Translational Application of microRNA Profiling to Detect Non-Small Cell Lung Cancers

by

Jennifer Gyoba

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in
Experimental Surgery

Department of Surgery
University of Alberta

Abstract

Lung cancer has the highest mortality rates of all cancers worldwide, with a 5-year survival rate less than 15%. Screening methods are in need for the high risk population, as lung cancer is asymptomatic in its early stages. Proper screening methods would allow earlier diagnosis and curative intent treatment. microRNAs (miRNAs) are small, non-coding strands of ribonucleic acid (RNA) that are shown to lead to carcinogenesis when dysregulated. They are stable and detectable in small quantities, thus are promising candidates for biomarkers. miRNAs are also expressed in a tissue specific manner and measurable in small quantities of different biological fluids.

In chapters 2 and 3, we show that in our nested case control study, a risk score analysis comparing miRNAs 21, 150, 210 and 223 in early stage non-small cell lung cancers (NSCLC) matched with similar age and smoking history controls, showed that miRNA profiling could be used as a screening method when measured in blood plasma. We also showed that pre-operative and post-operative NSCLC miRNA levels stay dysregulated 5-8 months post tumour resection, regardless of cancer recurrence or metastasis.

In chapter 4 we discuss the results, which demonstrate the benefits of using miRNAs as a screening method for NSCLC, but also that it is not viable to be used as a test on its own. We suggest, in order to improve miRNAs screening capabilities, that the test be combined with another method, such as low-dose computed tomography (CT) scanning, to improve early detection in the high-risk population.

Preface

This thesis is the original work of Jennifer Gyoba. No part of this thesis has been previously published in an academic journal. The research project received ethics approval from the Health Research Ethics Board of Alberta (HREBA) on June 8th, 2015 (HREBA.CC-15-0023). All research analysis and manuscript composition was conducted by Jennifer Gyoba, with the statistical analysis conducted by Dr. Sunita Ghosh.

Dedication

This thesis is dedicated to my sister, Lex, and my father, Markus. Thank you for all of the support.

Acknowledgements

I would like to thank my supervisors Dr. Eric Bédard and Dr. Wilson Roa for the opportunity to learn and grow through this MSc program. The amount of guidance, support, and patience afforded to me has not gone unrecognized, and I thank you both for all of the past and future opportunities this experience has given, and will give, me.

I would also like to thank Dr. Linghong Guo. He went out of his way to teach me everything I needed to know to complete this project and more. Dr. Guo went above and beyond to ensure that I would be successful. I would also like to thank Dr. Thomas Churchill. He always made time when I needed someone to talk to about my thesis or issues outside of my studies. I thank Chandra Strasbourg and Tracey Zawalusky for their help and guidance regarding conferences, preparing for my thesis, and information I needed to get to this point.

I thank my committee for being a part of my graduate studies journey. I also would like to thank all of the lab members that have come and gone over the years for their help and encouragement.

I would also like to thank the Alberta's Tomorrow Project for generously donating samples to achieve this research.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AC	Adenocarcinoma
Ago2	Argonaute2
AKT	Protein Kinase B
ALK	Anaplastic Lymphoma Kinase
ATP	Alberta's Tomorrow Project
AUC	Area Under the Curve
BAC	Bronchioloalveolar Carcinoma
C. elegans	*Caenorhabditis elegans*
cel-miR-39	*Caenorhabditis elegans* microRNA 39
C.I.	Confidence Interval
c-myc	Cellular Myelocytomatosis
COPD	Chronic Obstructive Pulmonary Disease
CRS	Combined Risk Score
CT	Computed Tomography
C_T	Cycle Threshold
DGCR8	DiGeorge Syndrome Critical Region 8
DNA	Deoxyribonucleic Acid
dsRNA	Double-stranded Ribonucleic Acid
EGFR	Epidermal Growth Factor Receptor
ELM4	Echinoderm Microtubule-associated Protein-like 4
JNK	c-Jun N-terminal Kinase
KRAS	Kirsten Rat Sarcoma
LCC	Large Cell Carcinoma
MAPK	Mitogen-activated Protein Kinases

MAP2K3	Mitogen-activated Protein Kinase Kinase 3
miRNAs	Micro-Ribonucleic Acids
mRNAs	Messenger Ribonucleic Acids
NLST	National Lung Screening Trial
NSCLC	Non-small Cell Lung Cancers
OR	Odds Ratio
PET	Positron Emission Tomography
Pol II	RNA Polymerase II
Pol III	RNA Polymerase III
qRT-PCR	Quantitative Reverse Transcriptase Polymerase Chain Reaction
RISC	RNA-induced Silencing Complex
RNA	Ribonucleic Acid
RNase	Ribonuclease
ROC	Receiver Operating Curve
RQ	Relative Expression
SCC	Squamous Cell Carcinoma
SCLC	Small Cell Lung Cancers
SD	Standard Deviation
SDS	Sequence Detection System
SPSS	Statistical Package for the Social Sciences
TNM	Tumour, Lymph Node, and Metastasis
TP53INP1	Tumour Protein-53-Induced-Nuclear-Protein 1
3'UTR	3'-untranslated region

Chapter 1: Introduction

1.1: Introduction of Translational Application of microRNA Profiling to Detect Non-Small Cell Lung Cancers

1.1.1: Lung Cancer Overview

In 2012, approximately 14.1 million new cancer cases were diagnosed worldwide. Lung cancer, the most commonly diagnosed type of cancer, was responsible for 1.83 million new cases, as well as 1.59 million deaths in 2012 (1). In Canada, it was estimated in 2016 that lung cancer was the second and third most commonly diagnosed cancer in women and men, respectively, with 14.1% of new female cancer cases and 14% of new male cancer cases being lung cancer. In addition, lung cancer in Canada has the highest mortality rates of approximately 26% for both genders (2).

The main risk factor for developing lung cancer is smoking tobacco, which is the main risk factor in 85% of all cases of lung cancer in Canada. Other risk factors include exposure to second-hand smoke, asbestos, radon, arsenic, diesel engine exhaust, mine dust, silica dust, and radiation. Other lung diseases, such as tuberculosis and chronic obstructive pulmonary disease (COPD), as well as a familial history of lung cancer are also risk factors (3).

There are two main groups of lung cancer, non-small cell lung cancer (NSCLC) and small-cell lung cancer (SCLC). NSCLC is the most common, making up approximately 85-90% of all lung cancer cases while SCLC is responsible for approximately 10-15% of all lung cancer cases (4).

Over 75% of all lung cancer cases are diagnosed in stages III/IV, which are both associated with very poor survival rates (2). In patients presenting with advanced lung cancer, treatment options decrease and the disease is treated to slow down progression, rather than with curative intent.

Given the high rates of initial presentation for patients with stage III/IV NSCLC, there is an obvious need for methods to assist in earlier diagnoses to allow for earlier treatment intervention.

1.1.2: Lung Cancer Screening

Screening high risk populations (smokers/ex-smokers, workers exposed to harmful chemicals, and age) potentially allows the lung cancer to be detected in its early stages and be treated more successfully, and in turn, prolongs the life of the patients. Currently, the protocols used are through computed tomography (CT) scans. Issues that accompany the use of CT scanning as a screening method for lung cancer are high false positive rates, repeated risk of exposure to radiation, and complications from follow-up tests after the CT scan (2). Therefore, a need exists to develop adjunct methods for screening the high risk population, or methods to use in conjunction with current screening protocols.

1.1.3: Use of Biomarkers as a Screening Method: microRNA

A new approach that is being investigated is the use of micro-ribonucleic acids (miRNAs) as biomarkers for different diseases. miRNAs are small, non-coding RNA strands that are prime candidates for biomarkers as they are tissue-specific and stable in small quantities. Increases and decreases of specific miRNAs are characteristic of the onset of certain types of cancer (5, 6, 7).

1.1.4: Thesis Aims

The aim of this thesis is to explore how measuring several miRNAs and creating a miRNA profile could have the ability to distinguish early stage NSCLC from healthy controls. Specifically, this involves the comparison of stage I/II NSCLC versus healthy controls and pre-operative versus post-operative NSCLC samples with or without recurrence, as well as the comparison of pre-operative samples to tumour tissue, in order to distinguish one from the other using miRNA profiling. We hypothesize that miRNA profiling has the predictive ability needed in a screening test to distinguish early stage NSCLC from healthy controls.

The thesis will continue further into Chapter 1 to explore lung cancer, miRNAs, and their potential as a non-invasive screening method for the high risk population. In Chapter 2 all methodologies of the projects' experiments will be explained, with Chapter 3 moving into the results of the thesis. In Chapter 4 a discussion will continue to evaluate the results of the project, as well as look into limitations and future directions that can come from these experiments. Lastly, Chapter 5 will conclude the thesis.

1.2: Clinical Manifestations of Lung Cancer:

1.2.1: Classifications of Lung Cancer

Lung cancer is separated into two categories – NSCLC and SCLC. Both categories split into several different cellular subtypes (Figure 1.1).

