The potential use of SIFT-MS as a breath diagnostic tool

by

Alexa Hryniuk

A Thesis Submitted to the Faculty of Biology in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

Department of Biology
Lakehead University
Thunder Bay, Ontario

June 2009

Acknowledgements

I would like to thank my supervisor, Dr. Brian Ross, for his guidance and support over the past two years. He was always very helpful whenever I encountered a problem and kept me motivate throughout the thesis writing process.

A large thanks is given to Susie Ostrowski for collecting cancer samples at the Thunder Bay Regional Health Sciences centre for this project. Without her help I would not have had access to samples for the cancer pilot project. Many thanks are given as well to the Celiac Association of Thunder Bay for taking an interest in our celiac pilot study and their help in recruitment by advertising to their members.

I am greatful for all of the support from my fellow graduate students for all their encouragement and feedback throughout the past two years. Neal Esau, Ryan Lehto, Hilary Sparkes, Robin Lecraw, Lisa McKee, Eric Misener, Christina Mol and Cristina Spalvieri were also always good for lots of laughs and a few pints whenever a break was needed.

To my family I would like to thank you for your support in the journey to completing my thesis. I could always count on Mom and Dad for motivation especially near the end when they would ask me daily "Are you finished yet?" I would also like to thank my best friend Kate Leroux for letting me talk science to her even though she very rarely understood what I was talking about but she always helped me through the tough times and helped to celebrate the successes.

From this experience, I feel as though I've grown as a student and acquired valuable skills to use in my future endeavours.

Table of contents

List of Figures

List of Tables

List of Abbreviations

CLL	Chronic lymphocytic leukemia
CO_2	Carbon dioxie
CT	Computerized axial tomography
EGD	esophagogastroduodenoscopy
FS	Full scan
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
IgA	Immunoglobulin A
IMS	Ion mobility spectrometry
m/z	mass to charge ratio
MIM	Multiple ion monitoring
NO	Nitric oxide
NSCLC	Non small cell lung cancer
NV	Normalized Values
O_2	Oxygen
PTR-MS	Proton transfer reaction mass spectrometry
SCLC	Small cell lung cancer
SIFT-MS	Selected ion flow tube mass spectrometry
SPEM	Solid phase microextraction
tTG	Tissue transglutaminase
VOCs	Volatile organic compounds

Abstract

The measurement of volatile chemicals in human breath has recently emerged as a non-invasive technique with the potential for the early diagnosis of disease. A benefit of early diagnosis is that treatment plans including lifestyle changes can be implemented earlier, which will lead to improved patient outcomes. A novel gas analysis technique called selected ion flow tube mass spectrometry (SIFT-MS) offers a more rapid breath analysis process without the need for calibration. The SIFT-MS technique is currently being investigated as to its potential to be used as a diagnostic tool for various diseases. Two pilot studies were conducted, the first with lung cancer subjects and the second with celiac subjects, to determine if possible biomarkers could be distinguished in those with a disease compared to healthy control subjects using the SIFT-MS technique. Samples were taken and analyzed by scanning the entire sample for differences in the disease groups compared to the controls and by selecting specific compounds to monitor for a difference between the study groups that have been proposed as potential biomarkers for lung cancer or celiac disease. Outliers were discovered in the first pilot study between certain cancer subjects and the healthy controls suggesting a possibility for volatile biomarker. In the second pilot study outliers were also found for three ions from the breath sample of a newly diagnosed celiac patient compared to those who have been diagnosed for six months or more and the healthy control group. Unfortunately, the samples size of the groups in the pilot studies were too small for in depth statistical analysis to be performed. SIFT-MS analysis was also complicated by the low chemical resolution of the technique which causes a lack of ability to differentiate between similarly sized ions. For example, isotopologue effects and adduct ions produced by precursor ions at times interfered with the detections of ions of interest, causing

incorrect calculations in the amounts of ions detected. Overall the SIFT-MS has shown that it is a technique that has potential as a diagnostic tool, however, larger sample groups need to be studied for both the disease groups to determine if any of the ions that seem to be unique to cancers or celiac disease could be biomarkers and to alleviate some difficulties associated with the analysis by SIFT-MS.

1.0 Introduction

The background research that is published regarding the subject areas relevant to this project is presented. This includes information about diagnostics and the newly emerging diagnostic techniques using breath analysis. The advantages and disadvantages of breath analysis are discussed as well as some of the current methods employed. Next the new SIFT-MS method of breath analysis will be presented. This method is of interest because of the potential to use SIFT-MS as a diagnostic tool for disease. Lastly, information and diagnostics for lung cancer and celiac disease will be discussed as they are the two diseases examined in the pilot studies that were performed to determine the effectiveness of SIFT-MS as a diagnostic tool and search for potential biomarkers unique to lung cancer and celiac disease.

1.1 Diagnostics

Diagnostics are crucial for identifying the presence of disease not only at the individual level but also at the population level. Correctly assessing the nature of disease, determining the appropriate course of treatment, monitoring the effects of interventions whether they are preventative or therapeutic and determining recurrence of existing disease are important features of any diagnostic test (Hay Burgess *et al.* 2006). For a diagnostic test to be useful it needs to be accurate, simple, and affordable for the intended population. Diagnostic tests also need to yield results fairly rapidly, so effective control measures such as treatment can be initiated, which can limit or prevent long-term complications. Diagnostic tests have a multitude of uses: patient management, screening of asymptomatic patients, surveillance, epidemiological studies, and studying effectiveness of treatments (Peeling *et al.* 2006). The basic performance characteristics of a good diagnostic test are the sensitivity and

the specificity of the test. Sensitivity is the probability that a patient with a specific disease will have a positive result using a diagnostic test, whereas, specificity is the probability that a person without the disease will have a negative result using a diagnostic test to determine disease (Peeling *et al.* 2006). Specificity and sensitivity are usually determined against a reference standard test commonly called a "gold standard" test. A reference standard is the best available approximation of a true result, commonly indicating a test method that is currently accepted as accurate to diagnose for a specific disease and it is used as a reference method for assessing the performance characteristics of new diagnostic tests (Urdea *et al.* 2006). It is difficult to evaluate new potential diagnostic tests when no reference standard is available for comparison. Operational characteristics are also significant to a diagnostic test. These include the time it takes to perform the test, the technical simplicity or ease of use, user accessibility and the stability of the test under user conditions. The technical simplicity will depend on the straightforwardness of acquiring and maintaining the equipment for the diagnostic test, how difficult it is to train staff to use the test and to interpret the results correctly, and the stability of the test under the expected conditions of use (Banoo *et al.* 2006). All of these characteristics are important for determining the setting in which the diagnostic test can be used and the amount of training required by the staff. Current diagnostic tools are largely inadequate for meeting the health needs in many countries around the world. These shortcomings include the inability to distinguish between diseases that present with similar symptoms, the invasiveness of the majority of tests, the ability to provide low-cost monitoring for treatment, and monitoring of disease recurrence (Hay Burgess *et al.* 2006). Recent technological developments have led to the proliferation of new, rapid diagnostic tests that hold promise for the improved management and control of various

diseases. An example of new diagnostic development is through breath analysis (Banoo *et al.* 2006).

1.1.1 Breath Analysis

Breath tests are among the least invasive of the available methods used for clinical diagnosis, disease state monitoring and environmental assessment (Cao and Duana 2006). It is only recently that breath analysis has become a popular area of study for clinical purposes. Breath analysis is still a very underdeveloped method of diagnosis and not widely practiced in the clinical setting (Cao and Duana 2006). There are currently a few types of breath tests which are being used as diagnostic tools, these include tests for increased nitric oxide (NO) in diagnosing airway inflammation and a breath test for diagnosis of Helicobacter pylori infection (Cao and Duana 2006). Many other applications using the breath analysis technique that are still at various stages of development, including the detection and monitoring of renal failure, early diagnosis of cancer, and bacterial infections (Rolla *et al.* 2008; Poli *et al.* 2008; Enderby *et al.* 2009).

The majority of the air an individual exhales is a mixture of nitrogen, oxygen, carbon dioxide, water vapour and inert gases. The remaining small fraction of breath volume consists of more than 1000 trace volatile organic compounds (VOCs) with concentrations that range from parts per billion to parts per trillion in volume (Risby and Solg 2006). The origin of these volatile compounds may be endogenous, generated in the body or may be exogenous, absorbed as contaminants from the environment (Risby and Solg 2006). The composition of volatile organic compounds varies between individuals and although over 1000 volatile organic compounds have been identified only a few are common to all humans (Cao and Duana 2006). These common volatile organic compounds, which include acetone,

ethanol, methanol and isoprene are products of everyday (core) metabolic processes and therefore, have the potential to be very informative for clinical diagnostics (Mukhopadhyay 2004). The main gases and the trace volatile organic compounds in breath exchange between the blood and the alveolar air at the blood-gas barrier in the lung. A known exception is NO, which is released into the airway from epithelial cells in the case of airway inflammation (Cao and Dunna 2006).

1.1.2 Breath Diagnosis

The origin of the substances found in human breath, either endogenous or exogenous, determines the type of applications that breath tests can perform. The diagnosis and monitoring of disease is achieved through the analysis of endogenous volatiles and the assessment of exposure to drugs and environmental pollutants by the study of exogenous volatiles (Cao and Duana 2006). Endogenous markers, which are commonly used for diagnostic purposes, are hydrocarbons (ie. ethane , pentane, isoprene), oxygen containing compounds (ie. acetone, acetaldehyde, ethanol), sulphur-containing compounds (ie. dimethylsulfide, ethyl mercaptan, and carbon disulfide) and nitrogen containing compounds (ie. ammonia, dimethylamine) (Meikesch *et al.* 2004). The detection of volatile compounds in breath for the purpose of diagnosis is not a novel idea. In ancient Greek times, physicians learned to deduce that the various aromas of human breath could provide clues to diagnose disease. For example physicians who detected the sweet fruity odor of acetone could diagnose patients with uncontrolled diabetes or if they detected a urine-like smell then kidney failure was suspected. Modern breath analysis had its beginning in the 1970s when Pauling *et al.* identified more than 200 compounds in human breath using gas chromatography (GC) (Miekisch *et al.* 2004). At that time there were many problems with the effectiveness of the

separation and identification of exhaled substances. These problems were alleviated in the 1980s and the 1990s with technical advances in analytical methods. The focus of breath analysis is presently on the physiologic meaning of breath substances and correlation of breath markers with patients' symptoms (Miekisch *et al.* 2004). It is important to understand the relationship between various breath compounds and physical conditions.

1.1.3 Advantages and Limitations of Breath Analysis

Breath tests provide clues to the composition of blood and physiological processes occurring in the body. Human alveolar breath contains many biomarkers derived from the blood by passive diffusion across the pulmonary alveolar membrane. Breath testing offers a new approach to the diagnosis of disease and the evaluation of common disorders (Banoo et al 2006). New advances in technology have led to a growth in the field of breath analysis. Breath analysis and diagnosis have a variety of advantages over more traditional diagnostic methods.

Some advantages of breath testing include:

• Breath testing is a non-invasive method that is easily repeated, and does not have the discomfort or embarrassment associated with blood and urine tests.

• Breath samples closely reflect the blood concentrations of biological compounds and may eliminate the need for the collection of blood samples, which requires more staff training and has a risk of potential disease transmission. Breath analysis could be a particularly advantageous alternative to blood testing in regular monitoring a patient disease progression or treatment.

• The composition of breath is a much less complicated mixture than blood or urine and complete analysis of all compounds present can be performed. Many analyses of blood or

13

urine require a sample work up, whereas breath samples do not.

• Breath analysis allows for quicker return of results compared to other diagnostic test such as blood or biopsy samples.

• Breath analysis provides direct information about the respiratory function of an individual that cannot be obtained by other means.

• Breath analysis can monitor the decay of volatile toxic substances in the body in real-time, which can aid in understanding their pharmacokinetics.

Despite the undoubted appeal of breath analysis, breath testing also has limitations:

• A considerable problem pertaining to the use of breath tests in a clinical setting is the lack of standardization of analytical methods and the wide variation in results obtained in different studies.

• Reliable forms of sample collection are important as most volatile compound concentrations in exhaled breath range from parts per billion to parts per trillion by volume. Storage, the stability of the sample and contamination are problems that need to be addressed when taking a breath sample. Also different analytical methods used to analyze the samples collected may give very different results.

• Another important issue is the high water content of breath samples, which may affect the separation, and detection of single compounds.

• Currently instrumentation used for breath analysis is expensive. At present, the most common breath analysis method, gas chromatography-mass spectrometry (GC-MS), requires bulky instrumentation, is time-consuming, and needs skilled operators to obtain accurate results.

• The lack of established links between breath substances and disease is also a problem. If

this link is not understood, the potential for a breath test to give an incorrect diagnostic result is high, but the breath analysis can still be used as a part of a collection of diagnostics tests to determine an accurate diagnosis (Cao and Dunna 2006).

1.1.4 Analytical methods used for breath analysis

It is a technical challenge to detect and quantify trace gas analysis in humid breath with accuracy for medical diagnosis. The earliest work in breath analysis was carried out using gas chromatography combined with mass spectrometry. More recently, breath analysis is being explored using novel mass spectrometry methods, such as ion mobility spectrometry and proton transfer reaction mass spectrometry, to detect and quantify a wide variety of volatile compounds. All of these techniques have their strengths and weaknesses and all contribute to furthering the development of breath analysis as a tool for medical diagnosis.

The earliest and still widely used analytical technique of gas chromatography coupled with mass spectrometry (GC-MS) continues to help the growth of the field of breath diagnostics. GC-MS has the ability to separate and identify the compounds in multi-component mixtures such as exhaled breath. This separation is accomplished by using known elution times of the compounds from the GC column and their characteristic fragmentation (Basanta *et al.* 2007). However, to achieve quantification of the volatiles in exhaled breath, the GC-MS must be calibrated using known compounds. Therefore, the investigator needs to have specific compounds of interest that can be monitored and compared by scanning a sample and quantifying unusual increases and decreases in the volatiles for a particular study group. Also, sample collection followed by injection into the column is used for the examination of breath samples, which makes GC-MS unusable for real-time analysis. Solid phase microextraction (SPEM) is the most common technique for sample collection, which is

then followed by automatic thermal desorption onto the GC column (Basanta *et al.* 2007). This procedure allows for improved detection sensitivity but the accuracy is diminished by uncertainties in the collection and desorption efficiencies. Despite these limitations of the GC-MS, this technique remains a main method for breath analysis but essentially remains an off-line analytical technique that quantifies trace gas concentrations some hours after the breath samples are taken.

Recently, new mass spectrometry techniques have been developed for more versatile breath analysis, which attempt to overcome some of the past difficulties with GC-MS. Ion mobility spectrometry (IMS) is a method of identifying trace gas by the mobilities and drift velocities of their gas phase ions in a buffer gas (Ruzsanyi *et al.* 2005). The different masses of the compounds and their geometric structures provide unique drift velocities for the different ion species found in exhaled breath. The precursor ions are formed by a radioactive source or by gas electrical discharges. Precursors and the sample are introduced in the buffer gas where complex ion chemistry proceeds to form characteristic ions of the trace gases. The arrival times of the product ions at the detector are then measured to characterize the ions in the gas sample (Ruzsanyi *et al.* 2005). IMS provides high sensitivity and low technical expenditure to achieve high-speed data acquisition (10-50 ms). Thus, the IMS is ideal to for online analysis but, due to complex ion-molecule reactions and poor mass resolution, this method is not good for identification of unknown compounds. IMS is used to identify mobility peaks that relate to unknown compounds in the breath from patients with a particular disease compared to normal breath samples (Ruzsanyi *et al.* 2005).

For trace gas identification and quantification in a breath sample proton transfer reaction mass spectrometry (PTR-MS) makes up for the inadequacies of ISM to identify

unknown compounds. PTR-MS has the potential to quantify trace gas samples to sub-ppb

levels (Hewitt *et al.* 2003). The breath samples are introduced into a drift tube where they

react exclusively with the precursor ion H_3O^+ to produce ion products by proton transfer and

these product ions are detected at the end of the drift tube by an analytical mass spectrometer

(Hewitt *et al.* 203). Product ion identification can be relatively simple, but with only one

precursor some ions can be misidentified such as those that are isobaric compounds or those

product ions that spontaneously lose an H_2O molecule. The carrier gas in the PTR-MS is the

actual breath to be analyzed therefore it is very moist and possible cluster ions

($H_3O^+ \cdot (H_2O)_{1,2,3}$) could form. However, PTR-MS uses a drift tube in which the ions are

constrained to move under the influence of an axial electric field and this inhibits the

formation of cluster ions making the product ion spectrum easier to interpret (Hewitt *et al.*

2003). The disadvantage of this method is that the kinetics of the ion reactions and the

reaction time are not well defined causing a difficulty in accurately quantifying the volatile

compound detected. Thus, absolute quantification of potential markers is not possible using

the PTR-MS. While the technique has produced interesting qualitative data, PTR-MS is

unlikely to form the future basis for clinically useful tests.

