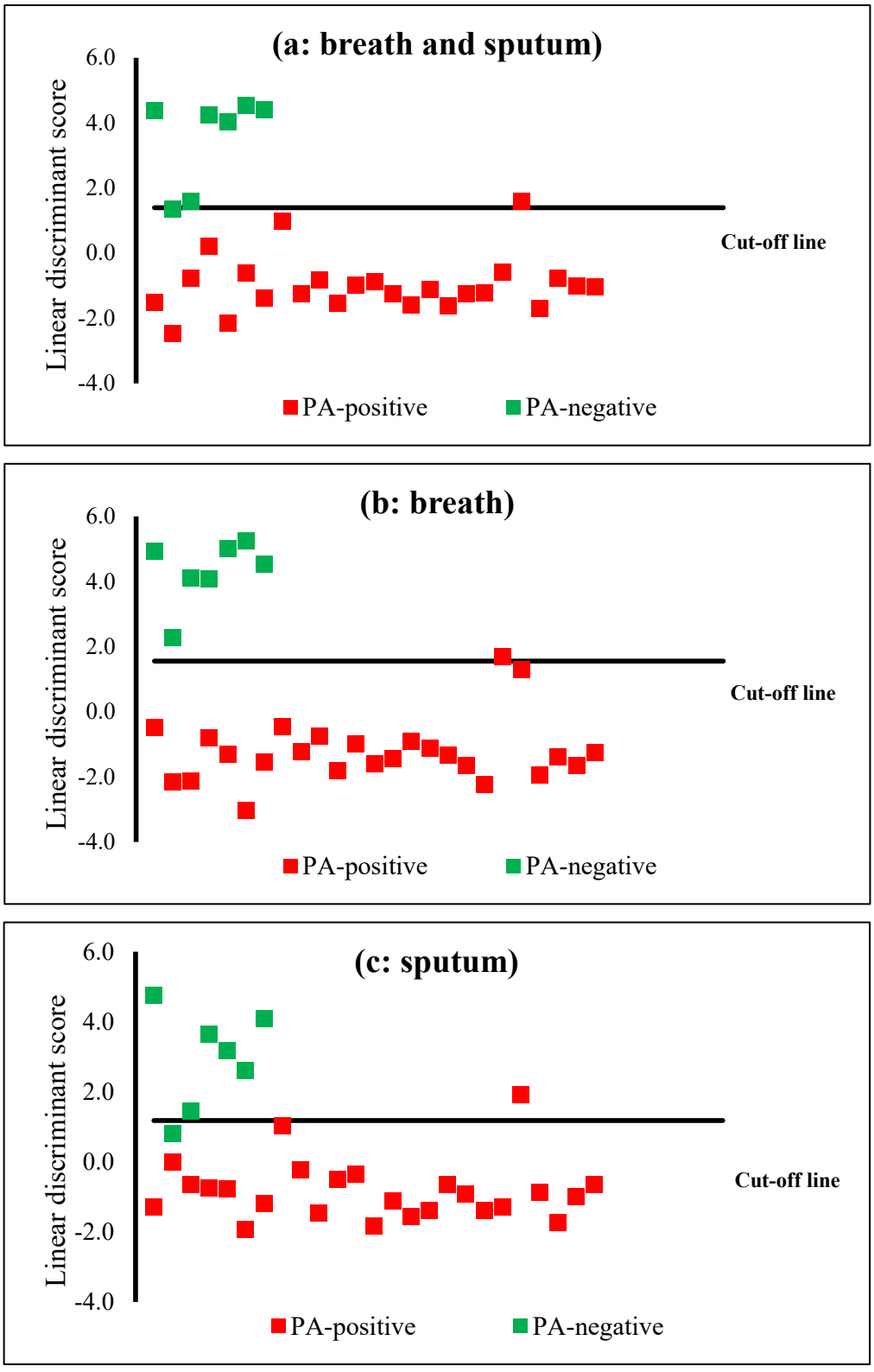


Figure 5.5: LD scores of subjects from two contrasting groups (PA-positive vs PA-negative) when LDAs were performed using the 24 VOCs common between sputum and breath profiles: (a) sputum and breath profiles applied together, (b) only breath profiles used, and (c) only sputum profiles were used.

The overall comparison of LD scores between three approaches (a, b, and c) exhibited similar patterns. However, breath profiles performed relatively better than sputum profiles in the LDA to classify two contrasting groups (PA-positive vs PA-negative). In particular, all CF subjects were classified correctly according to their infection status (PA-positive vs PA-negative) using breath profiles. In contrast, sputum profiles produced one misclassification from both the PA-positive and PA-negative groups. In addition, LOOCV performed for these two approaches also exhibited relatively better performance of breath profiles (total eight misclassifications: six from the PA-positive group + two from the PA-negative group) compared to a total of fourteen misclassifications obtained in the LOOCV performed on the classification model of sputum profiles (nine from the PA-positive group + five from the PA-negative group). The overall outcome of the LDAs performed in this chapter showed that the VOCs common between sputum and breath samples have relatively better classification accuracy than the entire profile of the VOCs. In addition, breath profiles exhibited relatively better classification accuracy than the sputum profiles.

5.3.5. VOCs common between sputum and breath samples: general characteristics and report on literature

Table 5.4 lists all the VOCs common between sputum and breath profiles (to note, this Table is a subset of **Table 5.2**). Aromatics, esters, and terpenoids are the three compound groups which contain the highest number of these VOCs (total 12, as four VOCs were from each of these three groups). Among the remainder of the VOCs, three are nitrogen containing compounds, two alcohols, two hydrocarbons, two ketones, and one compound from each of aldehyde, carboxylic acid, and halogenated compound groups.

Among these 24 VOCs common between sample types, only a few were reported previously in literature as related to lung infections in CF subjects. For instance, the first compound in the list, 2-ethyl-1-hexanol, is reported to be detected *in vitro* cultures of clinical isolates of *Aspergillus fumigatus* which is a common CF-related fungal species (Gerritsen et al., 2018). The next compound on the list, 1-propanol, is also reported in a study which analysed 36 genotypically different strains of *Pseudomonas aeruginosa* and reported 1-propanol in the headspace of most of the cultures (Shestivska et al., 2012). Additionally, propanoic acid is reported in several studies related to lung infections in CF subjects (Zang et al., 2017, Dryahina et al., 2016). For instance, Dryahina et al. (2016)

reported propanoic acid in the cultures of different CF related bacterial species such as *Pseudomonas aeruginosa, Staphylococcus aureus, Stenotrophomonas maltophilia and the Burkholderia cepacia complex* (Dryahina et al., 2016). The last compound in the list which is reported in a CF related study is octane; in that study, the authors co-cultured lung epithelial cell lines with *Pseudomonas aeruginosa* and reported elevated levels of octane from hydrogen peroxide treated cells, likely as a result of peroxidation of oleic acids (Lawal et al., 2018a).

Table 5.4: List of VOCs common between sputum and breath samples

Order	Compounds	Compound class
1	1-Hexanol, 2-ethyl-	Alcohol
2	1-Propanol	Alcohol
3	Octanal	Aldehyde
4	2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	Aromatic
5	Benzene	Aromatic
6	Benzene, 1,3-dimethyl-	Aromatic
7	Toluene	Aromatic
8	Propanoic acid	Carboxylic acid
9	Methyl propionate	Ester
10	Propane, 2,2-dimethoxy-	Ester
11	Propane, 2-methoxy-2-methyl-	Ester
12	Propanoic acid, 2-oxo-, methyl ester	Ester
13	Trichloromethane	Halogenated
14	Heptane, 2,2,4,6,6-pentamethyl-	Hydrocarbon
15	Octane	Hydrocarbon
16	Acetoin	Ketone
17	Cyclohexanone	Ketone
18	2-Oxazolidinone, 3-amino-5-(4-morpholinylmethyl)-	Nitrogen containing
19	Acetonitrile	Nitrogen containing
20	Morpholine, 4-octadecyl-	Nitrogen containing
21	α-Pinene	Terpenoid
22	Camphene	Terpenoid
23	Eucalyptol	Terpenoid
24	Limonene	Terpenoid

5.4. Conclusion

In this study, sputum and breath samples were collected from CF subjects (total 32) and profiled for VOCs using GC×GC-TOFMS. The aim was to analyse VOCs directly in the headspace of sputum samples and to compare the sputum VOC profiles with their corresponding breath VOC profiles. The rationale behind the aim was to study the potential transition of VOCs from sputum to breath and to identify any VOC biomarker/s associated with lung infections in CF subjects. A total of 132 VOCs were detected in this

study (100 from sputum and 56 from breath samples), while only 24 VOCs were common between both sample types. There was also significant difference in the composition of VOCs between the two sample types. For instance, sputum samples have higher percentage of alcohols, aldehydes, carboxylic acids, halogenated compounds, and ketones than corresponding breath samples. On the other hand, compound classes such as aromatics, esters, hydrocarbons, nitrogen containing compounds, and terpenoids demonstrated higher percentage values in breath samples compared to sputum samples. In addition, compound classes such as hydrocarbons, alcohols, and aldehydes exhibited very high percentage difference between the two sample types (sputum vs breath). For instance, only 6% of the VOCs detected in sputum samples were hydrocarbons compared to 20% in breath samples. However, sulfur containing compounds were the only compound class detected exclusively in sputum samples and were not present in any of the breath samples.

An attempt was made to differentiate between CF subjects with different lung infection status using their sputum and breath VOC profiles using multivariate analysis. To facilitate this process, CF subjects were divided into two groups according to their *P. aeruginosa (PA)* infection status (i.e. PA-positive vs PA-negative). Initially, PCA was performed to differentiate between the two groups using both sputum and breath profiles. However, it was not possible to allow clear distinction between the two contrasting groups using PCA. In the next step, LDA was performed as the supervised approach to allow differentiation between PA-positive and PA-negative CF subjects. The outcome of LDA showed that it is possible to correctly classify PA-positive and PA-negative CF subjects using both profiles with some limitations. However, the VOCs common between both profiles exhibited better classification accuracy than all of the VOCs combined. In addition, breath profiles provided better classification accuracy than their sputum counterpart. However, based on the overall findings of this study, it is important to consider both sample types while searching for biomarker/s associated with lung infections in CF subjects, as it will assist with the identification of the core VOCs which are common between both sample types.

Chapter 6: COMPARISON OF BREATH PROFILES FROM CF SUBJECTS WITH DIFFERENT LUNG INFECTIONS

Chapter 6: COMPARISON OF BREATH PROFILES FROM CF SUBJECTS WITH DIFFERENT LUNG INFECTIONS

6.1. Introduction

Diverse metabolic processes within the body produce a large number of volatile organic compounds (VOCs) that are released into the blood stream and partitioned in the breath once the blood reaches the lungs (Phillips et al., 1994, Boots et al., 2012). For this reason, certain metabolic disorders can significantly influence the composition of VOCs in human breath, and thus can provide important clues to diagnosis (Amann et al., 2007, Ratiu et al., 2019). For instance, subjects with uncontrolled diabetes can have an excessive amount of acetone in their breath which is responsible for the rotten apple-like smell (Kim et al., 2012, Owen et al., 1982). In addition to metabolic disorders, chronic infections can also result in the excretion of VOCs into human breath and can generate unique VOC patterns (Boots et al., 2012, Elmassry and Piechulla, 2020). For instance, Phillips et al. (2007, 2010, 2012) reported a unique pattern of breath VOCs which has allowed discrimination between subjects with/without active pulmonary tuberculosis (Phillips et al., 2012, Phillips et al., 2010, Phillips et al., 2007).

In the past decade, a small number of studies have focused on the VOCs present in breath samples from subjects suffering from cystic fibrosis (CF) (Barker et al., 2006, Enderby et al., 2009, Gilchrist et al., 2013, Kamboures et al., 2005, Neerincx et al., 2016b, Robroeks et al., 2010b, White et al., 2013, Rees et al., 2018). Among these studies, some performed targeted analysis of specific compounds/groups of compounds (e.g. volatile sulfur compounds (Kamboures et al., 2005, Barker et al., 2006) and identified a set of VOCs (Barker et al., 2006)) using gas chromatography-mass spectrometry (GC-MS). Other studies reported the application of selected-ion flow-tube mass spectrometry (SIFT-MS) to study the difference of hydrogen cyanide (Enderby et al., 2009, Gilchrist et al., 2013) and acetic acid vapour (Smith et al., 2016) in exhaled breath between CF subjects and healthy controls. Still other analytical techniques such as Proton-transfer-reaction time-of-flight mass spectrometry (PTR-TOFMS: (White et al., 2013)) and secondary electrospray ionization high-resolution mass spectrometry (SESI-HRMS: (Gaisl et al., 2018)) have been used to study the pattern of VOCs. However, there is only two studies

which has reported differentiation between CF subjects with different lung infections both applying thermal desorption-gas chromatography-mass spectrometry (TD-GC-MS) (Neerincx et al., 2016b, Robroeks et al., 2010b). Robroeks et al. (2010) reported a combination of 22 VOCs to differentiate between CF subjects and healthy controls. In the current thesis (chapter 4) we have further reduced the number of VOCs to 16 to demonstrate a distinction between CF subjects and healthy controls. In addition to differentiating between CF subjects and healthy controls, Robroeks et al. (2010) also reported successful discrimination between CF subjects with/without *Pseudomonas aeruginosa* (PA) infections; however, the authors failed to report the molecular identity of the VOCs which allowed such discrimination (Robroeks et al., 2010b). In contrast, Neerincx et al. (2016b) listed a set of nine VOCs important for discrimination between CF subjects with/without *Staphylococcus aureus* (SA) infections. However, Neerincx et al. (2016b) did not provide information of their statistical model which they claimed allowed differentiation between CF subjects with/without SA (Neerincx et al., 2016b).

In this study, breath samples were collected from CF subjects with different lung infections and a small number of CF subjects without any infectious species detected in their lungs and airways as confirmed using their sputum culture results. The non-targeted analysis of VOCs present in these breath samples was performed using comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC–TOFMS). Finally, multivariate analysis was performed to study the pattern of VOCs between CF subjects with different lung infections and those without any known lung infections detected in their sputum cultures.

6.2. Materials and Method

6.2.1. Sample collection

Breath samples were collected from CF subjects during their regular visit to an outpatient CF clinic in Westmead Hospital, NSW, Australia. The details of breath sampling are described in **Section 3.2**. In brief, participants were asked to sit and relax and provide a set of sputum samples (one for routine microbial culture and another for the profiling of VOCs) and a corresponding breath sample in a 1 L transparent Tedlar® bag (SKC Inc., USA). Tedlar® bags containing breath samples were transported to the University of Technology Sydney (UTS) for the extraction and analysis of VOCs within 72 h of collection to maintain sample integrity.

6.2.2. Extraction of VOCs

The VOCs present in breath samples were extracted using a three-phase (divinylbenzene/carboxen/polydimethylsiloxane) SPME fibre attached to a manual fibre holder (Supelco, Bellefonte, PA, USA). The optimisation of breath sampling is described in **Chapter 4**.

6.2.3. GC×GC-TOFMS analysis of samples

Samples were analysed using a Pegasus® 4D GC×GC–TOFMS (LECO, Castle Hill, NSW, Australia) equipped with a liquid nitrogen cryogenic quad jet modulator. Refer to **section 3.2.4** for details of the GC×GC–TOFMS analysis of breath samples.

6.2.4. Data processing and statistical analysis

A total of 65 different breath samples were collected from individual CF subjects and were analysed for VOCs using GC×GC–TOFMS. This provided a total of 181 different chromatograms, as most of the samples were analysed in triplicates (51×3=153), while some were analysed in duplicates (14×2=28). A computerised data acquisition and peak integration system (ChromaTOF® version 4.51.6.0) was used to process the raw data, while the identification of compounds was performed using the mass spectral library database provided by the USA National Institute of Standards and Technology (NIST, 2011). The alignment of peaks identified using the NIST database was performed using the Statistical Compare software feature in ChromaTOF®. Statistical Compare was performed with samples sorted into two classes: (i) breath (n=181) and (ii) bag blank (pre-cleaned Tedlar® bags filled with room air: n=10). This approach allowed the extraction of any VOCs which were not from the bag materials or the internal standard (deuterated chlorobenzene). A signal-to-noise (S/N) of 20 was also used to search for peaks not found during the initial peak finding step. In addition, a mass spectral match >60% was required for peaks to be identified as the same compound across chromatograms during alignment. After alignment, the analyte peak areas were normalized using the internal standard peak area.

Further statistical comparison was also carried out to identify the compounds with the highest between-class (i.e. "breath vs bag blank") variance by applying a Fisher Ratio (FR) calculated using the same Statistical Compare software feature in ChromaTOF®. This was performed by comparing the calculated FR to a critical F value (F_{crit}) obtained

through an F-test (Bean et al., 2015, Nizio et al., 2016). Analytes with higher FR values (or analytes absent from a class or only detected in a single sample in a class thus labelled as 'undefined') indicated compounds that statistically differed in abundance between the defined classes and were therefore retained. For all samples where an *F*-test was performed, compounds with FR above the F_{crit}, which includes those labelled as 'undefined', were exported as a *.csv file. The resulting *.csv file was imported into Microsoft Excel for the manual removal of chromatographic artefacts (i.e. column bleed, solvent peaks, and bag artefacts) and further statistical analysis.

6.2.5. Linear discriminant analysis

LDA was performed to allow differentiation between: (i) PA-positive vs PA-negative CF subjects and (ii) PA-positive/PA-negative vs CF subjects with no known lung infections as confirmed using their sputum culture results. This Chapter also utilised a larger set of CF breath profiles (total 65 CF subjects: 38 PA-positive vs 21 PA-negative vs 6 CF subjects with no known lung infections) to perform LDA. In this study, LDA was performed using RStudio (version R 3.5.3, 2019).

6.3. Results and Discussion

6.3.1. Infection status of CF subjects

Of the 65 CF subjects, 45 (69.2 %) were infected with different bacterial species. Of the 45 CF subjects with bacterial infections, 6 had *S. aureus*, 32 had *P. aeruginosa* (either mucoid/non-mucoid or both), 6 had both *S. aureus* and *P. aeruginosa*, and only 1 had *S. maltophilia*. Among the remaining 20 CF subjects, 3 had fungal infections, 11 had both bacterial and fungal infections, while 6 CF subjects had no species detected in their sputum culture (**Table 6.1)**.

Table 6.1: Lung infection status of individual CF subjects

Subject ID	Organisms	
	Bacteria	**Fungus**
[A] CF subjects with bacterial infections		
CF16	*S. aureus*	
CF36	*S. aureus*	
CF51	*S. aureus*	
CF56	*S. aureus*	
CF67	*S. aureus*	
CF68	*S. aureus*	
CF17	*S. maltophilia*	
CF03	*Mucoid P. aeruginosa*	
CF04	*Mucoid P. aeruginosa*	

CF06	*Mucoid P. aeruginosa*		
CF10	*Mucoid P. aeruginosa*		
CF12	*Mucoid P. aeruginosa*		
CF19	*Mucoid P. aeruginosa*		
CF22	*Mucoid P. aeruginosa*		
CF23	*Mucoid P. aeruginosa*		
CF25	*Mucoid P. aeruginosa*		
CF26	*Mucoid P. aeruginosa*		
CF27	*Mucoid P. aeruginosa*		
CF29	*Mucoid P. aeruginosa*		
CF35	*Mucoid P. aeruginosa*		
CF41	*Mucoid P. aeruginosa*		
CF44	*Mucoid P. aeruginosa*		
CF49	*Mucoid P. aeruginosa*		
CF50	*Mucoid P. aeruginosa*		
CF52	*Mucoid P. aeruginosa*		
CF60	*Mucoid P. aeruginosa*		
CF63	*Mucoid P. aeruginosa*		
CF64	*Mucoid P. aeruginosa*		
CF66	*Mucoid P. aeruginosa*		
CF02	*Non-mucoid P. aeruginosa*		
CF05	*Non-mucoid P. aeruginosa*		
CF40	*Non-mucoid P. aeruginosa*		
CF42	*Non-mucoid P. aeruginosa*		
CF65	*Non-mucoid P. aeruginosa*		
CF15	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa*	
CF28	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa*	
CF30	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa*	
CF43	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa*	
CF45	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa*	
CF20	*Mucoid P. aeruginosa*	*S. aureus*	
CF48	*Mucoid P. aeruginosa*	*S. aureus*	
CF53	*Mucoid P. aeruginosa*	*S. aureus*	
CF11	*Non-mucoid P. aeruginosa*	*S. aureus*	
CF39	*Non-mucoid P. aeruginosa*	*S. aureus*	
CF07	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa; S. aureus*	
[B] CF subjects with only fungal infection			
CF21			*A. xylosoxidans*
CF47			*A. xylosoxidans*
CF69			*T. mycotoxinivorans*
[C] CF subjects with both bacterial and fungal infections			
CF54	*Mucoid P. aeruginosa*		*A. fumigatus; A. niger*
CF24	*Mucoid P. aeruginosa*		*A. fumigatus*
CF58	*Mucoid P. aeruginosa*		*A. xylosoxidans*
CF18	*Non-mucoid P. aeruginosa*		*S. apiospermum*
CF14	*Non-mucoid P. aeruginosa*	*Mucoid P. aeruginosa*	*S. aurantiacum*
CF37	*Mucoid P. aeruginosa*	*S. aureus*	*A. fumigatus*
CF33	*N. farcinica*		*A. fumigatus*
CF46	*S. aureus*		*A. fumigatus*
CF57	*S. aureus*		*S. aurantiacum*
CF38	*S. aureus*	*S. maltophilia; β-Lactamase (-ve) H. influenzae*	*S. aurantiacum*
CF13	*S. aureus*		*T. mycotoxinivorans*
[D] CF subjects without any species detected in their sputum culture			
CF31	None		
CF32	None		
CF34	None		
CF55	None		
CF59	None		
CF61	None		

6.3.2. List of detected VOCs

Table 6.2 lists the VOCs detected in breath samples from CF subjects (n=65). Among these 72 VOCs detected, the highest numbers were classed as aromatics (18), followed by hydrocarbons (10), nitrogen containing compounds (10), alcohols (8), esters (8), ketones (8), terpenoids (5), halogenated compounds (3), a carboxylic acid (1) and a sulfur containing compound (1).