NSCLC accounts for 85-90% of all cases of lung cancer. Adenocarcinoma (AC), which usually occurs in the periphery of the lung, is the most common type of NSCLC. Types of adenocarcinoma are predominantly grouped via their microscopic appearance. The acinar subtype appears as tiny sac-like structures in cells, papillary subtype present as small, finger-like projections in cells, micropapillary subtype are even smaller finger-like projections in cells, and

the solid subtype appear as nest-like, thick structures. There is also a mixed type, where the tumour's microscopic appearance have a combination of different features. Another type of AC is Bronchioloalveolar Carcinoma (BAC), which is found within the respiratory airways and alveoli without invasion into the basement membrane. The cells originate in the alveoli, and tend not to spread outside of the lung, giving a better prognosis compared to other AC's (2).

There is also Squamous Cell Carcinoma (SCC), another type of NSCLC. It is primarily found in the hilum, or root, of the lung. They are categorized via their microscopic structure such as the papillary subtype, clear cell subtype wherein cells appear empty and clear, and basaloid subtype which appear as small round shaped cells. SCC tumours also tend to have mixed microscopic structures (2).

Large Cell Carcinoma (LCC) is categorized as a NSCLC which can have basaloid structured cells, and clear cell type. Other subtypes of NSCLC, which are much less common than those previously stated, are sarcomas (fibrosarcoma, leiomyosarcoma, and hemangiopericytoma), sarcomatoid carcinoma, lymphoma, and superior sulcus (Pancoast) tumours (2).

SCLC encompasses 10-15% of all cases of lung cancer. Categories included in this type are small cell carcinoma and combined small cell carcinoma. SCLC is comprised of small flat shaped cells that are spread throughout the affected area (2).

1.2.2: Diagnosis, Treatment and Prognosis of Lung Cancer

People are most commonly diagnosed with lung cancer because of complaints to their family doctor about their symptoms at an average age of 70 years. This leads to the patient being sent for other tests to confirm suspicions of disease. Tests often given are x-rays to look for abnormalities in the lungs, CT scans to image the size and location of the tumour, and positron

emission tomography (PET) scans to visualize the condition of surrounding lymph nodes. Other tests used are analyzing sputum from the lung for cancer cells and ultrasound. Issues with these methods are that none can conclusively diagnose lung cancer. The current and only protocol used to do this is via biopsy (2).

Different procedures can be used to obtain a tissue sample of the lesion or nodule found in imaging tests. The first is a needle biopsy, in which a needle goes through the chest wall percutaneously via CT guidance or fluoroscopy. Another is a transbronchial biopsy where the tissue sample is obtained through bronchoscopy. These two biopsy techniques are usually done with the patient awake or with light sedation. There are similar risks involved with these procedures such as pneumothorax, bleeding in the lungs, and infection (4).

Other procedures used are performed under general anesthesia. Thoracoscopic biopsy is a minimally invasive procedure with entry through the chest wall, where a tissue sample can be obtained during the removal of nodules. The last technique is an open biopsy, in which the surgeon makes incisions through the chest in order to obtain the sample. The risks for these techniques are blood loss or clots, infection, and the associated dangers with being put under general anesthesia. Also, patients that receive these type of biopsies require some hospital stay (4).

The available types of biopsy to diagnose lung cancer are invasive, and some patients may not be in a state to undergo the procedures due to other health issues. Another issue with invasive biopsies is the large amount performed unnecessarily. Imaging techniques that screen for lung cancer, such as CT scans or chest radiography, have high false positive rates. These false

positives lead to many biopsies being performed, increasing hospital wait times. Unfortunately, there are no other means to definitively diagnose lung cancer (4).

After diagnosis, doctors prepare a treatment regime for the patient. This is highly dependent on the type and stage of the lung cancer. In stage I, the treatment is primarily surgery to excise the affected tissue. In stage II, the size of the tumour influences how treatment is approached. For large tumours, radiation and chemotherapy may be used before surgery to decrease the size of the tumour, and increase overall the chances of complete removal of the mass during surgery. If the tumour is of moderate size, surgery may be attempted initially. In stage III lung cancer, surgery with chemotherapy and/or radiation are commonly used.

Stage IV lung cancer uses chemotherapy as its primary treatment. Concurrent chemotherapy may be used with radiation therapy, and targeted and maintenance chemotherapy may be used depending on the patients' disease characteristics. Surgery is not performed to assist in removal of the cancer as it has now metastasized, but may be used as palliative care (2).

The prognosis for lung cancer is inversely correlated to the stage at time of diagnosis. As previously stated, lung cancer has the highest mortality rate of all the cancers in Canada, as well as worldwide. In Canada, 19.1% of cases are diagnosed in stage I and 5.6%, 27.4%, and 47.9% of cases are diagnosed at stage II, III, and IV, respectively for NSCLC (2) (Table 1.1). Due to approximately 75% of patients being diagnosed in stages III or IV, it becomes apparent why treatment is rarely curative post diagnosis. Looking at data from the United States, we can see how many people survive at least five years after diagnosis with NSCLC. 49% and 45% of patients diagnosed at stages IA and IB, respectively, will survive past five years. When diagnosed in stages IIA and IIB, survival rates are 30% and 31%, respectively. In later stages of

lung cancer, survival rates past five year fall to 14%, 5%, and 1% in stages IIIA, IIIB, and IV, respectively (8) (Table 1.1).

1.2.3: Lung Cancer Screening

In order for a screening test to be effective, there are many criteria that should be achieved. The first being that the disease has serious consequences, or, that the targeted disease has high mortality rates (9). As previously stated, lung cancer has the highest mortality rates out of all of the cancers worldwide (1), fulfilling this requirement.

Another criteria is that there is a high prevalence of detectable disease in the screening population during the preclinical phase (9). The preclinical phase is defined as the time between the onset of the disease and the onset of symptoms, and in the case of lung cancer, the screening population would be people between the ages of 55-74 years with a smoking history of 30 pack-years that are currently smoking or have quit in the last 15 years (10, 11). Also, a good screening test should not detect symptoms that appear to be a disease (12). Currently, CT scans are the method used for screening but it is estimated that there are approximately 19 false positive scans per one true positive scan (10). Thus, both of these requirements need improvement.

The next requirement is that the screening test has high accuracy for detecting the disease in its preclinical phase. This is measured in sensitivity and specificity (Figure 1.2). Sensitivity is a measure to obtain how well a screening method can detect the disease from persons who are inflicted by the disease. When a test has a high sensitivity, it is more reliable with a negative result. Therefore, this is useful for ruling out disease from a population. When using specificity, one is measuring how well a screening method can detect a healthy person from a group of people without disease. In contrast to sensitivity, specificity is useful for ruling in disease. For a

screening test to be considered acceptable, it must have > 95% sensitivity if the specificity is ≤ 95%, or if the sensitivity is ≤ 95% the specificity must exceed 95% (12). It is also necessary that the test detects the disease before a "critical point" in its progression. Generally, when considering cancer, this would be perceived as before the cancer metastasizes (stages I-III) (12). In the specific case of lung cancer, stages I and II have 5-year survival rates varying between 30-49% then drop severely in stages III and IV (Table 1.1) (8). Currently, literature does not have a defined critical point for lung cancer, thus for simplicity purposes, stages I and II will be regarded as before the critical point of disease. Screening tests should also cause little danger to the participants and not increase morbidity of the disease (12). It is acceptable to have a small effect on the patient, such as small doses of radiation via CT scans, as long as it gives some benefit (13). The use of low-dose CT scans have a risk of development of cancer of approximately one in 25,000 (14-16).

The screening test used to detect a disease must also be affordable and available to the high risk population that requires it (12). CT scans have shown issues with availability to populations living in rural areas. In Canada, the estimated wait for someone to receive a CT scan is 3.7 weeks (17). There are not any studies looking at the costs specifically for CT screening and lung cancer, but, lung cancer does place a huge cost on the Canadian economy and healthcare system. The Conference Board of Canada performed an analysis in 2012 that showed lung diseases (including asthma and COPD) costs the government $12 billion a year. A large portion of the cost is due to the expensive treatments associated with lung cancer and the frequent hospitalization required for patients. It is predicted that this cost will double by the year 2030, showing the increasing pressure for better diagnosis and screening methods for the disease (18).

Screening tests must also improve patient outcomes. Thus, treatment must exist for the disease as well as it being more effective when the patient is asymptomatic (9, 12). Surgery, radiation and chemotherapy have shown to be overall successful in stages I and II, fulfilling this requirement. There are different methods that have been tested as possible screening techniques for lung cancer. The first is chest radiography, with several studies conducted in the 1970's and 80's. The United States National Cancer Institute conducted the Memorial Sloan-Kettering Study (19, 20) and the Johns Hopkins Study (21, 22). They used random control studies to evaluate the use of chest radiography and its effects on mortality rates in lung cancer.

An issue with chest radiography is the inability to locate the exact place of the lesion seen on the image, as it does not use cross-sectional imaging. Because these studies were not found to be significant, investigations into CT scanning commenced as an alternative method of screening for lung cancer.