1.2 SIFT-MS

A new technique for breath analysis, selected ion flow tube mass spectrometer (SIFT-

MS) has potential to be very beneficial in diagnosing disease because SIFT-MS can eliminate

several limitations that have plagued breath analysis in the past. Limitations that are

alleviated by SIFT-MS include long wait times for results, the difficulties associated with

high-water content in breath samples and overly technically demanding methods.

1.2.1 Background

The idea for the select ion flow tube technique was conceived and developed by Professor David Smith and associates to study reactions of compounds in the gas phase. In the beginning, the SIFT technique was used as a standard method for studying ion-neutral reactions such as interstellar gas phase reactions. The potential for the SIFT technique to be applied to the detection and analysis of trace gases was later realized through research conducted by Patrik Spanel and David Smith (Spanel and Smith 1996).

1.2.2 The principle of the SIFT-MS

SIFT-MS is a rapid analysis system that uses real-time detection. The SIFT-MS can detect trace volatile compounds in the presence of an abundance of much higher atmospheric gases by means of chemical ionization of these gases with charged precursor ions, normally H_3O^+, NO^+ and O_2^+ (Spanel et al. 1996; Smith and Spanel 2007). The precursor ions are created in a microwave discharge ion source. A specific precursor is selected by its mass-charge ratio (m/z) such as H_3O^+ with an m/z of 19 and extracted from a mixture of ionic species created in the source using a quadrupole mass filter. The selected precursor ions are then injected into a fast-flowing stream of inert carrier gas, usually pure helium, through a Venturi-type opening, into a flow tube approximately 30-100cm in length. The gas sample being tested is introduced into the flow tube through an air or breath sampler port just downstream from where the precursor ions are introduced (Spanel et al. 1996). The ions, either precursor or sample, do not fragment when they come into contact with the helium as long as they are injected at low energy. In the flow tube the precursor ions and the trace compounds in the sample react to produce product ions. The product ions flow towards the end of the flow tube where they exit via a pinhole approximately 0.3mm in diameter into a

differentially-pumped quadrupole mass spectrometer. After mass analysis the product ions are then quantified by an ion counting system. A mass spectra results and is recorded by an online computer for analysis and storage (Spanel *et al.* 1996; Smith and Spanel 2007). By knowing the reaction rate coefficient for the reaction between the precursor ion and the compound of interest, absolute quantification of the volatile compound can be analyzed in real time without the need for calibration standards (Spanel *et al.* 1996; Smith and Spanel 2007).

1.2.3 Applications of the SIFT-MS

The SIFT-MS can be employed to determine the rate coefficient for the reaction of the precursor ions with a known gas compound as well as the product ions of the reaction. To determine the rate coefficient, the known compound is introduced via the air sampler by a mass flow meter in controlled and measured amounts. The decline in the count rate of the precursor ions due to the reaction with the gas of interest as a function of the known gas flow rate into the flow tube is observed using the detection mass spectrometer and an ion counting system (Smith *et al.* 1998). The rate coefficient for the reaction is then calculated according to the procedure developed by Smith and Adams 1987. Many of the trace volatile compounds found in air and breath have already had their reaction kinetics calculated. A built-in database of rate coefficients and product ions obtained from many SIFT experiments has been developed as part of the analysis software for the SIFT-MS, which eliminates the need for calibration of the instrument.

Analysis of trace gas using the SIFT-MS is essentially the above mentioned in reverse. If the rate coefficient of the reaction between a precursor ion and the trace gas is known then amount of the trace gas in a sample of air or breath can be determined by the

SIFT-MS technique (Smith *et al.* 1998). The following is an example of the quantitative

analysis of the technique using H_3O^+ as a precursor ion and reacting it with acetone, a

common trace gas in exhaled breath. When the precursor ion, H_3O^+, reacts with acetone a

proton is transferred from H_3O^+ to the acetone molecule:

$$H_3O^+ + CH_3COCH_3 \longrightarrow H_2O + CH_3COCH_3 \cdot H^+$$
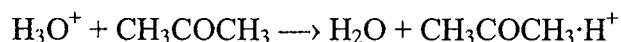
Proton transfers, like the one above, are called soft proton transfers. These reactions can

produce single or multiple ion products, but it is a simple procedure to determine the

branching ratios (Smith *et al.* 1998). This allows SIFT-MS to analyze complex mixtures of

trace gases and quantify the products. Therefore, when a breath sample is introduced into the

SIFT-MS, the soft proton transfer occurs which will result in some loss of H_3O^+ and the

production of protonated acetone ions, $CH_3COCH_3 \cdot H^+$. The count rate of the product ion is

directly proportional to the concentration of the trace gas, in this case acetone when it is

small enough to reduce the count rate of the primary ions slightly (Smith *et al.* 1998). This

factor is important for the SIFT-MS technique to be used in trace gas analysis because it

allows the concentration of various trace gases in a mixture that react with a product ion such

as H_3O^+, to be quantified from the observed count rates of each product ion species.

The calibration of the SIFT-MS is not necessary because the analysis software takes

into account all the necessary calculations when a sample is being introduced. The

introduction of a sample into the SIFT-MS allows a reaction to proceed resulting in some

loss in the amount of the precursor ions and the production of protonated product ions. The

SIFT-MS software uses the count rate of the precursor (I) at the downstream detector in the

presence of the ion being sampled and relates it to the count rate (I_0) in the absence of the ion

being observed to give the concentration of the ion being sampled [A]. The software uses to

the formula:

$$I = I_0 \exp - k[A]t = I_0 \exp - k[A] \frac{l + \varepsilon}{v_t}$$

where t is the time spent by the ion in the flow tube, l is the length of the reaction region of

the flow tube, ε is an end correction to l (about 1 cm), v_t is the ion flow velocity and k is the

reaction coefficient. The relationship between [A] in cm^{-3} and the flow rates of the helium

carrier gas, Φ_C, and the ion ,Φ_A, the carrier gas pressure, p_g (in torr) and the absolute

temperature, Tg, are monitored by the SIFT-MS software using the equation:

$$[A] = 3.54 \times 10^{16} p_g \frac{\Phi_A}{\Phi_C} \frac{273}{Tg}$$

where 3.54×10^{16} is the conversion coefficient between the gas pressure (torr) and the

concentration of molecules at 273 K. The SIFT-MS software holds all of the variables

constant while the sample is being taken to accurately determine the product count rates of

the ion to be detected ultimately giving a concentration of that ion in the sample without the

need for calibration standards.

1.2.4 Operational Modes

SIFT-MS instruments can operate in two modes: the full-scan (FS) mode and the

multiple ion monitoring (MIM) mode. The FS mode is a complete mass spectrum that is

obtained by having the detection mass quadrupole detect ions over a selected mass-charge

ratio range for a specific period of time, while a sample is being introduced into the

instrument at a steady flow. The count rates of the ions are calculated from the number of

counts and the total sampling time of each ion, then stored and displayed by an online

computer on a linear or a semi-logarithmic scale (Smith and Spanel 2005). The MIM mode

uses only the count rates of the precursor ions and specifically selected product ions to be

monitored in a gaseous sample introduced into the SIFT-MS. The count rates are achieved

by rapidly switching the detection mass spectrometer between the masses of all the precursor

ions and then product ions for a predetermined time interval (Smith and Spanel 2005).

Examples of the FS and MIM mode outputs can be seen in Figures 1 and 2.

1.2.5 Sampling procedures

Sampling procedures for the SIFT-MS vary depending on the type of gaseous sample

that is being tested. Ambient air is sampled by simply opening the sampler port and allowing

the air to flow into the instrument while the detection mass spectrometer is operating in either

the FS or the MIM mode (Spanel *et al.* 2006). FS and MIM spectra can also be obtained for

exhaled breath samples. Exhaled breath can be collected in collection bags or glass vessels,

which allows for extended sampling time until the sample is spent. Exhaled breath can also

be sampled by exhaling directly into the sampler port since online real-time analysis can

analyze several metabolites simultaneously with the MIM mode (Spanel *et al.* 2006). For the

individual donating the breath sample this is a simple and painless procedure that even those

in poor health can accomplish. Another method of sampling using the SIFT-MS is the

analysis of volatile compounds emitted by aqueous solutions such as urine, blood, food

products and cell cultures. These volatile compounds can be sampled by putting the liquid

sample in proximity to the sampler port where the sample and the ambient air will be tested

simultaneously or by containing the sample in a collapsible bag and allowing dry air or inert

gas into the bag to maintain atmospheric pressure (Spanel *et al.* 2006). For quantitative
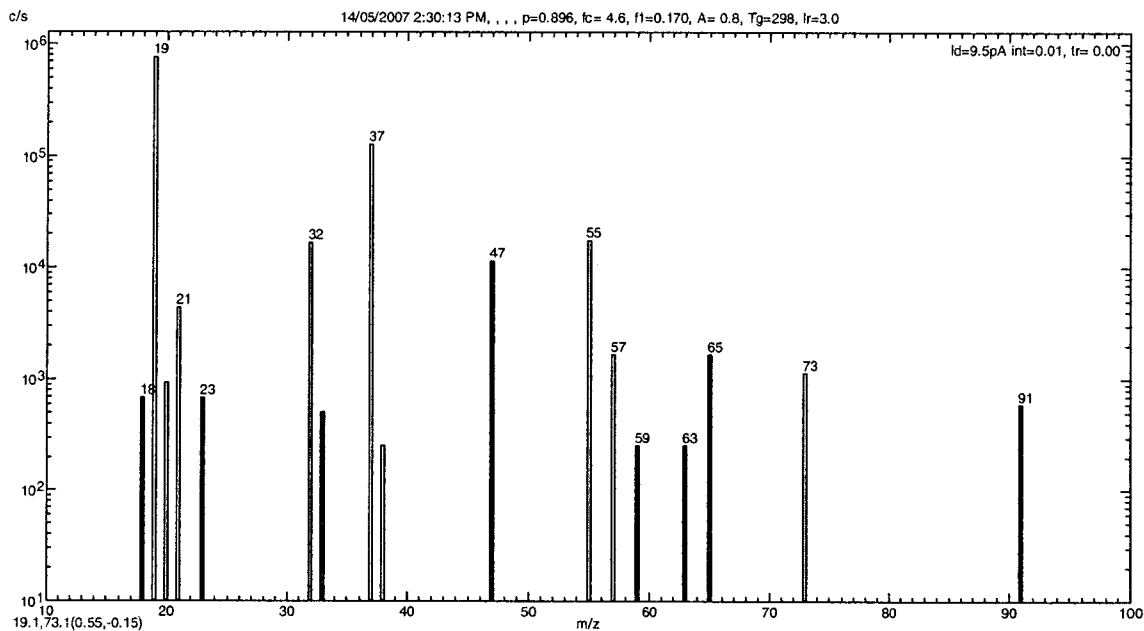
Figure 1. Example of a spectrum produced using the FS mode of the SIFT-MS. Ion signal intensities are given in counts per second, c/s, plotted against the mass-to-charge ratio, m/z. The numbers above the peaks refer to the ion mass of that peak in the spectra.
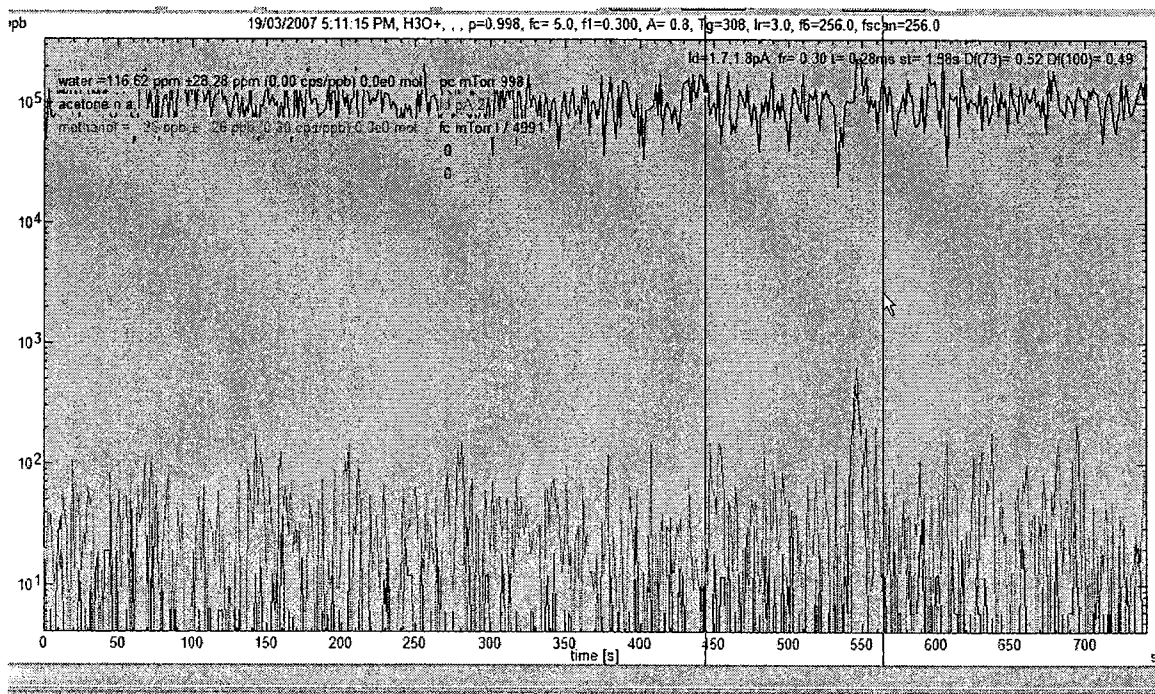
Figure 2. Example of a spectrum produced using the MIM mode of the SIFT-MS. Ion signal intensities are given in parts per billion, ppb, plotted against time, s. The compounds and their average detected over the period of time selected (interval between the two vertical lines) are given in the top left corner of the spectrum.

analysis of these aqueous samples, another method is to place the sample in a closed fixed volume container with a septum that is punctured by a needle attached directly to the sampler port of the SIFT-MS (Spanel et al 2006). All of these procedures can utilize the FS and the MIM modes to analyze the volatile compounds emitted by aqueous solutions.

1.2.6 Potential for Medical Diagnostics

The SIFT-MS technique offers the unique capability for the potential of early and rapid detection of a wide variety of diseases, infectious bacteria and patient conditions by differentiating between control and test groups. Diagnosis potential is achieved by identifying which masses and volatile compounds either diminish or increase in a particular disease. These compounds that are discovered contribute towards a successful classification of biomarkers in expired air for a particular disease state (Moorhead *et al.* 2007). The SIFT-MS technique uses real-time detection, which allows for faster analysis times and calibration using chemical standards is not required, as the concentrations of compounds under investigation are calculated using the known reaction rate as discussed earlier. Several studies have shown that the SIFT-MS has been successfully used to monitor normal physiology and pathophysiology of particular disease states. One notable result is the very high levels of ammonia that appear in the breath of patients with end-stage renal disease and the decrease in the ammonia to normal levels during haemodialysis sessions (Smith and Spanel 2007). The SIFT- MS technique would be very beneficial in pilot investigations as a non-invasive breath analysis tool to investigate physiological processes in humans, for clinical diagnosis and for therapeutic monitoring.

1.3 Pilot study #1

The first pilot study involved the collection of samples from lung cancer patients, which were compared to samples from healthy controls using the SIFT-MS method for breath analysis. Volatile compounds were monitored with the prospect of discovering compounds unique to those with lung cancer compared to those without that may potentially be useful as biomarkers for the disease.

1.3.1 Lung cancer

Lung cancer is a disease in which cells of the lung tissue begin to grow uncontrollably and form tumors. If these cells are not stopped from proliferating, tumor cells can grow to take over the lungs and metastasize to other areas of the body, which can lead to the eventual death of an individual.

1.3.2 Epidemiology

Lung cancer is the leading cause of cancer related death in Canada. It causes approximately 29% of cancer deaths in women compared to 22% among men (Jemal *et al.* 2008). There are important risk factors associated with lung cancer, the first being smoking. Cigarette smoke has been linked to over 80% of all cases of lung cancer. The risk of lung cancer rises with an increase in the number of cigarettes smoked per day as well as the number of years an individual has smoked (Toh 2009). Environmental factors are also associated with the potential development of lung cancer. Second hand smoke is a major risk factor for lung cancer in non-smokers since carcinogens are also present in the smoke inhaled by bystanders (Boffetta 2008). Other environmental risks that increase the occurrence of lung cancer are occupational exposure to substances such as asbestos, arsenic and polycyclic

hydrocarbons, a byproduct of fuel burning, and air pollution (Toh 2009). A family history of

lung disease is also a factor to consider when accessing an individual's potential of

developing lung cancer.