Table 6.2: List of VOCs detected in breath samples collected from CF subjects

Order	Compounds	Class
1	1-Hexanol, 2-ethyl-	Alcohols
2	1-Propanol	Alcohols
3	Ethanol	Alcohols
4	Ethanol, 2-(2-ethoxyethoxy)-	Alcohols
5	Ethanol, 2-(hexyloxy)-	Alcohols
6	1-Butanol	Alcohols
7	2-Propanol, 1-(2-butoxy-1-methylethoxy)-	Alcohols
8	Cyclohexanol, 5-methyl-2-(1-methylethyl)-, [1S-(1α,2α,5α)]-	Alcohols
9	Nonanal	Aldehydes
10	Octanal	Aldehydes
11	2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	Aromatics
12	Benzene	Aromatics
13	Benzene, 1,3-dichloro-	Aromatics
14	Benzene, 1,3-dimethyl-	Aromatics
15	Benzene, 1-methyl-3-(1-methylethyl)-	Aromatics
16	Ethylbenzene	Aromatics
17	Oxime-, methoxy-phenyl-_	Aromatics
18	Tetrahydrofuran	Aromatics
19	Toluene	Aromatics
20	Benzaldehyde	Aromatics
21	Benzene, 1-ethenyl-4-ethyl-	Aromatics
22	Benzene, 1-methyl-4-(1-methylethenyl)-	Aromatics
23	Furan, 2-pentyl-	Aromatics
24	Naphthalene	Aromatics
25	p-Xylene	Aromatics
26	Styrene	Aromatics
27	Propanoic acid	Carboxylic acids
28	Acetic acid, phenyl ester	Esters
29	Butyrolactone	Esters
30	Methyl propionate	Esters
31	Pentane, 3-ethyl-2,2-dimethyl-	Esters
32	Propane, 2,2-dimethoxy-	Esters
33	Propanoic acid, 2-oxo-, methyl ester	Esters
34	Acetic acid, methyl ester	Esters
35	Tridecanoic acid, methyl ester	Esters
36	Tetrachloroethylene	Halogenated
37	Trichloromethane	Halogenated
38	Butane, 1,1,3,4-tetrachloro-1,2,2,3,4,4-hexafluoro-	Halogenated
39	2,2,4,4-Tetramethyloctane	Hydrocarbons

40	Butane, 2,2,3,3-tetramethyl-	Hydrocarbons
41	Decane	Hydrocarbons
42	Heptane, 2,2,4,6,6-pentamethyl-	Hydrocarbons
43	Heptane, 5-ethyl-2,2,3-trimethyl-	Hydrocarbons
44	Octane	Hydrocarbons
45	Pentane, 2,2,3,3-tetramethyl-	Hydrocarbons
46	Tridecane	Hydrocarbons
47	Undecane	Hydrocarbons
48	n-Hexane	Hydrocarbons
49	2-Cyclohexen-1-one	Ketones
50	3-Hexanone, 2,5-dimethyl-4-nitro-	Ketones
51	Acetoin	Ketones
52	Cyclohexanone	Ketones
53	Isophorone	Ketones
54	Methyl Isobutyl Ketone	Ketones
55	2,3-Butanedione	Ketones
56	2-Propanone, 1-(acetyloxy)-	Ketones
57	1-Dodecanamine, N,N-dimethyl-	Nitrogen containing
58	2-Oxazolidinone, 3-amino-5-(4-morpholinylmethyl)-	Nitrogen containing
59	Acetonitrile	Nitrogen containing
60	Cyclohexane, isothiocyanato-	Nitrogen containing
61	Morpholine, 4-octadecyl-	Nitrogen containing
62	Urea, tetramethyl-	Nitrogen containing
63	4-Morpholinebutyric acid, α-methyl-α,α-diphenyl-	Nitrogen containing
64	Benzonitrile	Nitrogen containing
65	Cyclohexane, isocyanato-	Nitrogen containing
66	Propane, 2-isocyanato-2-methyl-	Nitrogen containing
67	1-Propene, 1-(methylthio)-	Sulfur containing
68	α-Pinene	Terpenoids
69	Camphene	Terpenoids
70	Eucalyptol	Terpenoids
71	Limonene	Terpenoids
72	α-Phellandrene	Terpenoids

6.3.3. Linear discriminant analysis

LDA performed in Chapter 5 showed that it is possible to classify with a high percentage accuracy the CF subjects according to their *Pseudomonas aeruginosa* infection status using breath VOCs profiles. However, the number of samples from each group was limited (25 PA-positive + 7 PA-negative CF samples). In this chapter (Chapter 6), we are able to utilise a larger set of breath profiles obtained from CF subjects (a total of 65 breath profiles) to allow comparison between PA-positive and PA-negative CF subjects. Among these 65 CF subjects, 38 were PA-positive, 21 were PA-negative, while 6 have no known lung infection as confirmed using their sputum culture results. Initially, the LDA was performed to classify PA-positive and PA-negative CF subjects using their complete breath VOCs profiles. To facilitate this process, each group was assigned a score on a

group measure; all PA-positive samples were assigned a value of 0, while all PA-negative samples were assigned a value of 1. LDA was then performed using the 72 VOCs detected in breath samples. As depicted in **Figure 6.1**, all CF subjects were well separated according to their PA-infection status. In particular, the PA-negative subjects (the green rectangles in **Figure 6.1** are located well above the LDA cut-off line). However, leave-one-out cross-validation (LOOCV) provided 16 misclassifications (out of 38) for the PA-positive group and 9 misclassifications (out of 21) for the PA-negative group.

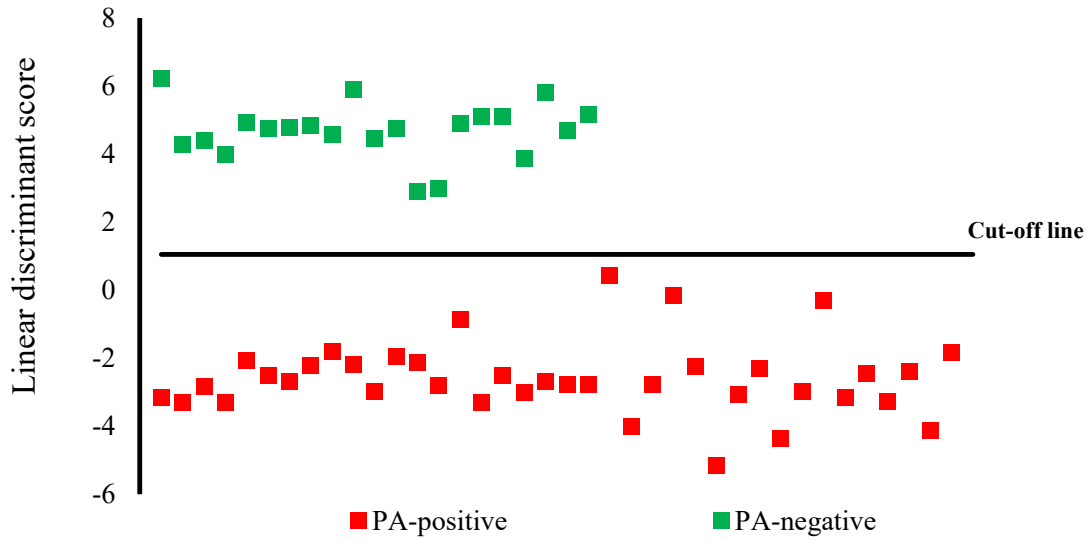


Figure 6.1: LD scores of subjects from PA-positive and PA-negative groups

After the first LDA, a second set of LDAs were performed to check whether it is possible to differentiate CF subjects with/without *Pseudomonas aeruginosa* from CF subjects with no known lung infection as confirmed using their sputum culture results. To facilitate this process, LDAs were performed in the following approaches: (a) PA-negative (n=21) vs No species CF subjects (n=6) and (b) PA-positive (n=38) vs No species CF subjects (n=6). As with the previous LDA, each group was assigned a score on a group measure. In this approach, (a) all PA-negative samples were assigned a value of 0, while CF subjects with no known lung infection were assigned a value 1. Similarly, in approach (b) all PA-positive samples were assigned a value 0, while CF subjects with no known lung infection were assigned a value of 1. **Figure 6.2** shows the linear discriminant (LD) scores obtained from the following comparisons: (a) PA-negative vs No species CF subjects and (b) PA-positive vs No species CF subjects.

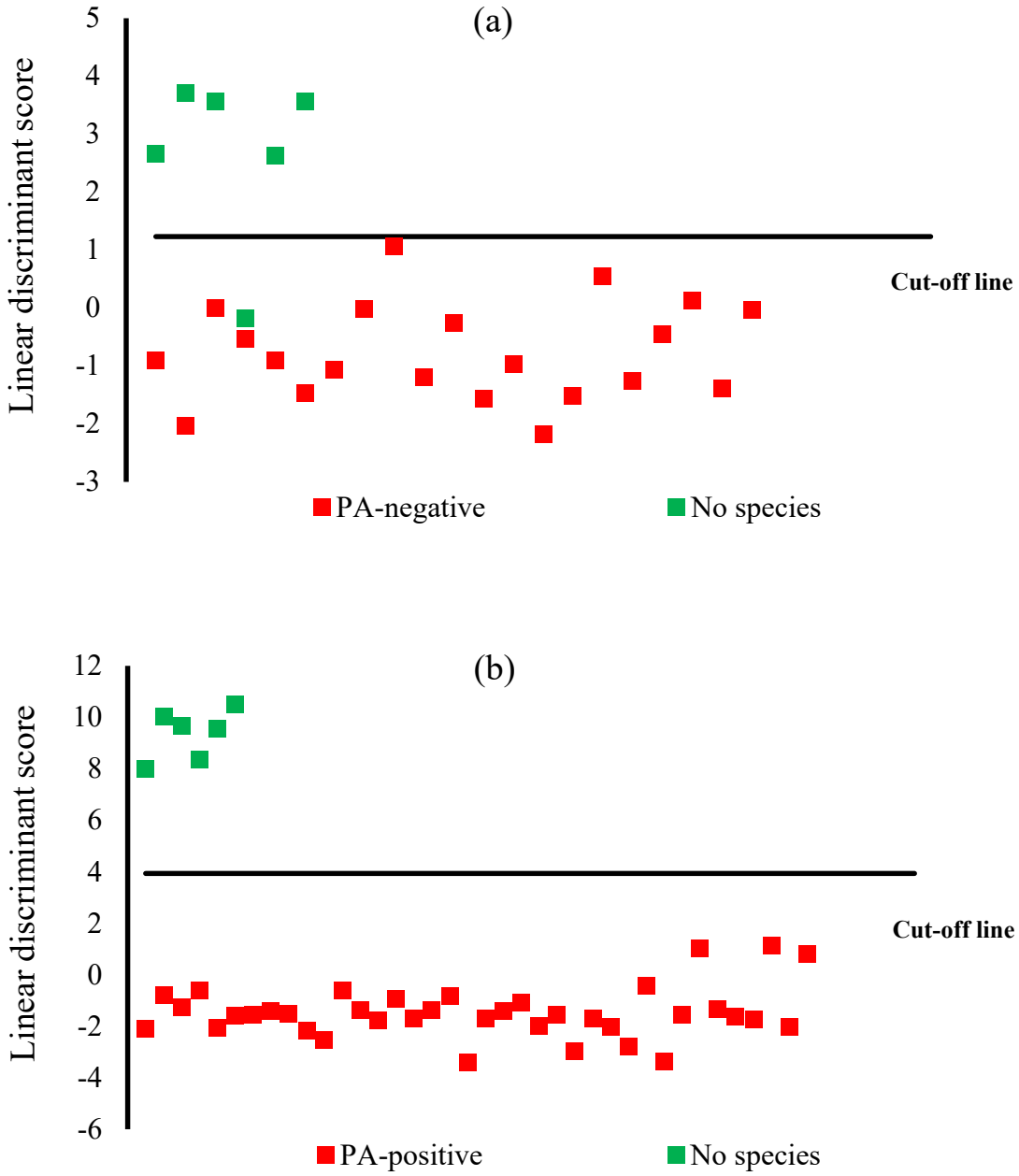


Figure 6.2: LD scores obtained from CF breath profiles: (a) PA-negative vs No species CF subjects and (b) PA-positive vs No species CF subjects.

In the LDA performed between PA-negative CF subjects and CF subjects with no known lung infection, one of the CF subjects without any infection was misclassified as PA-negative CF subject (the green rectangle below the cut-off line at **Figure 6.2a**). In contrast, there was no misclassification in the LDA performed between PA-positive and non-infected CF subjects (**Figure 6.2b)**. It was also important to notice that the separation between PA-positive and non-infected CF subjects is relatively stronger than the

separation between PA-negative and non-infected CF subjects. In particular, the LD scores for the six non-infected CF subjects in approach (b) (PA-positive vs non-infected CF subjects) ranged between 7.98 to 10.5, which is significantly higher than the LD scores for the six non-infected CF subjects (ranged between 2.62 to 3.71) obtained with approach (a) (PA-negative vs non-infected CF subjects). In addition, the LD scores for 21 PA-negative CF subjects (approach (a)) ranged between -2.18 to 1.06. The spread of this range is 3.24 (-2.18~1.06), which is also lower than the spread obtained from 38 PA-positive CF subjects with approach (b) (LD scores for PA-positive CF subjects ranged between -3.4 to 1.14 with spread value of 4.54 (-3.4~1.14)). The overall comparison between CF subjects with/without *Pseudomonas aeruginosa* infection and thus with no known lung infections using their breath VOCs profiles show that the separation between PA-positive and non-infected CF subjects are relatively more distinct than the separation between PA-negative and non-infected CF subjects.

6.3.4. The finding of this study in the context of relevant literature

In this study, adult CF subjects with/without *Pseudomonas aeruginosa* infection and those with no known lung infections (total 65 CF subjects: 38 PA-positive vs 21 PA-negative vs 6 CF subjects with no known lung infections) were differentiated using a total of 72 VOCs detected in their breath samples. To note, breath samples were collected in 1L Tedlar® bags, extracted using SPME, and analysed using GC×GC–TOFMS. The differentiation between samples was performed using linear discriminant analysis (LDA: model prepared using RStudio, version R 3.5.3, 2019).

To date, there are a number of studies which have reported the analysis of VOCs released from *in vitro* cultures of species associated with lung infections in CF subjects (Bean et al., 2012, Carroll et al., 2005, Filipiak et al., 2012, Goeminne et al., 2012, Labows et al., 1980, Savelev et al., 2011, Bean et al., 2015, Bean et al., 2016, Briard et al., 2016, Dryahina et al., 2016, Karami et al., 2017, Neerincx et al., 2016a, Neerincx et al., 2015, Nizio et al., 2016, Lawal et al., 2018c, Purcaro et al., 2018, Timm et al., 2018, Shestivska et al., 2015). Among these 18 studies: 4 have reported VOCs released from *Pseudomonas aeruginosa* cultures (Bean et al., 2012, Bean et al., 2015, Bean et al., 2016, Timm et al., 2018); 5 have reported VOCs from *Pseudomonas aeruginosa* and other species cultured separately (e.g. *Staphylococcus aureus, Burkholderia cepacia complex, Stenotrophomonas maltophilia, Haemophilus influenza* and etc.) (Dryahina et al., 2016,

Filipiak et al., 2012, Labows et al., 1980, Nizio et al., 2016); and 3 have reported VOCs from co-cultures of *Pseudomonas aeruginosa* and other CF related species (e.g. *Aspergillus fumigatus and respiratory syncytial virus*) (Briard et al., 2016, Neerincx et al., 2016a, Purcaro et al., 2018). Among the remaining studies (total 6), two have reported VOCs from *in vitro* cultures of different *Stenotrophomonas strains* (Shestivska et al., 2015) and from cultures of *Staphylococcus aureus, Escherichia coli and Candida albicans* (Karami et al., 2017). The other 4 studies reported VOCs detected in the following sample types: cultures of bacterial isolates from sputum and cough swab samples from CF subjects (Carroll et al., 2005, Neerincx et al., 2015) and sputum samples collected from CF subjects (Goeminne et al., 2012, Savelev et al., 2011).

All of these culture-based studies provide important information on the VOCs released from cultures/sputum samples associated with different CF lung infections. However, only two of these studies have attempted to differentiate between CF subjects based on their lung infections status (with/without *Pseudomonas aeruginosa*) using the VOCs detected in their sputum samples (Goeminne et al., 2012, Savelev et al., 2011). For instance, Savelev et al. (2011) analysed sputum samples collected from a total of 72 subjects (13 samples from CF subjects + 59 samples from non-CF bronchiectasis subjects). Among the 72 subjects, 32 had positive cultures of *Pseudomonas aeruginosa*, 12 had no pathogens, while the remaining 28 patients had other pathogens as confirmed using their sputum culture results. The authors reported a library of 17 compounds for the detection of *Pseudomonas aeruginosa* in sputum samples with a sensitivity of 91% and 88% specificity (Savelev et al., 2011). In a similar study, Goeminne et al. (2012) analysed a total of 28 sputum samples collected from CF subjects. Among these 28 CF subjects, 14 had *Pseudomonas aeruginosa* and 5 had no *Pseudomonas aeruginosa* as confirmed using their sputum culture results; the remaining 9 subjects had no history of having *Pseudomonas aeruginosa* infection. As reported by the authors, the differentiation between CF subjects with/without *Pseudomonas aeruginosa* infection produced a large number of false positives and false negatives. However, another model which they used to predict chronic *Pseudomonas aeruginosa* infection provided a relatively better outcome than the *Pseudomonas aeruginosa*-positive vs *Pseudomonas aeruginosa*-negative model (Goeminne et al., 2012).

Other than culture-based studies, there are several studies which reported combined analysis of VOCs from the cultures of CF-related species and also from the breath

samples collected from CF subjects (Kramer et al., 2015, Scott-Thomas et al., 2010, Shestivska et al., 2011). Among these studies, Scott-Thomas et al. (2010) analysed a single compound 2-aminoacetophenone in the cultures of different CF-related bacterial species (including *Pseudomonas aeruginosa*) and also in the breath samples collected from adult CF subjects (total 36: 16 *Pseudomonas aeruginosa*-positive vs 13 *Pseudomonas aeruginosa*-negative vs 17 healthy controls). The authors reported, higher 2-aminoacetophenone concentration in *Pseudomonas aeruginosa* culture (compared to other cultures) and also in the *Pseudomonas aeruginosa*-positive breath samples compared to *Pseudomonas aeruginosa*-negative and control breath (Scott-Thomas et al., 2010). Shestivska et al. (2011) also analysed a single compound methyl thiocyanate in *Pseudomonas aeruginosa* cultures and in breath samples collected from CF children with (n=19)/without (n=9) *Pseudomonas aeruginosa* and healthy control children (n=9). However, the authors reported no significant variation of methyl thiocyanate in breath samples collected from these three contrasting groups (Shestivska et al., 2011). In a recent study, Kramer et al. (2015) analysed VOCs from human epithelial cells infected with different CF-related species (including *Pseudomonas aeruginosa)* and breath samples from 9 adult CF subjects and 2 healthy controls. The authors also performed non-targeted analysis of VOCs and concluded that, rather than a single bio-marker,s a pattern/profile of VOCs may provide important clue in the diagnosis of lung infection in adult CF subjects.

A search of available literature also provided a total of 6 studies which has attempted to differentiate between CF subjects and healthy controls using only breath VOCs (Barker et al., 2006, Enderby et al., 2009, Gaisl et al., 2018, Kamboures et al., 2005, Smith et al., 2016, White et al., 2013). Among them, two of the early studies performed targeted analysis of certain breath VOCs between adult CF subjects and healthy controls and reported their concentration differences between these two contrasting groups (Barker et al., 2006, Kamboures et al., 2005). Enderby et al. (2009) analysed breath samples from children with CF (n=16) and with asthma (n=21) and reported elevated concentrations of hydrogen cyanide (HCN) in the breath of CF children compared to children with asthma (Enderby et al., 2009). Smith et al. (2016) reported elevated concentrations of acetic acid in the exhaled breath of CF subjects which was reported to be independent of their *Pseudomonas aeruginosa* infection status (Smith et al., 2016). The remaining studies were not conclusive while comparing breath samples from CF subjects and healthy

controls but provided tentative evidence that some exhaled breath features were different between CF subjects and healthy controls when conducting PTR-TOFMS (White et al., 2013) or SESI-HRTOFMS (Gaisl et al., 2018) based analysis of breath VOCs.

Finally, there are only three studies which have attempted to differentiate between CF subjects with different lung infection status using breath VOCs (Gilchrist et al., 2013, Neerincx et al., 2016b, Robroeks et al., 2010b). Gilchrist et al. (2013) analysed breath samples from CF subjects with/without *Pseudomonas aeruginosa* and reported the detection of HCN in mouth exhaled breath of both groups and only in the nose-exhaled breath of subjects with chronic *Pseudomonas aeruginosa* infection. Robroeks et al. (2010) analysed breath samples collected from a total of 105 children (48 CF + 57 controls). Among these 48 CF children, 49% were reported to have positive cultures with *Pseudomonas aeruginosa* in the past two years of that study. The authors reported a total of 22 VOCs which allowed differentiation between CF subjects and healthy controls with 100% accuracy. In addition, authors reported that it was possible to allow successful discrimination between CF subjects with/without *Pseudomonas aeruginosa* infections (Robroeks et al., 2010b). In the most recent study, Neerincx et al. (2016 b) analysed breath samples from CF subjects with (n=13)/without *Staphylococcus aureus* infection (n=5) and reported that it was possible to discriminate between *Staphylococcus aureus*-positive and *Staphylococcus aureus*-negative CF subjects with high sensitivity (100%) and specificity (80%) using their breath VOCs (Neerincx et al., 2016b).

Overall, an extensive search of the literature showed that there is no reported study which has successfully differentiated between adult CF subjects with/without *Pseudomonas aeruginosa* and those with no-known lung infections as confirmed using their sputum culture results. The only comparable study was conducted by Robroeks et al. (2010). However, the major difference between our study and the Robroeks et al. (2010) study is the age of the study group (children with CF-Robroeks et al. (2010) vs adult CF subjects in this study). In addition, Robroeks et al. (2010) allowed the following differentiations: (i) differentiation between CF subjects and healthy controls and (ii) CF subjects with/without *Pseudomonas aeruginosa* infection. In our study, we have allowed the following differentiations: (i) CF subjects with different lung infections and healthy controls (Chapter 4); (ii) CF subjects with/without *Pseudomonas aeruginosa* infection (Chapter 5 and Chapter 6); (iii) additional differentiations were allowed between CF

subjects with/without *Pseudomonas aeruginosa* and those with no-known lung infections as confirmed using their sputum culture results (Chapter 6).