The first study to show that low dose CT scans can decrease lung cancer mortality rates was the National Lung Screening Trial (NLST), a randomized control study comparing low dose CT scans with chest radiography by measuring mortality rates in current or previous heavy smokers. Overall, the study was a success showing that CT scans produced 20.3% less cancer related deaths than chest radiography, and a 6.7% decrease in overall mortality (23).

Despite the success of the NLST, further research has been conducted to investigate its results via external labs. In 2013, a team showed that there was an overdiagnosis of 18% in the low-dose CT scan group (24).

Currently, the American Cancer Society guidelines recommend low-dose CT scans for screening lung cancer in the high risk population. However, these guidelines only apply to individuals who are 55-74 years of age, and have a 30-pack year smoking history and currently smoke or have

quit within the past 15 years (10, 11). As previously stated, the average age of diagnosis for lung cancer is approximately 70 years old, making the 74 year old cut off affecting many potential patients.

Another study showed that CT scans have a low specificity of only 61% for lung cancer diagnosis. This is problematic for the healthcare system and the patients' wellbeing. It is estimated that there are approximately 19 false positive scans per one true positive scan using low-dose CT (25). This leads to a tremendous amount of unneeded investigations, costing the healthcare system large amounts and subsequent patient anxiety. The use of CT scans also brings up the risk of repeated radiation exposure, and the increased chances of developing cancer (10).

Additionally, PET scans have a specificity and sensitivity of nearly 90%, but this comes at a cost of high expense and does not alter the outcome of surgical resection for a high percentage of patients (26).

Therefore, there is a vital need for development of minimally invasive, cost-effective, easy to administer and approachable tests that can replace or be used in conjunction with CT scans to augment their low specificity value.

1.2.4: Lung Cancer Biomarkers

1.2.4.1: Epidermal Growth Factor Receptor (EGFR)

Epidermal Growth Factor Receptor (EGFR) is a transmembrane protein that, when activated, initiates a signaling pathway that regulates many cell functions such as cell growth, proliferation, differentiation, and survival (27-29). The main downstream pathways to the EGFR pathway are

mitogen-activated protein kinases (MAPK), Protein Kinase B (AKT), and c-Jun N-terminal Kinase (JNK) (30).

Mutations in EGFR has been well documented regarding different types of cancers, specifically mutations that lead to an overexpression of EGFR and uncontrolled cell division (31), such as colorectal (32), brain (33), head and neck (34), and others. For lung cancer specifically, in 80% of SCC cases there is an EGFR mutation (31).

EGFR's main use regarding lung cancer is in treatment protocols. Specifically, it has become a target for many drugs, leading to the development of EGFR inhibitors. Specifically, afatinib, brigatinib, erlotinib, gefitinb and icotinib (35, 36). There has also been development of a vaccine that targets EGFR mutations in NSCLC through raising antibodies specific to the mutated protein ensuring it cannot be stimulated (37, 38). Clinical trials separating patients into EGFR-positive and EGFR-negative groups, as well as pursuing a more specific treatment regime using EGFR inhibitors has shown a 60% response rate in NSCLC (39). Issues have arisen with the use of EGR inhibitors as many patients grow resistant to the drugs they are given making it difficult to control the growth of the malignant tumours over time (39).

1.2.4.2: Kirsten Rat Sarcoma (KRAS)

Kirsten rat sarcoma (KRAS) was identified in the Kirsten rat sarcoma virus, functioning as a proto-oncogene (40). When this proto-oncogene is activated, it results in a high amount of growth factor protein recruitment, resulting in overall cell proliferation (41). Mutations in the KRAS gene have shown implications in many different types of cancer, such as lung adenocarcinoma, leukemia, pancreatic cancer, and colorectal cancer (42-47).

In lung cancer specifically, there are not any approved drugs that target KRAS mutations. However, patients who exhibit both EGFR and KRAS mutations do not show a response to the drugs erlotinib and gefitinib compared to EGFR-positive patients alone (48-50).

1.2.4.3: Anaplastic Lymphoma Kinase (ALK)

Anaplastic Lymphoma Kinase (ALK) is an enzyme that plays different roles in brain development and the nervous system (51).

The mutation of ALK has been shown in many different types of cancer, specifically brain, esophageal, breast, colorectal, lung, and thyroid cancers (52-65). In NSCLC, 3-5% of cases are due to a gene fusion between ALK and Echinoderm microtubule-associated protein-like 4 (EML4) (66).

Development of ALK inhibitors is being investigated, looking particularly at crizotinib, certinib, and entrectinib in late stage and metastatic lung cancers (67, 68).

1.2.4.4: Autoantibodies

Autoantibodies are antibodies that are produced by patients that target mutated tumour proteins, and overexpression of these have been shown to occur in different cancers. Research has looked into the use of several autoantibodies, resulting in an autoantibody signature, looking at cases versus controls in NSCLC and SCLC. This resulted in a sensitivity of 40% and specificity of 93% in NSCLC and a sensitivity of 55% and specificity of 93% in SCLC (69-73).

1.2.5: Economic Effects of Lung Cancer

Lung cancer places a huge cost on the Canadian economy and healthcare system. The Conference Board of Canada performed an analysis in 2012 that showed lung diseases (including

asthma and COPD) cost the government $12 billion a year. A large portion of the cost is due to the expensive treatments associated with lung cancer and the frequent hospitalizations required for patients. It is predicted that this cost will double by the year 2030, showing the increasing pressure for better diagnostics and screening methods for the disease (18).

Overall, it is clear that lung cancer affects the Canadian economy greatly, and a way to decrease these effects is through early diagnosis of lung cancer for immediate treatment. A screening method that is reliable and accurate is required for this to be achieved. With early intervention, there would be fewer false positive diagnoses, thus lower costs to the healthcare system for surgical biopsies and unneeded treatment, and there would be better survival rates with fewer repeated hospital stays.

1.3: microRNA

1.3.1: microRNA Background

In 1993, scientists Lee et al. investigated the control of larval development in *Caenorhabditis elegans (C. elegans)*. They were looking to identify the process involving the inhibition of the gene *lin-14* by another gene, *lin-4*. *lin-14* is a gene that encodes the protein LIN-14 which is important in the development of the larvae. It is expressed or repressed at different times during development. *lin-4* is a gene that encodes what they thought was a protein that inhibits the translation of *lin-14* gene. However, they discovered instead that a small, non-coding RNA approximately 22 nucleotides long was being produced. It contained partially complementary sequences to the *lin-14* gene, suggesting that its function is to inhibit the translation of the LIN-14 protein (74). This was the first discovery of miRNA, although they did not know what the small protein yet was.

A few years later in 2000, another group characterized another small RNA, *let-7*. They found that *let-7* repressed *lin-41* in *C. elegans* (75). Research continued to show that these RNAs shared similar characteristics, suggesting that they are all part of a large class of small RNAs (76). This led to researchers coining the term "miRNA" to describe the newly discovered group. Further research investigating these small proteins has resulted in over 1000 miRNAs found in the human genome.

Over time, a miRNA annotation system has been created (77). The most common form seen is the standard "miR" prefix followed by a dash and a number, referencing that specific miRNA, for example, miR-21. At times, there are "families" of specific miRNAs, such as when two or more have nearly identical nucleotide sequences, excluding one or two nucleotides. These are annotated by adding a lower case letter at the end of the specific miRNA, for example, miR-146a and miR-146b. To assign a species to the specific miRNA being described, a prefix is added, for example, hsa-miR-21 for humans. Lastly, when two miRNAs originate from the same pre-miRNA, they are denoted with a -3p or -5p suffix, establishing which arm they were transcribed from in the nucleus (77). An example of this is miR-21-3p or miR-21-5p.

1.3.2: microRNA Biogenesis

In the nucleus, RNA polymerase II (Pol II) is a polymerase that binds to a promoter in order to encode a deoxyribonucleic acid (DNA) sequence. For miRNA, Pol II is the usual suspect for transcription (78), although RNA polymerase III (Pol III) also transcribes some miRNAs (79). When the polymerase binds to the promoter to transcribe the miRNA, a hairpin loop pri-miRNA is encoded. This hairpin is a double-stranded RNA (dsRNA), which can at times contain more than one miRNA precursor. The pri-miRNA is further recognized by the protein DiGeorge

Syndrome Critical Region 8 (DGCR8) and the enzyme Drosha. These two proteins associate in order to form the microprocessor complex. pri-miRNA is recognized by this complex due to its dsRNA structure, and the complex continues by cleaving the pri-miRNA by cutting the 3' and 5' ends of the hairpin, resulting in a new structure called pre-miRNA (80, 81), as depicted in Figure 1.3.

Exportin-5, a nucleocytoplasmic shuttler, exports the pre-miRNA out of the nucleus into the cytoplasm (82). Once the pre-miRNA has exited the nucleus, it is further cleaved by the enzyme Dicer. This cleavage step results in the miRNA-duplex, which is no longer a hairpin structure. The duplex has two strands of miRNA, both of which can be used. It is most common that one is degraded while the other continues to become a mature miRNA (83). In order for the miRNA to be functional, it must interact with RNA-induced silencing complex (RISC) and Argonaute 2 (Ago2), as noted in Figure 1.3. The miRNA requires these proteins for orientation to its messenger RNA (mRNA) target (84).