1.3.3 Signs and symptoms

Signs and symptoms of lung cancer vary depending on the individual and where and

how widespread the tumor is. The warning signs of lung cancer are not always present or

easy to identify. Approximately 25% of people who get lung cancer first discover the tumor

with a routine chest X-ray or CT scan as a small solitary mass called a coin lesion (Beckles *et

al.* 2003). These patients often report no symptoms at the time of cancer discovery. Other

individuals with lung cancer may exhibit nonspecific symptoms seen in many types of

cancers. Some of theses symptoms include weight loss, weakness, and fatigue.

Psychological symptoms are also common such as depression and mood changes (Beckles *et

al.* 2003). Symptoms specific to individuals with lung cancer that result from the invasion of

a tumour into the lung tissue may include the following: a cough, shortness of breath,

wheezing, chest pains, and coughing up blood (Spiro *et al.* 2007). If the cancer has invaded

nerves it may cause paralysis of the vocal cords leading to hoarseness or Pancoast's

syndrome, which is shoulder pain that travels down the outside of the arm. Tumor invasion

of the esophagus may lead to dysphagia, difficulty swallowing. The potential for collapse of

a portion of the lung may occur from a large tumor obstruction and cause infections such as

abscesses and pneumonia in the obstructed area (Spiro *et al.* 2007).

1.3.4 Pathology

Lung cancers, also known as bronchogenic carcinomas, are classified into two types: small cell lung cancers (SCLC) and non-small cell lung cancers (NSCLC) (Kalemkerian et al 1994). The classification is based upon the microscopic appearance of the tumor cells. SCLC and NSCLC grow and spread in different ways and may have different treatment options, therefore a distinction between these two cancer types is important.

SCLC comprise about 20% of lung cancers and are the more aggressive and rapidly growing of all lung cancers (Stahel 1994). SCLC often starts in the bronchi near the center of the chest, and they tend to spread widely throughout the body fairly early in the course of the disease, usually before symptoms appear. SCLC occurrence is strongly related to cigarette smoking, with only 1% of SCLC tumors occurring in nonsmokers (Stahel 1994). SCLC metastasize rapidly to many sites within the body and are most often discovered after they have spread extensively.

NSCLC account for about 80% of all lung cancers and are the most common type of lung cancers. NSCLC are slow growing tumors that are fairly non-aggressive (Ihde and Minna 1991). NSCLC can be divided into three main types that are named based upon the type of cells found in the tumor: Adenocarcinomas are the most commonly seen type of NSCLC. Adenocarcinomas are developed by both smokers and nonsmokers. Most adenocarcinomas arise in the outer, or peripheral areas of the lungs. Squamous cell carcinomas are the second most common NSCLC, accounting for approximately 30% of NSCLC. Squamous cell cancers arise most frequently in the central chest area in the bronchi. Lastly, large cell carcinomas, sometimes referred to as undifferentiated carcinomas, are the least common type of NSCLC. Large cell carcinomas start in the bronchi and are strongly

associated with smoking. Tumor mixtures of these different types of NSCLC can also be seen (Ihde and Minna 1991).

Other types of cancers can also arise in the lung, although these types of cancer are much less common than NSCLC and SCLC. Bronchial carcinoids account for up to 5% of lung cancers (Kalemkerian 1994). Typically, bronchial carcinoids are small tumors of 3-4 cm or less when discovered and occur most commonly in people under the age of 40. Carcinoid tumors are unrelated to cigarette smoking (Kalemkerian 1994). These tumors generally grow and spread more slowly than bronchogenic cancers, and most are detected early enough for surgical resection resulting in long-term remission.

1.3.5 Diagnosis

The diagnosis of lung cancer begins by using a chest X-ray or a computerized axial tomography scan (CT) to take images of the lungs to look for lesions or tumors that may indicate that an individual has lung cancer. These initial screenings are carried out when an individual exhibits symptoms of lung cancer although 25% of lung cancers are found when individuals receive chest X-rays or CT scans for an unrelated problem (Bach et al 2003).

The chest X-ray is the first diagnostic step when an individual presents with symptoms associated with lung cancer. The chest X-ray procedure often involves a view from the back to the front of the chest as well as a view from the side. Chest X-rays, like any X-ray procedure, expose the patient briefly to a minimum amount of diagnostic radiation (0.06 mSv) (Bach *et al.* 2003). Chest X-rays may reveal suspicious areas in the lungs such as calcified nodes or benign tumors that may mimic lung cancer but it is not possible from these images alone to determine if these areas are cancerous.

The second diagnostic step, a CT scan of the chest may be ordered when X-rays do not show an abnormality or do not yield sufficient information about the extent or location of a tumor. CT scans generate cross- sectional images of the body using X-ray procedures that combine multiple images with the aid of a computer. The images are taken at different angles around the body by a large donut-shaped X-ray machine (Manser *et al.* 2005). One advantage of CT scans is that they are more sensitive than standard chest X-rays in detecting lung nodules. Sometimes a contrast material is given intravenously prior to the CT scan to help define the organs and their positions (Manser *et al.* 2005).

Finally, to gain an accurate diagnosis of lung cancer more invasive techniques need to be utilized. These include broncoscopy and thoracic needle biopsy. Broncoscopy is the examination of the airways by visualization with a thin, fiberoptic probe that is inserted through the nose or mouth into the lungs (Kim *et al.* 2004). This procedure can be uncomfortable and it requires the patient to be sedated or anesthetized. A broncoscopy may reveal areas of tumor growth that can be biopsied for diagnosis of lung cancer by a pathlogist. This technique is especially useful for testing possible tumors in the central lung areas or in the larger airways (Kim at al. 2004). When a tumor is adequately visualized and adequately sampled by broncoscopy, an accurate cancer diagnosis is usually possible.

Another diagnostic technique for lung cancer is thoracic needle biopsy. Thoracic needle biopsy uses fine needle aspiration through the chest wall into the lower lobes of the lung to retrieve cells for diagnosis from tumor nodules in the lungs (Westcott *et al.* 1997). This technique uses radiological imaging for guidance and a small amount of local anesthetic is given at the insertion point prior to the procedure. Needle biopsies are particularly useful

in diagnosing lung cancer when the suspected tumor is peripherally located in the lung or not accessible to sampling by broncoscopy (Westcott *et al.* 1997).

1.3.6 Treatment

Treatment for lung cancer can involve surgical removal of the cancer, chemotherapy, or radiation therapy, as well as combinations of these treatments (Molina *et al.* 2008). The decision about which treatments will be the most effective for a given patient must be determined by the localization and extent of the tumor as well as the overall health status of the patient.

1.3.7 Current approaches of breath diagnostics for lung cancer

Researchers have been investigating the possibility of using volatile biomarkers to detect lung cancer in a clinical setting. Various studies have been carried out to discover unique volatiles in the breath samples from patients with lung cancer compared to those without. In 1985, Gordon *et al.* used GC-MS to identify several compounds unique in the expired air of those with lung cancer. Since then a number of other studies have been carried out in order to characterize specific compounds in breath samples from patients with and without lung cancer (O'Neil *et al.* 1988; Petri *et al.* 1988; Philips *et al.* 1999; Deng *et al.* 2004; Poli *et al.* 2005; Yu *et al.* 2005; Phillips *et al.* 2006; Chen X *et al.* 2007; Song G *et al.* 2009). From these studies many combinations of compounds have been proposed as biomarkers. Common volatiles suggested throughout the research for detection of lung cancer are various aldehydes, specifically formaldehyde, acetaldehyde, hexanal and heptanal, propanol, isoprene and alkanes. Also considerable evidence supports the hypothesis that oxidative stress may be associated with lung cancer and the increase in the byproducts of

30

these reactions may be indicative of lung cancer (Phillips *et al.* 1999; Deng *et al.* 2004; Chen X *et al.* 2007). Byproducts of oxidative stress include aldehydes, isoprostanes, hydrocarbons and hydroperoxides (Halliwell and Gutteridge 2003). As well, a recent study involving the SIFT-MS has shown acetaldehyde being detected into the headspace of lung cancer cells in vitro (Smith *et al.* 2003). All of these previous studies have discussed the potential for various volatiles to be biomarkers for lung cancer but the limitation in this research is that only a few clinical samples were studied. From the volatiles mentioned in the past studies for breath diagnosis of lung cancer, a list of target compounds was established to specifically monitor compounds with the SIFT-MS from the lung cancer samples taken. The target compounds consist mostly of aldehydes since they seem to be a prominent group of potential biomarkers discussed most of the papers. In addition to potential target compounds, a more speculative total spectrum of ions over the range of 10 to 200 m/z was also obtained to view any possible changes between the cancer subjects and healthy individuals. Identification and detection of a biomarker with SIFT-MS for the purpose of early non-invasive detection of lung cancer would help to get those affected into treatment programs earlier and reduce the mortality of the disease.

1.4 Pilot Study #2

The second pilot study involved comparing breath samples from those with celiac disease to samples taken from a healthy control group using the SIFT-MS as the method for breath analysis. Volatiles were monitored for differences between the two groups with the potential for specific markers to be identified that would identify those with the disease compared to those without.

1.4.1 Celiac Disease

Celiac disease, also known as gluten sensitive entropy and nontropical sprue, is a prevalent autoimmune disorder that is triggered by the ingestion of gluten. The immune response to the gluten causes inflammation and villous atrophy in the small intestine resulting in severe malnutrition in susceptible individuals.

1.4.2 Epidemiology

Celiac disease was once considered a rare disease of the pediatric population. It is now believed that one percent of the population in the United States, as well as other countries around the world, have celiac disease (Alaedini *et al.* 2005). While celiac disease can present at any age, it is more predominant in the female population compared to the male population by a ratio of 3 to 1 (Briani *et al.* 2008). The factors involved in the development of celiac disease are the consumption of gluten products in conjunction with a genetic predisposition for the disease. It is not completely understood how gluten sensitivity begins, however, the majority of patients suffering with celiac disease express human leukocyte antigen HLA-DQ2 or HLA-DQ8, which facilitates the immune response to gluten (Briani *et al.* 2008). These cells present the immunogenic gluten peptides to T-cell in the small intestine, which cause an inflammatory response, damaging villi.

1.4.3 Signs and symptoms

Celiac disease presents differently in every individual. Symptoms of celiac disease may appear as early as infancy, but celiac disease can develop in late adulthood. Symptoms may or may not occur in the digestive system. For example, one individual might suffer

from chronic diarrhea and abdominal pain while another individual would have depression

and irritability (Alaedini *et al.* 2005).

Classical symptoms in adults include:
- Recurring abdominal pain and bloating
- Chronic diarrhea
- Weight loss
- Pale, foul smelling stool
- Unexplained anemia
- Gas
- Bone pain
- Depression
- Behaviour changes
- Muscle cramps
- Joint pain
- Tingling and numbness in the legs
- Missed menstrual periods
- Dermatitis herpetiformis
- Aphthus ulcers

Classic symptoms in children include:
- Delayed growth
- Failure to thrive as infants
- Seizures
- Tooth discoloration or loss of enamel
- Earaches
- Irritability
- Mental development issues

Most of the symptoms are related to the malnutrition of the individual since they are not

getting the appropriate nutrients due to the destruction of the villi in the small intestine from

the reaction to the ingestion of gluten. Some people with celiac disease may not have any

symptoms (Alaedini *et al.* 2005). This can happen because a portion of the small intestine is

undamaged and able to absorb enough nutrients to prevent symptoms. However, these

people may develop secondary diseases such as arthritis, osteoporosis, thyroid disease, type 1

diabetes, liver disease, colon vascular disease and Sjogren's syndrome instead of the

traditional symptoms of celiac disease (Alaedini *et al.* 2005).

33

1.4.4 Pathology

In patients with celiac disease, the proximal mucosa of the small intestine becomes abnormal when gluten is ingested. This abnormality in the small intestine is characterized by stunted villi or the absence of villi, called flat mucosa (Polanco 2008). With villous atrophy there is elongation and hyperplasia of intestinal crypts and cells of the villi appear cuboidal instead of columnar. Also the proliferative compartments of the crypts expand causing an increased rate of epithelial cell loss in the intestine (Polanco 2008). It is believed that the innate immune system is the precursor to the mucosal changes that are viewed in celiac disease. The antigen presenting cells expressing HLA-DQ2 and HLA-DQ8 have an increased affinity for the gluten peptides. After binding with gluten peptides the antigen presenting cells can activate CD4+ T cells specific to gluten in the lamina propria of the small intestine. Activation of the T cells is accompanied by the production of cytokines, which in turn produces inflammation and villi damage by the release of metalloproteinases by fibroblasts and inflammatory cells (Briani *et al.* 2008; Alaedini *et al.* 2005). All intestinal structural damage resolves on gluten withdrawal but can recur if gluten is reintroduced to the diet.

1.4.5 Diagnosis

No single test has been approved as the standard for diagnosis of celiac disease. However, serum antibody testing for endomysial antibody immunoglobulin A and small bowel biopsy have been shown to be highly sensitive and specific in making a correct diagnosis, especially in patients with symptoms indicating celiac disease and those with increased risk such as individuals with a family history of celiac disease or an associated autoimmune disorder (Koning *et al.* 2005). Diagnostic testing must be done while the patient

is still on a gluten diet because the small intestine will regenerate villi when gluten is not ingested and therefore diagnosis will not be accurate. Patients with persistent gastrointestinal symptoms such as diarrhea, malabsorption, unexplained weight-loss, and abdominal pain should be evaluated for celiac disease, as well as patients presenting with premature osteoporosis, unexplained anemia and unexplained liver abnormalities. Currently patients who are asymptomatic are not recommended for screening, because in this group a higher number of false positives arise, leading to unnecessary testing, and psychological, emotional and financial burdens (Koning *et al.* 2005).

The first test performed in the diagnosis pathway for celiac disease is the serologic test. Serology tests study the fluid portion of the blood, serum, for its antibody content. The most common antibodies used in serologic testing for celiac disease are serum immunoglobulin A (IgA) endomysial antibodies and IgA tissue transglutaminase (tTG) antibodies. Most studies have found the sensitivity and specificity of using IgA endomysial and tTG antibodies to be greater then 95% (Setty *et al.* 2008). However, the sensitivity depends of the mucosal involvement, meaning the degree to which the villi are atrophied. TTG antibodies are the autoantigens recognized by the endomysial antibody and therefore it is not necessary to test for both IgA and tTG antibodies. The tTg serum test is less costly as it uses an enzyme linked immunosorbent assay and is recommended as the single serologic test for celiac disease in primary care facilities (Setty et al 2008).

Testing for HLA phenotypes DQ2 and DQ8 may also be useful in determining a diagnosis of celiac disease. HLA-DQ2 or HLA-DQ8 or both are found in 99% of the population of celiac patients and in less then 40% of the non-celiac disease population

(Briani *et al.* 2008). If these genetic markers are absent then it is very unlikely that celiac disease is present.

Another issue with only using the serum test for celiac disease diagnosis is the possibility of celiac patient also having IgA deficiency. IgA deficiency occurs when immunoglobulin A (IgA), a type of antibody that protects against infections of the mucous membranes lining the mouth, airways, and digestive tract is not present (Setty *et al.* 2008). The serum levels of IgA are undetectable in the presence of normal serum levels of IgG and IgM. IgA deficiency is the most common of the primary antibody deficiencies. If a patient has celiac disease, but suffers from this antibody deficiency, false negatives can occur. It is recommended to test for IgA deficiency if serum tTG is negative and celiac disease is highly suspected. Even though the serum testing is very specific and sensitive, because celiac disease does not have a high prevalence in North American society (approximately 1 percent of the population) the occurrence of a false positive result is very high (Setty *et al.* 2008). Therefore, due to the possibility for both false positive and false negative results serum tests cannot be relied upon as a diagnostic test for celiac disease. However, a positive marker can indicate the need for further testing with small bowel biopsy or capsule endoscopy and negative markers should never prevent further testing if suspicion of celiac disease is high.

Despite the increase in sensitivity of serum testing, a small bowel biopsy is required to confirm the diagnosis of celiac disease in the majority of patients. Biopsy is recommended for individuals who have a negative result from the serological test or if a physician strongly suspects celiac disease (Presutti *et al.* 2007). Small intestinal biopsy is performed by taking a small portion of the lining of the small intestine either by colonoscopy or esophagogastroduodenoscopy (EGD). During a colonoscopy, a flexible tube with fiberoptic

capabilities and surgical attachments called an endoscope is inserted through the rectum, through the colon, and into the end part of the small intestine, whereas, during an EGD, the endoscope is inserted through the mouth or nose and into the upper gastrointestinal tract to take a small sample of tissue (Presutti *et al.* 2007). Small bowel biopsies will demonstrate histopathological finding in the small intestinal mucosa such as intra-epithelial lymphocytosis, mucosal changes that may vary from partial to total villous atrophy, subtle crypt lengthening or increased epithelial lymphocytes for a positive diagnosis of celiac disease. These characteristics may be patchy and mucosal abnormalities on intestinal biopsy may be missed (Presutti *et al.* 2007). Therefore it is recommended that at least four tissue samples are obtained to avoid false-negative results from the endoscopic biopsy.