The finding of the current study showed that it is possible to differentiate between CF subjects with different lung infections from non-infected CF subjects using their breath VOCs profiles. However, in this study, we were only able to compare CF subjects with/without *Pseudomonas aeruginosa* infection with those with no known lung infection. It is important to perform further studies to allow comparison between CF subjects infected with other major CF-related species (e.g. *Staphylococcus aureus*, *Burkholderia cepacia, and Stenotrophomonas maltophilia*) and non-infected CF subjects using their breath VOCs profiles. In addition, future studies would benefit from an increased number of samples from the non-infected CF group.

6.4. Conclusion

In this study, breath samples were collected from CF subjects with/without *Pseudomonas aeruginosa* infection and thus with no known lung infections as confirmed using their sputum culture results. The profiling of VOCs present in breath samples was performed using GC×GC–TOFMS which identified a total of 72 VOCs in all breath profiles. LDAs were then performed using these VOCs to differentiate between CF subjects with/without *Pseudomonas aeruginosa* infection and thus with no known lung infection. Using LDA, it was possible to classify the majority of PA-positive (n=38) and PA-negative (n=21) CF subjects correctly. In addition, the final set of LDAs performed between CF subjects with/without *Pseudomonas aeruginosa* and non-infected CF subjects showed that it is also possible to allow separation between both PA-positive and non-infected CF subjects and PA-negative and non-infected CF subjects using their breath VOCs profiles. However, overall findings of the LDAs showed that the separation between PA-positive and non-infected CF subjects are relatively more distinct than the separation between PA-negative and non-infected CF subjects.

Chapter 7:
CONCLUSIONS AND FUTURE DIRECTIONS

Chapter 7: CONCLUSIONS AND FUTURE DIRECTIONS

7.1. Conclusions

The aim of this thesis was to develop an analytical technique for detecting the VOCs present in both sputum and breath samples, and to apply that technique to the detection and identification of certain lung infections in CF subjects. To facilitate this process, samples were collected from both CF subjects (sputum and breath) and healthy controls (only breath) and profiled for VOCs using comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC–TOFMS).

The second chapter of this thesis reported the application of a SPME-GC×GC-TOFMS method for the optimal extraction and chemical characterisation of VOCs present in the headspace of sputum samples collected from CF subjects. In addition, complex GC×GC data processing and multivariate analysis (e.g. principal component analysis: PCA) performed in this study was also introduced. The results demonstrated that the applied method was effective to extract, detect and identify the VOCs associated with CF sputum. In addition, PCA performed using the VOCs detected in the sputum samples exhibited accurate clustering samples in the scores plots.

In the next chapter, the performance of two common breath sampling devices (Bio-VOC™ vs Tedlar® bag) and VOC-extraction techniques (SPME vs ST-extraction) were evaluated. The comparison between sampling devices showed relatively better performance of the Tedlar® bags compared to the Bio-VOC™ breath sampler. Although, samples collected using both devices allowed clear differentiation between exhaled human breath and background air, samples collected in Tedlar® bags exhibited better precision in terms of triplicate analysis and were thus selected as the sampling device for further breath sampling. Further comparison between extraction techniques (SPME vs ST-extraction) using Tedlar® bag samples demonstrated that both SPME and ST-extraction was able to distinguish between breath samples collected from CF subjects and healthy controls. However, SPME allowed relatively better separation among CF subjects in terms of their breath VOC profiles, thus the combination of Tedlar® bag sampling and SPME was selected as the optimal technique for the analysis of further breath samples collected in this study.

After sample collection and method optimisation, comparisons were carried out between CF subjects with lung infections and heathy controls without lung infections using their breath VOC profiles. Among the CF subjects, 59 were infected with different lung infections, while 6 had no known lung infections as confirmed using their sputum culture results. In contrast, all 51 control subjects recruited in this study were healthy individuals with no known health complications including lung infections. The univariate comparison of VOCs between the two contrasting study groups showed that CF breath profiles have relatively higher average numbers and abundance of VOCs than their control counterpart. A set of 16 VOCs was identified, while linear discriminant analysis (LDA) performed using these VOCs allowed classification between subjects with/without lung infections. In particular, controls without lung infections were classified with 98% accuracy, while the accuracy of classification for lung-infected CF subjects was 92%. The 16 VOCs discriminating the groups were common between both groups (infected vs non-infected). However, most of these VOCs were detected more frequently with higher concentrations in infected CF subjects. An increased concentration and prevalence of discriminating VOCs in CF breath is possibly due to their lung infection status. However, differences of breath VOCs between infected CF and non-infected controls can also be influenced by numerous other factors (i.e. the health status of CF subjects, antibiotics to treat lung infections, bronchodilators, inhaled corticosteroids, special CF diet, lifestyle, etc.).

In the next stage of the study, the goal was set to understand the source of breath-borne VOCs, and in particular any potential transition of important VOCs from CF sputum to CF breath. Hence, matching sputum and breath samples were collected from 32 CF subjects and profiled for VOCs using the same analytical technique. CF subjects were divided into two different groups according to their lung infection status as confirmed using sputum culture results *(i.e. Pseudomonas aeruginosa-positive and Pseudomonas aeruginosa-negative CF subjects)*. Notably, a total of 132 VOCs were detected in this study, which is higher than the total number of VOCs detected in all the CF and control breath samples (total 81 VOCs). The overall large numbers of VOCs were contributed mostly by the sputum samples. In particular, a total of 100 VOCs were detected in the sputum samples while only 56 were detected in the corresponding breath samples. Only 24 VOCs were common between both sample types. In addition, the composition of VOCs were also significant between the two sample types. For instance, sputum samples demonstrated a higher percentage of alcohols, aldehydes, carboxylic acids, halogenated

compounds, and ketones while the chemical composition of the corresponding breath samples was dominated by compound classes such as aromatics, esters, hydrocarbons, nitrogen containing compounds, and terpenoids. Further evaluation of sputum and breath profiles in terms of their capability to allow differentiation between PA-positive (n=25) and PA-negative (n=7) CF subjects using linear discriminant analysis (LDA) highlighted the ability of both sample types to allow differentiation between two contrasting groups. However, the VOCs common between sample types had better classification accuracy than the whole profile of the VOCs. In addition, breath profiles exhibited relatively better classification accuracy than the sputum profiles.

In the final chapter of this study, LDA was performed to allow further separation between CF subjects with/without *Pseudomonas aeruginosa* infection utilising a larger set of breath profiles obtained from CF subjects (a total of 65 breath profiles: 38 were PA-positive, 21 were PA-negative, while 6 had no known lung infection. LDA performed between PA-positive and PA-negative CF subjects showed clear separation between groups. In addition, further LDAs between CF subjects with/without *Pseudomonas aeruginosa* infection and thus with no known lung infections showed that it is also possible to allow separation between both PA-positive and non-infected CF subjects and PA-negative and non-infected CF subjects using their breath VOCs profiles.

The findings of this thesis show that it is possible to allow differentiation between CF subjects with different lung infections and healthy controls without any known health complications including lung infections using a set of 16 VOCs detected in their breath samples (Chapter 4). In addition, it was possible to allow differentiation between CF subjects with/without *Pseudomonas aeruginosa* infection using sputum (Chapter 5) and/or breath VOCs (Chapter 5 and Chapter 6). Finally, this is the first study which has reported the differentiation between adult CF subjects with/without *Pseudomonas aeruginosa* infection and those with no known lung infections (as confirmed using their sputum culture results) using their breath VOCs profiles. Overall, this thesis provided important insights into the diagnosis of lung infection/s in adult CF subjects using their breath VOCs profiles. However, due to the limited number of samples in this study, the comparisons were only possible between CF subjects with/without *Pseudomonas aeruginosa* infection with those with no known lung infection. To extend the knowledge in this area of research, it would be important to perform further studies allowing comparison between adult CF subjects with no known lung infection and those infected

with other major CF-related infectious species (e.g. *Staphylococcus aureus*, *Burkholderia cepacia, and Stenotrophomonas maltophilia*). Further studies in this area of research would also benefit from an increased number of samples from the non-infected CF group.

7.2. Future directions

In this thesis, a rapid and non-invasive analytical technique (based on GC×GC-TOFMS) was developed for the chemical characterisation of VOCs present in both sputum and breath samples of CF subjects. However, breath samples were also analysed from healthy participants as a control group and a set of 16 VOCs was identified which allowed discrimination between groups (i.e. CF subjects with lung infections vs healthy controls without lung infections). A future study could exclusively analyse these 16 VOCs using real-time measurement techniques (e.g. PTR-TOFMS) for rapid screening of subjects with/without lung infections. Another approach could investigate the targeted analysis of these VOCs using relatively simpler instrumentation (in comparison to GC×GC-TOFMS) such as a portable GC-FID for rapid screening of lung infections in clinical setups. This approach could widen the applicability of this technique to infections such as pulmonary tuberculosis (TB) whereby practitioners need to detect a particular species or need a yes/no answer rapidly (e.g. TB positive/PB-negative).

In addition, the developed technique was able to allow differentiation between CF subjects with/without *Pseudomonas aeruginosa* and thus without any active lung-infections although a much larger study population is required, especially in the non-infected CF group. A future study applying the analytical technique developed in this thesis and exploring breath profiles from a larger/equal number of PA-positive/PA-negative vs non-infected CF subjects would provide a more conclusive answer as to whether it is possible to allow clear differentiation between CF subjects with/without active lung infections. As it is very challenging to obtain breath and/or sputum samples from CF subjects without any lung infections from a CF clinic (as subjects normally do not visit CF clinics frequently if they have no lung infections), further studies would consider collecting samples from places other than the clinic (e.g. home or workplace of CF subjects). However, this will require extensive ethics approvals and a dedicated study population who is willing to donate their time for a longitudinal study.

Finally, the findings of this study indicate that the differences in breath VOC profiles between groups (CF subjects with lung infections and healthy controls without any lung

infections) could be due to their lung infection status, but the probability of influence from exogenous sources (e.g. CF medication, diet, and lifestyle) is very significant. In this study, it was not possible to consider all of these parameters and exclude CF subjects based on such criteria. For instance, 80% of the CF subjects recruited in this study were treated with different antibiotics during the study period. In addition, almost all of the CF subjects were treated with bronchodilators, while some were also treated with inhaled corticosteroids. Selecting any of these factors as exclusion criteria would rarely provide sufficient numbers of CF subjects to study as the maintenance of CF is a lifelong process and subjects often undergo extensive treatments due to their health status. However, it is recommended that future research is conducted focusing only on these factors (e.g. medications, diet, and any other health complications in CF subjects) which could potentially impact the breath VOC profile of CF subjects. Importantly, any such study will require a significant number of participants. For instance, if only antibiotics are considered as a discriminatory factor, at least a certain number of participants with CF are required who are treated with different antibiotics such as Azithromycin, Flucloxacillin, Bactrim, Colomycin, Tobramycin, Minocycline, Ciprofloxacin, Rifaximin, and Ceftazidime. A much simpler and initial approach would be the study of the impact of these antibiotics on *Pseudomonas Aeruginosa* (or other) cultures to observe any change of VOCs released from these cultures due to the use of different antibiotics.

REFERENCES

REFERENCES

ACFDR 2012. 15th ANNUAL REPORT AUSTRALIAN CYSTIC FIBROSIS DATA REGISTRY, Link: https://shop.cysticfibrosis.org.au/media/wysiwyg/CF-Australia/medical-documents/ACFDR_2012/ACFDR_Annual_Report_2012r.pdf.

ACFDR 2016. AUSTRALIAN CYSTIC FIBROSIS DATA REGISTRY ANNUAL REPORT 2016, Link: https://www.cysticfibrosis.org.au/getmedia/a3b28200-caeb-4c5a-ad15-98c71a8c7dc8/ACFDR-2016-Annual-Report-Final-Copy-Single-Page-Version.pdf.aspx.

AL-ALOUL, M., MILLER, H., ALAPATI, S., STOCKTON, P., LEDSON, M. & WALSHAW, M. 2005. Renal impairment in cystic fibrosis patients due to repeated intravenous aminoglycoside use. *Pediatric pulmonology,* 39**,** 15-20.

ALEIX, M. M. & AVINASH, C. K. 2001. PCA versus LDA. *IEEE Transactions on PAMI,* 23**,** 228-233.

AMANN, A., MIEKISCH, W., PLEIL, J., RISBY, T. & SCHUBERT, J. 2010. Methodological issues of sample collection and analysisofexhaledbreath. *European Respiratory Monograph 49: Exhaled Biomarkers*.

AMANN, A., POUPART, G., TELSER, S., LEDOCHOWSKI, M., SCHMID, A. & MECHTCHERIAKOV, S. 2004. Applications of breath gas analysis in medicine. *International Journal of Mass Spectrometry,* 239**,** 227-233.

AMANN, A. & SMITH, D. 2005. *Breath Analysis for Clinical Diagnosis and Therapeutic Monitoring:(With CD-ROM)*, World Scientific.

AMANN, A., SPANEL, P. & SMITH, D. 2007. Breath analysis: the approach towards clinical applications. *Mini reviews in medicinal chemistry,* 7**,** 115-129.

ARIS, R. M., MERKEL, P. A., BACHRACH, L. K., BOROWITZ, D. S., BOYLE, M. P., ELKIN, S. L., GUISE, T. A., HARDIN, D. S., HAWORTH, C. S., HOLICK, M. F., JOSEPH, P. M., O'BRIEN, K., TULLIS, E., WATTS, N. B. & WHITE, T. B. 2005. Guide to Bone Health and Disease in Cystic Fibrosis. *The Journal of Clinical Endocrinology & Metabolism,* 90**,** 1888-1896.

ARSLAN, F. N., KOLK, A. & JANSSEN, H. G. 2019. Methods for one–and two–dimensional gas chromatography with flame ionization detection for

identification of Mycobacterium tuberculosis in sputum. *Journal of Chromatography B,* 1124**,** 204-217.

ARTHUR, C. L. & PAWLISZYN, J. 1990. Solid phase microextraction with thermal desorption using fused silica optical fibers. *Analytical chemistry,* 62**,** 2145-2148.

BALINT, B., KHARITONOV, S., HANAZAWA, T., DONNELLY, L., SHAH, P., HODSON, M. & BARNES, P. 2001. Increased nitrotyrosine in exhaled breath condensate in cystic fibrosis. *European Respiratory Journal,* 17**,** 1201-1207.

BAPTISTA, I., SANTOS, M., RUDNITSKAYA, A., SARAIVA, J. A., ALMEIDA, A. & ROCHA, S. M. 2019. A comprehensive look into the volatile exometabolome of enteroxic and non-enterotoxic Staphylococcus aureus strains. *The international journal of biochemistry & cell biology,* 108**,** 40-50.

BARATLOO, A., HOSSEINI, M., NEGIDA, A. & EL ASHAL, G. 2015. Part 1: simple definition and calculation of accuracy, sensitivity and specificity.

BARBEN, J. U., DITCHFIELD, M., CARLIN, J. B., ROBERTSON, C. F., ROBINSON, P. J. & OLINSKY, A. 2003. Major haemoptysis in children with cystic fibrosis: a 20-year retrospective study. *Journal of Cystic Fibrosis,* 2**,** 105-111.

BARKER, M., HENGST, M., SCHMID, J., BUERS, H.-J., MITTERMAIER, B., KLEMP, D. & KOPPMANN, R. 2006. Volatile organic compounds in the exhaled breath of young patients with cystic fibrosis. *European Respiratory Journal,* 27**,** 929-936.

BEAN, H. D., DIMANDJA, J.-M. D. & HILL, J. E. 2012. Bacterial volatile discovery using solid phase microextraction and comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry. *Journal of Chromatography B,* 901**,** 41-46.

BEAN, H. D., HILL, J. E. & DIMANDJA, J.-M. D. 2015. Improving the quality of biomarker candidates in untargeted metabolomics via peak table-based alignment of comprehensive two-dimensional gas chromatography–mass spectrometry data. *Journal of Chromatography A,* 1394**,** 111-117.

BEAN, H. D., REES, C. A. & HILL, J. E. 2016. Comparative analysis of the volatile metabolomes of Pseudomonas aeruginosa clinical isolates. *Journal of breath research,* 10**,** 047102.

BEAN, H. D., ZHU, J., SENGLE, J. C. & HILL, J. E. 2014. Identifying methicillin-resistant Staphylococcus aureus (MRSA) lung infections in mice via breath

analysis using secondary electrospray ionization-mass spectrometry (SESI-MS). *Journal of breath research,* 8**,** 041001.

BEAUCHAMP, J., HERBIG, J., GUTMANN, R. & HANSEL, A. 2008. On the use of Tedlar® bags for breath-gas sampling and analysis. *Journal of breath research,* 2**,** 046001.

BECCARIA, M., BOBAK, C., MAITSHOTLO, B., MELLORS, T. R., PURCARO, G., FRANCHINA, F. A., REES, C. A., NASIR, M., BLACK, A. & HILL, J. E. 2018a. Exhaled human breath analysis in active pulmonary tuberculosis diagnostics by comprehensive gas chromatography-mass spectrometry and chemometric techniques. *J Breath Res,* 13**,** 016005.

BECCARIA, M., BOBAK, C., MAITSHOTLO, B., MELLORS, T. R., PURCARO, G., FRANCHINA, F. A., REES, C. A., NASIR, M., BLACK, A. & HILL, J. E. 2018b. Exhaled human breath analysis in active pulmonary tuberculosis diagnostics by comprehensive gas chromatography-mass spectrometry and chemometric techniques. *Journal of breath research,* 13**,** 016005.

BECCARIA, M., MELLORS, T. R., PETION, J. S., REES, C. A., NASIR, M., SYSTROM, H. K., SAIRISTIL, J. W., JEAN-JUSTE, M.-A., RIVERA, V. & LAVOILE, K. 2018c. Preliminary investigation of human exhaled breath for tuberculosis diagnosis by multidimensional gas chromatography–Time of flight mass spectrometry and machine learning. *Journal of Chromatography B,* 1074**,** 46-50.

BECKSTROM, A. C., HUMSTON, E. M., SNYDER, L. R., SYNOVEC, R. E. & JUUL, S. E. 2011. Application of comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry method to identify potential biomarkers of perinatal asphyxia in a non-human primate model. *Journal of chromatography A,* 1218**,** 1899-1906.

BILTON, D. 2008. Cystic fibrosis. *Medicine,* 36**,** 273-278.

BLANCA, M. J., ALARCÓN, R., ARNAU, J., BONO, R. & BENDAYAN, R. 2017. Non-normal data: Is ANOVA still a valid option? *Psicothema,* 29**,** 552-557.

BLAU, H., LINNANE, B., CARZINO, R., TANNENBAUM, E.-L., SKORIC, B., ROBINSON, P. J., ROBERTSON, C. & RANGANATHAN, S. C. 2014. Induced sputum compared to bronchoalveolar lavage in young, non-expectorating cystic fibrosis children. *Journal of Cystic Fibrosis,* 13**,** 106-110.

BLONDEAU, K., DUPONT, L., MERTENS, V., VERLEDEN, G., MALFROOT, A., VANDENPLAS, Y., HAUSER, B. & SIFRIM, D. 2008. Gastro-oesophageal reflux and aspiration of gastric contents in adult patients with cystic fibrosis. *Gut,* 57**,** 1049-1055.

BOOTS, A. W., VAN BERKEL, J. J., DALLINGA, J. W., SMOLINSKA, A., WOUTERS, E. F. & VAN SCHOOTEN, F. J. 2012. The versatile use of exhaled volatile organic compounds in human health and disease. *Journal of breath research,* 6**,** 027108.

BOROWITZ, D., BAKER, R. D. & STALLINGS, V. 2002. Consensus report on nutrition for pediatric patients with cystic fibrosis. *Journal of pediatric gastroenterology and nutrition,* 35**,** 246-259.

BOS, L. D., STERK, P. J. & SCHULTZ, M. J. 2013a. Volatile metabolites of pathogens: a systematic review. *PLoS Pathog,* 9**,** e1003311.

BOS, L. D., STERK, P. J. & SCHULTZ, M. J. 2013b. Volatile metabolites of pathogens: a systematic review. *PLoS pathogens,* 9**,** e1003311.

BOUCHER, R. C. 2007. Airway surface dehydration in cystic fibrosis: pathogenesis and therapy. *Annu. Rev. Med.,* 58**,** 157-170.

BRIARD, B., HEDDERGOTT, C. & LATGÉ, J.-P. 2016. Volatile compounds emitted by Pseudomonas aeruginosa stimulate growth of the fungal pathogen Aspergillus fumigatus. *MBio,* 7**,** e00219-16.

BUNGE, M., ARAGHIPOUR, N., MIKOVINY, T., DUNKL, J., SCHNITZHOFER, R., HANSEL, A., SCHINNER, F., WISTHALER, A., MARGESIN, R. & MÄRK, T. D. 2008. On-line monitoring of microbial volatile metabolites by proton transfer reaction-mass spectrometry. *Applied and environmental microbiology,* 74**,** 2179-2186.

BUSZEWSKI, B., KĘSY, M., LIGOR, T. & AMANN, A. 2007. Human exhaled air analytics: biomarkers of diseases. *Biomedical chromatography,* 21**,** 553-566.

CALDEIRA, M., PERESTRELO, R., BARROS, A., BILELO, M., MORETE, A., CAMARA, J. & ROCHA, S. 2012. Allergic asthma exhaled breath metabolome: a challenge for comprehensive two-dimensional gas chromatography. *Journal of Chromatography A,* 1254**,** 87-97.

CARPAGNANO, G. E., BARNES, P. J., FRANCIS, J., WILSON, N., BUSH, A. & KHARITONOV, S. A. 2004. Breath condensate pH in children with cystic

fibrosis and asthma: a new noninvasive marker of airway inflammation? *CHEST Journal,* 125**,** 2005-2010.

CARROLL, W., LENNEY, W., WANG, T., ŠPANĚL, P., ALCOCK, A. & SMITH, D. 2005. Detection of volatile compounds emitted by Pseudomonas aeruginosa using selected ion flow tube mass spectrometry. *Pediatric pulmonology,* 39**,** 452-456.

CARVALHO, B. M. C. 2014. In vitro screening of the anticancer activity of marine and soil-derived fungi extracts and compounds used alone and combined with doxorubicin: evaluation of the anticancer properties of fungi extracts and compounds alone and in combination with doxorubicin in lung cancer cells.

CFFA. 2016. *Cystic Fibrosis Federation Australia, https://www.cysticfibrosis.org.au/* [Online]. [Accessed 2020].

CHAN, L. W., ANAHTAR, M. N., ONG, T.-H., HERN, K. E., KUNZ, R. R. & BHATIA, S. N. 2020. Engineering synthetic breath biomarkers for respiratory disease. *Nature Nanotechnology***,** 1-9.