1.3.3: microRNA and Gene Regulation

Overall, miRNAs main function is gene regulation. miRNAs target mRNAs through complementary nucleotide sequences. RISC and Ago proteins orient miRNAs to their target mRNAs. This is done by imperfect base pairing in the 3'-untranslated region (3'-UTR) of the miRNA, leading to translational repression of the mRNA or target degradation via deadenylation (85-87).

Due to the size of miRNAs compared to mRNAs, some miRNAs have several target mRNAs, causing great difficulty in identifying all of its targets (87-90). Several attempts to create

algorithms to identify these targets through base pair matching has shown some success but with high rates of false positive findings (91-93).

The most established mechanism of miRNA gene regulation is through mRNA cleavage via Ago proteins. These proteins have small RNA binding modules that allow the miRNAs to attach, cleave, and deadenylate the mRNAs, overall decreasing their function in the cell (94).

Although there is great difficulty is understanding the exact mechanisms of miRNA and its targets, there are substantial amounts of research done on its roles in lung cancer, which will be discussed later in this chapter.

1.3.4: Quantification of microRNAs

The most common way to measure miRNAs is by quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) and/or RT-PCR analysis (95). It can be measured from many biological tissues such as blood, sputum, tissue, etc. miRNAs are more sensitive than mRNAs, and therefore many precautions must be taken to get true measurements of the proteins. First, miRNAs must be analyzed in a ribonuclease (RNAse) free environment because of its small nucleotide length, as it is easily degraded. miRNAs are required to be kept on ice to ensure that there is not further miRNA breakdown. Precautions must also be taken when storing miRNAs. It is recommended that before miRNAs are isolated from a biological fluid such as blood, it should be frozen at -80˚C as soon as possible (preferably within 4 hours post blood withdrawal) and not thawed until miRNA isolation is to be performed. Once this has occurred, the isolated miRNAs can be kept frozen at -80˚C to ensure it does not breakdown until qRT-PCR can be performed. Generally, researchers are advised to keep small volumes of the isolated miRNAs in aliquots to ensure that there are not repeated freeze-thaw cycles occurring (96, 97).

Microarrays can also be used to quantify miRNA and work by hybridizing the miRNA with probes to different targets that the miRNA can attach to. This allows for a relative measurement. Microarrays allow for hundreds or even thousands of targets to be used in order to measure miRNA (97, 98).

1.4: Clinical Applications of microRNA

1.4.1: microRNA Dysregulation and Disease

miRNAs have been associated with many diseases when they are dysregulated due to mutations. Examples of this are hereditary progressive hearing loss (99), hereditary keratoconus (100), growth defects (101), cardiomyopathies (102-104), cardiogenesis and cardiac conductance (105, 106), heart hypertrophic growth response (107, 108), atherosclerosis (109, 110), and kidney disease (111).

The first disease associated with miRNA dysregulation was leukemia (112).

Further research has found that many miRNAs have links with other types of cancer when dysregulated, leading to some being referred as oncomirs (112). miRNA profiling has been studied for colorectal cancer using blood plasma in early stage patients (113, 114), with specific miRNAs being associated with specific subtypes (115). miR-21 has also shown interaction with the tumour repressor gene mitogen-activated protein kinase kinase 3 (MAP2K3), often found in hepatocellular carcinoma (116), allowing treatment decisions to be made regarding specific gene mutations. Other miRNAs miR-21, 494, and 1973 may be biomarkers that allow predictions of disease responses in Hodgkin's lymphoma when measured in blood plasma (117). Research has gone into investigating miRNAs and breast cancer, as well. miR-205 has shown to be involved in

inhibiting metastasis in breast cancer (118) and miRNAs 141, 200a, 200b, 200c, and 429 have shown to be downregulated in the disease (119).

It has also been shown that miRNAs affect how a cancer develops. This was done through mice engineering, in which the mice produced excess cellular myelocytomatosis (c-Myc). c-Myc is a protein that is implicated in many forms of cancer in its mutated form. By producing excess c-Myc, miRNAs were further measured and compared to how long the mice survived. Mice with high amounts of miRNAs connected to lymphoma developed the disease faster, as well as died faster than mice with lower amounts of miRNA (112).

1.4.2: microRNA in Lung Cancer

1.4.2.1: microRNA in Lung Cancer via Sputum

Several studies investigating a range of miRNAs in sputum to differentiate lung cancer from healthy controls have been performed. These reports have looked at one to several miRNAs at once, including miR-21, 31, 126, 139, 143, 155, 200b, 205, 210, 372, 375, 429, 486, and 708 (6, 11, 120-127). These early studies resulted in sensitivities ranging from 61.5-100% and specificities from 80-100%.

But, in these studies many different protocols were used, such as different internal controls, how lung cancer diagnoses were confirmed, miRNA assays, and if there was utility of training or validation cohorts. In addition to this, many studies did not have appropriate sample sizes.

From these studies, the miRNAs that are able to differentiate lung cancer patients from healthy controls were miRNA-21 (121, 124-127), 210 (6, 11, 122, 123, 125-127), 31 (6, 11, 122, 125), and 155 (121, 126, 127). Consequently, these miRNAs have not been tried together in one

profile, leading to the need for more investigation into this area. Sputum could be a useful medium for measuring miRNAs, as expression levels do not change 7 days post collection (11).

1.4.2.2: microRNA in Lung Cancer via Whole Blood

Studies investigating whole blood miRNA profiling in lung cancer is extremely lacking, with only four studies it total being performed. miR-10b, 21, 190b, 328, 630, 942, and 1284 were measured in these studies, yielding sensitivities ranging between 70-88%, and specificities ranging between 76.9-100% (128-131).

Unfortunately, no overlap is found between the studies, as all looked at different combinations of miRNAs. Also, sample sizes were not sufficient in these studies. miRNA profiling in whole blood requires much more investigation to ensure validity of results.

1.4.2.3: microRNA in Lung Cancer via Blood Serum

miRNA profiling using serum has been greatly investigated in comparison to other biological fluids. miR-7, 15b, 20a, 24, 25, 27b, 125a-5p, 125b, 126, 205, 210, 145, 152, 193a-3p, 194, 199a-5p, 214, 221, 222, 223, 320, 483-5p, 574-5p, 652, 660, and 1254 were all found to be dysregulated when comparing lung cancer patients to healthy controls. Sensitivities in these studies range from 69-100% and specificities range between 66.4-93.4% (132-141).

There was not coherence between the studies in regards to which miRNAs were investigated, with no common miRNAs analyzed. Many of the studies show promising results, but there requires consensus on which miRNAs used in a profile are best able to differentiate lung cancer patients with healthy controls.

1.4.2.4: microRNA in Lung Cancer via Blood Plasma

miR-20a, 21, 210, 145, 155, 182, 197, 221, 223, 486-5p, 944, and 3662 were found to be significantly dysregulated when comparing lung cancer patients with healthy controls in blood plasma. Sensitivities yielded a wide range between 67-91.7% and specificities between 68-96.6% (142-150).

There were overlaps found between the studies regarding which miRNAs were investigated, with miR-21(142, 144, 145, 147, 148), 155 (143, 147), 210 (142, 144), and 486-5p (142, 144) being the most common miRNAs found to be dysregulated.

As previously stated, much more research is needed in order to validate these results although they seem promising, similarly to the previous biological fluids. A study that takes the most significant players when differentiating lung cancer patients from healthy controls is needed to create a test that can aid in screening the high risk population.

1.5: Thesis Hypotheses

The main hypothesis of this thesis is that the use of miRNA profiling can distinguish early stage NSCLC from healthy controls using blood plasma. This would distinguish the change of miRNA levels after lung cancer resection with dependence on cancer recurrence.

The thesis will proceed into chapter 2, where the methodology of the project as well as the materials used will be thoroughly explained, continuing into chapter 3 to go over the results of the project. Chapter 4 will follow to discuss the results, as well as the limitations and implications of the project, and end with conclusions in chapter 5.

Stage of lung cancer	Patients diagnosed according to stage[1] (%)	Patients' 5-year survival rates[2] (%)
Ia/b	19.1	49/45
IIa/b	5.6	30/31
IIIa/b	27.4	14/5
IV	47.9	1

Table 1.1: Patient Statistics According to Lung Cancer Stage for Diagnosis and 5-year Survival Rates

Rates regarding patient diagnosis according to lung cancer stage and 5-year survival rates. 5-year survival rates measure the amount of patients with lung cancer that survive 5-years post diagnosis. There is a noticeable increase in the amount of patients being diagnosed as lung cancer stages advance, and a corresponding decrease in the patients' 5-year survival rates.

[1] Statistics obtained from the Canadian Cancer Society (2016). Rates for stage subtypes not given.

[2] Statistics obtained from the American Cancer Society (2012). Rates for stage subtypes given and thus separated by a slash.