Endoscopic biopsy is neither 100 percent sensitive nor totally specific and other clinical abnormalities such as infection, enteritis, bacterial growth and lymphoma may have a similar histopathological appearance to celiac disease. Under circumstances in which the small bowel biopsy is questionable, performing a capsule endoscopy may be helpful in distinguishing intestinal abnormalities in the jujenum or ileum beyond the reach of the standard endoscope. Capsule endoscopy utilizes a capsule containing a camera the size and shape of a pill to visualize the gastrointestinal tract (Culliford *et al.* 2005). The capsule is ingested and eight abdominal leads are attached to the upper, mid and lower abdomen and a belt containing a data recorder and a battery pack is attached to the waist of the patient. Approximately 8 hours after ingestion the data belt recorder is removed and the data is transferred to a computer where it can be visualized (Culliford *et al.* 2005). Capsule endoscopy has been shown to detect major pathologic changes of the small intestine but may not detect minor changes which resemble normal villous patterns. More testing is needed to

identify if capsule endoscopy could play a role in diagnosing celiac patients with positive

serologic test but negative biopsies (Spada *et al.* 2008).

1.4.6 Treatment

There is no cure for celiac disease. The nutrient malnutrition and symptoms of celiac

disease can be averted by adhering to a gluten free diet which means avoiding eating all

foods that contain gluten (See and Murray 2006). For the majority of people, following a

gluten free diet will stop symptoms, heal existing intestinal damage and prevent further

damage. Improvements begin almost immediately after starting the gluten free diet and the

villi of the small intestine regenerate and regain their normal function in 3 to 6 months (See

and Murray 2006). The gluten free diet needs to be maintained for the rest of the celiac

sufferer's life. Eating even a small amount of gluten will result in damage to the intestine.

Depending on a person's age at time of diagnosis, some problems such as delayed growth

and tooth deterioration may not improve (Niewinski 2008). There are also a small

percentage of people with celiac disease who do not improve on the gluten free diet. These

individuals often have severe damage to their small intestine that does not heal even after

gluten has been eliminated and will need supplements because they are still not absorbing

sufficient nutrients. Drug treatments are currently being investigated to help those with

unresponsive celiac disease (Niewinski 2008).

1.4.7 Current approaches of breath diagnostics for celiac disease

Currently there are no other studies being conducted on the topic of breath diagnostic for

celiac disease. However, other gastrointestinal problems, such as bacterial infections, have

been investigated to determine if a breath test could be generated to distinguish those with an

infection from those without. A ^{13}C-urea breath test in the detection of *Helicobacter pylori* infection has been developed and is currently in clinical use (Gisbert and Pajares 2004). Therefore, if a method for the detection of a bacterial infection in the intestines can be developed there is potential for a gastrointestinal disease to also be detected in the same manner. In celiac disease, due to the autoimmune response to the ingestion of gluten, carbohydrates are incompletely or not absorbed in the small intestine and then are fermented by intestinal bacteria. Carbohydrate fermentation in the gut can cause unwanted gastrointestinal effects, especially in those who ingest a large amount of carbohydrates (Grabitske and Slavin 2008). These gastrointestinal effects include many of the symptoms seen in those with celiac disease. The fermentation by the various bacteria that inhabit the intestines produces various alcohols and short chain organic acids (Mandelstam *et al.* 1986). Therefore, for this pilot study various alcohols were combined into a database to be specifically monitor by SIFT-MS for their amount in the breath of celiac patients compare to healthy controls. If volatile compounds are found to distinguish between people with and without celiac disease it could lead to a more rapid diagnostic test which will allow for those with the disease to gain access to treatment more rapidly.

2. Objectives

The purpose of this study was to determine if the novel technique selected ion flow tube mass spectrometry (SIFT-MS) could be employed as a diagnostic tool for breath analysis of various diseases and to determine if potential biomarkers could be detected for lung cancer and celiac disease. In pilot study #1, the SIFT-MS will use full scans of breath samples from lung cancer patients and compare those to samples from healthy individuals and use the ion monitoring technique to monitor the amount specific aldehydes, isoprene and propanol between the sample groups. In pilot study #2, samples from individuals suffering with celiac disease will be compared to samples from healthy control subjects using the SIFT-MS full scan mode and the amounts of various alcohols will be monitored and compared between the two sample groups.

3. Hypothesis

The SIFT-MS will be a beneficial technique as a diagnostic tool for the detection of disease as it can detect differences in the concentration of volatiles at the ppb level. SIFT-MS will aid in the identification of compounds unique to lung cancer and celiac disease that can be used as biomarkers to distinguish those with or without a disease. Aldehydes such as acetaldehyde, hexanal and heptanal will be increased in the samples from those with lung cancer compared to samples from the healthy controls. Whereas, various alcohols should be higher in sample from celiac patients compared to samples from healthy individual due to the excess fermentation of unabsorbed carbohydrates in the intestines of celiac sufferers. Also the full scan from both studies may reveal other ions that differ between the disease groups and the healthy individuals that may not yet be known as potential biomarkers.

4. Methods

4.1 Ethics

Before beginning the cancer and the celiac studies, ethics approval was obtained for both studies. For the cancer research study, ethics approval was granted by the Thunder Bay Regional Sciences Centre Ethics Team in August 2007 and ethics approval was also granted by the Research Ethics Board at Lakehead University on October 9, 2007. Ethics approval was granted for the celiac research study by the Research Ethics Board at Lakehead University on May 18, 2008. The subjects were asked to provide samples of their breath. This is an entirely non-invasive procedure, which carries minimal risk to the subjects enrolled in the study. The risk of cross-infection between study subjects was very low since disposable collection devices were used throughout the study. There was no benefit to the subjects who took part in these studies but the studies may lead to a distinct biomarker or biomarkers that can be used for non invasive detection of lung cancer and celiac disease. A final report was submitted to the Research Ethics Board upon completion of the projects.

4.2 Subject selection

Subjects were recruited by invitation or advertisement (see Recruitment) according to the following inclusion and exclusion criteria for lung cancer and celiac disease. Each subject was asked to complete a questionnaire that collected required information, which confirmed that they were within the inclusion and exclusion criteria.

Lung cancer:

The inclusion criteria required that the subjects be over 40 years of age and have a diagnosis of lung cancer or, for the control group, be exhibiting no clinical symptoms of the disease. The

exclusion criteria required that the subjects have not been diagnosed with a pulmonary disease or any other major disease such as heart disease or diabetes.

Celiac Disease:

The inclusion criteria required that the subjects have celiac disease or gluten intolerance or, for the control group, not to exhibit any clinical symptoms of celiac disease or gluten intolerance. The exclusion criteria required that the subjects have not been diagnosed with a gastrointestinal disease or any other major disease.

4.3 Recruitment

Lung Cancer

Persons with lung cancer were recruited by invitation by Dr. Gehman at the Thunder Bay Medical Centre. Unfortunately due to low recruitment rates and staff changes Dr. Gehman's office no longer wished to participate in recruiting lung cancer patients. To continue recruiting, Dr. Vergidis (Northern Ontario School of Medicine and Thunder Bay Regional Health Science Centre) and his nurse Susie Ostrowski from the Thunder Bay Regional Sciences Centre began to recruit cancer patients at the Northern Ontario Cancer Centre. Because of the variety of cancers treated at the centre, the focus changed from specifically recruiting patients with lung cancer to recruiting all cancer patients that fit the inclusion criteria to potentially discover similar volatile compounds being expelled by the cancer subjects versus healthy control subjects. Subjects who fit the criteria had the research project introduced to them and then were given a copy of the Subject Information Sheet to read. The researcher then went through the Subject Information Sheet with the subject, explaining verbally the information contained in the sheet. The subject was asked if they understood the information presented to them and whether they wished to ask any questions. If the subject wanted to take part then they were asked to sign and date the

consent form witnessed by the researcher and fill out a short questionnaire. The subject retained the copy of the Subject Information Form to take away with them and if for whatever reason they wished to withdraw from the study the patient could contact the researchers from the contact information on the Subject Information Sheet.

To try to age and gender match those that gave cancer samples with the healthy control group recruitment for the healthy controls took place at the Thunder Bay 55 Plus Centre. Laurie Biggar, the supervisor for the centre, was contacted about the study and once she had reviewed all the information about the research, she allowed recruitment at the centre. A booth was set up in the main lobby of the Thunder Bay 55 Plus Centre one day a week and if people were interested in participating the research project they received the Subject Information form to read and keep, signed the consent form witnessed by the researcher and filled out the questionnaire. The subjects in the healthy control group could also withdraw at any point by contacting the researchers.

Celiac Disease

Celiac sufferers were recruited by advertisement to the Thunder Bay chapter of the Canadian Celiac Association at chapter meetings, on their website and in their newsletter. Anyone suffering from celiac disease could participate in the research project. If subjects were interested they contacted the researchers. The subjects were given information on the study and were asked to sign a consent form and fill out a questionnaire. Healthy control subjects were recruited by invitation by the researchers to try to age and gender match the subjects. The healthy control subjects were given all the same information as the celiac subjects and were also asked to sign a consent form and fill out a questionnaire. All subjects could withdraw any time from the study by contacting the researchers.

4.4 Sample collection

Samples in the cancer study were collected from 12 participants (7 women and 5 men) with 9 healthy control participants (5 women and 4 men). Samples in the celiac study were collected from 10 participants (4 women and 6 men) with 10 healthy control participants (5 women and 5 men). Sample collection for all subjects was conducted using the same approach. Subjects were asked to sit and breath normally for about a minute. Subjects were then asked to inhale normally and then exhale fully into a disposable 3L Tedlar collection bag via a disposable PVC tube. The subject exhaled into the bag until approximately 2.5L of gas had been collected. If after one breath, the collection bag was not filled, then the subject was asked to exhale into the bag again until the desired amount was obtained. It was noted how many breaths were needed to fill the collection bag by each participant. The bag was then sealed and returned to the laboratory for analysis. A second 1L bag was inflated with ambient air using a BioVOC Breath Sampler with a volume of 130ml for comparison with levels of the same compounds found in breath to control for the effect of ambient air contamination. The BioVOC Breath sampler was filled by pulling back on the plunger, which then was attached to the collection bag and the ambient air was forced into the bag by the plunger. This procedure was repeated until the required volume had been collected. In the laboratory the bags were heated to 40°C for 20 minutes in an incubator to minimise condensation effects. The time between sample collection and analysis varied between half and hour and 3 hours. This was due to the time it took to pick up and transport the sample from the collection site to the lab and how quickly I was contacted after the sample collection took place. Samples for the cancer subjects were collected at either the Thunder Bay Medical Centre or at the Northern Ontario Cancer Centre. Healthy control subjects for the cancer study were collected at the Thunder Bay 55 Plus Centre. All celiac

subjects and healthy controls for the celiac study were collected in the laboratory at the Northern Ontario School of Medicine building at the Lakehead University campus. Chemical levels are determined by the analysis of mass spectra, using the SIFT-MS, derived from the analysis of each sample.

4.5 SIFT-MS

The SIFT-MS analysis was run using a Profile 3 instrument from Instrument Science, Crewe, UK. Analysis of each breath sample was conducted in two ways. The first analysis was conducted in the full-scan (FS) mode while the second analysis was conducted using the multiple ion monitoring (MIM) mode. Before analyzing a sample the SIFT-MS was warmed up for approximately 30 minutes. All instruments settings, injection and detection settings, were checked and formatted appropriately. Precursor ion counts were determined for each precursor: H_3O^+, NO^+ and O_2^+. The ion purity of the precursor ions was determined and adjusted so that each precursor had an ion purity of 90% or higher. All samples (cancer samples, celiac samples and healthy control samples) were analyzed by the same method in the FS mode. The FS mode cycles through H_3O^+, then NO^+, and lastly O_2^+, automatically saving each precursor scans after it is complete. The full scan mode parameters were set up in the MS control (Mass spectrum) box. The range was set to scan from 10 m/z to 200 m/z on all three precursors. Three scans using each precursor were performed for 120 seconds each and then the ions were averaged. The same parameters were run on the ambient air collected as a background control. A machine blank was also performed by doing a full scan, cycling through the three precursors, with the direct sampling valve turned off. The second analysis performed on the breath samples collected was completed in the MIM mode of the SIFT-MS. The kinetic library was used to quantify levels of all gases by monitoring the H_3O^+, NO^+ and O_2^+ precursors and their specific product ions as

previously reported (Spanel 1996). Specific databases were set up to monitor compounds of

interest for subjects participating in the cancer study and those participating in the celiac study.

Table 1 and 2 displays the compounds analyzed in each database along with the precursor used,

the kinetic rate coefficient, and the products with mass to charge ratios produced. Each

compound was analyzed for 90 seconds from the breath samples. The absolute concentrations of

the compounds were calculated using the software supplied with the SIFT-MS with ionic

diffusion being accounted for by using an approximation instead of the measured values (Smith

1987). The background ambient air samples were analyzed in the same manner and a machine

blank was performed by turning off the direct sampling for each of the compounds analyzed.

The downstream mass spectrometer timing parameters used were flyback: 0.02s, wait: 0.02s,

precursor: 0.1s, product: 0.1s. Compounds selected to build the databases were determined by

information gathered from previous studies. The cancer database was built from studies that

have previously reported detecting organic volatiles such as acetaldehyde in the headspace of

lung cancer cell cultures and the headspace of blood taken from cancer patients (Deng et al.

2004). The database for the celiac study was built based on the assumption that there is an

increase in fermentation in the gut for those with celiac disease, which would produce an

increase in alcohol products and the excess products may be excreted from the body as volatiles

in exhaled breath.

4.6 Anonymity, confidentiality and Data Storage

Personal information which could be used to identify the study subjects was only

recorded on the questionnaire. Therefore, personal information can only be accessed by the

researchers. A subject number was assigned to each subject and it is the subject number that was

used to identify subjects in all other documents and computer files. No information will be

revealed in the dissemination of results which could be used to identify study subjects and only aggregate data will be used in the dissemination of results. If reference to a specific subject is required in order to properly describe the research data, only study subject numbers will be used. Personal information in written form was stored in a locked filing cabinet located in the medical school. Personally identifiable information was not stored electronically. A backup of all computer files was made weekly and stored in a secure cabinet.

4.7 Statistical Analysis

Values for the ions calculated using the FS mode were normalized before any statistical analysis was performed. The non-parametric Mann-Whitney U test was used to assess whether the two independent breath samples observed, samples from individuals with the disease and those without, came from the same distribution. Values for the ions obtained using the MIM mode were also compared. The mean was calculated and error bars represent the standard deviation of the mean. An indepth statistical analysis was not done because of the limited sample size.

Table 1. Summary of compounds in the lung cancer database detected using the MIM mode of the SIFT-MS for the cancer study

Compound	Precursor	Rate Coefficient*	Products (m/z)
Water	H_3O^+	1.3	$H_3O^+ \cdot H_2O$ (37), $H_3O^+ \cdot 2H_2O$ (55), $H_3O^+ \cdot 3H_2O$ (73)
Propanol	H_3O^+	2.7	$C_3H_7^+$ (43)
Acetaldehyde	H_3O^+	3.7	$C_2H_5O^+$ (45), $C_2H_5O^+ \cdot 2H_2O$ (81)
Formaldehyde	H_3O^+	2.7	CH_3O^+ (31)
Acetonitrile	H_3O^+	3.7	$C_2H_4N^+$ (42), $C_2H_4N^+ \cdot H_2O$ (60)
Acetone	H_3O^+	3.9	$C_3H_7O^+$ (59), $C_3H_7O^+ \cdot 2H_2O$ (77)
Propanal	NO^+	3.0	$C_3H_5O^+$ (57), $C_3H_5O^+ \cdot H_2O$ (75)
Pentanal	NO^+	3.0	$C^5H_9O^+$ (85), $C_5H_9O^+ \cdot H_2O$ (103)
Hexanal	NO^+	3.1	$C_6H_{11}O^+$ (99), $C_6H_{11}O^+ \cdot H_2O$ (117)
Heptanal	NO^+	3.1	$C_7H_{13}O^+$ (113), $C_7H_{13}O^+ \cdot H_2O$ (131)
Isoprene	NO^+	1.7	$C_5H_8^+$ (68)
Acetic Acid	NO^+	1.5	$NO^+ \cdot C_2H_4O_2$ (90)
Butanal	O_2^+	3.1	$C_2H_4O^+$ (44), $C_4H_8O^+$ (72)

* Reference for rate coefficients (Spanel *et al.* 1997)

Table 2. Summary of compounds in the celiac database detected using the MIM mode of the SIFT-MS for the celiac study.