CHAPARRO, C., MAURER, J., GUTIERREZ, C., KRAJDEN, M., CHAN, C., WINTON, T., KESHAVJEE, S., SCAVUZZO, M., TULLIS, E. & HUTCHEON, M. 2001. Infection with Burkholderia cepacia in cystic fibrosis: outcome following lung transplantation. *American journal of respiratory and critical care medicine,* 163**,** 43-48.

CHIPPENDALE, T. W., GILCHRIST, F. J., ŠPANĚL, P., ALCOCK, A., LENNEY, W. & SMITH, D. 2014. Quantification by SIFT-MS of volatile compounds emitted by in vitro cultures of S. aureus, S. pneumoniae and H. influenzae isolated from patients with respiratory diseases. *Analytical Methods,* 6**,** 2460-2472.

CLEMENT, A., TAMALET, A., LEROUX, E., RAVILLY, S., FAUROUX, B. & JAIS, J.-P. 2006. Long term effects of azithromycin in patients with cystic fibrosis: a double blind, placebo controlled trial. *Thorax,* 61**,** 895-902.

COLLINS, F. S. 1992. Cystic fibrosis: molecular biology and therapeutic implications. *Science,* 256**,** 774.

COLOMBO, C., BATTEZZATI, P. M., CROSIGNANI, A., MORABITO, A., COSTANTINI, D., PADOAN, R. & GIUNTA, A. 2002. Liver disease in cystic fibrosis: a prospective study on incidence, risk factors, and outcome. *Hepatology,* 36**,** 1374-1382.

CONWAY, S., MORTON, A., OLDROYD, B., TRUSCOTT, J., WHITE, H., SMITH, A. & HAIGH, I. 2000. Osteoporosis and osteopenia in adults and adolescents with cystic fibrosis: prevalence and associated factors. *Thorax,* 55**,** 798-804.

DADAMIO, J., VAN DEN VELDE, S., LALEMAN, W., VAN HEE, P., COUCKE, W., NEVENS, F. & QUIRYNEN, M. 2012. Breath biomarkers of liver cirrhosis. *Journal of Chromatography B,* 905**,** 17-22.

DALLÜGE, J., BEENS, J. & UDO, A. 2003. Comprehensive two-dimensional gas chromatography: a powerful and versatile analytical tool. *Journal of Chromatography A,* 1000**,** 69-108.

DALLÜGE, J., VREULS, R. J., BEENS, J. & BRINKMAN, U. A. T. 2002. Optimization and characterization of comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection (GC× GC–TOF MS). *Journal of Separation Science,* 25**,** 201-214.

DASENBROOK, E. C., MERLO, C. A., DIENER-WEST, M., LECHTZIN, N. & BOYLE, M. P. 2008. Persistent methicillin-resistant Staphylococcus aureus and rate of FEV1 decline in cystic fibrosis. *American journal of respiratory and critical care medicine,* 178**,** 814-821.

DAVIDSON, D. J., DORIN, J. R., MCLACHLAN, G., RANALDI, V., LAMB, D., DOHERTY, C., GOVAN, J. & PORTEOUS, D. J. 1995. Lung disease in the cystic fibrosis mouse exposed to bacterial pathogens. *Nature genetics,* 9**,** 351-357.

DE BOECK, K., MALFROOT, A., VAN SCHIL, L., LEBECQUE, P., KNOOP, C., GOVAN, J., DOHERTY, C., LAEVENS, S. & VANDAMME, P. 2004. Epidemiology of Burkholderia cepacia complex colonisation in cystic fibrosis patients. *European Respiratory Journal,* 23**,** 851-856.

DE JONG, P., NAKANO, Y., LEQUIN, M., MAYO, J., WOODS, R., PARE, P. & TIDDENS, H. 2004. Progressive damage on high resolution computed tomography despite stable lung function in cystic fibrosis. *European Respiratory Journal,* 23**,** 93-97.

DENG, C., ZHANG, J., YU, X., ZHANG, W. & ZHANG, X. 2004a. Determination of acetone in human breath by gas chromatography–mass spectrometry and solid-phase microextraction with on-fiber derivatization. *Journal of Chromatography B,* 810**,** 269-275.

DENG, C., ZHANG, X. & LI, N. 2004b. Investigation of volatile biomarkers in lung cancer blood using solid-phase microextraction and capillary gas

chromatography–mass spectrometry. *Journal of Chromatography B,* 808**,** 269-277.

DOLCH, M., FREY, L., HORNUSS, C., SCHMOELZ, M., PRAUN, S., VILLINGER, J. & SCHELLING, G. 2008. Molecular breath-gas analysis by online mass spectrometry in mechanically ventilated patients: a new software-based method of CO2-controlled alveolar gas monitoring. *Journal of breath research,* 2**,** 037010.

DÖRING, G. & CONWAY, S. P. 2008. Osteoporosis in cystic fibrosis. *Jornal de pediatria,* 84**,** 1-3.

DRYAHINA, K., SOVOVÁ, K., NEMEC, A. & ŠPANĚL, P. 2016. Differentiation of pulmonary bacterial pathogens in cystic fibrosis by volatile metabolites emitted by their in vitro cultures: Pseudomonas aeruginosa, Staphylococcus aureus, Stenotrophomonas maltophilia and the Burkholderia cepacia complex. *Journal of breath research,* 10**,** 037102.

EDWARDS, M., MOSTAFA, A. & GÓRECKI, T. 2011. Modulation in comprehensive two-dimensional gas chromatography: 20 years of innovation. *Analytical and bioanalytical chemistry,* 401**,** 2335-2349.

EFRATI, O., BARAK, A., MODAN-MOSES, D., AUGARTEN, A., VILOZNI, D., KATZNELSON, D., SZEINBERG, A., YAHAV, J. & BUJANOVER, Y. 2003. Liver cirrhosis and portal hypertension in cystic fibrosis. *European journal of gastroenterology & hepatology,* 15**,** 1073-1078.

ELMASSRY, M. M. & PIECHULLA, B. 2020. Volatilomes of bacterial infections in humans. *Frontiers in Neuroscience,* 14.

EMMONS, R. V., TAJALI, R. & GIONFRIDDO, E. 2019. Development, optimization and applications of thin film solid phase microextraction (TF-SPME) devices for thermal desorption: A comprehensive review. *Separations,* 6**,** 39.

ENDERBY, B., SMITH, D., CARROLL, W. & LENNEY, W. 2009. Hydrogen cyanide as a biomarker for Pseudomonas aeruginosa in the breath of children with cystic fibrosis. *Pediatric pulmonology,* 44**,** 142-147.

EQUI, A., BALFOUR-LYNN, I., BUSH, A. & ROSENTHAL, M. 2002. Long term azithromycin in children with cystic fibrosis: a randomised, placebo-controlled crossover trial. *The Lancet,* 360**,** 978-984.

FILIPIAK, W., BEER, R., SPONRING, A., FILIPIAK, A., AGER, C., SCHIEFECKER, A., LANTHALER, S., HELBOK, R., NAGL, M. & TROPPMAIR, J. 2015.

Breath analysis for in vivo detection of pathogens related to ventilator-associated pneumonia in intensive care patients: a prospective pilot study. *Journal of breath research,* 9**,** 016004.

FILIPIAK, W., SPONRING, A., BAUR, M. M., FILIPIAK, A., AGER, C., WIESENHOFER, H., NAGL, M., TROPPMAIR, J. & AMANN, A. 2012. Molecular analysis of volatile metabolites released specifically by Staphylococcus aureus and Pseudomonas aeruginosa. *BMC microbiology,* 12**,** 1.

FILIPIAK, W., SPONRING, A., FILIPIAK, A., BAUR, M., AGER, C., WIESENHOFER, H., MARGESIN, R., NAGL, M., TROPPMAIR, J. & AMANN, A. 2013. Volatile Organic Compounds (VOCs) Released by Pathogenic Microorganisms in vitro: Potential Breath Biomarkers. *Volatile Biomarkers: Non-Invasive Diagnosis in Physiology and Medicine***,** 463.

FINAMORE, P., SCARLATA, S. & INCALZI, R. A. 2019. Breath analysis in respiratory diseases: State-of-the-art and future perspectives. *Expert review of molecular diagnostics,* 19**,** 47-61.

FLUME, P. A., STRANGE, C., YE, X., EBELING, M., HULSEY, T. & CLARK, L. L. 2005a. Pneumothorax in cystic fibrosis. *CHEST Journal,* 128**,** 720-728.

FLUME, P. A., YANKASKAS, J. R., EBELING, M., HULSEY, T. & CLARK, L. L. 2005b. Massive hemoptysis in cystic fibrosis. *CHEST Journal,* 128**,** 729-738.

FORBES, S., TROOBNIKOFF, A., UELAND, M., NIZIO, K. & PERRAULT, K. 2016. Profiling the decomposition odour at the grave surface before and after probing. *Forensic science international,* 259**,** 193-199.

FRANCHINA, F. A., PURCARO, G., BURKLUND, A., BECCARIA, M. & HILL, J. E. 2019. Evaluation of different adsorbent materials for the untargeted and targeted bacterial VOC analysis using GC× GC-MS. *Analytica chimica acta,* 1066**,** 146-153.

FRYSINGER, G. S. & GAINES, R. B. 2002. Forensic analysis of ignitable liquids in fire debris by comprehensive two-dimensional gas chromatography. *Journal of Forensic Science,* 47**,** 471-482.

FU, X. A., LI, M., KNIPP, R. J., NANTZ, M. H. & BOUSAMRA, M. 2014. Noninvasive detection of lung cancer using exhaled breath. *Cancer medicine,* 3**,** 174-181.

GAISL, T., BREGY, L., STEBLER, N., GAUGG, M. T., BRUDERER, T., GARCÍA-GÓMEZ, D., MOELLER, A., SINGER, F., SCHWARZ, E. I. & BENDEN, C.

2018. Real-time exhaled breath analysis in patients with cystic fibrosis and controls. *Journal of breath research,* 12**,** 036013.

GERRITSEN, M., BRINKMAN, P., ESCOBAR, N., BOS, L., DE HEER, K., MEIJER, M., JANSSEN, H., DE COCK, H., WÖSTEN, H. A. & VISSER, C. 2018. Profiling of volatile organic compounds produced by clinical Aspergillus isolates using gas chromatography–mass spectrometry. *Medical mycology,* 56**,** 253-256.

GILCHRIST, F. J., BRIGHT-THOMAS, R. J., JONES, A. M., SMITH, D., ŠPANĚL, P., WEBB, A. K. & LENNEY, W. 2013. Hydrogen cyanide concentrations in the breath of adult cystic fibrosis patients with and without Pseudomonas aeruginosa infection. *Journal of breath research,* 7**,** 026010.

GOEMINNE, P. C., VANDENDRIESSCHE, T., VAN ELDERE, J., NICOLAI, B. M., HERTOG, M. L. & DUPONT, L. J. 2012. Detection of Pseudomonas aeruginosa in sputum headspace through volatile organic compound analysis. *Respiratory research,* 13**,** 87.

GÓRECKI, T., YU, X. & PAWLISZYN, J. 1999. Theory of analyte extraction by selected porous polymer SPME fibres. *Analyst,* 124**,** 643-649.

GRASEMANN, H., KNAUER, N., BUSCHER, R., HUBNER, K., DRAZEN, J. M. & RATJEN, F. 2000. Airway nitric oxide levels in cystic fibrosis patients are related to a polymorphism in the neuronal nitric oxide synthase gene. *American journal of respiratory and critical care medicine,* 162**,** 2172-2176.

GRASEMANN, H., MICHLER, E., WALLOT, M. & RATJEN, F. 1997. Decreased concentration of exhaled nitric oxide (NO) in patients with cystic fibrosis. *Pediatric pulmonology,* 24**,** 173-177.

HAHN, A., WHITESON, K., DAVIS, T. J., PHAN, J., SAMI, I., KOUMBOURLIS, A. C., FREISHTAT, R. J., CRANDALL, K. A. & BEAN, H. D. 2020. Longitudinal Associations of the Cystic Fibrosis Airway Microbiome and Volatile Metabolites: A Case Study. *Frontiers in Cellular and Infection Microbiology,* 10**,** 174.

HAMEED, S., MORTON, J. R., JAFFÉ, A., FIELD, P. I., BELESSIS, Y., YOONG, T., KATZ, T. & VERGE, C. F. 2010. Early glucose abnormalities in cystic fibrosis are preceded by poor weight gain. *Diabetes care,* 33**,** 221-226.

HARSHMAN, S. W., MANI, N., GEIER, B. A., KWAK, J., SHEPARD, P., FAN, M., SUDBERRY, G. L., MAYES, R. S., OTT, D. K. & MARTIN, J. A. 2016. Storage stability of exhaled breath on Tenax TA. *Journal of breath research,* 10**,** 046008.

HEIJERMAN, H. 2008. Cystic Fibrosis. Third Edition. *European Respiratory Journal,* 31**,** 482-482.

HELMIG, D. & VIERLING, L. 1995. Water adsorption capacity of the solid adsorbents Tenax TA, Tenax GR, Carbotrap, Carbotrap C, Carbosieve SIII, and Carboxen 569 and water management techniques for the atmospheric sampling of volatile organic trace gases. *Analytical Chemistry,* 67**,** 4380-4386.

HØIBY, N., CIOFU, O. & BJARNSHOLT, T. 2010. Pseudomonas aeruginosa biofilms in cystic fibrosis. *Future microbiology,* 5**,** 1663-1674.

HORVÁTH, I., DONNELLY, L. E., KISS, A., KHARITONOV, S. A., LIM, S., FAN CHUNG, K. & BARNES, P. J. 1998a. Combined use of exhaled hydrogen peroxide and nitric oxide in monitoring asthma. *American journal of respiratory and critical care medicine,* 158**,** 1042-1046.

HORVÁTH, I., DONNELLY, L. E., KISS, A., PAREDI, P., KHARITONOV, S. A. & BARNES, P. J. 1998b. Raised levels of exhaled carbon monoxide are associated with an increased expression of heme oxygenase-1 in airway macrophages in asthma: a new marker of oxidative stress. *Thorax,* 53**,** 668-672.

IQBAL, M. A. & KIM, K.-H. 2014. Generation of sub-ppb level vapor phase mixtures of biogenic volatile organic compounds from liquid phase standards and stepwise characterization of their volatilization properties by thermal desorption–gas chromatography–mass spectrometry. *Journal of Chromatography A,* 1373**,** 149-158.

IQBAL, M. A., KIM, K.-H. & AHN, J. H. 2014a. Monoterpenes Released from Fruit, Plant, and Vegetable Systems. *Sensors,* 14**,** 18286-18301.

IQBAL, M. A., KIM, K.-H., SZULEJKO, J. & CHO, J. 2014b. An assessment of the liquid–gas partitioning behavior of major wastewater odorants using two comparative experimental approaches: liquid sample-based vaporization vs. impinger-based dynamic headspace extraction into sorbent tubes. *Analytical and Bioanalytical Chemistry,* 406**,** 643-655.

ISLES, A., MACLUSKY, I., COREY, M., GOLD, R., PROBER, C., FLEMING, P. & LEVISON, H. 1984. Pseudomonas cepacia infection in cystic fibrosis: an emerging problem. *The Journal of pediatrics,* 104**,** 206-210.

JAREÑO-ESTEBAN, J. J., MUÑOZ-LUCAS, M. Á., CARRILLO-ARANDA, B., MALDONADO-SANZ, J. Á., DE GRANDA-ORIVE, I., AGUILAR-ROS, A., CIVERA-TEJUCA, C., GUTIÉRREZ-ORTEGA, C. & CALLOL-SÁNCHEZ, L.

M. 2013. Volatile organic compounds in exhaled breath in a healthy population: effect of tobacco smoking. *Archivos de Bronconeumología (English Edition),* 49**,** 457-461.

JENKINS, C. L. & BEAN, H. D. 2019. Influence of media on the differentiation of Staphylococcus spp. by volatile compounds. *Journal of Breath Research,* 14**,** 016007.

JOHANSSON, M. E., SJÖVALL, H. & HANSSON, G. C. 2013. The gastrointestinal mucus system in health and disease. *Nature Reviews Gastroenterology and Hepatology,* 10**,** 352-361.

JØRGENSEN, K. M., WASSERMANN, T., JOHANSEN, H. K., CHRISTIANSEN, L. E., MOLIN, S., HØIBY, N. & CIOFU, O. 2015. Diversity of metabolic profiles of cystic fibrosis Pseudomonas aeruginosa during the early stages of lung infection. *Microbiology,* 161**,** 1447-1462.

JÜNGER, M., VAUTZ, W., KUHNS, M., HOFMANN, L., ULBRICHT, S., BAUMBACH, J. I., QUINTEL, M. & PERL, T. 2012. Ion mobility spectrometry for microbial volatile organic compounds: a new identification tool for human pathogenic bacteria. *Applied microbiology and biotechnology,* 93**,** 2603-2614.

KÄLIN, N., CLAAß, A., SOMMER, M., PUCHELLE, E. & TÜMMLER, B. 1999. ΔF508 CFTR protein expression in tissues from patients with cystic fibrosis. *The Journal of clinical investigation,* 103**,** 1379-1389.

KAMBOURES, M., BLAKE, D., COOPER, D., NEWCOMB, R., BARKER, M., LARSON, J., MEINARDI, S., NUSSBAUM, E. & ROWLAND, F. 2005. Breath sulfides and pulmonary function in cystic fibrosis. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 15762-15767.

KARAMI, N., MIRZAJANI, F., REZADOOST, H., KARIMI, A., FALLAH, F., GHASSEMPOUR, A. & ALIAHMADI, A. 2017. Initial study of three different pathogenic microorganisms by gas chromatography-mass spectrometry. *F1000Research,* 6.

KARL, T., PRAZELLER, P., MAYR, D., JORDAN, A., RIEDER, J., FALL, R. & LINDINGER, W. 2001. Human breath isoprene and its relation to blood cholesterol levels: new measurements and modeling. *Journal of Applied Physiology,* 91**,** 762-770.

KHAN, T. Z., WAGENER, J. S., BOST, T., MARTINEZ, J., ACCURSO, F. J. & RICHES, D. 1995. Early pulmonary inflammation in infants with cystic fibrosis. *American journal of respiratory and critical care medicine,* 151**,** 1075-1082.

KIM, K.-H., JAHAN, S. A. & KABIR, E. 2012. A review of breath analysis for diagnosis of human health. *TrAC Trends in Analytical Chemistry,* 33**,** 1-8.

KING, J., MOCHALSKI, P., KUPFERTHALER, A., UNTERKOFLER, K., KOC, H., FILIPIAK, W., TESCHL, S., HINTERHUBER, H. & AMANN, A. 2010. Dynamic profiles of volatile organic compounds in exhaled breath as determined by a coupled PTR-MS/GC-MS study. *Physiological measurement,* 31**,** 1169.

KNOBEL, Z., UELAND, M., NIZIO, K. D., PATEL, D. & FORBES, S. L. 2018. A comparison of human and pig decomposition rates and odour profiles in an Australian environment. *Australian Journal of Forensic Sciences***,** 1-16.

KOEHLER, T., ACKERMANN, I., BRECHT, D., UTESCHIL, F., WINGENDER, J., TELGHEDER, U. & SCHMITZ, O. J. 2020. Analysis of volatile metabolites from in vitro biofilms of Pseudomonas aeruginosa with thin-film microextraction by thermal desorption gas chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry***,** 1-12.

KOLK, A., HOELSCHER, M., MABOKO, L., JUNG, J., KUIJPER, S., CAUCHI, M., BESSANT, C., VAN BEERS, S., DUTTA, R. & GIBSON, T. 2010. Electronic-nose technology using sputum samples in diagnosis of patients with tuberculosis. *Journal of Clinical Microbiology,* 48**,** 4235-4238.

KRAMER, R., SAUER-HEILBORN, A., WELTE, T., GUZMAN, C., HÖFLE, M. & ABRAHAM, W.-R. 2015. A rapid method for breath analysis in cystic fibrosis patients. *European Journal of Clinical Microbiology & Infectious Diseases,* 34**,** 745-751.

KÜNTZEL, A., FISCHER, S., BERGMANN, A., OERTEL, P., STEFFENS, M., TREFZ, P., MIEKISCH, W., SCHUBERT, J. K., REINHOLD, P. & KÖHLER, H. 2016. Effects of biological and methodological factors on volatile organic compound patterns during cultural growth of Mycobacterium avium ssp. paratuberculosis. *Journal of breath research,* 10**,** 037103.

LABOWS, J. N., MCGINLEY, K. J., WEBSTER, G. & LEYDEN, J. 1980. Headspace analysis of volatile metabolites of Pseudomonas aeruginosa and related species by gas chromatography-mass spectrometry. *Journal of Clinical Microbiology,* 12**,** 521-526.

LANCELLOTTI, L., D'ORAZIO, C., MASTELLA, G., MAZZI, G. & LIPPI, U. 1996. Deficiency of vitamins E and A in cystic fibrosis is independent of pancreatic function and current enzyme and vitamin supplementation. *European journal of pediatrics,* 155**,** 281-285.

LANNEFORS, L., BUTTON, B. M. & MCILWAINE, M. 2004. Physiotherapy in infants and young children with cystic fibrosis: current practice and future developments. *Journal of the royal society of medicine,* 97**,** 8.

LAWAL, O., AHMED, W. M., NIJSEN, T. M., GOODACRE, R. & FOWLER, S. J. 2017. Exhaled breath analysis: a review of 'breath-taking'methods for off-line analysis. *Metabolomics,* 13**,** 110.

LAWAL, O., KNOBEL, H., WEDA, H., BOS, L. D., NIJSEN, T. M., GOODACRE, R. & FOWLER, S. J. 2018a. Volatile organic compound signature from co-culture of lung epithelial cell line with Pseudomonas aeruginosa. *Analyst,* 143**,** 3148-3155.

LAWAL, O., KNOBEL, H., WEDA, H., NIJSEN, T. M. E., GOODACRE, R., FOWLER, S. J., AHMED, W. M., ARTIGAS, A., BANNARD-SMITH, J., BOS, L. D. J., CAMPRUBI, M., COELHO, L., DARK, P., DAVIE, A., DIAZ, E., GOMA, G., FELTON, T., FOWLER, S. J., GOODACRE, R., KNOBEL, H., LAWAL, O., LEOPOLD, J.-H., NIJSEN, T. M. E., VAN OORT, P. M. P., POVOA, P., PORTSMOUTH, C., RATTRAY, N. J. W., RIJNDERS, G., SCHULTZ, M. J., STEENWELLE, R., STERK, P. J., VALLES, J., VERHOECKX, F., VINK, A., WEDA, H., WHITE, I. R., WINTERS, T., ZAKHARKINA, T. & THE BREATHDX, C. 2018b. TD/GC–MS analysis of volatile markers emitted from mono- and co-cultures of Enterobacter cloacae and Pseudomonas aeruginosa in artificial sputum. *Metabolomics,* 14**,** 66.