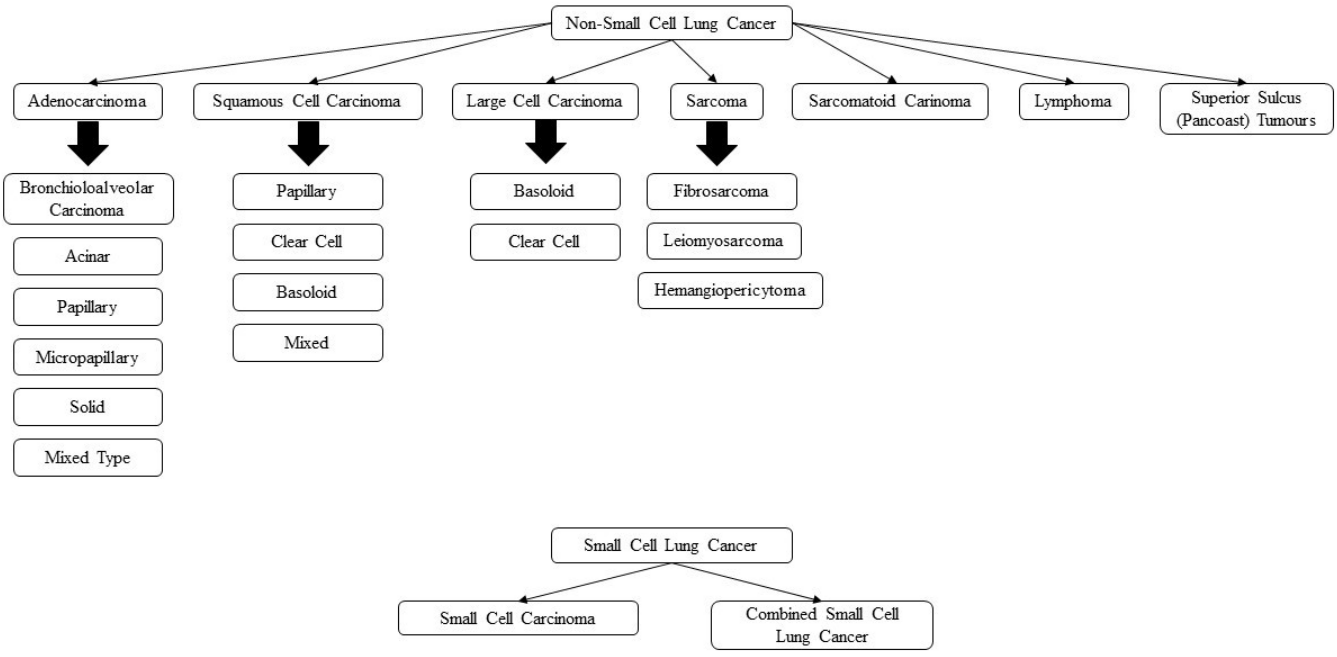


Figure 1.1: Types of Lung Cancer

Non-small cell lung cancer (NSCLC) comprises 85-90% of all lung cancers. Subtypes of NSCLC are adenocarcinoma, squamous cell carcinoma, large cell carcinoma, sarcoma, sarcomatoid carcinoma, lymphoma and superior sulcus tumours. Cell structure classifications are listed beneath each subtype. Small cell lung cancer (SCLC) encompasses 10-15% of all lung cancer cases, with subtypes small cell carcinoma and combined small cell carcinoma.

$$Sensitivity = \frac{\#\ true\ positives}{\#\ true\ positives + \#\ false\ negatives}$$

$$Specificity = \frac{\#\ true\ negatives}{\#\ true\ negatives + \#\ false\ positives}$$

Figure 1.2: Sensitivity and Specificity Equations

Sensitivity is a measure for ruling out disease from a population. It is calculated by dividing the number of true positives by the sum of the number of true positives and false negatives. Specificity is a measure for ruling in disease from a population. It is calculated by dividing the number of true negatives by the sum of the number of true negatives and false positives. Sensitivity and specificity are presented as percentages.

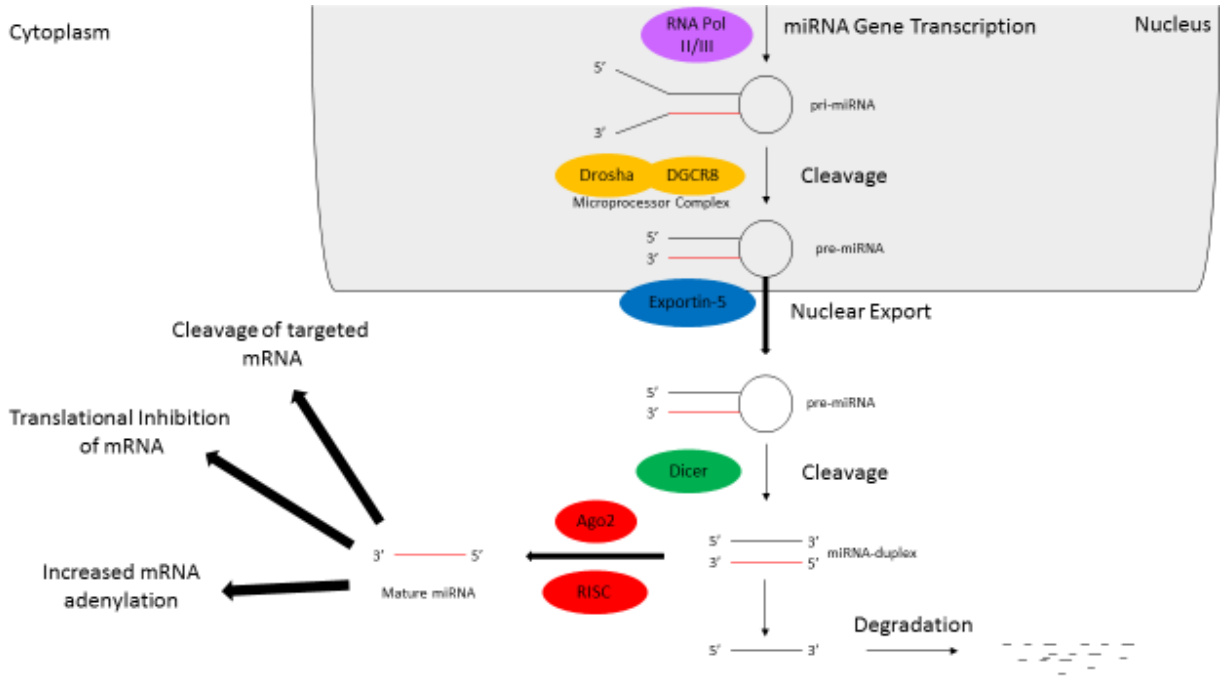


Figure 1.3: microRNA Biogenesis

miRNA biogenesis starts in the nucleus with RNA polymerase II/III (Pol II/III) binding to the promotor of a DNA sequence, transcribing a double stranded hairpin loop pri-miRNA. DiGeorge Syndrome Critical Region 8 (DGCR8) associates with Drosha creating the microprocessor complex and cleaves the pri-miRNA into pre-miRNA. The pre-miRNA is exported from the nucleus into the cytoplasm via exportin-5 and cleaved by Dicer, resulting in a miRNA-duplex. The duplex contains two strands of miRNA, with one commonly being degraded and the other interacting with an RNA-induced silencing complex (RISC) and Argonaute 2 (Ago2), resulting in a mature miRNA that can continue its function in the cell.

Chapter 2: Materials and Methodology

2.1: Experimental Design

2.1.1: Determination of Sample Size

Sample size was determined through the recommendation of a study involving cancer biomarkers for early stage detection. This study developed statistical guidelines through normal approximation for 95% confidence intervals and binomial distribution. The authors recommended that to ensure biases were controlled, the disease required the same control conditions in the study and the screening populations, and also that the test be based off of a combination of biomarkers.

The study should use sensitivity and specificity as a means to measure the ability for the biomarkers to be accurate. Overall, the study recommended a nested case-control analysis containing 110 controls and 70 cases (151).

In our study, we used 4 miRNAs in combination to detect NSCLC, and these specific biomarkers have been researched and found to be dysregulated in NSCLC when measured in blood plasma in previous studies (142-149). Our statistical analysis measured sensitivity and specificity in addition to a risk score analysis. Lastly, the target screening population for lung cancer are persons between the ages 55-74 with extensive smoking history that are still smokers or have quit within the last 15 years, as recommended by the NLST study (23). Our study has attempted to follow these criteria, as the patients included have extensive smoking history. Due to availability of samples, some patients were out of the age range previously mentioned, and how long a patient had been an ex-smoker at the time of sample collection was not known.

Therefore, in this study 110 controls and 70 cases were included in the analysis.

2.1.2: Determination of microRNAs

2.1.2.1: Previous Experimental Findings

In order to determine which miRNAs were to be included in the analysis, our lab looked at our past research findings along with current literature.

In a previous study using NSCLC sputum samples, a cluster analysis was performed (126). Sputum samples were analyzed and qRT-PCR was performed measuring 10 miRNAs in 4 NSCLC cases and 4 controls. miR-21, 92, 143, 145, 155, 182, 205, 210, 372, and let-7a were measured and through hierarchical clustering, miR-21, 143, 155, 210, and 372 were found to be significant biomarkers when comparing cases and controls (Figure 2.1) (126).