Compound	Precursor	Rate Coefficient*	Products (m/z)
Water	H_3O^+	1.3	$H_3O^+ \cdot H_2O$ (37), $H_3O^+ \cdot 2H_2O$ (55), $H_3O^+ \cdot 3H_2O$ (73)
Ethanol	H_3O^+	2.7	$C_2H_7O^+$ (47), $C_2H_7O^+ \cdot H_2O$ (65), $C_2H_7O^+ \cdot 2H_2O$ (83)
Methanol	H_3O^+	2.4	CH_5O^+ (33), $CH_5O^+ \cdot H_2O$ (51), $CH_5O^+ \cdot 2H_2O$ (69)
Propanol	H_3O^+	2.7	$C_3H_7^+$ (43)
Pentanol	H_3O^+	2.8	$C_5H_{11}^+$ (71), $C_5H_{11}^+ \cdot H_2O$ (89), $C_5H_{11}^+ \cdot 2H2O$ (107)
Butanol	H_3O^+	2.8	$C_4H_9^+$ (57), $C_4H_9^+ \cdot H_2O$ (75), $C_4H_9^+ \cdot 2H_2O$ (93)
Hexanol	H_3O^+	2.9	$C_6H_{13}^+$ (85)
Ethanol	NO^+	1.2	$C_2H_5O^+$ (45)
Propanol	NO^+	2.3	$C_3H_7O^+$ (59)
Pentanol	NO^+	2.5	$C_5H_{11}^+$ (71), $C_5H_{11}O^+$ (87)
Butanol	NO^+	2.2	$C_4H_9^+$ (57), $C_4H_9O^+$ (73)
Hexanol	NO^+	2.2	$C_6H_{13}O^+$ (101)

* Reference for rate coefficients (Spanel and Smith 1997)

5. Results

5.1 Direct sampling versus sampling using Tedlar bags.

Firstly, a comparison was made between two SIFT-MS sampling methods, direct and bag sample collection, because not all individuals that participate in the pilot study are healthy enough to come into the lab, the alternative method of sampling using the Tedlar bag will be necessary. Direct sampling into the sampling involves the subject exhaleing through a mouth piece directly into the machine and bag sample collection involves the subject exhaling into a Tedlar bag and then the gas in the bag is sampled via the sampling tube. Samples were taken from 10 subject volunteers and acetone and isoprene were the two compounds monitored for the comparison of the two sampling methods. In Fig. 3 it can be seen that both acetone and isoprene exhibit a linear correlation between direct sampling and sampling using Tedlar bags with an average ratio between measurements of 1 to 1. This allows for the use of the Tedlar bag sample method and direct sampling when appropriate without concern that there will be discrepancies in the results due to the sampling method.

5.2 The relationship between the breath values collected from subjects and the background ambient air

Next the relationship between the values of specific compounds in breath samples given by participants and the background ambient values were investigated in both the cancer and the celiac studies. This analysis was performed to determine the relationships between counts for volatiles found in the samples compared to their levels in the background ambient air. As illustrated in Fig. 4, there are three different outcomes that
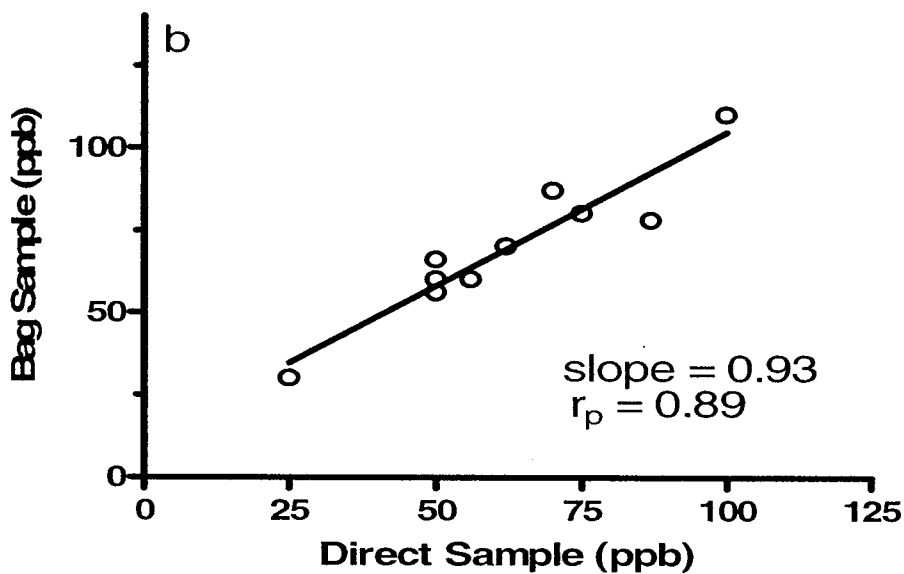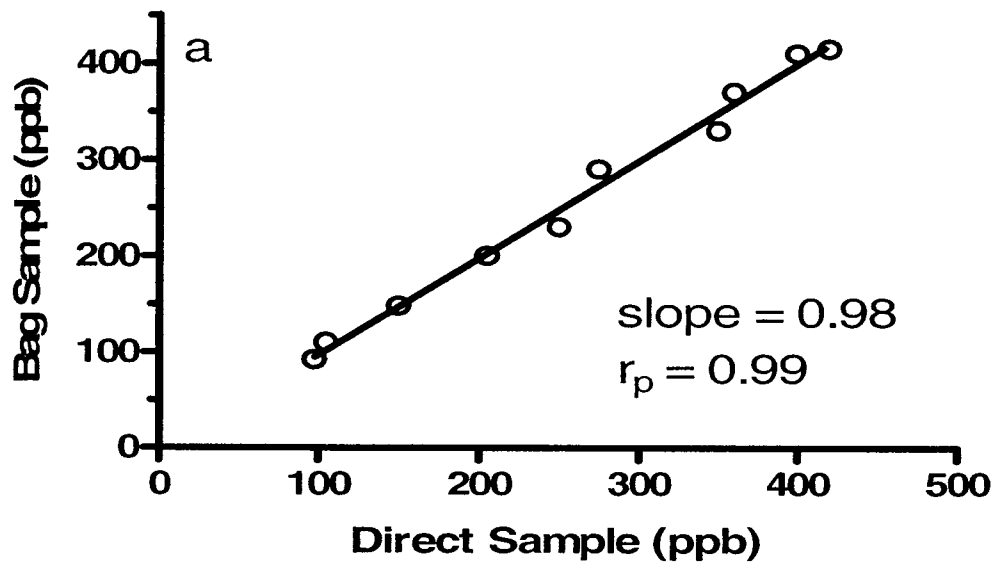
Figure 3. Comparison of direct and bag sampling methods for acetone (a) and isoprene (b) levels measured using SIFT-MS. The line represent the 'best-fit' linear regression line. The average ratio between the two measurement methods i.e. the slope of the regression line was 0.98 for acetone (a) and 0.93 for isoprene (b). The correlation coefficients were 0.99 for acetone and 0.89 for isoprene.
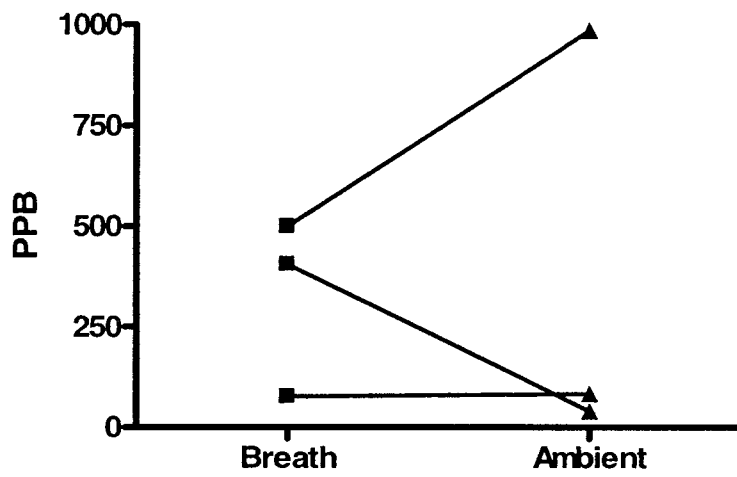
Figure 4. The three different relationships between breath sample values and ambient air values when using the SIFT-MS.Figure 4.

can occur between the values from the samples given by participants and the background ambient air values. The first was that the counts from the subject samples can be higher then those detected in the background ambient air. The ion product of acetone, 59 m/z, with the precursor H_3O^+ was used to illustrate this relationship. The average ppb from the samples was 406 where as the average ppb of the ambient air was 40. The second outcome that occurred was that counts from the subject samples were the same as the ambient air. The ion product used to display this relationship was 43 m/z, propanol, with the precursor H_3O^+. The average ppb for propanol in breath samples was 77. Similarly, the average for propanol in ambient air was 83. The third outcome was that the counts from the subject samples were lower than the ambient air. The ion product of ethanol, 43 m/z, with the precursor H_3O^+ was used to describe the relationship. The average ppb from breath samples was 498 compared to the average from ambient air that was 985. From this data it can be seen that subtracting the background values from the values of the breath samples given by subjects would results in a range from positive to negative values. This range of positive to negative values can make it difficult for analysis comparison to be completed therefore based on these results the remainder of the analysis of the results of the cancer and celiac disease studies will use the subject sample values without subtracting the background values. The background values will still be noted during the analysis and discussed.

5.3 Cancer full scan (FS) mode scans

The FS mode scans in the cancer pilot study were analyzed to determine if there were any compounds that showed any differences between the sample groups that could help lead to volatile marker of the disease. Scans of all the breath samples collected from

subjects with cancer were compared with those taken from healthy control subjects that were approximately age and gender matched. Scans on each precursor (H_3O^+, NO^+ and O_2^+) were done in triplicate for the range of 10 m/z to 200 m/z and the data obtained from the scans were then averaged and normalized before analysis was performed. To normalize the data the precursor counts were calculated and then scan values of each ion were divided by the value of the precursor counts. Most of the new normalized values were very small and therefore all values were multiplied by 1000000 to make them easier to display. Data from the ambient air was prepared in the same manner as the data from the samples collected from the study subjects. Mann-Whitney U test was performed on the averaged normalized values from the scans to detect any significant differences between the cancer subjects and the healthy control subjects ($p< 0.05$). There were two values of note that showed a distinct difference when comparing the cancer group to the healthy controls: 77 m/z detected using the H_3O^+ precursor and 50 m/z detected using the NO^+ precursor. With further analysis, ion 77m/z was found to be much lower in the breath samples from cancer subjects OCS05 and OCS10 with normalized values (NV) of 9 and 11 respectively, compared to the other samples with an average of 230 NV (Figure 5a). Background levels for ion 77 m/z gave an average of 8 NV indicating that the subject OCS05 and OCS10 were not producing this ion compared to all other samples subjects. Ion 50 m/z was seen to be produced in higher amounts from cancer subject OCS02, 121 NV, compared to an average of 9 NV from all other sample subjects (Figure 5 b.). Background levels for ion 50 m/z were much lower compared to subject OCS02, average of 13 NV, indicating that subject OCS02 is producing this compound at higher levels compared to all other subjects.

Figure 5. Outliers discovered between various patients within the cancer subject group. Ion 77m/z detected using the precursor H_3O^+ has two cancer subjects with low values compared to the other study subjects (a) while ion 55 m/z detected using the precursor NO+ has one cancer subject giving a high ion count value compared to other subjects (b). Outliers are identified by being circled.

5.4 Comparison of specific compound values for cancer subjects and healthy controls

Next the data collected using the MIM mode of the SIFT-MS was analyzed. The compounds analyzed in the MIM mode were compared to view any potential differences in the amount detected in breath samples from cancer patients compared with samples from healthy control subjects. This analysis was used to determine if specific compounds thought to be increased in those with lung cancer could act as biomarkers for the disease when analyzed with the SIFT-MS. The mean values for each compound were compared between the two groups and as illustrated in Figure 6, the only compound to have a significant difference ($p < 0.05$) between the cancer subjects and the healthy controls was acetone. The mean acetone levels from the cancer subject groups was 215 ppb while the mean acetone levels form the healthy control subjects was 479 ppb. Background levels of acetone for the cancer group and the healthy control group were negligible, mean values of 16 ppb and 12 ppb respectively. Acetone has been known to vary depending on the time since a person's last meal and their calorie intake, which could be the factor for the difference seen between sample groups.
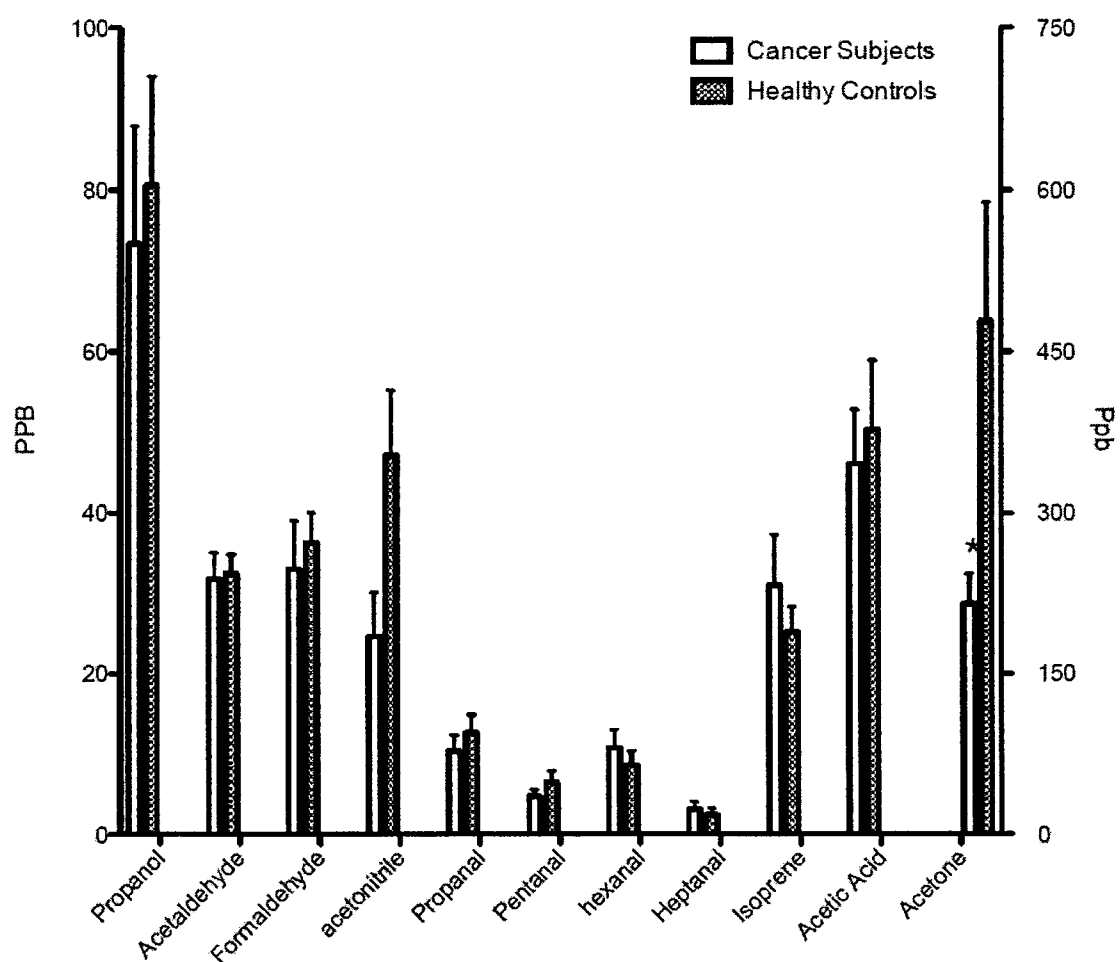
Figure 6. Comparison of mean values between cancer subjects and healthy controls for compounds detected using the MIM mode of the SIFT-MS. Acetone is the only value that showed a significant difference between groups ($p < 0.05$). Error bars represent the standard error of the mean. Note the mean values of acetone are plotted on the right axis.

5.5 Celiac full scan (FS) mode scans

The data from the celiac pilot study was performed in the same manner as the cancer pilot study. First the FS mode scans were analyzed to determine if a broad scan of ions could reveal differences in the expression of some of these ions in samples from celiac patients compared to healthy controls. This analysis may lead to the discovery of potential ions to be investigated as biomarkers. Scans of all the breath samples collected from subjects suffering from celiac disease were compared with those taken from healthy control subjects that were approximately age and gender matched. Scans on each precursor (H_3O^+, NO^+ and O_2^+) were done in triplicate and the data obtained from the scans were then averaged and normalized before analysis was performed. Normalized data was prepared in the same manner as the data from the cancer FS mode scan as previously described. Data from the ambient air was prepared in the same manner as the data from the samples collected form the study subjects. Celiac subjects were compared to healthy control subjects by performing Mann-Whitney U non-parametric analysis on the averaged normalized data from the scans ($p< 0.05$). There were no significant results found between the celiac and healthy control subjects.