LAWAL, O., MUHAMADALI, H., AHMED, W. M., WHITE, I. R., NIJSEN, T. M., GOODACRE, R. & FOWLER, S. J. 2018c. Headspace volatile organic compounds from bacteria implicated in ventilator-associated pneumonia analysed by TD-GC/MS. *Journal of breath research,* 12**,** 026002.

LIBARDONI, M., STEVENS, P., WAITE, J. H. & SACKS, R. 2006. Analysis of human breath samples with a multi-bed sorption trap and comprehensive two-dimensional gas chromatography (GC× GC). *Journal of Chromatography B,* 842**,** 13-21.

LIPUMA, J. J. 2005. Update on the Burkholderia cepacia complex. *Current opinion in pulmonary medicine,* 11**,** 528-533.

LIPUMA, J. J. 2010. The changing microbial epidemiology in cystic fibrosis. *Clinical microbiology reviews,* 23**,** 299-323.

LOURENÇO, C. & TURNER, C. 2014. Breath analysis in disease diagnosis: methodological considerations and applications. *Metabolites,* 4**,** 465-498.

LYCZAK, J. B., CANNON, C. L. & PIER, G. B. 2002. Lung infections associated with cystic fibrosis. *Clinical microbiology reviews,* 15**,** 194-222.

MACKENZIE, T., GIFFORD, A. H., SABADOSA, K. A., QUINTON, H. B., KNAPP, E. A., GOSS, C. H. & MARSHALL, B. C. 2014. Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the cystic fibrosis foundation patient registry. *Annals of internal medicine,* 161**,** 233-241.

MACKIE, A., THORNTON, S. & EDENBOROUGH, F. 2003. Cystic fibrosis-related diabetes. *Diabetic medicine,* 20**,** 425-436.

MAISONNEUVE, P., FITZSIMMONS, S. C., NEGLIA, J. P., CAMPBELL, P. W. & LOWENFELS, A. B. 2003. Cancer risk in nontransplanted and transplanted cystic fibrosis patients: a 10-year study. *Journal of the National Cancer Institute,* 95**,** 381-387.

MAISONNEUVE, P., MARSHALL, B. & LOWENFELS, A. 2007. Risk of pancreatic cancer in patients with cystic fibrosis. *Gut,* 56**,** 1327-1328.

MARCO, E. & GRIMALT, J. O. 2015. A rapid method for the chromatographic analysis of volatile organic compounds in exhaled breath of tobacco cigarette and electronic cigarette smokers. *Journal of Chromatography A,* 1410**,** 51-59.

MARIA, P. 1996. Evaluation of Anasorb CMS and comparison with Tenax TA for the sampling of volatile organic compounds in indoor and outdoor air by breakthrough measurements. *Analyst,* 121**,** 303-307.

MARRIOTT, P. & SHELLIE, R. 2002. Principles and applications of comprehensive two-dimensional gas chromatography. *TrAC Trends in Analytical Chemistry,* 21**,** 573-583.

MASSIE, J., CURNOW, L., GAFFNEY, L., CARLIN, J. & FRANCIS, I. 2010. Declining prevalence of cystic fibrosis since the introduction of newborn screening. *Archives of disease in childhood***,** archdischild172916.

MATSUI, H., GRUBB, B. R., TARRAN, R., RANDELL, S. H., GATZY, J. T., DAVIS, C. W. & BOUCHER, R. C. 1998. Evidence for periciliary liquid layer depletion,

not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. *Cell,* 95**,** 1005-1015.

MCCAFFREY, C. A., MACLACHLAN, J. & BROOKES, B. I. 1994. Adsorbent tube evaluation for the preconcentration of volatile organic compounds in air for analysis by gas chromatography–mass spectrometry. *Analyst,* 119**,** 897-902.

MCKEON, D., DAY, A., PARMAR, J., ALEXANDER, G. & BILTON, D. 2004. Hepatocellular carcinoma in association with cirrhosis in a patient with cystic fibrosis. *Journal of Cystic Fibrosis,* 3**,** 193-195.

MELLORS, T. R. 2018. *Using breath to diagnose and monitor M. tuberculosis infection A Thesis Submitted to the Faculty.* Dartmouth College Hanover, New Hampshire.

MIEKISCH, W., KISCHKEL, S., SAWACKI, A., LIEBAU, T., MIETH, M. & SCHUBERT, J. K. 2008. Impact of sampling procedures on the results of breath analysis. *Journal of breath research,* 2**,** 026007.

MIEKISCH, W. & SCHUBERT, J. K. 2006. From highly sophisticated analytical techniques to life-saving diagnostics: Technical developments in breath analysis. *TrAC Trends in Analytical Chemistry,* 25**,** 665-673.

MIEKISCH, W., SCHUBERT, J. K. & NOELDGE-SCHOMBURG, G. F. 2004. Diagnostic potential of breath analysis—focus on volatile organic compounds. *Clinica chimica acta,* 347**,** 25-39.

MILLONIG, G., PRAUN, S., NETZER, M., BAUMGARTNER, C., DORNAUER, A., MUELLER, S., VILLINGER, J. & VOGEL, W. 2010. Non-invasive diagnosis of liver diseases by breath analysis using an optimized ion–molecule reaction-mass spectrometry approach: a pilot study. *Biomarkers,* 15**,** 297-306.

MOCHALSKI, P., KING, J., UNTERKOFLER, K. & AMANN, A. 2013. Stability of selected volatile breath constituents in Tedlar, Kynar and Flexfilm sampling bags. *Analyst,* 138**,** 1405-1418.

MOCHALSKI, P., WZOREK, B., ŚLIWKA, I. & AMANN, A. 2009. Suitability of different polymer bags for storage of volatile sulphur compounds relevant to breath analysis. *Journal of Chromatography B,* 877**,** 189-196.

MOHLER, R. E., DOMBEK, K. M., HOGGARD, J. C., PIERCE, K. M., YOUNG, E. T. & SYNOVEC, R. E. 2007. Comprehensive analysis of yeast metabolite GC× GC–TOFMS data: combining discovery-mode and deconvolution chemometric software. *Analyst,* 132**,** 756-767.

MONDELLO, L., TRANCHIDA, P. Q., DUGO, P. & DUGO, G. 2008. Comprehensive two-dimensional gas chromatography-mass spectrometry: A review. *Mass spectrometry reviews,* 27**,** 101-124.

MONTUSCHI, P., KHARITONOV, S. A., CIABATTONI, G., CORRADI, M., VAN RENSEN, L., GEDDES, D. M., HODSON, M. E. & BARNES, P. J. 2000. Exhaled 8-isoprostane as a new non-invasive biomarker of oxidative stress in cystic fibrosis. *Thorax,* 55**,** 205-209.

MONTUSCHI, P., PARIS, D., MELCK, D., LUCIDI, V., CIABATTONI, G., RAIA, V., CALABRESE, C., BUSH, A., BARNES, P. J. & MOTTA, A. 2011. NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis. *Thorax***,** thoraxjnl-2011-200072.

NARASIMHAN, L., GOODMAN, W. & PATEL, C. K. N. 2001. Correlation of breath ammonia with blood urea nitrogen and creatinine during hemodialysis. *Proceedings of the National Academy of Sciences,* 98**,** 4617-4621.

NASIR, M., BEAN, H. D., SMOLINSKA, A., REES, C. A., ZEMANICK, E. T. & HILL, J. E. 2018. Volatile molecules from bronchoalveolar lavage fluid can 'rule-in'Pseudomonas aeruginosa and 'rule-out'Staphylococcus aureus infections in cystic fibrosis patients. *Scientific reports,* 8**,** 1-11.

NEERINCX, A., GEURTS, B., HABETS, M., BOOIJ, J., VAN LOON, J., JANSEN, J., BUYDENS, L., VAN INGEN, J., MOUTON, J. & HARREN, F. 2016a. Identification of Pseudomonas aeruginosa and Aspergillus fumigatus mono-and co-cultures based on volatile biomarker combinations. *Journal of breath research,* 10**,** 016002.

NEERINCX, A., GEURTS, B., VAN LOON, J., TIEMES, V., JANSEN, J., HARREN, F., KLUIJTMANS, L., MERKUS, P., CRISTESCU, S. & BUYDENS, L. 2016b. Detection of Staphylococcus aureus in cystic fibrosis patients using breath VOC profiles. *J Breath Res,* 10**,** 046014.

NEERINCX, A. H., MANDON, J., VAN INGEN, J., ARSLANOV, D. D., MOUTON, J. W., HARREN, F. J., MERKUS, P. J. & CRISTESCU, S. M. 2015. Real-time monitoring of hydrogen cyanide (HCN) and ammonia (NH3) emitted by Pseudomonas aeruginosa. *Journal of breath research,* 9**,** 027102.

NIH. 2016. *US National Library of Medicine, https://ghr.nlm.nih.gov/gene/CFTR* [Online]. [Accessed].

NIXON, P. A., ORENSTEIN, D. M., KELSEY, S. F. & DOERSHUK, C. F. 1992. The prognostic value of exercise testing in patients with cystic fibrosis. *New England Journal of Medicine,* 327**,** 1785-1788.

NIZIO, K. & FORBES, S. 2018. Developing a Method for the Collection and Analysis of Burnt Remains for the Detection and Identification of Ignitable Liquid Residues Using Body Bags, Dynamic Headspace Sampling, and TD-GC× GC-TOFMS. *Separations,* 5**,** 46.

NIZIO, K., PERRAULT, K., TROOBNIKOFF, A., UELAND, M., SHOMA, S., IREDELL, J., MIDDLETON, P. & FORBES, S. 2016. In vitro volatile organic compound profiling using GC× GC-TOFMS to differentiate bacteria associated with lung infections: a proof-of-concept study. *Journal of breath research,* 10**,** 026008.

NIZIO, K. D., MGINITIE, T. M. & HARYNUK, J. J. 2012. Comprehensive multidimensional separations for the analysis of petroleum. *Journal of Chromatography A,* 1255**,** 12-23.

O'SULLIVAN, B. P. & FREEDMAN, S. D. 2009. Cystic fibrosis. *The Lancet,* 373**,** 1891-1904.

OLIVER, A., CANTÓN, R., CAMPO, P., BAQUERO, F. & BLÁZQUEZ, J. 2000. High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. *Science,* 288**,** 1251-1253.

OOI, C. Y., DORFMAN, R., CIPOLLI, M., GONSKA, T., CASTELLANI, C., KEENAN, K., FREEDMAN, S. D., ZIELENSKI, J., BERTHIAUME, Y. & COREY, M. 2011. Type of CFTR mutation determines risk of pancreatitis in patients with cystic fibrosis. *Gastroenterology,* 140**,** 153-161.

OWEN, D. 1965. The power of Student's t-test. *Journal of the American Statistical Association,* 60**,** 320-333.

OWEN, O., TRAPP, V., SKUTCHES, C., MOZZOLI, M., HOELDTKE, R., BODEN, G. & REICHARD, G. 1982. Acetone metabolism during diabetic ketoacidosis. *Diabetes,* 31**,** 242-248.

PANI, O. & GÓRECKI, T. 2006. Comprehensive two-dimensional gas chromatography (GC× GC) in environmental analysis and monitoring. *Analytical and bioanalytical chemistry,* 386**,** 1013-1023.

PANKOW, J. F., LUO, W., MELNYCHENKO, A. N., BARSANTI, K., ISABELLE, L. M., CHEN, C., GUENTHER, A. B. & ROSENSTIEL, T. N. 2012. Volatilizable

Biogenic Organic Compounds (VBOCs) with two dimensional Gas Chromatography-Time of Flight Mass Spectrometry (GC× GC-TOFMS): sampling methods, VBOC complexity, and chromatographic retention data.

PAREDI, P., KHARITONOV, S. A., LEAK, D., SHAH, P. L., CRAMER, D., HODSON, M. E. & BARNES, P. J. 2000. Exhaled ethane is elevated in cystic fibrosis and correlates with carbon monoxide levels and airway obstruction. *American journal of respiratory and critical care medicine,* 161**,** 1247-1251.

PAULING, L., ROBINSON, A. B., TERANISHI, R. & CARY, P. 1971. Quantitative analysis of urine vapor and breath by gas-liquid partition chromatography. *Proceedings of the National Academy of Sciences,* 68**,** 2374-2376.

PAWLISZYN, J. 1999. *Applications of solid phase microextraction*, Royal Society of Chemistry.

PAWLISZYN, J. & PEDERSEN-BJERGAARD, S. 2006. Analytical microextraction: current status and future trends. *Journal of chromatographic science,* 44**,** 291-307.

PEREIRA, J., PORTO-FIGUEIRA, P., CAVACO, C., TAUNK, K., RAPOLE, S., DHAKNE, R., NAGARAJARAM, H. & CÂMARA, J. S. 2015. Breath analysis as a potential and non-invasive frontier in disease diagnosis: an overview. *Metabolites,* 5**,** 3-55.

PERL, T., JÜNGER, M., VAUTZ, W., NOLTE, J., KUHNS, M., BORG-VON ZEPELIN, M. & QUINTEL, M. 2011. Detection of characteristic metabolites of Aspergillus fumigatus and Candida species using ion mobility spectrometry–metabolic profiling by volatile organic compounds. *Mycoses,* 54**,** e828-e837.

PERRAULT, K. A., NIZIO, K. D. & FORBES, S. L. 2015. A comparison of one-dimensional and comprehensive two-dimensional gas chromatography for decomposition odour profiling using inter-year replicate field trials. *Chromatographia,* 78**,** 1057-1070.

PESESSE, R., STEFANUTO, P. H., SCHLEICH, F., LOUIS, R. & FOCANT, J. F. 2019. Multimodal chemometric approach for the analysis of human exhaled breath in lung cancer patients by TD-GC × GC-TOFMS. *Journal of Chromatography B*.

PHILLIPS, C., MAC PARTHALÁIN, N., SYED, Y., DEGANELLO, D., CLAYPOLE, T. & LEWIS, K. 2014. Short-term intra-subject variation in exhaled volatile organic compounds (VOCs) in COPD patients and healthy controls and its effect on disease classification. *Metabolites,* 4**,** 300-318.

PHILLIPS, M. 2008. Breath Test for Oral Malodor. Google Patents.

PHILLIPS, M., BASA-DALAY, V., BLAIS, J., BOTHAMLEY, G., CHATURVEDI, A., MODI, K. D., PANDYA, M., NATIVIDAD, M. P. R., PATEL, U. & RAMRAJE, N. N. 2012. Point-of-care breath test for biomarkers of active pulmonary tuberculosis. *Tuberculosis,* 92**,** 314-320.

PHILLIPS, M., BASA-DALAY, V., BOTHAMLEY, G., CATANEO, R. N., LAM, P. K., NATIVIDAD, M. P. R., SCHMITT, P. & WAI, J. 2010. Breath biomarkers of active pulmonary tuberculosis. *Tuberculosis,* 90**,** 145-151.

PHILLIPS, M., CATANEO, R. N., CHATURVEDI, A., KAPLAN, P. D., LIBARDONI, M., MUNDADA, M., PATEL, U. & ZHANG, X. 2013. Detection of an extended human volatome with comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry. *PloS one,* 8**,** e75274.

PHILLIPS, M., CATANEO, R. N., CONDOS, R., ERICKSON, G. A. R., GREENBERG, J., LA BOMBARDI, V., MUNAWAR, M. I. & TIETJE, O. 2007. Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis,* 87**,** 44-52.

PHILLIPS, M., CATANEO, R. N., DITKOFF, B. A., FISHER, P., GREENBERG, J., GUNAWARDENA, R., KWON, C. S., TIETJE, O. & WONG, C. 2006. Prediction of breast cancer using volatile biomarkers in the breath. *Breast cancer research and treatment,* 99**,** 19-21.

PHILLIPS, M., GLEESON, K., HUGHES, J. M. B., GREENBERG, J., CATANEO, R. N., BAKER, L. & MCVAY, W. P. 1999a. Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study. *The Lancet,* 353**,** 1930-1933.

PHILLIPS, M., GREENBERG, J. & AWAD, J. 1994. Metabolic and environmental origins of volatile organic compounds in breath. *Journal of clinical pathology,* 47**,** 1052-1053.

PHILLIPS, M., HERRERA, J., KRISHNAN, S., ZAIN, M., GREENBERG, J. & CATANEO, R. N. 1999b. Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications,* 729**,** 75-88.

PHILLIPS, M., HERRERA, J., KRISHNAN, S., ZAIN, M., GREENBERG, J. & CATANEO, R. N. 1999c. Variation in volatile organic compounds in the breath of normal humans. *Journal of Chromatography B: Biomedical Sciences and Applications,* 729**,** 75-88.

PIERCE, K. M., HOGGARD, J. C., HOPE, J. L., RAINEY, P. M., HOOFNAGLE, A. N., JACK, R. M., WRIGHT, B. W. & SYNOVEC, R. E. 2006. Fisher ratio method

applied to third-order separation data to identify significant chemical components of metabolite extracts. *Analytical Chemistry,* 78**,** 5068-5075.

PLEIL, J. D. & LINDSTROM, A. B. 1995. Collection of a single alveolar exhaled breath for volatile organic compounds analysis. *American journal of industrial medicine,* 28**,** 109-121.

POLI, D., CARBOGNANI, P., CORRADI, M., GOLDONI, M., ACAMPA, O., BALBI, B., BIANCHI, L., RUSCA, M. & MUTTI, A. 2005. Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross sectional and nested short-term follow-up study. *Respiratory research,* 6**,** 1.

POLI, D., GOLDONI, M., CORRADI, M., ACAMPA, O., CARBOGNANI, P., INTERNULLO, E., CASALINI, A. & MUTTI, A. 2010. Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME–GC/MS. *Journal of Chromatography B,* 878**,** 2643-2651.

PROMPANYA, C. 2018. Study of bioactive secondary metabolites from the marine sponges and marine sponge-associated fungi.

PURCARO, G., NASIR, M., FRANCHINA, F. A., REES, C. A., ALIYEVA, M., DAPHTARY, N., WARGO, M. J., LUNDBLAD, L. K. & HILL, J. E. 2019. Breath metabolome of mice infected with Pseudomonas aeruginosa. *Metabolomics,* 15**,** 10.

PURCARO, G., REES, C. A., MELVIN, J. A., BOMBERGER, J. M. & HILL, J. E. 2018. Volatile fingerprinting of Pseudomonas aeruginosa and respiratory syncytial virus infection in an in vitro cystic fibrosis co-infection model. *Journal of breath research,* 12**,** 046001.

RAMSEY, B. W., FARRELL, P. M. & PENCHARZ, P. 1992. Nutritional assessment and management in cystic fibrosis: a consensus report. The Consensus Committee. *The American journal of clinical nutrition,* 55**,** 108-116.

RANINEN, K. J., LAPPI, J. E., MUKKALA, M. L., TUOMAINEN, T.-P., MYKKÄNEN, H. M., POUTANEN, K. S. & RAATIKAINEN, O. J. 2016. Fiber content of diet affects exhaled breath volatiles in fasting and postprandial state in a pilot crossover study. *Nutrition Research,* 36**,** 612-619.

RATIU, I.-A., BOCOS-BINTINTAN, V., MONEDEIRO, F., MILANOWSKI, M., LIGOR, T. & BUSZEWSKI, B. 2019. An optimistic vision of future: diagnosis of bacterial infections by sensing their associated volatile organic compounds. *Critical Reviews in Analytical Chemistry***,** 1-12.

RATJEN, F. & DÖRING, G. 2003. Cystic fibrosis. *The Lancet,* 361**,** 681-689.

REES, C. A., BURKLUND, A., STEFANUTO, P.-H., SCHWARTZMAN, J. D. & HILL, J. E. 2018. Comprehensive volatile metabolic fingerprinting of bacterial and fungal pathogen groups. *Journal of breath research,* 12**,** 026001.

REYNOLDS, J. C., BLACKBURN, G. J., GUALLAR-HOYAS, C., MOLL, V., BOCOS-BINTINTAN, V., KAUR-ATWAL, G., HOWDLE, M. D., HARRY, E., BROWN, L. J. & CREASER, C. S. 2010. Detection of volatile organic compounds in breath using thermal desorption electrospray ionization-ion mobility-mass spectrometry. *Analytical chemistry,* 82**,** 2139-2144.

REYNOLDS, J. C., JIMOH, M. A., GUALLAR-HOYAS, C., CREASER, C. S., SIDDIQUI, S. & THOMAS, C. L. P. 2014. Analysis of human breath samples using a modified thermal desorption: gas chromatography electrospray ionization interface. *Journal of breath research,* 8**,** 037105.

ROBROEKS, C. M., ROOZEBOOM, M. H., DE JONG, P. A., TIDDENS, H. A., JÖBSIS, Q., HENDRIKS, H. J., YNTEMA, J. B. L., BRACKEL, H. L., VAN GENT, R. & ROBBEN, S. 2010a. Structural lung changes, lung function, and non-invasive inflammatory markers in cystic fibrosis. *Pediatric allergy and immunology,* 21**,** 493-500.

ROBROEKS, C. M., ROSIAS, P. P., VAN VLIET, D., JÖBSIS, Q., YNTEMA, J. B. L., BRACKEL, H. J., DAMOISEAUX, J. G., DEN HARTOG, G. M., WODZIG, W. K. & DOMPELING, E. 2008. Biomarkers in exhaled breath condensate indicate presence and severity of cystic fibrosis in children. *Pediatric Allergy and Immunology,* 19**,** 652-659.

ROBROEKS, C. M., VAN BERKEL, J. J., DALLINGA, J. W., JÖBSIS, Q., ZIMMERMANN, L. J., HENDRIKS, H. J., WOUTERS, M. F., VAN DER GRINTEN, C. P., VAN DE KANT, K. D. & VAN SCHOOTEN, F.-J. 2010b. Metabolomics of volatile organic compounds in cystic fibrosis patients and controls. *Pediatric Research,* 68**,** 75-80.

RÖCK, F., BARSAN, N. & WEIMAR, U. 2008. Electronic nose: current status and future trends. *Chemical reviews,* 108**,** 705-725.