A following hierarchical clustering analysis with another 24 NSCLC cases and 4 controls was performed using the 5 significant miRNAs 21, 143, 155, 210, and 372 (126). This analysis found a sensitivity of 83.3% and specificity of 100% when comparing cases and controls (Figure 2.1). Despite the significant results, there were still issues with the study. Both analyses contained patients with early and late stage NSCLC and had small sample sizes. Another issue encountered was the difficulty in sputum collection. Some samples from patients had little sputum, due to difficulty coughing up enough sputum for analysis.

Out of the significant miRNAs found in our previous study, miR-21, 155, and 210 had previously been shown in past literature to be dysregulated in NSCLC using blood plasma, thus we chose to include these 3 miRNAs in our analysis (142-149). miR-223 was also included, as it

has also been investigated in previous literature (146, 148, 149), but has not been measured in conjunction to the other 3 miRNAs in one profile.

2.1.2.2: microRNA 21

In previous literature, miR-21 is one of the most frequently found miRNAs to be dysregulated in many different cancers. This is likely due to the proteins interaction with the AKT (152-154) and MAPK (155) pathways.

The AKT pathway is a signal transduction pathway promoting cell survival and growth in response to extracellular signals. Activation of this leads to cell growth and proliferation (152, 153). miR-21 has been shown to inhibit this activity, and it is hypothesized that decreased amounts of miR-21 can lead to over activation of the pathway (154).

Activation of the MAPK pathway alters mRNA translation and can activate several transcription factors, similarly to miRNAs. Although little research has investigated these interactions, there is hypothesized interactions between miR-21 and the MAPK pathway (155).

Overall, miR-21 has been associated with many cancers including breast, ovary, cervix, colon, lung, liver, brain, esophageal, prostate, pancreas, and thyroid (156-164).

2.1.2.3: microRNA 155

Research has shown that decreases in miR-155 can trigger oncogenic cascades. For example, a study showed that tumour protein-53-induced-nuclear-protein 1 (TP53INP1) is silenced by miR-155, leading to tumour growth (165).

Despite the minimal research investigating miR-155 as a biomarker in different cancers, studies have shown it to be dysregulated in thyroid, breast, colon, cervical, and lung cancer (166).

2.1.2.4: microRNA 210

The specific cellular interactions of miR-210 have been less researched compared to miR-21 and 155. Despite this, miR-210 has been linked to the hypoxia pathway (167) and has been found to be dysregulated in lung cancer (142, 145).

2.1.2.5: microRNA 223

miR-223 has been shown to be dysregulated in liver, blood, lymphatic, ovarian, and lung cancers (168-173). Its dysregulation has also been associated with higher tumour burden, disease aggressiveness, and poor prognosis (169, 170). The specific cellular interactions of miR-223 in the cell have not been well described.

2.1.3: Determination of Biological Fluid

As previously stated, the sputum miRNA study performed in our lab showed issues with sputum collection, as many of the samples received would contain mainly saliva, with little to no sputum tissue (126). Due to this, we looked into the use of a different biological fluids for miRNA measurement. Blood plasma was ultimately chosen due to a combination of factors. First, blood plasma has been investigated by many different studies, and has been shown to be a reliable way to measure miRNA in different types of cancer. Also, when compared to using whole blood, a very small amount of plasma is required for proper miRNA isolation. Lastly, the availability of blood plasma in early stage NSCLC was prevalent in tissue banks for the amount of samples required for our study.

2.2: Acquisition of Samples

2.2.1: Tissue Bank Protocols

2.2.1.1: Lung Cancer Biospecimen Resource Network

The Lung Cancer Biospecimen Resource Network (LCBRN) shipped two batches of blood plasma samples to the University of Alberta. The first contained 20 patient samples (pre-operative blood plasma, 5-8 month post-operative blood plasma, and tumour tissue) and the second contained 50 patient samples (pre-operative and post-operative blood plasma). These samples were collected from the Medical University of South Carolina, the University of Virginia, and the Washington University in St. Louis and stored at the University of Virginia.

Within 4 hours, blood plasma was collected from the patient and stored in -80˚C. All patient labels were created and attached to blood vacutainer tubes prior to blood collection. After blood collection, tubes were centrifuged at 1300 x g for 10 minutes at room temperature to separate blood phases for blood plasma collection. Blood plasma was pipetted into 0.5 mL aliquots and placed into a -80˚C freezer until being sent to the University of Alberta to avoid freeze-thaw cycles (174, 175).

The collected tumour tissue was frozen within 30-60 minutes after the specimen was removed from the patient and all sample containers were pre-labelled. Post tumour resection, clinical personnel obtained tissue samples for histologic quality control. Once histology had been confirmed, the remaining tissue was sectioned and blotted with tissue to remove excess blood and bodily fluids. Tissue was subsequently weighed and placed in containers. Then, these containers were snapped frozen via liquid nitrogen bath for at least 1 minute, then stored in -80˚C (174, 175).

All patient samples were stage I/II NSCLC with extensive smoking history and no prior history of cancer.

2.2.1.2: Conservant Bio

20 blood plasma control samples were sent to the University of Alberta by Conservant Bio, located in the Hudson-Alpha Institute for Bio-Technology in Huntsville, Alabama.

All blood plasma samples were collected and frozen in -80˚C within 2 hours of patient blood collection. Post blood sample collection, samples were centrifuged at 2000 x g for 15 minutes at 4˚C. Plasma was then pipetted and aliquoted into 0.5 mL tubes. Samples were kept frozen until being sent to the University of Alberta.

All control samples had no history of malignancy, as well as an extensive smoking history.

2.2.1.3: Alberta's Tomorrow Project

The Alberta's Tomorrow Project (ATP) is a Canadian database collecting demographic and biological information on participants, and sent 90 control blood plasma samples to the University of Alberta for this study.

Blood samples were collected into pre-labelled collection tubes and frozen at -80˚C within 2 hours of collection. Prior to the freezing of samples, tubes of blood were centrifuged at 1500 x g at room temperature for 10 minutes. Blood plasma was pipetted into 1 mL aliquots and put into a -80˚C freezer until being sent to the University of Alberta.

All participants had no past history of cancer, and had an extensive smoking history.

2.2.2: University of Alberta Specimen Storage

All samples received from the LCBRN, Conservant Bio., and ATP were sent via Fedex in packages filled with dry ice. Upon arrival at the University of Alberta, all packages were opened

in a sterile virus hood and handled with proper gown and glove protection to ensure safety. All shipments had dry ice remaining in the packages on arrival. Once packages were opened, all samples were inspected to ensure that samples did not thaw, spill or have loose lids. Samples were subsequently stored on the 5th floor in the Katz building at the University of Alberta in -80˚C freezers until specimen analysis occurred.

2.3: microRNA Isolation

2.3.1: Blood Plasma

Blood plasma miRNA isolation was performed on the 5th floor of the Katz building at the University of Alberta via recommended instructions supplied by Qiagen's miRNeasy Serum/Plasma Kit (Applied Biosystems, USA) inside a sterile virus hood. 6 samples were analyzed at a time to ensure short wait times between steps. Plasma was retrieved from the -80˚C freezer and thawed using a 37˚C water bath for 1-2 minutes.

150 µl of blood plasma was transferred into a 1.5 mL Eppendorf tube followed by the addition of 750 µl of the phenol/guanidine-based QIAzol Lysis Reagent. The tube was capped and vortexed resulting in a lysate solution allowing removal of genomic DNA to improve efficiency of RNA extraction. The lysate solution was then incubated at room temperature (22˚C) for 5 minutes.

Following incubation, 2.6 µl of synthetic *Caenorhabditis elegans* miR-39 (cel-miR-39) was added to the lysate solution as an endogenous spike-in control. Next, 150 µl of chloroform was added and vortexed into the lysate solution for 15 seconds producing a homogenate followed by 2-3 minutes of room temperature incubation.

The solution was then centrifuged for 15 minutes at 12,000 x g at 4˚C to separate the homogenate into three phases (Figure 2.2). The top phase is a colourless aqueous liquid that contains RNA molecules, the middle is a white precipitate phase containing DNA, and the bottom is a red phase containing proteins. The upper aqueous phase was collected and transferred to a new, clean Eppendorf tube and the other phases were discarded. The approximate volume of the upper aqueous phase was 450 µl. 675 µl of 100% ethanol was added and mixed through pipetting the solution up and down several times. The addition of ethanol provided the binding conditions for RNA molecules.

miRNA isolation progressed as 700 µl of the upper aqueous and ethanol solution was pipetted into an RNeasy MinElute spin column (Figure 2.2) and centrifuged at 8000 x g for 15 seconds at room temperature. The solution was pushed through the membrane on the spin column, collecting RNA, and the flow-through was collected in a removable column beneath the membrane. The liquid flow-through was subsequently discarded. The remaining 425 µl of upper aqueous and ethanol solution was inserted into the same spin column from the previous step, and again centrifuged to collect RNA on the spin column membrane, discarding the solution flow-through.