5.5.1 Higher ion values in patients newly diagnosed with celiac disease

After finding no significance from the comparison between samples from celiac subjects and healthy controls the values from the celiac samples were scanned for any outliers that might indicate a difference between a patient who has been newly diagnosed with celiac disease and patients who have been diagnosed and on a gluten free diet for a more than a six month period. Subject Celiac S05 had been diagnosed with celiac disease less than a week prior to giving a breath sample. From the data collected in the FS mode

there were three ions that showed an increase in the newly diagnosed patient compared to the other celiac subjects and healthy controls when normalized values were compared. These ions were 60 m/z, 99 m/z and 117 m/z as illustrated in Figure 7. Ion 60 m/z gave a value of 58 NV, 99 m/z gave a value of 72 NV, and 117 m/z gave a value of 50 NV in the newly diagnosed celiac patient compared to the average values, 4 NV, 13 NV and 4 NV respectively, from other celiac subjects and control subjects. Background values were also negligible as they were similar to the average levels of the celiac subject diagnosed for more then 6 months and the healthy control subjects. There were no outliers in the background values. No statistical analysis was possible because the sample size of the two groups being compared is so small; 1 newly diagnosed patient, less than a week since diagnosis, and 8 celiac patients who have been diagnosed for more than six months.

5.6 Comparison of specific volatile between celiac subjects and healthy controls

Similar to the cancer pilot study, the celiac pilot study used a database of specific compounds to be compared between the celiac group and the healthy control group. This analysis was performed to determine if any of these known compounds were exhaled in different quantities by celiac patient compared to those without the disease and if so would they have the potential to be developed as a specific biomarker of the disease. Each compounds was monitored for 90 seconds and the values obtained were calculated using the SIFT-MS softwear with ionic diffusion being accounted for by using an approximation instead of the measured values. Figure 8 shows the average ppb values of water, propanol, pentanol, butanol and hexanol detected in breath samples from celiac subjects compared to the average ppb values from healthy control samples. There was no significant difference detected in the values obtained from the two groups. Significant
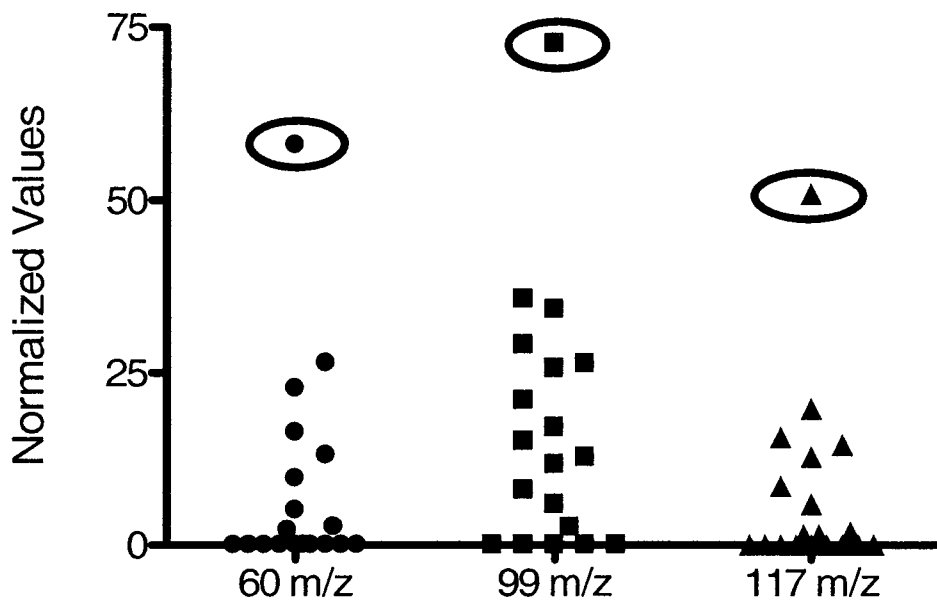
Figure 7. Ion counts for 60 m/z, 99m/z and 177 m/z with the precursor NO+ using the FS mode of the SIFT-MS. The outlier for each ion belongs to the same celiac subject (Celiac S05) that was diagnosed with celiac disease less then a week before giving a breath sample. Outliers identified by being circled.
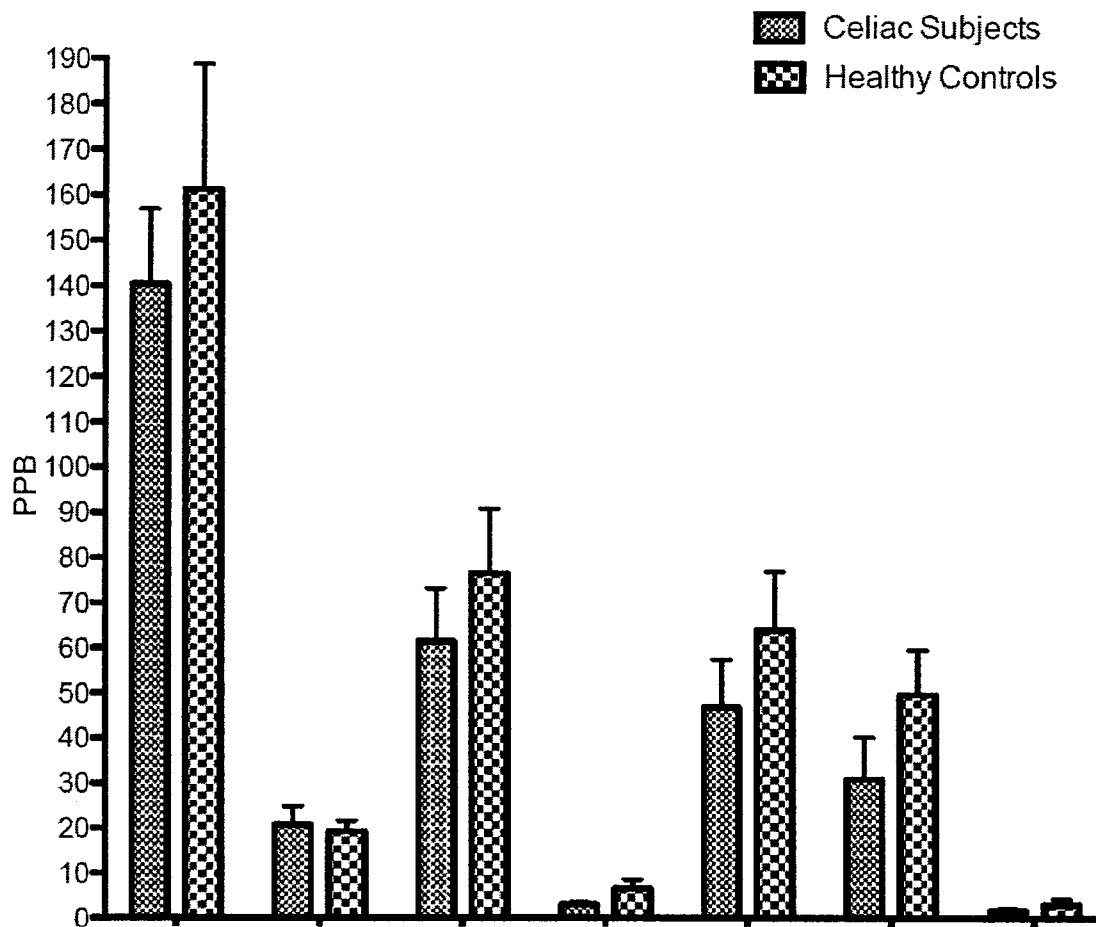
Figure 8. Comparison of mean values between celiac subjects and healthy controls for compounds detected using the MIM mode of the SIFT-MS. None of these compounds showed any significant difference between the samples. Error bars represent the standard error of the mean.

differences however, were detected for ethanol and methanol between the two sample groups using the Mann-Whitney U test with a $p < 0.05$ (Figure 9). The mean value for ethanol was 1104 ppb from the celiac patient samples compared to the higher mean value of 1933 ppb from the healthy control samples. Methanol showed the same relationship with a mean value of 479 from celiac samples compared to 1032 from healthy controls. Background mean levels of ethanol and methanol were then compared to levels detected in the samples. The mean background levels for ethanol were higher than the sample levels, background mean of 2602 ppb compared to a sample mean of 1104 ppb for celiac patients (Figure 10a) and a background mean of 3130 ppb compared to 1933 ppb for healthy controls (Figure 10b). Whereas the background mean values for methanol were similar to those from the celiac and healthy control samples, 440 to 497 and 995 to 1032 respectively (Figure 10c and 10d). This would indicate that ethanol levels may have potential for differentiating those with celiac disease and those without, compared to methanol levels in which sample and background levels are the same indicating the difference might be due to the differences in the ambient air values on the days when different samples were taken.
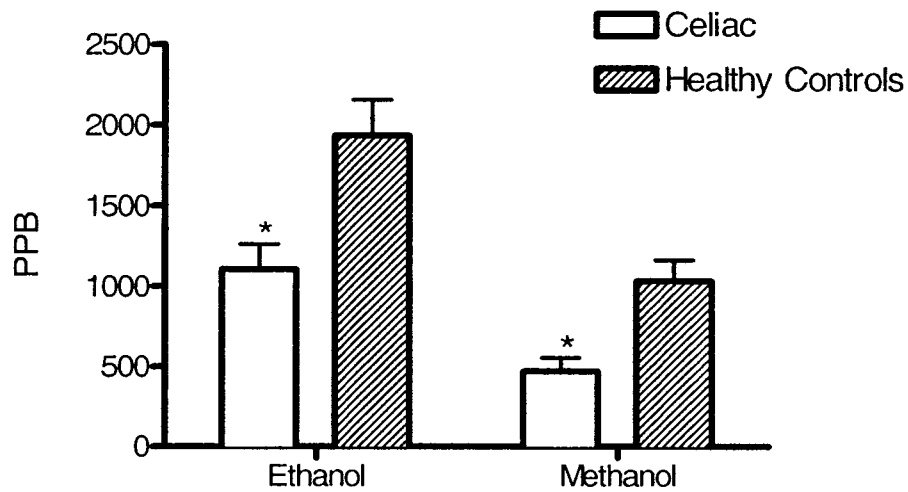
Figure 9. Comparison of mean values between celiac subjects and healthy controls for ethanol and methanol detected using the MIM mode of the SIFT-MS. Significant difference is seen between the celiac group and the healthy controls (p< 0.05). Error bars represent the standard error of the mean.
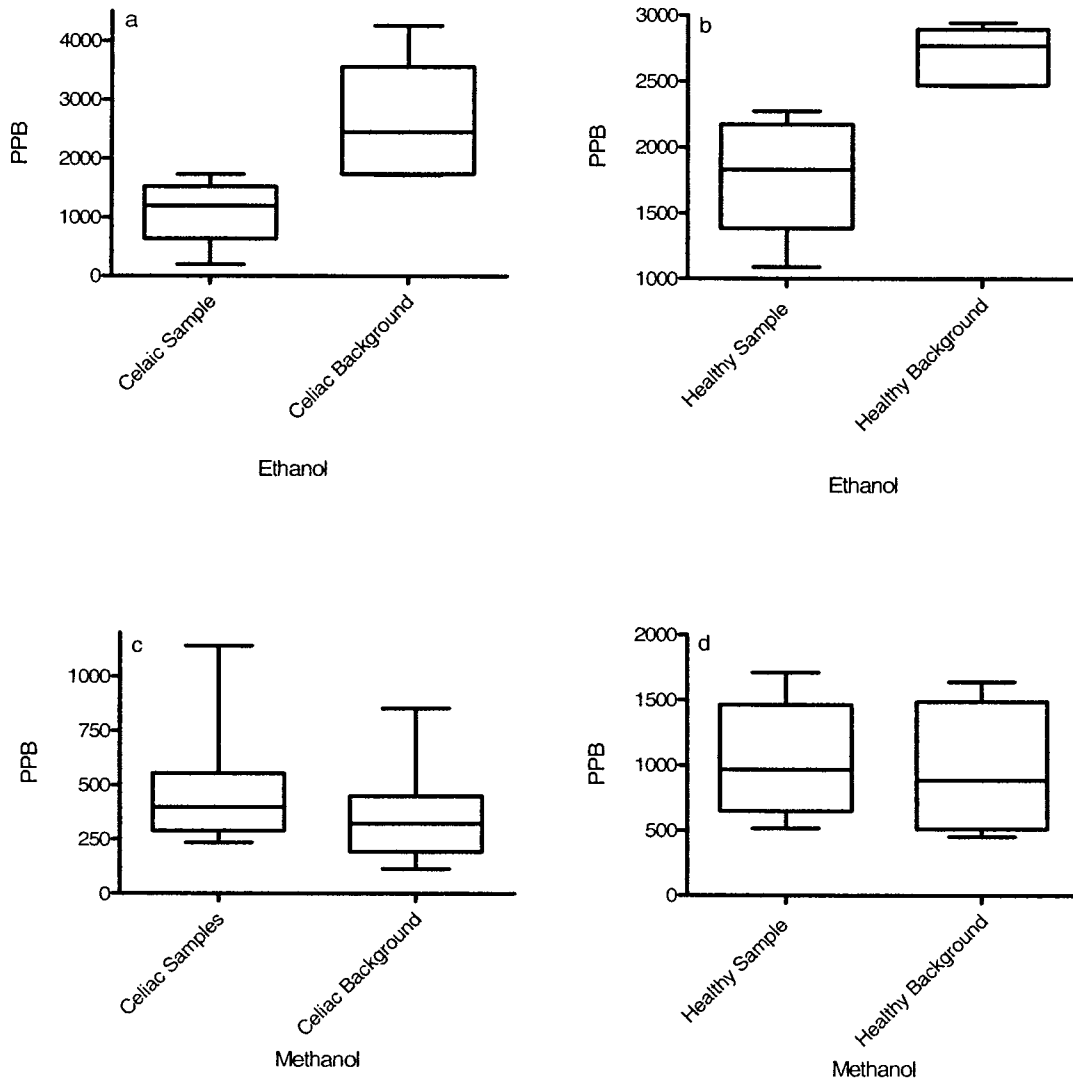
Figure 10. Comparison of sample values to the background values for ethanol and methanol. All box plots a-d represent the lower quartile, median and upper quartile while the whiskers represent the maximum and minimum values of the range.

5.6.1 Effect of an isotopologue on butanol concentrations

Butanol was another compound that was specifically tested for using the MIM mode of analysis. Butanol was monitored for 90 seconds and the values were calculated in the same manner as the other compound analyzed in the MIM mode. The values for butanol with some of the celiac subjects showed a significant increase when run with the NO^+ precursor compared with the H_3O^+ precursor (Table 3). Butanol produces the product ion 57 m/z when using the H_3O^+ precursor and this means that 55 m/z (H_3O and 2 H_2O) can interfere with the detection of butanol since about 0.2% of 55 m/z will appear at 57 m/z due to ^{18}O. This problem has already been corrected for in the database by subtracting the interference. Butanol produces the product ion 75 m/z when using the NO^+ precursor but if there is a higher H_3O^+ count then normal when using the NO^+ precursor then possible interference can come from m/z of 73 (H_3O and 3 H_2O) that appear as 75 m/z with ^{18}O because no correction has been applied to eliminate the interference.

Table 3. Summary of the values for butanol detected with the precursors H_3O^+ and NO^+ from celiac patients using the MIM mode of the SIFT-MS.

Celiac Subject	Butanol with precursor H_3O^+	Butanol with precursor NO^+
Celiac S01	31	30
Celiac S02	50	74
Celiac S03	25	43
Celiac S04	37	130
Celiac S05	62	1096
Celiac S06	90	955
Celiac S07	36	53
Celiac S08	148	322
Celiac S09	50	109
Celiac S10	85	519

6. Discussion

Pilot studies #1 and #2 both had similar major findings in terms of the usefulness of using SIFT-MS as a diagnostic tool for disease. There were difficulties in both pilot studies that are attributed to method analysis that need to be corrected before SIFT-MS can be reliable in the detection of specific biomarkers of disease such as isotopologues and artifact ions. Pilot study #1 gave insight into possible markers associated with various cancers while pilot study #2 shows the potential that biomarkers for celiac disease are more likely to be found in people who have been recently diagnosed with the disease.