ROLON, M., BENALI, K., MUNCK, A., NAVARRO, J., CLEMENT, A., TUBIANA-RUFI, N., CZERNICHOW, P. & POLAK, M. 2001. Cystic fibrosis-related diabetes mellitus: clinical impact of prediabetes and effects of insulin therapy. *Acta Paediatrica,* 90**,** 860-867.

ROMMENS, J. M., IANNUZZI, M. C., KEREM, B.-S., DRUMM, M. L., MELMER, G., DEAN, M., ROZMAHEL, R., COLE, J. L., KENNEDY, D. & HIDAKA, N. 1989. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science,* 245**,** 1059-1065.

ROSENFELD, M., GIBSON, R. L., MCNAMARA, S., EMERSON, J., BURNS, J. L., CASTILE, R., HIATT, P., MCCOY, K., WILSON, C. B. & INGLIS, A. 2001. Early pulmonary infection, inflammation, and clinical outcomes in infants with cystic fibrosis. *Pediatric pulmonology,* 32**,** 356-366.

ROSIAS, P. P., ROBROEKS, C. M., VAN DE KANT, K. D., RIJKERS, G. T., ZIMMERMANN, L. J., VAN SCHAYCK, C. P., HEYNENS, J. W., JÖBSIS, Q. & DOMPELING, E. 2010. Feasibility of a new method to collect exhaled breath condensate in pre-school children. *Pediatric Allergy and Immunology,* 21**,** e235-e244.

RUST, L. 2018. *Odour profiling of blood training aids for blood-detection dogs using comprehensive two-dimensional gas chromatography (GCxGC).*

SAIMAN, L., MARSHALL, B. C., MAYER-HAMBLETT, N., BURNS, J. L., QUITTNER, A. L., CIBENE, D. A., COQUILLETTE, S., FIEBERG, A. Y., ACCURSO, F. J. & CAMPBELL III, P. W. 2003. Azithromycin in patients with cystic fibrosis chronically infected with Pseudomonas aeruginosa: a randomized controlled trial. *Jama,* 290**,** 1749-1756.

SAMPAT, A., LOPATKA, M., SJERPS, M., VIVO-TRUYOLS, G., SCHOENMAKERS, P. & VAN ASTEN, A. 2016. Forensic potential of comprehensive two-dimensional gas chromatography. *TrAC Trends in Analytical Chemistry,* 80**,** 345-363.

SANCHEZ, J. M. & SACKS, R. D. 2006. Development of a multibed sorption trap, comprehensive two-dimensional gas chromatography, and time-of-flight mass spectrometry system for the analysis of volatile organic compounds in human breath. *Analytical chemistry,* 78**,** 3046-3054.

SAVELEV, S., PERRY, J., BOURKE, S., JARY, H., TAYLOR, R., FISHER, A., CORRIS, P., PETRIE, M. & DE SOYZA, A. 2011. Volatile biomarkers of Pseudomonas aeruginosa in cystic fibrosis and noncystic fibrosis bronchiectasis. *Letters in applied microbiology,* 52**,** 610-613.

SCHLEICH, F. N., ZANELLA, D., STEFANUTO, P.-H., BESSONOV, K., SMOLINSKA, A., DALLINGA, J. W., HENKET, M., PAULUS, V.,

GUISSARD, F. & GRAFF, S. 2019. Exhaled volatile organic compounds are able to discriminate between neutrophilic and eosinophilic asthma. *American journal of respiratory and critical care medicine,* 200**,** 444-453.

SCHUBERT, J. K., MIEKISCH, W., GEIGER, K. & NÖLDGE–SCHOMBURG, G. F. 2004. Breath analysis in critically ill patients: potential and limitations. *Expert review of molecular diagnostics,* 4**,** 619-629.

SCOTT-THOMAS, A. J., SYHRE, M., PATTEMORE, P. K., EPTON, M., LAING, R., PEARSON, J. & CHAMBERS, S. T. 2010. 2-Aminoacetophenone as a potential breath biomarker for Pseudomonas aeruginosa in the cystic fibrosis lung. *BMC pulmonary medicine,* 10**,** 56.

SHELLIE, R., MARRIOTT, P. & CORNWELL, C. 2001. Application of comprehensive two-dimensional gas chromatography (GC× GC) to the enantioselective analysis of essential oils. *Journal of separation science,* 24**,** 823-830.

SHESTIVSKA, V., DRYAHINA, K., NUNVÁŘ, J., SOVOVÁ, K., ELHOTTOVÁ, D., NEMEC, A., SMITH, D. & ŠPANĚL, P. 2015. Quantitative analysis of volatile metabolites released in vitro by bacteria of the genus Stenotrophomonas for identification of breath biomarkers of respiratory infection in cystic fibrosis. *Journal of breath research,* 9**,** 027104.

SHESTIVSKA, V., NEMEC, A., DŘEVÍNEK, P., SOVOVÁ, K., DRYAHINA, K. & ŠPANĚL, P. 2011. Quantification of methyl thiocyanate in the headspace of Pseudomonas aeruginosa cultures and in the breath of cystic fibrosis patients by selected ion flow tube mass spectrometry. *Rapid Communications in Mass Spectrometry,* 25**,** 2459-2467.

SHESTIVSKA, V., ŠPANĚL, P., DRYAHINA, K., SOVOVÁ, K., SMITH, D., MUSILEK, M. & NEMEC, A. 2012. Variability in the concentrations of volatile metabolites emitted by genotypically different strains of P seudomonas aeruginosa. *Journal of applied microbiology,* 113**,** 701-713.

SIDHU, H., HOPPE, B., HESSE, A., TENBROCK, K., BROMME, S., RIETSCHEL, E. & PECK, A. B. 1998. Absence of Oxalobacter formigenes in cystic fibrosis patients: a risk factor for hyperoxaluria. *The Lancet,* 352**,** 1026-1029.

SINAASAPPEL, M., STERN, M., LITTLEWOOD, J., WOLFE, S., STEINKAMP, G., HEIJERMAN, H. G. M., ROBBERECHT, E. & DÖRING, G. 2002. Nutrition in patients with cystic fibrosis: a European Consensus. *Journal of Cystic Fibrosis,* 1**,** 51-75.

SMITH, D., SOVOVÁ, K., DRYAHINA, K., DOUŠOVÁ, T., DŘEVÍNEK, P. & ŠPANĚL, P. 2016. Breath concentration of acetic acid vapour is elevated in patients with cystic fibrosis. *Journal of breath research,* 10**,** 021002.

SMITH, D., WANG, T., SULÉ-SUSO, J., ŠPANĚL, P. & HAJ, A. E. 2003. Quantification of acetaldehyde released by lung cancer cells in vitro using selected ion flow tube mass spectrometry. *Rapid communications in mass spectrometry,* 17**,** 845-850.

STAFLER, P., DAVIES, J. C., BALFOUR-LYNN, I. M., ROSENTHAL, M. & BUSH, A. 2011. Bronchoscopy in cystic fibrosis infants diagnosed by newborn screening. *Pediatric pulmonology,* 46**,** 696-700.

STEFANUTO, P.-H., PERRAULT, K., LLOYD, R., STUART, B., RAI, T., FORBES, S. & FOCANT, J.-F. 2015. Exploring new dimensions in cadaveric decomposition odour analysis. *Analytical Methods,* 7**,** 2287-2294.

STEFANUTO, P.-H., ZANELLA, D., VERCAMMEN, J., HENKET, M., SCHLEICH, F., LOUIS, R. & FOCANT, J.-F. 2020. Multimodal combination of GC × GC-HRTOFMS and SIFT-MS for asthma phenotyping using exhaled breath. *Scientific Reports,* 10**,** 16159.

STEINKAMP, G., WIEDEMANN, B., RIETSCHEL, E., KRAHL, A., GIELEN, J., BÄRMEIER, H., RATJEN, F. & GROUP, E. B. S. 2005. Prospective evaluation of emerging bacteria in cystic fibrosis. *Journal of cystic fibrosis,* 4**,** 41-48.

STRANDVIK, B., GRONOWITZ, E., ENLUND, F., MARTINSSON, T. & WAHLSTRÖM, J. 2001. Essential fatty acid deficiency in relation to genotype in patients with cystic fibrosis. *The Journal of pediatrics,* 139**,** 650-655.

SUMNER, L. W., AMBERG, A., BARRETT, D., BEALE, M. H., BEGER, R., DAYKIN, C. A., FAN, T. W.-M., FIEHN, O., GOODACRE, R. & GRIFFIN, J. L. 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics,* 3**,** 211-221.

TANG, Z., LIU, Y. & DUAN, Y. 2015. Development of solid-phase microextraction fibers based on multi-walled carbon nanotubes for pre-concentration and analysis of alkanes in human breath. *Journal of Chromatography A,* 1425**,** 34-41.

TATE, S., MACGREGOR, G., DAVIS, M., INNES, J. & GREENING, A. 2002. Airways in cystic fibrosis are acidified: detection by exhaled breath condensate. *Thorax,* 57**,** 926-929.

THORN, R. M. S., REYNOLDS, D. M. & GREENMAN, J. 2011. Multivariate analysis of bacterial volatile compound profiles for discrimination between selected species and strains in vitro. *Journal of Microbiological Methods,* 84**,** 258-264.

TIMM, C. M., LLOYD, E. P., EGAN, A., MARINER, R. & KARIG, D. 2018. Direct Growth of Bacteria in Headspace Vials Allows for Screening of Volatiles by Gas Chromatography Mass Spectrometry. *Frontiers in microbiology,* 9**,** 491.

TRANCHIDA, P. Q., DONATO, P., CACCIOLA, F., BECCARIA, M., DUGO, P. & MONDELLO, L. 2013. Potential of comprehensive chromatography in food analysis. *TrAC Trends in Analytical Chemistry,* 52**,** 186-205.

TRIVEDI, B. P. 2013. Doorway to a Cure for Cystic Fibrosis. *Discover*, p.1.

TUNNEY, M. M., FIELD, T. R., MORIARTY, T. F., PATRICK, S., DOERING, G., MUHLEBACH, M. S., WOLFGANG, M. C., BOUCHER, R., GILPIN, D. F. & MCDOWELL, A. 2008. Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *American journal of respiratory and critical care medicine,* 177**,** 995-1001.

VAN DER SCHEE, M., FENS, N., BRINKMAN, P., BOS, L., ANGELO, M., NIJSEN, T., RAABE, R., KNOBEL, H., VINK, T. & STERK, P. 2012. Effect of transportation and storage using sorbent tubes of exhaled breath samples on diagnostic accuracy of electronic nose analysis. *Journal of breath research,* 7**,** 016002.

VAN DER SCHEE, M. P., PAFF, T., BRINKMAN, P., VAN AALDEREN, W. M. C., HAARMAN, E. G. & STERK, P. J. 2015. Breathomics in lung disease. *Chest,* 147**,** 224-231.

VAN STEE, L. & BRINKMAN, U. T. 2016. Peak detection methods for GC× GC: an overview. *TrAC Trends in Analytical Chemistry,* 83**,** 1-13.

VESELOVA, M., PLYUTA, V. & KHMEL, I. 2019. Volatile Compounds of Bacterial Origin: Structure, Biosynthesis, and Biological Activity. *Microbiology,* 88**,** 261-274.

VON MÜHLEN, C., ZINI, C. A., CARAMÃO, E. B. & MARRIOTT, P. J. 2006. Applications of comprehensive two-dimensional gas chromatography to the characterization of petrochemical and related samples. *Journal of Chromatography A,* 1105**,** 39-50.

WANG, C., KE, C., WANG, X., CHI, C., GUO, L., LUO, S., GUO, Z., XU, G., ZHANG, F. & LI, E. 2014. Noninvasive detection of colorectal cancer by analysis of exhaled breath. *Analytical and bioanalytical chemistry,* 406**,** 4757-4763.

WATERS, V., ATENAFU, E. G., LU, A., YAU, Y., TULLIS, E. & RATJEN, F. 2013. Chronic Stenotrophomonas maltophilia infection and mortality or lung transplantation in cystic fibrosis patients. *Journal of Cystic Fibrosis,* 12**,** 482-486.

WEHINGER, A., SCHMID, A., MECHTCHERIAKOV, S., LEDOCHOWSKI, M., GRABMER, C., GASTL, G. A. & AMANN, A. 2007. Lung cancer detection by proton transfer reaction mass-spectrometric analysis of human breath gas. *International Journal of Mass Spectrometry,* 265**,** 49-59.

WESTHOFF, M., LITTERST, P., FREITAG, L., URFER, W., BADER, S. & BAUMBACH, J. I. 2009. Ion mobility spectrometry for the detection of volatile organic compounds in exhaled breath of patients with lung cancer: results of a pilot study. *Thorax,* 64**,** 744-748.

WHITE, I. R., WILLIS, K. A., WHYTE, C., CORDELL, R., BLAKE, R. S., WARDLAW, A. J., RAO, S., GRIGG, J., ELLIS, A. M. & MONKS, P. S. 2013. Real-time multi-marker measurement of organic compounds in human breath: towards fingerprinting breath. *Journal of breath research,* 7**,** 017112.

WILDE, M. J., CORDELL, R. L., SALMAN, D., ZHAO, B., IBRAHIM, W., BRYANT, L., RUSZKIEWICZ, D., SINGAPURI, A., FREE, R. C. & GAILLARD, E. A. 2019. Breath analysis by two-dimensional gas chromatography with dual flame ionisation and mass spectrometric detection–Method optimisation and integration within a large-scale clinical study. *Journal of Chromatography A,* 1594**,** 160-172.

WILSON, H. & MONSTER, A. 1999. New technologies in the use of exhaled breath analysis for biological monitoring. *Occupational and environmental medicine,* 56**,** 753-757.

WOOLFENDEN, E. 2010a. Sorbent-based sampling methods for volatile and semi-volatile organic compounds in air. Part 2. Sorbent selection and other aspects of optimizing air monitoring methods. *Journal of Chromatography A,* 1217**,** 2685-2694.

WOOLFENDEN, E. 2010b. Sorbent-based sampling methods for volatile and semi-volatile organic compounds in air: Part 1: Sorbent-based air monitoring options. *Journal of Chromatography A,* 1217**,** 2674-2684.

WOOLFENDEN, E. & MCCLENNY, W. 1999. Compendium Method TO-17. Determination of volatile organic compounds in ambient air using active sampling onto sorbent tubes. *Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air*, 17-28.

XU, M., TANG, Z., DUAN, Y. & LIU, Y. 2015. GC-based Techniques for Breath Analysis: Current Status, Challenges and Prospects. *Critical Reviews in Analytical Chemistry*, 00-00.

XUE, R., DONG, L., ZHANG, S., DENG, C., LIU, T., WANG, J. & SHEN, X. 2008. Investigation of volatile biomarkers in liver cancer blood using solid-phase microextraction and gas chromatography/mass spectrometry. *Rapid communications in mass spectrometry,* 22, 1181-1186.

YANKASKAS, J. R., MALLORY, G. B. & COMMITTEE, C. 1998. Lung transplantation in cystic fibrosis: consensus conference statement. *Chest,* 113, 217-226.

YU, H., XU, L. & WANG, P. 2005. Solid phase microextraction for analysis of alkanes and aromatic hydrocarbons in human breath. *Journal of Chromatography B,* 826, 69-74.

ZANG, X., MONGE, M. E., MCCARTY, N. A., STECENKO, A. A. & FERNÁNDEZ, F. M. 2017. Feasibility of early detection of cystic fibrosis acute pulmonary exacerbations by exhaled breath condensate metabolomics: a pilot study. *Journal of Proteome Research,* 16, 550-558.

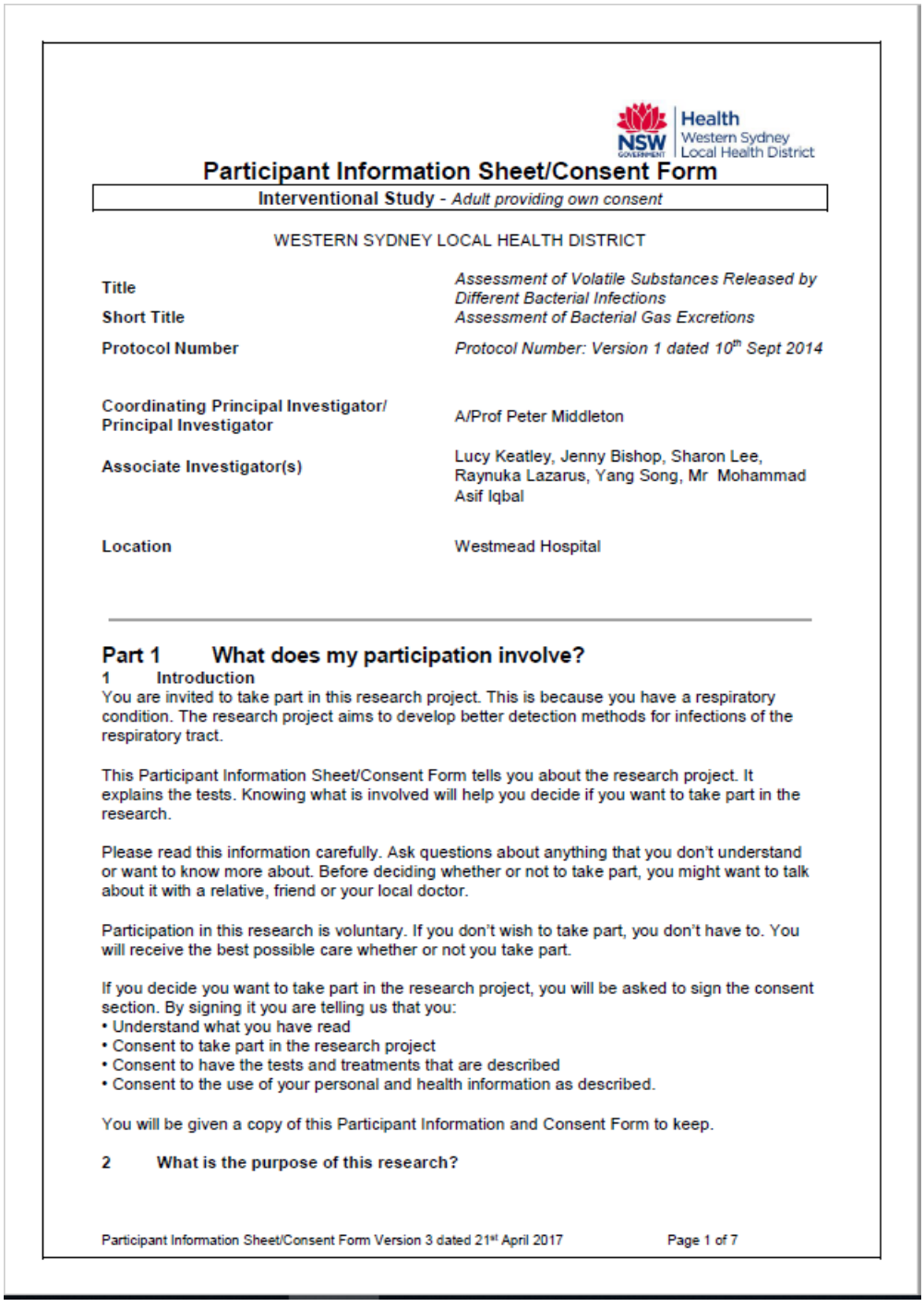
ZECHMAN, J. M. & LABOWS JR, J. N. 1985. Volatiles of Pseudomonas aeruginosa and related species by automated headspace concentration–gas chromatography. *Canadian journal of microbiology,* 31, 232-237.

ZEMEL, B. S., JAWAD, A. F., FITZSIMMONS, S. & STALLINGS, V. A. 2000. Longitudinal relationship among growth, nutritional status, and pulmonary function in children with cystic fibrosis: analysis of the Cystic Fibrosis Foundation National CF Patient Registry. *The Journal of pediatrics,* 137, 374-380.

ZIELENSKI, J., ROZMAHEL, R., BOZON, D., KEREM, B.-S., GRZELCZAK, Z., RIORDAN, J. R., ROMMENS, J. & TSUI, L.-C. 1991. Genomic DNA sequence of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Genomics,* 10, 214-228.

APPENDICES

APPENDIX A: CF SUBJECTS INFORMATION SHEET AND CONSENT FORM



Health
Western Sydney
Local Health District

Participant Information Sheet/Consent Form

Interventional Study - *Adult providing own consent*

WESTERN SYDNEY LOCAL HEALTH DISTRICT

Title	*Assessment of Volatile Substances Released by Different Bacterial Infections*
Short Title	*Assessment of Bacterial Gas Excretions*
Protocol Number	*Protocol Number: Version 1 dated 10th Sept 2014*
Coordinating Principal Investigator/ Principal Investigator	A/Prof Peter Middleton
Associate Investigator(s)	Lucy Keatley, Jenny Bishop, Sharon Lee, Raynuka Lazarus, Yang Song, Mr Mohammad Asif Iqbal
Location	Westmead Hospital

Part 1 What does my participation involve?

1 Introduction

You are invited to take part in this research project. This is because you have a respiratory condition. The research project aims to develop better detection methods for infections of the respiratory tract.

This Participant Information Sheet/Consent Form tells you about the research project. It explains the tests. Knowing what is involved will help you decide if you want to take part in the research.

Please read this information carefully. Ask questions about anything that you don't understand or want to know more about. Before deciding whether or not to take part, you might want to talk about it with a relative, friend or your local doctor.

Participation in this research is voluntary. If you don't wish to take part, you don't have to. You will receive the best possible care whether or not you take part.

If you decide you want to take part in the research project, you will be asked to sign the consent section. By signing it you are telling us that you:
- Understand what you have read
- Consent to take part in the research project
- Consent to have the tests and treatments that are described
- Consent to the use of your personal and health information as described.

You will be given a copy of this Participant Information and Consent Form to keep.

2 What is the purpose of this research?

Participant Information Sheet/Consent Form Version 3 dated 21st April 2017 Page 1 of 7

The purpose of this study is to determine if examination of the air that you breathe out can identify if there are bacteria in your lungs, similar to the routine examination of the mucous you cough up.
This research is being conducted in co-ordination with the University of Technology Sydney. This research project has been designed to make sure the researchers interpret the results in a fair and appropriate way and avoids study doctors or participants jumping to conclusions.

3 What does participation in this research involve?

What Will Happen on the Study?

If you agree to participate in this study, you will be asked to sign the Participant Information and Consent Form. The study will be conducted during one of your clinic visits to Westmead Hospital.

If you agree to participate, you will first be asked to provide the investigator with a sputum sample. This is something you have most likely done before.