For RNA purification, 700 µl of the guanidinium thiocyanate and ethanol solution Qiagen Buffer RWT was added onto the spin column. The column was then centrifuged for 15 seconds at 8000 x g and flow-through was discarded. 500 µl of the buffer solution containing ethanol and Qiagen Buffer RPE was added to the spin column and centrifuged for 15 seconds at 8000 x g followed by discarding the flow-through. This was done to remove traces of salts from the spin column and membrane. To finish washing the spin column membrane, 500 µl of 80% ethanol was added and centrifuged for 2 minutes at 8000 x g and flow-through was discarded. The spin column was

then placed onto a new, dry and clean bottom flow-through collection column and centrifuged again at 8000 x g for 5 minutes. This was done to dry the spin column membrane and to avoid RNA contamination.

In order to collect the RNA contents, the spin column membrane was placed onto a new, labelled collection tube. 50 µl of RNase-free water was pipetted directly onto the spin column membrane and centrifuged at full speed for 1 minute. This was repeated, resulting in a total collection of 100 µl of miRNA containing elute. The elute was then stored in a -20˚C freezer until qRT-PCR was performed.

2.3.2: Tumour Tissue

Lung tumour tissue was removed from the -80˚C storage and 20 mg was measured for miRNA isolation, as instructed for such tissue in Qiagen's miRNeasy Mini Kit (Applied Biosystems, USA). The 20 mg of tissue was shredded and disrupted through the addition of 700 µl of Qiazol Lysis Reagent for 40 seconds until the sample was uniformly homogenous. The solution was then incubated at room temperature (22˚C) for 5 minutes to promote the dissociation of nucleoprotein complexes. This allowed RNA molecules to be available for isolation. After the addition of 140 µl of chloroform to the homogenous solution, it was vigorously shaken for 15 seconds, followed by room temperature incubation for 2-3 minutes. The solution was then transferred to a pre-labelled 1.5 mL Eppendorf tube.

The homogenous chloroform solution was centrifuged for 15 minutes at 12,000 x g at 4˚C, resulting in a top, colourless aqueous phase containing RNA molecules, a middle white phase containing DNA, and a bottom red organic phase containing proteins. 350 µl of the upper RNA containing phase was collected and transferred into a new collection tube with 525 µl of 100%

ethanol followed by mixing through pipetting the solution up and down several times. The middle and bottom phases were discarded.

700 µl of the aqueous phase and ethanol solution was then pipetted into a spin column and centrifuged at 8000 x g for 15 seconds at room temperature, passing through the spin column membrane into the bottom collection tube (Figure 2.2). Flow-through was then discarded. This was repeated with the other 175 µl and the same spin column. The RNA on the spin column membrane was then purified by the addition of 700 µl of Qiagen's Buffer RWT solution, then centrifuged for 15 seconds at 8000 x g followed by discarding of the flow-through. 500 µl of Qiagen's Buffer RPE was added to the spin column and centrifuged for 15 seconds at 8000 x g to collect salts and wash the spin column and membrane. This was repeated with another 500 µl of Buffer RPE except for 2 minutes of centrifuge at 8000 x g. After discarding flow-through, the spin column was centrifuged again for 1 minute at full speed to allow for drying of the membrane to avoid residual carryover of the previous buffers used to clean the membrane.

A new, pre-labelled collection tube was then placed underneath the spin column and 50 µl of RNase-free water was added directly onto the membrane, then centrifuged for 1 minute at 8000 x g. This step was repeated without changing the collection tube, allowing for 100 µl of elute solution to be collected for subsequent qRT-PCR analysis. Between obtaining the elute and qRT-PCR, the elute was stored at -20˚C.

2.4: Quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR) Analysis

qRT-PCR was performed first through RNA reverse transcriptase (RT). The Taqman microRNA Reverse Transcriptase Kit (Applied Biosystems, USA) and the Taqman microRNA Assay's

specific stem-loop primers (Applied Biosystems, USA) were used. Step-loop primers of miR-21, 155, 210, 223, and cel-miR-39 were used for miRNA measurement.

A 96-well RT-PCR plate was prepped for RT reaction. Columns 1 and 12 were left empty, and each row contained isolated miRNA from 2 patients, resulting in 16 samples being analyzed at once. Each well contained 2.33 µl of RT Master Mix (0.5 µl of 10x RT Buffer 0.05 µl dNTP Mix, 0.064 µl of RNase Inhibitor, 0.33 µl of Multiscribe RT, and 1.39 µl of RNase-free water), 1 µl of stem-loop specific miRNA primer (miR21, 155, 210, 223, or cel-miR-39), and 1.67 µl of isolated blood plasma miRNA elute. Each well contained a total volume of 5 µl (Figure 2.3).

All steps of adding RT Master Mix, miRNA primer, and elute were carried while plates and materials were on ice. After all RT-PCR plate wells were loaded, the plate was sealed and mixed through vortexing, then centrifuged down for 2 minutes at 12,000 x g in a suspended-arm centrifuge. Plates were inspected for air bubbles, and centrifuged again if they had occurred. RT reaction was done following the manufacturer protocol using the StepOnePlus RT-PCR Instrument (Applied Biosystems, USA) located on the basement floor at the Cross Cancer Institute (Edmonton, AB). The plate was incubated for 30 minutes at 16˚C, 30 minutes at 42˚C, 5 minutes at 85˚C, and then held at 4˚C until plate was removed from apparatus. RT products were subsequently stored at -20˚C.

When qRT-PCR could begin, the RT product containing plate was removed from the -20˚C freezer and diluted by adding 5 µl of RNase-free water. A new 96-well RT-PCR plate was set up by adding 10 µl of 1X Taqman Universal PCR Master Mix, 4 µl of RNase-free water, and 1 µl of stem-loop miRNA specific primer and probe (miR-21, 155, 210, 233, and cel-miR-39) into each well. 4 µl of the RT dilute from the first plate was then pipetted into the corresponding well of

the second plate, resulting in a total well volume of 20 µl (Figure 2.3). The RT-PCR plate was then sealed and vortexed, then centrifuged for 2 minutes at 12,000 x g using a suspended-arm centrifuge. Plates were inspected for air bubbles, then centrifuged again if they had occurred. qRT-PCR was performed using a 7900HT Fast Real-Time PCR System (Applied Biosystems, USA) located on the 3rd floor at the Cross Cancer Institute (Edmonton, AB). The reaction occurred through an incubation of 10 minutes at 95˚C followed by 40 cycles of 95˚C for 15 seconds and 60˚C for 1 minute. Cycle threshold (C_T) was defined as the number of PCR cycles required for a fluorescent signal to be higher than baseline variability, determined through Sequence Detection System (SDS) 2.3 software (Applied Biosystems, USA).

2.5: Statistical Analysis

2.5.1: microRNA Relative Expression Calculation

miRNA relative expression (RQ) was measured through normalizing of the C_T values measured through qRT-PCR to the cel-miR-39 spiked-in control. This was done by calculating:

$$RQ = 2^{-\Delta C_T}, \text{ where } \Delta C_T = C_T(\text{miRNA of interest}) - C_T(\text{cel-miR-39})$$

Due to extreme values, which will be discussed in further chapters, RQ values were log transformed prior to hierarchical clustering and binary logistic regression analysis.

2.5.2: Hierarchical Clustering

Hierarchical clustering was performed using the Statistical Package for the Social Sciences (SPSS) version 15 software (IBM SPSS Inc., USA). miR-21, 155, 210 and 223 RQ values were used to evaluate the cases and controls using average linkage and correlation similarity. Cluster

analysis was performed using the nearest-neighbour method and intervals were measured by squared Euclidean distance.

2.5.3: Risk Score Analysis via Binary Logistic Regression

Binary logistic regression was performed using the SPSS (version 15) software (IBM SPSS Inc., USA) to develop a combined risk score (CRS) in order to determine a patients risk of having lung cancer. Receiver operating curve (ROC) analysis was used to determine risk category cut-off values based off of sensitivity and specificity. The cut-off value was then used to dichotomize a patients' CRS in high and low risk categories of having lung cancer. Probabilities were calculated as an odds ratio (OR) to portray the probable risk of having lung cancer if a patient has a risk score above the cut-off value.

A combined risk score is calculated as:

$$CRS = (OR_{miR\text{-}21})(RQ_{miR\text{-}21}) \pm (OR_{miR\text{-}155})(RQ_{miR\text{-}155}) \pm (OR_{miR\text{-}210})(RQ_{miR\text{-}210}) \pm (OR_{miR\text{-}223})(RQ_{miR\text{-}223})$$

CRS for an individual patient is calculated through substituting their specific RQ for each miRNA measured through qRT-PCR. The product of the OR and RQ are either added to the CRS because they are upregulated in cases compared to controls or are subtracted from the CRS because they are downregulated in cases compared to controls.