6.1 The contribution of ambient air

The relationship of ambient air samples and breath samples were compared before any analysis began. Analysis of breath compounds also present in ambient air is made difficult since the ambient air values cannot simply be subtracted from sample values. This is due to the fact that exhaled breath levels are not additive to exogenous and endogenous concentrations. An example of this process is seen with the inhalation and exhalation of oxygen (O_2) and carbon dioxide (CO_2). There is approximately 20% O_2 and 0.04% CO_2 in the air that is inhaled by an individual. Oxygen then diffuses through membranes in the lungs into the circulatory system where it binds to hemoglobin and then is distributed to cells throughout the body where it is used in the process of cellular respiration. After the process of cellular respiration the waste product CO_2 is produced and returned to the lung for expulsion. When individuals exhale the percentage of O_2 decreases to 15% and the percentage of CO_2 increases to 4.5%. (Piiper and Sheid 1971; West 1974) If the ambient air was subtracted from the expired air the resulting amount of oxygen would be a negative number indicating that the relationship between ambient, alveolar

and exhaled O_2 levels is extremely complex. If the breath levels are close to the levels detected in the ambient air then reliable quantification of breath abundance is very difficult. For reliable quantification of a compound the difference in the amount of that compound in an exhaled breath sample needs to be 3 or more times greater then the amount detected in ambient air (Smith and Spanel 2007). Since the relationships are very complex, the detected values from samples were first compared between subjects with a disease state and the healthy control subjects and if there was a significant difference found between the two groups for an ion then the background levels of that ion were also analyzed.

A difficulty that was encountered while conducting the cancer study is that samples were not all collected in the same location. The breath samples that were collected from the various cancer subjects were taken in the Thunder Bay Medical Centre and the Northern Ontario Regional Cancer Centre, located at the Thunder Bay Regional Sciences Centre. The breath samples collected from the healthy control subjects were taken at the Thunder Bay 55 Plus Centre. The first location is a building that contains offices that physicians use for the purpose of meeting with patients while the second is a division in a hospital and the third is a seniors' community centre. These three locations will have their own unique ambient air values because of differences in age and location of the buildings, cleaning products used, equipment and furniture, and ventilation (Johnson and Ettinger 1991; Wallace *et al.* 1987; Tichenor and Mason 1988). Therefore, any significant differences discovered either between the cancer subjects and the healthy control group or within the cancer group should be very carefully compared to the background ambient air of the locations the samples were taken. It would have been difficult to collect all of the samples used in the cancer study in the same location since some individuals with cancer may not be in good enough health to come into the lab and people are more

agreeable to volunteer when they can give a sample right away instead of making an appointment to give the sample in a different location another day. Furthermore, a useful and practical clinical test must be robust enough to allow samples to be taken at more or less any location. The complication of having samples taken at different locations was eliminated in the celiac study as all subjects, celiac and healthy controls came into the lab at the Northern Ontario School of Medicine to give breath samples so all samples had equivalent ambient air to compare against.

6.2 Difficulties with recruitment

The low recruitment rates presented problems in analysis in both the cancer and the celiac studies. After Dr. Gehman decided not to participate in the recruitment of lung cancer patients, it became very difficult to recruit only lung cancer patients for the study. Therefore, the cancer study expanded to include all cancer types with the help of Dr. Vergidis recruiting cancer patients that fit the criteria from his professional practice. The difficulty with including all types of cancer is that they all have different epidemiology and pathology. This complicates finding of a profile of volatiles that is similar to all types of cancer that could provide a biomarker for breath diagnosis. It would have been ideal if another physician who specializes in lung cancer had agreed to participate in recruitment of patients because then the study group could have remained only those individuals that have lung cancer. Another factor contributing to the difficulty in recruiting cancer patients was that a second party and not a primary investigator was recruiting patients for the cancer study. Since the second party recruiter did not have any other role in the cancer study other then recruitment, it was not always a top priority to continuously recruit subjects for the study. There was also a period of time during the study when the recruiter at the Northern Ontario Regional Cancer Centre unfortunately became very ill and

during that time no recruitment was performed. A potential solution to the recruitment problem could have been to have the physician ask the patients if they would potentially participate in the study and then have the primary researcher speak with the patient and take the breath sample if the patient agreed to participate in the study. This would have required that the researcher be present at the Northern Ontario Regional Cancer Centre during all clinics. The presence at the Cancer Centre would then remind the physician helping with the study to present it to their patients and having the researcher take the sample would alleviate other staff having to take the sample which disrupts their normal work routine and may interfere patient related care.

Recruitment rates were low in the celiac study but this was expected since celiac disease is currently underdiagnosed in all populations and has a prevalence rate of one percent in the population. The majority of the celiac subjects recruited had been diagnosed for several years prior to recruitment and adhere specifically to a gluten free diet. Since these individuals do not ingest gluten the small intestine chemistry and function should be the same as a healthy individual. Therefore, people suspected of celiac disease or those that have been newly diagnosed, within a week, would be better study subjects. For further studies involving breath analysis with celiac patients it would be beneficial to collaborate with a gastroenterologist who would see patients during the process of diagnosis of celiac disease and be able to help recruit those patients who have not yet started treatment of a gluten free diet and therefore still have an altered state of the small intestine.

6.3 The causes of variation of acetone in breath samples

In the pilot study #1 acetone was the only volatile compound of those monitored using the MIM mode of the SIFT-MS that showed a significant difference between the cancer subjects and healthy controls. Acetone was chosen as a volatile compound of interest to monitor because it

has been known since the 1950's that increased levels of acetone in the breath is a biomarker of diabetes. (Tassopoulos *et al.* 1968). Acetone was routinely monitored since it is an easily measureable and well defined breath compound which is useful as a check of the integrity and sampling and measurement process. The range of values for acetone was significantly larger in the healthy control group than in the celiac group. It has been previously shown that the measured values of acetone in healthy volunteers had a range from 238 ppb to 832 ppb when measured using the SIFT-MS (Turner *et al.* 2006). Also previous SIFT-MS studies have shown that breath acetone levels vary depending on caloric intake and that high breath acetone levels due to fasting are quickly lowered by eating (Turner *et al.* 2006; Smith *et al.* 1999). The variation of acetone in the breath due to eating and fasting could very well be the cause of the decrease in acetone levels in cancer patients compared to the acetone levels of the healthy controls. The majority of the breath samples collected from cancer patients were taken either early morning or early afternoon. Therefore it could be deduced that the cancer patients giving the samples had just recently had a meal before being asked to participate in the study and this would result in the lower levels of acetone. Whereas, the healthy control subjects gave breath samples between 10:30am and noon, before a midday meal could be consumed and several hours after breakfast, which may explain why their levels of acetone are higher. Another factor that has an effect on the levels of acetone in breath is gender but this confound should not have been a factor in the increase of acetone levels between cancer subjects and healthy controls as the two groups were gender and age matched accordingly. (Turner *et al.* 2006)

6.4 No significant differences in aldehydes or isoprene between cancer and healthy subjects

Other then acetone none of the other compounds analyzed in the MIM mode of the SIFT-MS showed any significant difference that would indicate a potential biomarker when samples

from cancer subjects and healthy control subjects were compared. Potential biomarkers for lung cancer were proposed in previous studies on volatile compounds from breath samples of cancer patients, headspace of lung cancer cell cultures and headspace of blood from cancer patients. Common potential biomarkers that were described in these studies include a selection of various aldehyde compounds (Chen *et al.* 2007; Phillips *et al.* 2007; Deng *et al.* 2004; Phillips *et al.* 2003). Therefore, these compounds were chosen to be specifically monitored to look for increases of their volatile levels in people with lung cancer in contrast to healthy control subjects. Unfortunately recruitment problems that occurred in the course of the study resulted in not enough lung cancer samples being collected for a comparison to be accomplished. These compounds were still monitored for the cancer group when samples from multiply cancer types began to be collected. The values collected were compared between the various cancer samples to determine if there might be a compound that was noticeably higher or lower in a subject that may need further investigation. It is not surprising that as a group, the samples from cancer patients yielded no significant results because of the mixture of different cancer subjects that participated in the study. The cancers originated in different tissues, not all cancer tumors will grow and progress at the same rate and some were more severe than others. All of these factors contributed to the variation in the results of the cancer group with no unique pattern emerging from the compounds that had been proposed as potential biomarkers.

6.5 Significant outliers discovered in cancer sample group

Due to the mixture of a variety of types of cancer among the cancer subject group there was not a strong volatile compound profile that distinguished those people who had cancer from the healthy control group. However, on further analysis of the cancer group two ions were identified as noticeably different in a couple cancer patients compared to others in their group

and to the healthy controls. Using the FS mode of the SIFT-MS ion 77 m/z was detected to be less abundant in two cancer subject when analyzed with the H_3O^+ precursor while analysis with the NO^+ precursor showed an increase in the ion 50 m/z in one of the cancer subject compared to all other subjects.

The decrease in the ion 77 m/z was detected in the cancer subjects OCS05 and OCS10. Both of these subjects were women in their 60s who were suffering from breast cancer. The women were also having chemotherapy treatment. There was no information on the type and stage of breast cancer of the two women. Studies involving volatile biomarkers in breath as a predictor for breast cancer are currently being conducted. A combination of five volatile organic compounds, 2-propanol, 2,3-dihydro-1-phenyl-4(1H)-quinazolinone, 1-phenyl-ethanone, heptanal, and isopropyl myristate, have been shown to predict breast cancer by Phillips *et al.* 2006. These proposed ions were seen to be elevated in breast cancer patient whereas in this study ion 77m/z levels were much lower compared to all of the other subjects, cancer or healthy control. Subjects OCS05 and OCS10 seem to not be producing any 77 m/z as the levels detected in the breath samples are similar to background levels (8 NV). All other subjects exhaled an average ion count of 294 NV with ion 77 m/z. It is doubtful that this decrease was a result of the chemotherapy treatments as other cancer subject who gave samples were also on a chemotherapy regiment. Perhaps there is a metabolic process that has been disrupted by the growth of the breast cancer that is causing a decrease in the ion 77 in subjects OCS05 and OSC10. To determine if this compound is unique to individuals with breast cancer a more comprehensive study with a large test subject group is needed.

The increase of ion 50 m/z when analyzed on the NO^+ precursor was observed in the sample given by subject OCS02. This subject is an 85 year old male with chronic lymphocytic

leukemia (CLL). CLL is a cancer of the white blood cells that is a disease of older adults in which the majority of sufferers are men (Boelens *et al.* 2009). CLL is a disease that is easily diagnosed by a blood test (Chiorazzi *et al.* 2005) that is not overly invasive, therefore the development of a breath test for diagnosis for CLL may not be necessary. Since this type cancer is in the circulatory system, if there is an increased amount of a compound being produced due to the cancer the most likely route of excretion would be exhalation. The compound can easily move from the blood into the alveolar air across the blood air barrier and then be exhaled from the body. This helps to explain how there can be an increase in a compound from a CLL patient compared to other cancer subjects since their diseases are located deeper in different tissues in the body. More studies are required to determine if the ion present at 50 m/z could be a biomarker in breath for people with CLL.

6.6 Outliers identified in patients newly diagnosed with celiac disease

In the pilot study #2 the celiac samples did not shown any significant difference when compared to the healthy control subjects however on further analysis of the results there were three ions that were noticeably higher in subject Celiac S05 in comparison to the other celiac subjects and the healthy controls. These ions are 60 m/z, 99 m/z and 117 m/z. Background levels of these ions are negligible since they are almost at undetectable levels (less then 5 NV), therefore it can be concluded that these increased levels are being excreted from the body and not inhaled from the environment.

Currently there is no compound in the SIFT-MS database that has been defined to produce 60 m/z as a product from the reaction with the precursor NO^+. A possible volatile compound that may produce a product of 60 m/z is C_2H_7NO: ethanolamine. Ethanolamine has a molecular weight of 61 but when it reacts with NO^+ the products ions that would result and be

detected are NOH and $C_2H_6NO^+$ (m/z of 60), which could be the ion detected. Ethanolamine is an organic compound that is both a primary amine and a primary alcohol and has a high volatility. Ethanolamine is the second most abundant head group for phospholipids in biological membranes as well as a head group on plasmalogens. Plasmalogens are constituents of many tissues in the human body and bacteria such as *E. coli* and *Veillonella*, both of which are normal gut flora found in the small intestine. Bacteria normally found in the small intestine have a higher percentage of ethanolamine plasminogens when compared to other bacteria that can cause an infection and it has been determined that the increased amount of ethanolamine plasminogen causes a bacteria to lose its ability to invade its host. (Harnett and Harnett 1999) A possible explanation for the production of 60 m/z if the parent compound is ethanolamine is that when the autoimmune reaction to gluten occurs there is destruction of the biological membranes of cells in the small intestine and the normal gut flora. This membrane destruction may cause an increase in the amount of ethanolamine in the small intestine, the excess amount of ethanolamine may then be transported to the lung via the circulatory system for excretion, which could then be detected by breath analysis.

The ions of 99 m/z and 117 m/z have been characterized as the products $C_6H_{11}O^+$ and the hydrate $C_6H_{11}O^+$ H_2O for the reaction of hexanal with the precursor NO^+. Normally the association of water with the reaction products of molecules with the NO^+ precursor is slow but because there is H_3O^+ circulating, as the ion purity of NO^+ was not 100 percent and a high percentage of water vapour was present in the sample, the water molecules can react to form a hydrate. However, without further study it cannot be certain that the ions of 99 m/z and 117 m/z are increased levels of hexanal.

More subjects will need to be sampled to detect and confirm if an increase in the ions 60 m/z, 99 m/z and 117 m/z are unique to individuals newly diagnosed with celiac disease as well as studies need to be conducted to determine the exact molecule producing these products. A study group that consists of individuals suspected to have celiac disease and individuals that have been diagnosed for less then 2 months would be ideal subjects to conduct an investigation regarding the increase in the above mentioned compounds. These individuals will potentially still suffer from some symptoms of celiac disease and the villi in the small intestine will not have yet returned to normal. It takes approximately 3 to 6 months, depending on the severity of damage that the small intestine has suffered, for the villi to regrow and normal function to return (Setty *et al.* 2008.) Therefore, any volatile compounds associated with celiac disease and the damage to the small intestine have the potential to be detected at this stage of the disease. Celiac patients that have been diagnosed and treated for more than six months would not show the same volatile profile due to their disease being under control and normal function of the small intestine would have returned.

6.7 Differences in ethanol and methanol between celiac and healthy control groups

Ethanol and methanol were the only two compounds detected using the MIM mode of the SIFT-MS to show any significant differences between the celiac subject group and the healthy control group. The background values of these two molecules were then analyzed and compared to values obtained from the samples. Ethanol showed a much larger increase in background levels compared to the ethanol levels in the sample, which could suggest that some of the ethanol is being absorbed and used by the body. In comparison the background values of methanol were the same as the levels detected from the subject samples. This implies that the ambient methanol was being inhaled and then exhaled by the participants without any more or less methanol being

excreted by the subjects themselves. Due to this observation it can be concluded that methanol in fact does not exhibit a difference between the celiac subjects and the control group but that the differences observed were due to the variation in the ambient air values on different collection days.

Recently, it has been discovered that due to the increased sensitivity of the SIFT-MS, adduct ions composed of major gases present in air (N_2, O_2) and the precursor ions are now being detected at the ppb level in times of a few seconds (Spanel and Smith 2009). This can be troublesome, especially, with the detection of ethanol because the adduct ion $H_3O^+N_2$ can be detected instead of ethanol when sampling for ethanol. If the presence of these ions is not recognized, they can be improperly assigned to compounds that are not present. Also, in humid exhaled breath the signal levels of the adduct ions are greatly reduced which could explain the large difference between the breath levels of ethanol and ethanol levels in the ambient air (Spanel and Smith 2009). The reduction is the result of the adduct ions rapid reaction with water molecules in ligand switching reactions that generate the ions $H_3O^+H_2O$, NO^+H_2O and $O_2^+\cdot H_2O$, and these ions are routinely accounted for by in the analysis of trace gases (Spanel *et al.* 2006). This phenomenon could explain the decreased values of ethanol in breath samples compared to background levels. If the adduct ions are recognized and excluded from the ethanol values then the results would most likely be similar to those seen with methanol, that it is the differences in the ambient air causing the difference between groups and not a difference in the amount exhaled between celiac and control subjects. Even though the values of ethanol detected exhibited a significant difference between the celiac subjects and the control group more experiments will need to be conducted to determine if the significant decrease in ethanol levels are due to celiac disease, to SIFT-MS instrument function or to some other factor.