You will be given a small container and asked to cough and spit any mucous in to the container.

Your sample will be sent to the University of Technology Sydney for analysis.

After you have provided this sample, we will ask you to provide a sample of your breath. You will be asked to sit quietly and breathe through a mouthpiece for 10 minutes. The machine collects your breath for the 10 minutes and then we will send this sample off to be analysed at the University of Technology Sydney.

You may also be asked to complete a short questionnaire, which will ask you questions about your history of respiratory infections and the condition of your lungs at the time of testing. This will take no more than 10 minutes to complete.

In addition, the researchers would like to have access to your medical record to obtain information relevant to the study.

Participation in this study will not cost you anything, nor will you be paid to participate. You may be reimbursed for any reasonable travel, parking, meals and other expenses associated with the research project visit.

4 Other relevant information about the research project
Approximately 40 persons will be asked to participate in the study. It will be performed in collaboration with persons at the University of Technology Sydney, and participants will be invited from Westmead Hospital only.

5 Do I have to take part in this research project?
Participation in any research project is voluntary. If you do not wish to take part, you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.
Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect your routine treatment, your relationship with those treating you or your relationship with Westmead Hospital or The Western Sydney Local Health District.

6 What are the alternatives to participation?
You do not have to take part in this research project to receive treatment at this hospital. Currently available methods for diagnosis will always be available to you. Further information can be provided by your study doctor if you wish.

7 What are the possible benefits of taking part?

The study aims to improve the currently available methods of detection of bacterial infections of the respiratory tract; however it may not directly benefit you.

8 What are the possible risks and disadvantages of taking part?
All medical procedures involve some risk of injury. In addition, there may be risks associated with this study that are presently unknown or unforeseeable. In spite of all reasonable precautions, you might develop medical complications from participating in this study. There are no known risks associated with this trial.

There may be side effects that the researchers do not expect or do not know about and that may be serious. Tell your study doctor immediately about any new or unusual symptoms that you get.

9 Pregnancy and Contraception
There is no risk associated with pregnancy and providing sputum samples or breath samples,

10 What will happen to my test samples?
Collection of sputum is necessary for this study; its purpose is to provide information about the bacteria in your mucous. Once your sample has been analysed it will be sterilised by autoclave (exposure to high pressure steam) and disposed of in a biological waste bin.

If you decide to participate in the study, you will be assigned a "Study ID" (eg SEBC001). In this manner your identity will remain anonymous to all persons outside the research project. Your sputum sample will be labelled with your Study ID and the date of collection. All sputum and breath samples taken throughout this study will be for research, and are not for the purposes of routine care/monitoring.

11 Can I have other treatments during this research project?
Yes, you are permitted to take all of your usual medications to treat your medical conditions.

12 What if I withdraw from this research project?
If you decide to withdraw from the project, please notify a member of the research team before you withdraw. This notice will allow that person or the research supervisor to discuss any health risks or special requirements linked to withdrawing.

If you do withdraw your consent during the research project, the study doctor and relevant study staff will not collect additional personal information from you, although personal information already collected will be retained to ensure that the results of the research project can be measured properly and to comply with law. You should be aware that data collected by the sponsor up to the time you withdraw will form part of the research project results. If you do not want them to do this, you must tell them before you join the research project.

13 Could this research project be stopped unexpectedly?
No, it is expected that we will study 40 participants.

14 What happens when the research project ends?
If you give us your permission by the consent document, we plan to discuss/publish the results at scientific conferences and publish them in scientific journals.
In any publication, information will be provided in such a way that you cannot be identified.
Results of the study will be provided to you once it is complete, if you wish. The study is expected to take approximately a year, but may take more or less time. The bacteria present in your sputum will be provided to your physician.

Part 2 How is the research project being conducted?

15 What will happen to information about me?

By signing the consent form you consent to the study doctor and relevant research staff collecting and using personal information about you for the research project. Any information obtained in connection with this research project that can identify you will remain confidential. You will be assigned a "Study ID" (eg SEBC001); all subsequent study documentation will use this ID. In this way only those persons involved in the research study will be able to identify you. Your information will only be used for the purpose of this research project and it will only be disclosed with your permission, except as required by law.

In accordance with relevant Australian and New South Wales privacy and other relevant laws, you have the right to request access to your information collected and stored by the research team. You also have the right to request that any information with which you disagree be corrected. Please contact the study team member named at the end of this document if you would like to access your information.

16 Complaints and compensation

If you suffer any injuries or complications as a result of this research project, you should contact the study team as soon as possible and you will be assisted with arranging appropriate medical treatment. If you are eligible for Medicare, you can receive any medical treatment required to treat the injury or complication, free of charge, as a public patient in any Australian public hospital.

In the event of loss or injury, the parties involved in this research project have agreed to assist you in obtaining any compensation you are entitled to under the law.

17 Who is organising and funding the research?

This research is being funded by the Department of Respiratory & Sleep Medicine at Westmead hospital. In addition, if knowledge acquired through this research leads to discoveries that are of commercial value to the Western Sydney Local Health District and/or the University of Technology Sydney, the study doctors or their institutions, there will be no financial benefit to you or your family from these discoveries.

19 Who has reviewed the research project?

All research in Australia involving humans is reviewed by an independent group of people called a Human Research Ethics Committee (HREC). The ethical aspects of this research project have been approved by the HREC of the Western Sydney Local Health District.

20 Further information and who to contact

The person you may need to contact will depend on the nature of your query.

If you want any further information concerning this project or if you have any medical problems which may be related to your involvement in the project (for example, any side effects), you can contact the principal study doctor on 9845 6797(office hours) or 9845 5555 then ask to page A/Prof Peter Middleton (after hours) or any of the following people:

Clinical contact person

Name	*A/Prof Peter Middleton*
Position	*Staff Specilaist*
Telephone	*8890 6797*
Email	*peter.middleton@sydney.edu.au*

For matters relating to research at the site at which you are participating, the details of the local site complaints person are:

Complaints contact person

Name	*Patient Representative*
Telephone	*8890 7014*
Email	wslhd-pals-mail@health.nsw.gov.au

If you have any complaints about any aspect of the project, the way it is being conducted or any questions about being a research participant in general, then you may contact:

Reviewing HREC approving this research and HREC Executive Officer details

Reviewing HREC name	*Western Sydney Local Health District*
HREC Executive Officer	*Kellie Hansen*
Telephone	*8890 8183*
Email	Wslhd-researchoffice@health.nsw.gov.au

Local HREC Office contact (Single Site -Research Governance Officer)

Name	Margaret Piper
Position	Research Governance Officer
Telephone	8890 9634
Email	wslhd-rgo@health.nsw.gov.au

APPENDIX B: HEALTHY CONTROLS INFORMATION SHEET AND CONSENT FORM

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

PARTICIPANT INFORMATION SHEET
Chemical Profiling of Exhaled Breath from Healthy Participants Using GCxGC-TOFMS
(HREC REF: ETH17-1806)

WHO IS DOING THE RESEARCH?

My name is Mohammad Asif Iqbal and I am a PhD student at UTS. My supervisor is Prof. Shari Forbes.

WHAT IS THIS RESEARCH ABOUT?

Cystic fibrosis (CF) is the most common lethal genetic disorder in Australia. Due to this disorder, CF patients develop an excessive amount of thick and sticky mucus within their lungs and airways. This leads to frequent and repeated lung infections causing irreversible damage. For this reason, CF patients have to undergo regular health checks for bacterial lung infections. The aim of this study is to develop a non-invasive technique for the diagnosis of bacterial lung infections in CF patients based on breath analysis. In this process, breath samples will be collected from both CF patients and healthy controls and compounds in these samples will be analysed using comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry. The obtained data will be analysed to identify compounds that can reveal bacterial lung infections in CF patients.

WHY HAVE I BEEN ASKED?

You have been invited to participate in this study because you represent our healthy control group with no known health complications. Sampling from both CF patients and healthy controls will help us to answer the following questions:

- Can we differentiate between CF patients and healthy humans through breath analysis?
- Is it possible to identify bacterial lung infections in CF patients though breath analysis?
- How precise is our breath analysis technique?

IF I SAY YES, WHAT WILL IT INVOLVE?

If you decide to participate, you will be asked to sign a consent form confirming that you are capable of providing breath samples and do not have any of the following health complications/history:

- Cystic fibrosis
- Known to have:
  - Pulmonary bacterial infection
  - Any respiratory condition (e.g. Asthma)
  - Halitosis
  - Poor dentition/gingivitis/gum disease
  - Cardiovascular disease
  - Chronic anaemia
  - Panic disorders
  - Any digestive disease (e.g. Irritable Bowel Syndrome)
- Eaten a meal within 1 hr
- Regular smoker

In addition, we will ask you to provide the following information:

- Your age
- Sex
- Anthropometrics (i.e. height and weight)

If no exclusion criteria is present and you are willing to provide us the above mentioned information, I will invite you to provide a total of three breath samples (one every two months). For each sample, you have to breathe through a disposable mouthpiece for ~2 – 5 minutes. The sample will be collected into a 1 L Tedlar breath sampling bag. This technique is non-invasive and should take only 5-10 minutes. The total time you have to spend is 15-30 minutes (for three sampling sessions).

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

ARE THERE ANY RISKS/INCONVENIENCE?

There is no known health risk associated with this trial. However, you may feel uncomfortable about providing breath samples. The investigator/research student involved in this study will provide necessary guidelines and accompany you during the whole sampling period to avoid any discomfort. In addition, sampling will be done in a relaxed and stepwise manner to avoid exhaustion.

DO I HAVE TO SAY YES?

Participation in this study is voluntary. It is completely up to you whether or not you decide to take part.

WHAT WILL HAPPEN IF I SAY NO?

If you decide not to participate, it will not affect your relationship with the researchers or the University of Technology Sydney. If you wish to withdraw from the study once it has started, you can do so at any time without having to give a reason, by contacting either Mohammad Asif Iqbal or Prof. Shari Forbes. If you withdraw from the study, the results of your breath sample/s will be erased from our computerised data system and the collected documents will also be destroyed according to the regulation.

CONFIDENTIALITY

By signing the consent form you consent to the research team collecting and using personal information about you for the research project. All this information will be treated confidentially. We will not collect any identifiable information (e.g. name and date of birth) and all collected data will be kept securely. The questionnaire collected from you will be stored in a locked drawer in an office area. The only people to have regular access will be the investigator/research student of this project. The results will be recorded in the computer attached to the instrument used for analysis and then transferred to the researcher's computer for data analysis. This computer system is highly secure and password protected with very limited assess. In addition, we also plan to publish the results as journal articles. However, in any publication, information will be provided in such a way that participants cannot be identified.

WHAT IF I HAVE CONCERNS OR A COMPLAINT?

If you have concerns about the research that you think I, Mohammad Asif Iqbal or my supervisor can help you with, please feel free to contact me at MohammadAsif.Iqbal@student.uts.edu.au or Prof. Shari Forbes at Shari.Forbes@uts.edu.au

You will be given a copy of this form to keep.

NOTE:

This study has been approved by the University of Technology Sydney Human Research Ethics Committee [UTS HREC]. If you have any concerns or complaints about any aspect of the conduct of this research, please contact the Ethics Secretariat on ph.: +61 2 9514 2478 or email: Research.Ethics@uts.edu.au, and quote the UTS HREC reference number. Any matter raised will be treated confidentially, investigated and you will be informed of the outcome.

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Title: Chemical Profiling of Exhaled Breath from Healthy Participants Using GCxGC-TOFMS

Chief Investigator: Professor Shari Forbes

Research student: Mohammad Asif Iqbal

Subject ID: _____ Date: _____

Study Questionnaire

Has this study been explained to you?

Yes ☐ No ☐ Comment _____

Have you read the consent form?

Yes ☐ No ☐ Comment _____

Have you signed the consent form prior to any study procedures?

Yes ☐ No ☐ Comment _____

Questions posed to you are answered by _____

(Print Researcher's Name)

Inclusion/Exclusion Criteria

Inclusion Criteria

Are you over 18 years old and able to provide informed consent? Yes ☐ No ☐

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Title: Chemical Profiling of Exhaled Breath from Healthy Participants Using GCxGC-TOFMS

Chief Investigator: Professor Shari Forbes

Research student: Mohammad Asif Iqbal

Subject ID: _____ Date: _____

Exclusion Criteria

Please respond to each of the following questions.

Do you have any of the following conditions?

Cystic fibrosis	Yes ☐	No ☐
Any pulmonary bacterial infection	Yes ☐	No ☐
Other respiratory conditions (e.g. asthma)	Yes ☐	No ☐
Clinical history of halitosis	Yes ☐	No ☐
Poor dentition/gingivitis/gum disease	Yes ☐	No ☐
Cardiovascular disease	Yes ☐	No ☐
Chronic anaemia	Yes ☐	No ☐
Panic disorders	Yes ☐	No ☐
Any digestive disease (e.g. irritable bowel syndrome)	Yes ☐	No ☐
An inability to provide breath samples	Yes ☐	No ☐
Have you eaten a meal within the previous hour	Yes ☐	No ☐
Are you a regular smoker	Yes ☐	No ☐

Additional Information Required

Age at time of testing		
Gender		
Anthropometrics	Height	
	Weight	

UTS UNIVERSITY OF TECHNOLOGY SYDNEY

CONSENT FORM
Chemical Profiling of Exhaled Breath from Healthy Participants Using GCxGC-TOFMS
(HREC REF: ETH17-1806)

I _____ *[participant's name]* agree to participate in the research project [**HREC REF: ETH17-1806**] being conducted by Mohammad Asif Iqbal and Professor Shari Forbes at the UTS City Campus (15 Broadway, NSW, 2007; Phone: 9514 1717).

I have read the Participant Information Sheet or someone has read it to me in a language that I understand.

I understand the purposes, procedures and risks of the research as described in the Participant Information Sheet.

I have had an opportunity to ask questions and I am satisfied with the answers I have received.

I freely agree to participate in this research project as described and understand that I am free to withdraw at any time without affecting my relationship with the researchers or the University of Technology Sydney.

I understand that I will be given a signed copy of this document to keep.

I agree that the research data gathered from this project may be published in a form that:
☐ Does not identify me in any way
☐ May be used for future research purposes *

* We would like to store your information for future use in research projects that are an extension of this research project. In all instances your information will be treated confidentially.

I am aware that I can contact Mohammad Asif Iqbal and Professor Shari Forbes if I have any concerns about the research.

_____ ____/____/____
Name and Signature [participant] Date

_____ ____/____/____
Name and Signature [researcher or delegate] Date

APPENDIX C: CAS NUMBER OF ALL CHEMICALS NAMED IN THIS STUDY

Compound name	CAS number
1,2-Benzenedicarboxylic acid, bis(2-methylpropyl) ester	84-69-5
1,2-Benzenedicarboxylic acid, butyl 2-ethylhexyl ester	85-69-8
1,3-Cyclohexadiene, 5-(1,5-dimethyl-4-hexenyl)-2-methyl-	495-60-3
1,5-Hexadien-3-ol	924-41-4
1-Butanol	71-36-3
1-Butanol, 3-methyl-	123-51-3
1-Dodecanamine, N,N-dimethyl-	112-18-5
1H,1H,2H,2H-Perfluorooctan-1-ol	647-42-7
1-Hexanol	111-27-3
1-Hexanol, 2-ethyl-	104-76-7
1H-Pyrazole, 4,5-dihydro-5,5-dimethyl-4-isopropylidene-	41077-76-3
1-Methyl-2-pyrrolidinone	872-50-4
1-Nonadecanamine, N,N-dimethyl-	49859-87-2

1-Nonen-4-ol	35192-73-5
1-Octen-3-ol	3391-86-4
1-Octen-3-one	4312-99-6
1-Pentadecanamine, N,N-dimethyl-	17678-60-3
1-Pentanol	71-41-0
1-Penten-3-ol	616-25-1
1-Propanol	71-23-8
1-Propanol, 2-(2-hydroxypropoxy)-	106-62-7
1-Propanol, 2,2'-oxybis-	65997-33-3
1-Propene, 1-(methylthio)-	42848-06-6
1-Tridecanamine, N,N-dimethyl-	19047-96-2
1-Undecene	821-95-4
2,2,4,4-Tetramethyloctane	62183-79-3
2,3-Butanedione	431-03-8
2,3-Hexanedione	3848-24-6
2,3-Octanedione	585-25-1

2,3-Pentanedione	600-14-6
2,4,7,9-Tetramethyl-5-decyn-4,7-diol	126-86-3
2',4'-Dihydroxy-3'-methylacetophenone	10139-84-1
2,6-Bis(1,1-dimethylethyl)-4-(1-oxopropyl)phenol	14035-34-8
2,6-Dimethyldecane	13150-81-7
2-Butanone	78-93-3
2-Cyclohexen-1-one	930-68-7
2-Heptanone	110-43-0
2-Heptanone, 6-methyl-	928-68-7
2-Nonanone	821-55-6
2-Octanone	111-13-7
2-Octenal	2548-87-0
2-Oxazolidinone, 3-amino-5-(4-morpholinylmethyl)-	43056-63-9
2-Pentanone	107-87-9
2-Piperidinone	675-20-7
2-Propanol, 1-(2-butoxy-1-methylethoxy)-	29911-28-2

2-Propanone, 1-(acetyloxy)-	592-20-1
3-Hexanone, 2,5-dimethyl-4-nitro-	59906-54-6
3-Methylheptyl acetate	72218-58-7
3-Octanone	06-68-3
4-Methoxy-4-methyl-2-pentanol	141-73-1
4-Morpholinebutyric acid, α-methyl-α,α-diphenyl-	3626-55-9
7-Octen-2-ol, 2,6-dimethyl-	18479-58-8
Acetamide, N-methyl-	79-16-3
Acetic acid	64-19-7
Acetic acid ethenyl ester	108-05-4
Acetic acid, methyl ester	79-20-9
Acetic acid, phenyl ester	830-03-5
Acetoin	513-86-0
Acetonitrile	75-05-8
Acetyl valeryl	96-04-8
Aziridine, 1-ethenyl-	5628-99-9

Benzaldehyde	532-28-5
Benzene	71-43-2
Benzene, 1-(1,5-dimethyl-4-hexenyl)-4-methyl-	644-30-4
Benzene, 1,2,3,5-tetramethyl-	527-53-7
Benzene, 1,2,3-trimethyl-	526-73-8
Benzene, 1,3-dichloro-	541-73-1
Benzene, 1,3-dimethyl-	108-38-3
Benzene, 1-ethenyl-4-ethyl-	3454-07-7
Benzene, 1-ethyl-3-methyl-	622-96-8
Benzene, 1-ethyl-4-methyl-	535-77-3
Benzene, 1-methyl-3-(1-methylethyl)-	535-77-3
Benzene, 1-methyl-4-(1-methylethenyl)-	1195-32-0
Benzeneacetaldehyde	122-78-1
Benzenemethanol, α,α-dimethyl-	617-94-7
Benzofuran, 4,5,6,7-tetrahydro-3,6-dimethyl-	17957-94-7
Benzoic acid, methyl ester	93-58-3

Benzonitrile	100-47-0
Butanal, 3-methyl-	590-86-3
Butane, 1,1,3,4-tetrachloro-1,2,2,3,4,4-hexafluoro-	423-38-1
Butane, 2,2,3,3-tetramethyl-	594-82-1
Butanenitrile, 3-methyl-	625-28-5
Butanoic acid	107-92-6
Butanoic acid, 2-methyl-	116-53-0
Butanoic acid, 3-methyl-	503-74-2
Butylated Hydroxytoluene	128-37-0
Butyrolactone	96-48-0
Camphene	79-92-5
Carbamic acid, N-[1,1-bis(trifluoromethyl)ethyl]-, 4-(1,1,3,3-tetramethylbutyl)phenyl ester	296242-69-8
Carvone	99-49-0
Cyclohexane, isocyanato-	3173-53-3
Cyclohexane, isothiocyanato-	1122-82-3
Cyclohexanol, 3,3,5-trimethyl-	116-02-9

Cyclohexanol, 5-methyl-2-(1-methylethyl)-, [1S-(1α,2α,5α)]-	15356-70-4
Cyclohexanone	108-94-1
Cyclohexanone, 2-methyl-5-(1-methylethenyl)-, trans-	7764-50-3
Decane	124-18-5
Decane, 2,5-dimethyl-	17302-37-3
Dimethyl sulfone	67-71-0
Dimethyl trisulfide	3658-80-8
Disulfide, dimethyl	624-92-0
DL-3-Methyl-2-butanol	598-75-4
DL-Glyceraldehyde	56-82-6
Dodecane	112-40-3
Ethanol	64-17-5
Ethanol, 2-(2-ethoxyethoxy)-	111-90-0
Ethanol, 2-(hexyloxy)-	112-25-4
Ethanol, 2-butoxy-	11-76-2
Ethanone, 1-[1-(4-amino-1,2,5-oxadiazol-3-yl)-5-methyl-1H-1,2,3-triazol-4-yl]-2-morpholino-	-

Ethyl Acetate	141-78-6
Ethylbenzene	100-41-4
Eucalyptol	470-82-6
Furan, 2,5-dihydro-	1708-29-8
Furan, 2-pentyl-	3777-69-3
Heptane, 2,2,4,6,6-pentamethyl-	13475-82-6
Heptane, 5-ethyl-2,2,3-trimethyl-	62199-06-8
Hexadecane	544-76-3
Hexanal	66-25-1
Hydrazine	7803-57-8
Hydroxyurea	127-07-1
Indole	120-72-9
Isophorone	78-59-1
Levomenthol	2216-51
Limonene	5989-27-5
Methyl Isobutyl Ketone	108-10-1

Methyl isovalerate	556-24-1
Methyl propionate	554-12-1
Methylene chloride	75-09-2
Morpholine, 4-octadecyl-	16528-77-1
N-(tert-Butoxycarbonyl)glycine	4530-20-5
Naphthalene	91-20-3
n-Hexane	110-54-3
N-Morpholinomethyl-isopropyl-sulfide	77422-34-5
Nonanal	124-19-6
Octanal	124-13-0
Octane	111-65-9
Octane, 2,4,6-trimethyl-	62016-37-9
o-Cymene	527-84-4
o-Menthan-8-ol	498-81-7
Oxime-, methoxy-phenyl-	2475-92-5
p-Cymene	99-87-6

Pentadecanoic acid, 14-methyl-, methyl ester	5129-60-2
Pentanal, 2-methyl-	123-15-9
Pentane, 2,2,3,3-tetramethyl-	7154-79-2
Pentane, 3-ethyl-2,2-dimethyl-	16747-32-3
Phenylethyl Alcohol	60-12-8
Phosphonic acid, (p-hydroxyphenyl)-	33795-18-5
Phthalic acid, isobutyl tridec-2-yn-1-yl ester	-
Propanal, 2-methyl-	35730-34-8
Propane, 1-(methylthio)-	3877-15-4
Propane, 2,2-dimethoxy-	77-76-9
Propane, 2-isocyanato-2-methyl-	1609-86-5
Propane, 2-methoxy-2-methyl-	1634-04-4
Propanoic acid	79-09-4
Propanoic acid, 2-methyl-, 1-(1,1-dimethylethyl)-2-methyl-1,3-propanediyl ester	1472-87-3
Propanoic acid, 2-methyl-, 2,2-dimethyl-1-(2-hydroxy-1-methylethyl)propyl ester	74367-33-2
Propanoic acid, 2-methyl-, 3-hydroxy-2,4,4-trimethylpentyl ester	74367-34-3

Propanoic acid, 2-oxo-, methyl ester	600-22-6
p-Xylene	106-42-3
Pyrazine, 2,5-dimethyl-	123-32-0
Pyrazine, methyl-	109-08-0
Pyridine	110-86-1
Pyrrole	109-97-7
Styrene	100-42-5
Sulfide, allyl methyl	10152-76-8
Tetrachloroethylene	127-18-4
Tetradecane	629-59-4
Tetrahydrofuran	109-99-9
Thiocyanic acid, methyl ester	556-64-9
Toluene	108-88-3
Trichloromethane	67-66-3
Tricyclo[5.2.2.0(2,6)]undecan-11-one-8,9-dicarboxylic anhydride, 3-[(2-methoxyethoxy)methoxy]-2-methyl-	67634-20-2
Tridecane	629-92-5

Tridecanoic acid, methyl ester	1731-88-0
Undecane	1120-21-4
Undecane, 2-methyl-	7045-71-8
Urea, tetramethyl-	632-22-4
α-Bourbonene	5208-59-3
α-Phellandrene	99-83-2
α-Pinene	7785-70-8
α-Terpineol	8000-41-7
γ-Terpinene	99-86-5

APPENDIX D: ANTHROPOMETRIC DATA OF CF SUBJECTS

Order	*Patient ID*	*Age (years)*	*Sex*	*Height (cm)*	*Weight (kg)*	*Asthma*	*Smoking*	*Complications*	*Latest Exacerbation (Months ago)*
1	CF02	22	M	187	64	NO	NO	NO	14
2	CF03	27	F	168	74	NO	NO	NO	5
3	CF04	25	M	178	92	NO	NO	NO	14
4	CF05	19	F	165	52.5	NO	NO	NO	0
5	CF06	28	M	178	56.9	YES	NO	NO	0
6	CF07	37	F	172	91	NO	NO	NO	6
7	CF10	25	M	182	70.5	NO	NO	NO	14
8	CF11	21	M	175	74.9	NO	NO	NO	7
9	CF12	44	M	177	63	NO	NO	NO	4
10	CF13	20	M	175	59.8	NO	NO	NO	1
11	CF14	32	F	178	69	NO	NO	NO	6
12	CF15	37	F	163	64.4	YES	NO	NO	5
13	CF16	22	M	174	54.4	NO	NO	NO	2
14	CF17	21	M	177	68.8	YES	NO	NO	19
15	CF18	32	F	178	69.7	NO	NO	NO	7
16	CF19	23	M	180	77.6	NO	NO	NO	6
17	CF20	59	M	180	78	NO	NO	NO	45
18	CF21	32	F	160	55	NO	NO	NO	6
19	CF22	25	M	178	91.7	NO	NO	NO	18
20	CF23	27	F	168	75.9	NO	NO	NO	8
21	CF24	26	M	182	70.5	NO	NO	NO	16
22	CF25	45	M	177	61.4	NO	NO	NO	2
23	CF26	33	M	162	72.6	YES	NO	NO	3
24	CF27	18	F	162	48.9	YES	NO	NO	12

25	CF28	20	F	151	39.2	NO	NO	NO	1
26	CF29	32	M	172	60.8	YES	NO	NO	13
27	CF30	24	M	170	73.3	NO	NO	NO	5
28	CF31	19	M	176.4	75.5	NO	NO	NO	46
29	CF32	18	F	165	65.3	NO	NO	NO	0
30	CF33	24	M	165	59.6	NO	NO	NO	5
31	CF34	22	F	163	61.1	NO	NO	NO	1
32	CF35	36	M	161	62.7	NO	NO	NO	1
33	CF36	32	F	163.7	52.2	NO	NO	NO	0
34	CF37	19	M	171	49.4	NO	NO	NO	2
35	CF38	19	M	182	73.7	NO	NO	NO	23
36	CF39	21	M	167	55.1	NO	NO	NO	2
37	CF40	19	M	162	93.3	NO	NO	NO	5
38	CF41	19	F	163	49.5	YES	NO	NO	1
39	CF42	32	F	156	58.3	NO	NO	NO	38
40	CF43	18	M	176.5	62.7	NO	NO	NO	0
41	CF44	31	F	160	57	NO	NO	NO	1
42	CF45	32	M	165	71	NO	EX	NO	35
43	CF46	30	F	177	69	NO	NO	NO	15
44	CF47	23	F	163	61.4	NO	NO	NO	2
45	CF48	20	M	160	57.5	NO	NO	NO	22
46	CF49	45	M	177	60.5	NO	NO	YES	5
47	CF50	33	M	171	65.5	NO	NO	NO	22
48	CF51	19	F	165	65.3	NO	NO	NO	0
49	CF52	18	M	174	61.4	NO	NO	NO	0
50	CF53	20	M	167	66.9	NO	NO	YES	14
51	CF54	26	M	182	70.5	NO	NO	NO	26
52	CF55	32	F	165	65.1	NO	NO	NO	0
53	CF56	26	F	166	71.6	NO	NO	NO	0

54	CF57	20	M	182	73	NO	NO	NO	0
55	CF58	55	F	158	56	NO	NO	NO	78
56	CF59	40	M	162	94.5	NO	NO	NO	0
57	CF60	28	F	166	77.5	NO	NO	NO	12
58	CF61	29	M	183	59	NO	NO	NO	17
59	CF63	26	M	167	60	NO	NO	NO	2
60	CF64	22	F	170	57.5	NO	NO	NO	52
61	CF65	28	M	160	60.1	NO	NO	NO	5
62	CF66	21	F	158	49.6	NO	NO	NO	0
63	CF67	34	M	162	72.6	YES	NO	YES	8
64	CF68	35	F	146	45.8	YES	NO	NO	64
65	CF69	24	M	168	68.6	NO	NO	NO	28

APPENDIX E: ANTHROPOMETRIC DATA OF HEALTHY CONTROLS

Order	Sample code	Age	Sex	Height	Weight
1	CB01	30	M	170.2	70
2	CB02	30	F	165	62
3	CB03	32	M	176	75
4	CB04	27	F	160	50
5	CB05	29	M	185	80
6	CB06	29	F	168	60
7	CB07	29	M	177.8	85
8	CB08	30	M	170.2	70
9	CB09	40	F	167	57.5
10	CB10	23	M	170	105
11	CB11	29	F	165	60
12	CB12	25	M	180	70
13	CB13	28	M	175.3	78.5
14	CB14	29	M	198	95
15	CB15	33	F	170	60
16	CB16	34	M	177.8	80
17	CB17	29	M	175	80
18	CB18	31	F	167	58
19	CB19	25	M	170	60
20	CB20	27	F	173	65
21	CB21	25	F	158	58
22	CB22	44	M	183	78
23	CB23	26	F	157.5	54
24	CB24	36	M	182	70
25	CB25	33	F	169	59

26	CB26	33	M	178	64
27	CB27	33	M	177	84
28	CB28	25	M	176	77
29	CB29	26	M	167	59
30	CB30	40	F	163	105
31	CB31	24	F	165	58
32	CB32	21	F	172	60
33	CB33	23	M	180	72
34	CB34	23	F	168	60
35	CB35	23	F	165	60
36	CB36	24	F	167	72
37	CB37	56	F	158	70
38	CB38	28	F	168	50
39	CB39	30	M	175	105
40	CB40	21	M	180	69
41	CB41	48	M	182	78
42	CB42	35	M	186	83
43	CB43	32	M	168	50
44	CB44	25	M	180.3	60
45	CB45	26	M	175	68
46	CB46	35	F	160	50
47	CB47	23	M	165	85
48	CB48	23	F	165	65
49	CB49	41	M	175	78
50	CB50	23	F	159	60
51	CB51	23	F	170	66

APPENDIX F: LUNG INFECTION STATUS OF INDIVIDUAL CF SUBJECTS

Order	**Patient ID**	***Organisms***		
1	CF37	*Mucoid P. aeruginosa*	*S. aureus*	*A. fumigatus*
2	CF54	*Mucoid P. aeruginosa*	*A. fumigatus*	*A. niger*
3	CF14	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	*S. aurantiacum*
4	CF07	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	*S. aureus*
5	CF24	*Mucoid P. aeruginosa*	*A. fumigatus*	
6	CF46	*S. aureus*	*A. fumigatus*	
7	CF58	*Mucoid P. aeruginosa*	*A. xylosoxidans*	
8	CF33	*A. fumigatus*	*N. farcinica*	
9	CF15	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	
10	CF28	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	
11	CF30	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	
12	CF43	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	
13	CF45	*Mucoid P. aeruginosa*	*Non-mucoid P. aeruginosa*	
14	CF18	*Non-mucoid P. aeruginosa*	*S. apiospermum*	

15	CF57	*S. aureus*	*S. aurantiacum*	
16	CF11	*Non-mucoid P. aeruginosa*	*S. aureus*	
17	CF20	*Mucoid P. aeruginosa*	*S. aureus*	
18	CF39	*Non-mucoid P. aeruginosa*	*S. aureus*	
19	CF48	*Mucoid P. aeruginosa*	*S. aureus*	
20	CF53	*Mucoid P. aeruginosa*	*S. aureus*	
21	CF38	*S. aureus*	*S. maltophilia; S. aurantiacum; β-Lactamase (-ve) H. influenzae*	
22	CF13	*S. aureus*	*T. mycotoxinivorans*	
23	CF02	*Non-mucoid P. aeruginosa*		
24	CF21	*A. xylosoxidans*		
25	CF47	*A. xylosoxidans*		
26	CF03	*Mucoid P. aeruginosa*		
27	CF04	*Mucoid P. aeruginosa*		
28	CF06	*Mucoid P. aeruginosa*		
29	CF10	*Mucoid P. aeruginosa*		

30	CF12	*Mucoid P. aeruginosa*		
31	CF19	*Mucoid P. aeruginosa*		
32	CF22	*Mucoid P. aeruginosa*		
33	CF23	*Mucoid P. aeruginosa*		
34	CF25	*Mucoid P. aeruginosa*		
35	CF26	*Mucoid P. aeruginosa*		
36	CF27	*Mucoid P. aeruginosa*		
37	CF29	*Mucoid P. aeruginosa*		
38	CF35	*Mucoid P. aeruginosa*		
39	CF41	*Mucoid P. aeruginosa*		
40	CF44	*Mucoid P. aeruginosa*		
41	CF49	*Mucoid P. aeruginosa*		
42	CF50	*Mucoid P. aeruginosa*		
43	CF52	*Mucoid P. aeruginosa*		
44	CF60	*Mucoid P. aeruginosa*		
45	CF63	*Mucoid P. aeruginosa*		

46	CF64	*Mucoid P. aeruginosa*		
47	CF66	*Mucoid P. aeruginosa*		
48	CF05	*Non-mucoid P. aeruginosa*		
49	CF40	*Non-mucoid P. aeruginosa*		
50	CF42	*Non-mucoid P. aeruginosa*		
51	CF65	*Non-mucoid P. aeruginosa*		
52	CF16	*S. aureus*		
53	CF36	*S. aureus*		
54	CF51	*S. aureus*		
55	CF56	*S. aureus*		
56	CF67	*S. aureus*		
57	CF68	*S. aureus*		
58	CF17	*S. maltophilia*		
59	CF69	*T. mycotoxinivorans*		
60	CF31	None		
61	CF32	None		

62	CF34	None		
63	CF55	None		
64	CF59	None		
65	CF61	None		

APPENDIX G: DETAILS OF MEDICATIONS FOR INDIVIDUAL CF SUBJECTS

Order	***Patient ID***	***Antibiotics***			***Bronchodilators***		
					Short acting	***Long acting***	***Inhaled corticosteroids***
1	CF33	Azithromycin	Bactrim	Colomycin; Minocycline	Salbutamol	Flutiform	Alvesco
2	CF35	Azithromycin	Colomycin	Ciprofloxacin	Salbutamol	Symbicort	Alvesco
3	CF45	Azithromycin	Flucloxacillin	Colomycin	Salbutamol	Seretide	None
4	CF61	Azithromycin	Ceftazidime	Ciprofloxacin	Salbutamol	Flutiform	None
5	CF27	Azithromycin	Bactrim		Salbutamol	Flutiform	None
6	CF34	Azithromycin	Bactrim		Salbutamol	Seretide	None
7	CF63	Azithromycin	Bactrim		None	Flutiform	None
8	CF14	Azithromycin	Colomycin		Salbutamol	Seretide	None
9	CF15	Azithromycin	Colomycin		Salbutamol	Symbicort	None
10	CF18	Azithromycin	Colomycin		Salbutamol	Symbicort	None

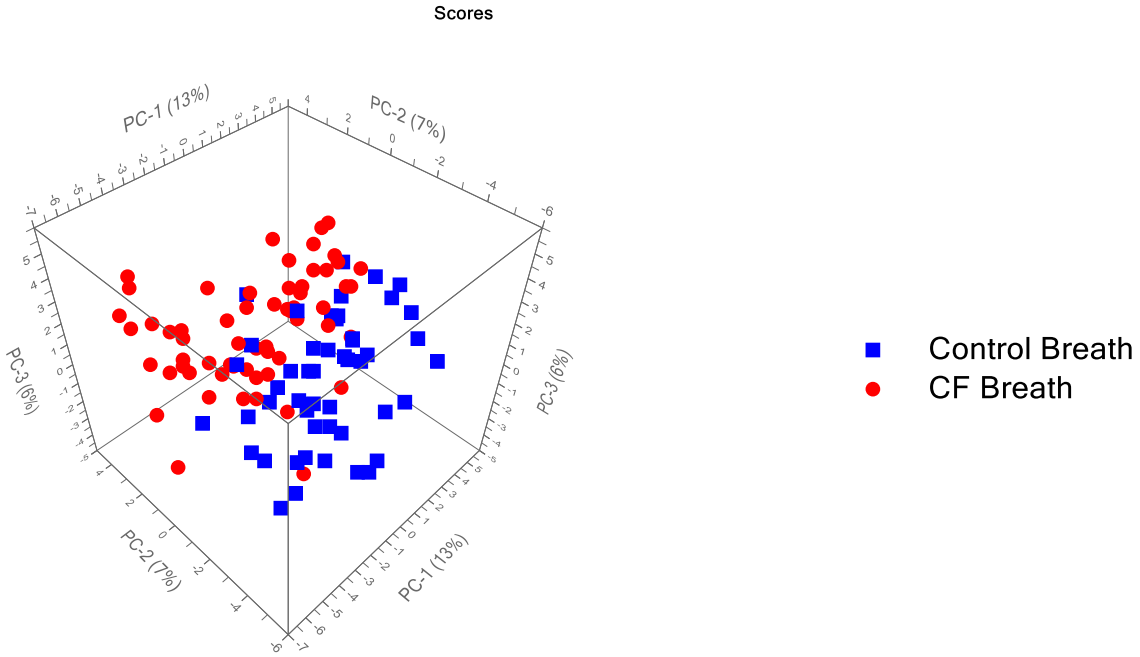
11	CF29	Azithromycin	Colomycin		Salbutamol	Symbicort	None
12	CF39	Azithromycin	Colomycin		Salbutamol	Flutiform	Alvesco
13	CF43	Azithromycin	Colomycin		Salbutamol	Seretide	None
14	CF50	Azithromycin	Colomycin		Salbutamol	Symbicort	None
15	CF65	Azithromycin	Colomycin		Salbutamol	Symbicort	None
16	CF66	Azithromycin	Colomycin		Salbutamol	Flutiform	None
17	CF07	Azithromycin	Flucloxacillin		Salbutamol	Seretide	None
18	CF42	Azithromycin	Flucloxacillin		Salbutamol	Seretide	None
19	CF44	Azithromycin	Flucloxacillin		Salbutamol	Seretide	None
20	CF03	Azithromycin	Tobramycin		Salbutamol	Seretide	None
21	CF41	Azithromycin	Tobramycin		Salbutamol	Symbicort	Alvesco
22	CF47	Bactrim	Tobramycin		Salbutamol	Seretide	None
23	CF55	Flucloxacillin	Tobramycin		Salbutamol	None	None
24	CF36	Flucloxacillin			Salbutamol	Flutiform	None

25	CF02	None			Salbutamol	Flutiform	None
26	CF04	Azithromycin			Salbutamol	Seretide	None
27	CF05	Azithromycin			Salbutamol	Symbicort	None
28	CF06	Azithromycin			Salbutamol	Seretide	None
29	CF10	Azithromycin			None	Symbicort	None
30	CF11	Azithromycin			Salbutamol	Symbicort	Nonr
31	CF12	Tobramycin			Salbutamol	Flutiform	None
32	CF13	Azithromycin			Salbutamol	Flutiform	Alvesco
33	CF16	Azithromycin			None	Flutiform	None
34	CF17	None			Salbutamol	Flutiform	None
35	CF19	None			None	Flutiform	None
36	CF20	Azithromycin			Salbutamol	Seretide	None
37	CF21	None			Salbutamol	Flutiform	None
38	CF22	Azithromycin			Salbutamol	Symbicort	None

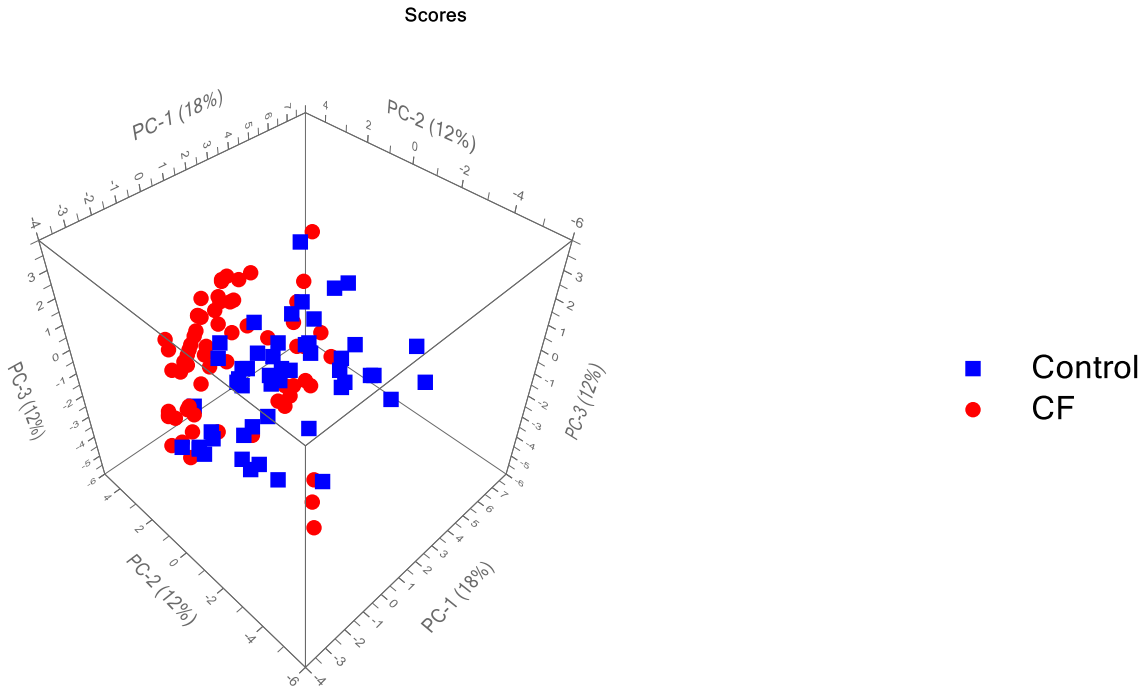
39	CF23	Azithromycin			Salbutamol	Seretide	None
40	CF24	Azithromycin			Salbutamol	Seretide	None
41	CF25	None			Salbutamol	Seretide	None
42	CF26	Azithromycin			Salbutamol	Seretide	None
43	CF28	Azithromycin			Salbutamol	Seretide	None
44	CF30	Azithromycin			Salbutamol	Flutiform	None
45	CF31	None			Salbutamol	Symbicort	None
46	CF32	None			Salbutamol	Symbicort	None
47	CF37	Rifaximin			Salbutamol	Seretide	None
48	CF38	Azithromycin			None	None	None
49	CF40	None			None	Symbicort	None
50	CF46	None			Salbutamol	Symbicort	None
51	CF48	Azithromycin			Salbutamol	Seretide	None
52	CF49	Colomycin			Salbutamol	Flutiform	None

53	CF51	None			Salbutamol	Flutiform	None
54	CF52	Tobramycin			None	None	None
55	CF53	Bactrim			Salbutamol	Seretide	None
56	CF54	Azithromycin			None	Flutiform	None
57	CF56	Azithromycin			Salbutamol	Flutiform	None
58	CF57	None			Salbutamol	Seretide	None
59	CF58	Bactrim			Salbutamol	Seretide	None
60	CF59	None			Salbutamol	Seretide	None
61	CF60	Azithromycin			Salbutamol	Symbicort	Alvesco
62	CF64	Azithromycin			Salbutamol	Seretide	None
63	CF67	Azithromycin			Salbutamol	Seretide	None
64	CF68	None			Salbutamol	Seretide	None
65	CF69	Azithromycin			Salbutamol	Flutiform	Alvesco

APPENDIX H: THREE DIMENSIONAL (3D) SCORES PLOT OF THE PCAS PERFORMED IN CHAPTER 4



Corresponding 3D scores plot of Figure 4.2: 3D PCA scores plot of breath profiles obtained from CF subjects and healthy controls. The plot was prepared using 81 VOCs detected in breath profiles.



Corresponding 3D scores plot of Figure 4.3: 3D PCA scores plot of breath profiles obtained from CF subjects and healthy controls. The plot was prepared using 26 VOCs which were statistically significant between study groups as determined using Student's *t*-test.