6.8 Isotopologue effects

Isotopologues are molecules that differ only in their isotopic composition, that is they will have at least one atom with a different number of neutrons to its parent. An example of an isotopologue is heavy oxygen water where the oxygen molecule is ^{18}O instead of the usual ^{16}O. Precise and accurate measurements of isotopologue distributions in biological molecules are needed for determination of isotope effects during analysis in metabolic and biophysical studies (Cassano *et al.* 2007; Wang *et al.* 2007). It is suspected that an isotopologue effect caused the increased levels of butanol that were detected using the MIM mode of the SIFT-MS. Butanol produces the product ion 57 m/z when using the H_3O^+ precursor. The water product of 55 m/z (H_3O and 2 H_2O) can interfere with the detection of butanol since about 0.2% of 55 m/z will appear at 57 m/z due to ^{18}O causing an isotopologue. This problem has already been corrected in the database by subtracting the interference from the ion product 57 m/z. Butanol gas produces the product ion 75 m/z when using the NO^+ precursor but because of the high water content in the breath sample and the ion purity of the NO^+ precursor not being 100 percent an isotopologue effect may occur. This interference is due to H_2O product of 73 m/z (H_3O and 3 H_2O) that appears as 75 m/z with ^{18}O because no correction has been applied to eliminate this interference. Since the isotopologue effect was corrected with the H_3O^+ precursor the same principle of correction could be applied to the NO^+ precursor. The amount of isotopologue produced due to the high water content can be directly subtracted in the kinetic library by calculating the percentage of ^{18}O that will be produced but the percent due to ion purity would have to be calculated and subtracted separately after compound detection from the recorded ion purities of the precursor NO^+ during the experiment. The precursor ion purities were calculated before all samples were analyzed by the SIFT-MS and the instrument was tuned up to give an ion purity of

90% or higher for all precursors. The increase in butanol by what appears to be an isotopologue effect may have also originated because during the analysis the precursor ion purity of NO+ decreased without being detected. The detection of an isotopologue effect requires the continuous monitoring of the precursor ion purity while analyzing a sample with the SIFT-MS. Ideally precursor ion purities should be at 100 percent where no isotopologue effects of water would appear, although, this is not always possible to achieve.

7. Conclusions

The SIFT-MS technique of breath analysis has the potential to be a diagnostic tool for the detection of various diseases. However, there are still some difficulties in the analysis methods that need to be alleviated. The appearance of adduction ions formed by the precursor ions due to the increased sensitivity of the SIFT-MS machine need to be considered and incorporated into the database for compounds specifically monitored by in the MIM mode. These adduction ions should also be taken into account when interpreting results from a full scan of a sample. Another difficulty to overcome is the isotopologue effects that can happen with certain compounds such as butanol. Due to the increase in water in the samples because of the humidity of breath the formation of isotopologues can cause an incorrect measurement of the amount of certain compounds found in the breath samples. Corrections for the isotopologues can be inserted to rectify their effect and give the accurate measurement of the compound affected.

The SIFT-MS was able to identify a few ions in both of the pilot studies that could potentially be used as biomarkers for various cancers and celiac disease. Due to the low recruitment into the studies more research needs to be conducted with larger groups to confirm with any significance if the ions that differ in the cancer subjects will be unique to the specific cancers. The same is true for the ions that differed in the newly diagnosed celiac patients compare to celiac patients diagnosed for more then six months and the healthy controls. Even though, there were problems in the recruitment of patients and detection of volatile compounds, the SIFT-MS has the potential to be a beneficial diagnostic tool in the area of breath analysis and detection of disease.

References

Alaedini A, Green PH: Narrative review: celiac disease: understanding a complex autoimmune disorder. Ann Intern Med 2005; 142: 289–298.

Bach PB, Kelley MJ, Tate RC, McCrory DC: Screening for lung cancer: a review of the current literature. Chest. 2003; 123 Suppl 1: 72S-82S.

Banoo S, Bell D, Bossuyt P, Herring A, Mabey D, Poole F, Smith PG, Sriram N, Wongsrichanalai C, Linke R, O'Brien R, Perkins M, Cunningham J, Matsoso P, Nathanson CM, Olliaro P, Peeling RW, Ramsay A; Evaluation of diagnostic tests for infectious diseases: general principles. Nat Rev Microbiol. 2006; 4(12 Suppl): S20-32.

Basanta M, Koimtzis T, Singh D, Wilson I, Thomas CL: An adaptive breath sampler for use with human subjects with an impaired respiratory function. Analyst. 2007; 132: 153-163.

Beckles MA, Spiro SG, Colice GL, Rudd RM: Initial evaluation of the patient with lung cancer: symptoms, signs, laboratory tests, and paraneoplastic syndromes. Chest. 2003; 123 Suppl 1: 97S-104S.

Boelens J, Lust S, Vanhoecke B, Offner F: Chronic lymphocytic leukaemia. Anticancer Res. 2009; 29: 605-615.

Boffetta P: Human cancer from environmental pollutants: the epide-miological evidence. *Mutat Res*. 2006; 608: 157-162.

Briani C, Samaroo D, Alaedini : Celiac disease: From gluten to autoimmunity. Autoimmun Rev. 2008; 7: 644-650.

Cao W, Duana Y:Breath Analysis: Potential for Clinical Diagnosis and Exposure Assessment. Clinical Chemistry. 2006; 52: 800-811.

Cassano AG, Wang B, Anderson DR, Previs S, Harris ME, Anderson VE: Inaccuracies in selected ion monitoring determination of isotope ratios obviated by profile acquisition: nucleotide 18O/16O measurements. Anal Biochem. 2007; 367: 28-39.

Chen X, Xu F, Wang Y, Pan Y, Lu D, Wang P, Ying K, Chen E, Zhang W: A study of the volatile organic compounds exhaled by lung cancer cells in vitro for breath diagnosis. Cancer. 2007; 110: 835-844.

Chiorazzi N, Rai KR, Ferrarini M: Chronic lymphocytic leukemia. N. Engl. J. Med. 2005; 352: 804–815.

Culliford A, Daly J, Diamond B, Rubin M, Green PH. The value of wireless capsule endoscopy in patients with complicated celiac disease. Gastrointest Endosc. 2005; 62: 55-61.

Deng C, Zhang X, Li N: Investigation of volatile biomarkers in lung cancer blood using solid-phase microextraction and capillary gas chromatography-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci. 2004; 808: 269-277.

Enderby B, Smith D, Carroll W, Lenney W: Hydrogen Cyanide as a biomarker for Psuedomonas aeruginosa in the breath of children with cystic fibrosis. Pediatr Pulmonol. 2009 [epub ahead of print]

Gordon SM, Szidon JP, Krotoszynski BK, Gibbons RD, O'Neill HJ: Volatile organic compounds in exhaled air from patients with lung cancer. Clin Chem. 1985; 31: 1278-1282.

Grabitske HA, Slavin JL: Low-digestible carbohydrates in practice. J Am Diet Assoc. 2008; 108: 1677-81.

Hakalehto E, Humppi T, Paakkanen H: Dualistic acidic and neutral glucose fermentation balance in small intestine: Simulation in vitro. Pathophysiology. 2008; 15: 211-20.

Halliwell B and Gutteridge JMC. (2003) Free radicals in biology and medicine. (Third edition) New York: Oxford University Press.

Harnett W, Harnett MM: Phosphorylcholine: friend or foe of the immune system? Immunol Today. 1999; 20: 125-129.

Hay Burgess DC, Wasserman J, Dahl CA: Global health diagnostics. Nature. 2006; 444 Suppl 1:1-2.

Hewitt CN, Hayward S, Tani A: The application of proton transfer reaction-mass spectrometry (PTR-MS) to the monitoring and analysis of volatile organic compounds in the atmosphere. J Environ Monit. 2003; 5: 1-7.

Ihde DC, Minna JD: Non-small cell lung cancer. Part I: Biology, diagnosis, and staging. Curr Probl Cancer. 1991; 15: 61-104.

Jemal A, Siegel R, Ward E: Cancer statistics, 2008. *CA Cancer J Clin*. 2008; 58 :71-96.

Johnson PC, Ettinger RA: Heuristic model for predicting the intrusion rate of contaminant vapors into buildings. Environ. Sci. Technol. 1991; 25: 1445–1452.

Kalemkerian GP: Biology of lung cancer. Curr Opin Oncol. 1994; 6: 147-55.

Kim K, Rice TW, Murthy SC: Combined bronchoscopy, mediastinoscopy, and thoracotomy for lung cancer: who benefits? *J Thorac Cardiovasc Surg*. 2004; 127: 850-856.

Koning F, Gilissen L, Wijmenga C: Gluten: a two-edged sword. Immunopathogenesis of celiac

disease. Springer Semin Immunopathol. 2005; 27: 217-232.

Mandelstam J, McQuillen K, Dawes I. (1986) Biochemistry of bacterial growth (Third Edition) London: Blackwell Scientific Publications.

Manser RL, Irving LB, de Campo MP, Abramson MJ, Stone CA, Pedersen KE, Elwood M, Campbell DA: Overview of observational studies of low-dose helical computed tomography screening for lung cancer. Respirology. 2005; 10: 97-104.

Miekisch W, Schubert JK, Noeldge-Schomburg G: Diagnostic potential of breath analysis— focus on volatile organic compounds. Clinica Chimica Acta. 2004; 347: 25-39.

Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA: Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. Mayo Clin Proc. 2008; 83: 584-94.

Moorhead KT, Lee D, Chase JG, Moot AR, Ledingham KM, Scotter J, Allardyce RA, Senthilmohan ST, Endre Z: Classifying algorithms for SIFT-MS technology and medical diagnosis. Comput Methods Programs Biomed. 2008; 89: 226-238.

Mukhopadhyay R :Don't waste your breath. Researchers are developing breath tests for diagnosing diseases, but how well do they work? Anal Chem. 2004; 76: 273–276.

Niewinski MM: Advances in celiac disease and gluten-free diet. J Am Diet Assoc. 2008; 108: 661-672.

O'Neil HJ, Gordon SM, O'Neil HM, Gibbons RD, Szidon JP: A computerized classification technique for screening for the presence of breath biomarker in lung cancer. Clin Chem. 1988; 34: 1613-1618.

Peeling RW, Smith PG, Bossuyt PM: A guide for diagnostic evaluations. Nat Rev Microbiol. 2006; 4(9 Suppl): S2-6.

Phillips M, Greenberg J: Method for the collection and analysis of volatile compounds in the breath. J. Chromatogr. 1991; 564: 242-249.

Phillips M, Cataneo RN, Cummin AR, Gagliardi AJ, Gleeson K, Greenberg J, Maxfield RA, Rom WN: Detection of lung cancer with volatile markers in the breath. Chest. 2003; 123: 2115-2123.

Phillips M, Cataneo RN, Ditkoff BA, Fisher P, Greenberg J, Gunawardena R, Kwon CS, Tietje O, Wong C: Prediction of breast cancer using volatile biomarkers in the breath. Breast Cancer Res Treat. 2006; 99: 19-21.

Phillips M, Altorki N, Austin JH, Cameron RB, Cataneo RN, Greenberg J, Kloss R, Maxfield RA, Munawar MI, Pass HI, Rashid A, Rom WN, Schmitt P: Prediction of lung cancer using volatile biomarkers in breath. Cancer Biomark. 2007; 3: 95-109.

Piiper J, Scheid P: Respiration: alveolar gas exchange. Annu Rev Physiol. 1971; 33: 131-54.

Polanco I. Celiac disease. J Pediatr Gastroenterol Nutr. 2008; 47: S3-S6.

Poli D, Goldoni M, Caglieri A, Ceresa G, Acampa O, Carbognani P, Rusca M, Corradi M: Breath analysis in non small cell lung cancer patients after surgical tumour resection. Acta Biomed. 2008; 79: 64-72.

Presutti RJ, Cangemi JR, Cassidy HD, Hill DA. Celiac disease. Am Fam Physician. 2007; 76: 1795-1802.

Risby TH, Solg SF: Current status of clinical breath analysis. Appl. Phys. B. 2006; 85: 421–426.

Rolla G, Bruno M, Bommarito L, Heffler E, Ferrero N, Petrarulo M, Bagnis C, Bugiani M, Guida G: Breath analysis in patients with end-stage renal disease: effects of haemodialysis. Eur J Invest. 2008; 38: 728-733.

Ruzsanyi V, Baumbach JI, Sielemann S, Litterst P, Westhoff M, Freitag L: Detection of human metabolites using multi-capillary columns coupled to ion mobility spectrometers. J. Chromatogr., A. 2005; 1084: 145–151.

See J, Murray JA: Gluten-free diet: the medical and nutrition management of celiac disease. Nutr Clin Pract. 2006; 21: 1-15.

Setty M, Hormaza L, Guandalini S: Celiac disease: risk assessment, diagnosis, and monitoring. Mol Diagn Ther. 2008; 12 :289-298

Smith D, Adams NG: The selected ion flow tube (SIFT): studies of ion neutral reactions. Adv. atom mol Phys. 1987; 24: 1-49.

Smith D, Spanel P, Thompson JM, Rajan B, Cocker J, Rolfe P: The Selected Ion Flow Tube Method for Workplace Analyses of Trace Gases in Air and Breath: Its Scope, Validation, and Applications Applied occupational and environmental hygiene 1998; 13: 817-823.

Smith D, Spanel P, Davies S: 1999 Trace gases in breath of healthy volunteers when fasting and after a protein- calorie meal: a preliminary study J. Appl. Physiol. 1999; 87: 1584–1588.

Smith D, Wang T, Sulé-Suso J, Spanel P, El Haj A: Quantification of acetaldehyde released by lung cancer cells in vitro using selected ion flow tube mass spectrometry. Rapid Commun Mass Spectrom. 2003; 17: 845-50.

Smith D, Spanel P: Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. Mass Spectrometry Reviews 2005; 24: 661– 700.

Smith D, Spanel P: The challenge of breath analysis for clinical diagnosis and therapeutic

monitoring. The Analyst. 2007; 132: 390-396.

Song G, Qin T, Liu H, Xu GB, Pan YY, Xiong FX, Gu KS, Sun GP, Chen ZD: Quantitative breath analysis of volatile organic compounds of lung cancer patients. Lung Cancer. 2009 Apr 29. [Epub ahead of print]

Spada C, Riccioni ME, Urgesi R, Costamagna G: Capsule endoscopy in celiac disease. World J Gastroenterol 2008; 14: 4146-4151.

Spanel P, Rolfe P, Rajan B, Smith D: The selected ion flow tube (SIFT) - a novel technique for biological monitoring. Ann. Occup. Hyg. 1996; 40: 615-626.

Spanel P and Smith D: SIFT studies of the reactions of H3O+, NO+ and O2+ with a series of alcohols. Int. J. Mass Spectrom. Ion Processes. 1997; 167/168: 375-388.

Spanel P, Ji Y and Smith D: SIFT studies of the reactions of H3O+, NO+ and O2+ with a series of aldehydes and ketones. Int. J. Mass Spectrom. Ion Processes. 1997; 165: 25-37.

Spanel P, Dryahina K, Smith D: A general method for the calculation of absolute trace gas concentrations in air and breath from selected ion flow tube mass spectrometry data. Int. J. Mass Spectrom. 2006; 249/250: 230-239.

Spanel P, Smith D: Influence of weakly bound adduct ions on breath trace gas analysis by selected ion flow tube mass spectrometry (SIFT-MS). Int. J Mass Spectrom. 2009; 280: 128-135.

Spiro SG, Gould MK, Colice GL: Initial evaluation of the patient with lung cancer: symptoms, signs, laboratory tests, and paraneoplastic syndromes: ACCP evidenced-based clinical practice guidelines (2nd edition). Chest. 2007; 132 Suppl 3:149S-160S.

Stahel RA: Biology of lung cancer. Lung Cancer. 1994; 10 Suppl 1: S59-65.

Tassopoulos CN, Barnett D, Fraser TR: Breath-acetone and blood-sugar measurements in diabetes. Lancet. 1969; 1: 1282-1286.

Tichenor, BA, Mason, MA: Organic emissions from consumer products and building materials to the indoor environment, Journal of the Air Pollution Control Association. 1988; 38: 264–268.

Toh CK: The changing epidemiology of lung cancer. Methods Mol Biol. 2009; 472:3 97-411.

Turner C, Spanel P, Smith D: A longitudinal study of ammonia, acetone and propanol in the exhaled breath of 30 subjects using selected ion flow tube mass spectrometry, SIFT-MS. Physiol Meas. 2006; 27: 321-337.

Urdea M, Penny LA, Olmsted SS, Giovanni MY, Kaspar P, Shepherd A, Wilson P, Dahl CA, Buchsbaum S, Moeller G, Hay Burgess DC: Requirements for high impact diagnostics in the developing world. Nature. 2006; 444 Suppl 1: 73-79.

Wallace LA, Pellizzari ED, Leaderer B, Zelon H, Sheldon LS: Emissions of volatile organic compounds from building materials and consumer products. Atmos. Environ. 1987; 21: 385-391.

Wang B, Sun G, Anderson DR, Jia M, Previs S, Anderson VE: Isotopologue distributions of peptide product ions by tandem mass spectrometry: quantitation of low levels of deuterium incorporation. Anal Biochem. 2007; 367: 40-48.

West JB: Blood flow to the lung and gas exchange. Anesthesiology. 1974; 41: 124-38.

Westcott JL, Rao N, Colley DP: Transthoracic needle biopsy of small pulmonary nodules. Radiology. 1997; 202: 97-103.