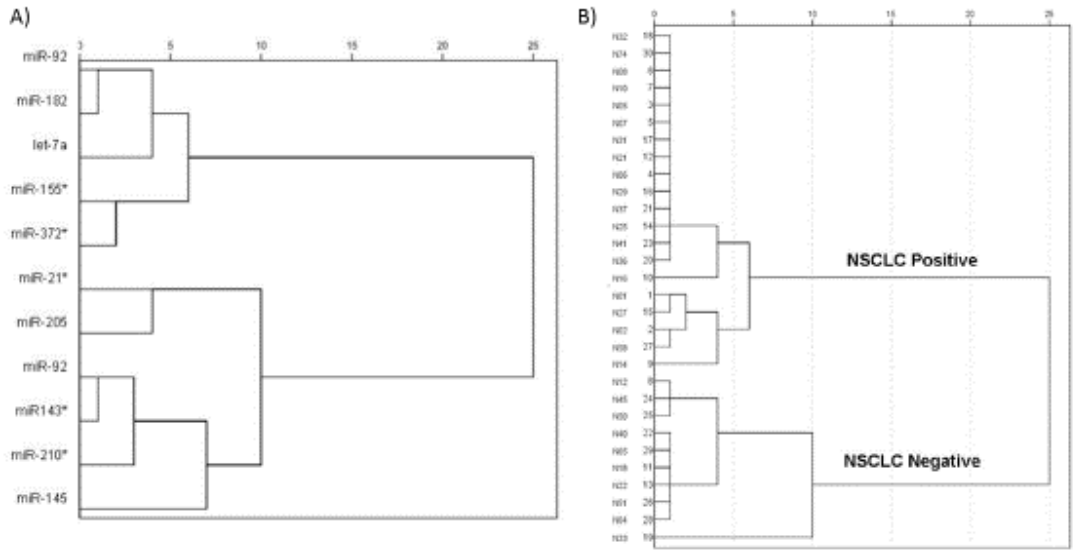


Figure 2.1: Non-Small Cell Lung Cancers (NSCLC) microRNA Hierarchical Clustering Analysis in Sputum

A) Dendogram from hierarchical cluster analysis of 4 NSCLC cases and 4 controls using the miRNAs 21, 92, 143, 145, 155, 182, 205, 210, 372, and let-7a using sputum.

B) Dendogram from hierarchical cluster analysis in sputum using miR-21, 143, 155, 210 and 372 were measuring again in another 24 NSCLC cases and 4 controls demonstrating two clusters of samples that are NSCLC positive and NSCLC negative.

* These images were originally published by Roa et al. (126).

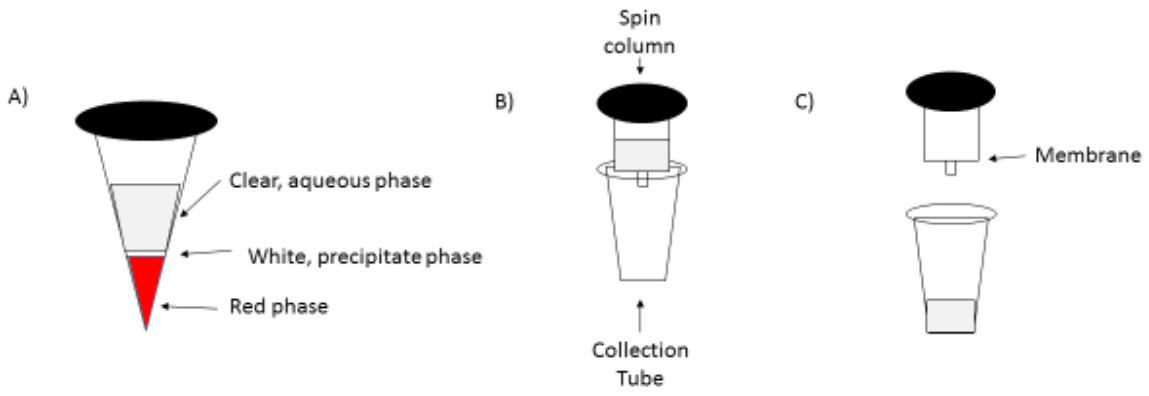


Figure 2.2: microRNA Isolation: Homogenate Phases and Spin Column Diagram

A) Diagram depicting the phases after homogenate is centrifuged for 15 minutes at 12,000 x g at 4˚C. Top, clear aqueous phase contain RNA molecules. Middle, white precipitate phase contains DNA. Bottom, red phase contains organic proteins. Top phase is collected for miRNA isolation and other phases are discarded.

B) RNeasy MinElute spin column. Liquids are loaded into the spin column, which is placed on top of the collection tube before centrifuge occurs.

C) RNeasy MinElute spin column. After centrifuge occurs, liquid has passed through the spin column membrane, emptying into the collection tube underneath. Liquid is then discarded from the collection tube, and placed back underneath the spin column.

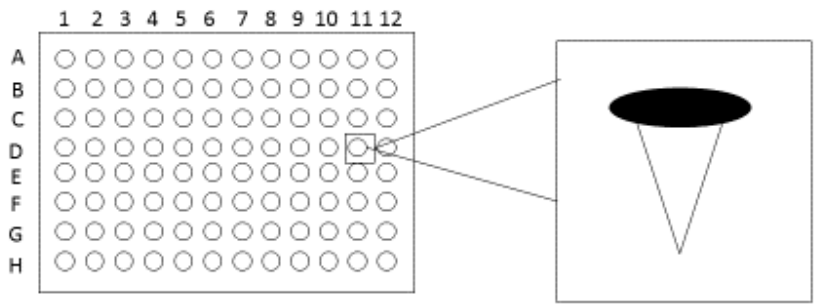


Figure 2.3: Quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR) Plate Setup

RT reaction: Columns 1 and 12 were not used. Each row contained isolated miRNA from 2 patients (example, row A in column 2-6 contained elute from one patient and elute from another patient in row A, columns 7-11). Each well had a total volume of 5 µl comprised of 2.33 µl of RT Master Mix (0.5 µl of 10x RT Buffer 0.05 µl dNTP Mix, 0.064 µl of RNase Inhibitor, 0.33 µl of Multiscribe RT, and 1.39 µl of RNase-free water), 1 µl of stem-loop specific miRNA primer (miR-21, 155, 210, 223, or cel-miR-39), and 1.67 µl of isolated blood plasma miRNA elute.

qRT-PCR reaction: Columns 1 and 12 were not used. RT reaction products were taken and inserted into the qRT-PCR plate corresponding to each well. Each well had a total volume of 20 µl containing 10 µl of 1X Taqman Universal PCR Master Mix, 4 µl of RNase-free water, 1 µl of stem-loop miRNA specific primer and probe (miR-21, 155, 210, 233, and cel-miR-39), and 4 µl of the RT products.

Chapter 3: Results

3.1: Demographics

3.1.1: Patient Demographic Overview

Patient demographics were organized into cases and controls (Table 3.1). In total, there were 179 patients in the study, with 69 cases and 110 controls (before exclusion). Age was compared by finding the median and range with 61 (range=36), and 63 (range=34) for cases and controls, respectively. Gender representation in the sample set shows that there is a higher amount of males than females in the cases group, with 43 males (62.3%) and 26 females (37.7%), than the controls group, as it has an even ratio of males and females.

Smoking status was similar across the two groups, with 22 (31.9%), and 42 (38.2%) patients being current smokers at time of diagnosis and sample retrieval in cases and controls, respectively. Smoking history was measured through pack-years, and the resulting calculations showed a higher smoking history average in the cases group with a mean 53.5 pack-years (standard deviation (SD)=31.9), in comparison to 30.8 pack-years (SD=10.8) in controls.

Focusing in on the cases group, there were three categories separating the patients according to tumour histology. 34 (49.3%) patients had AC, 31 (44.9%) had SCC, and 4 (5.8%) were put together into the "other" category. This last category has a case of AC with mixed features, adenosquamous carcinoma, BAC, and non-small cell carcinoma.

Each case was also categorized according to stage, through both the overall stages (I-IV) and the tumour, lymph node, and metastasis (TNM) classification of malignant tumours. In overall staging, the patients most common stages were IA and IIA with 25 (36.2%) and 19 (27.5%)

patients, respectively. Stages IB and IIB had 13 (18.8%) and 12 (17.4%) patients, respectively. Using TNM staging, the group was distributed into 8 groups: T1aN0M0, T1aN1M0, T1bN0M0, T1bN1M0, T2aN1M0, T2bN0M0, T2bN1M0, and T3N0M0. 17 (24.6%) patients were in the group T1aN0M0, 1 (1.4%) in T1aN1M0, 8 (11.6%) in T1bN0M0, 4 (5.8%) in T1bN1M0, 24 (34.9%) in T2aN1M0, 5 (7.2%) in T2bN0M0, 1 (1.4%) in T2bN1M0, and 9 (13.1%) patients in the T2N0M0 group.

A mean of 3.67 cm (SD=2.58) was found when looking at the greatest dimension of tumour in the cases group. Then, using the mean tumour dimension, patients were split into two groups of tumour size. The first group being less than 3.7 cm with 44 (64.7%) patients and the second being greater than or equal to 3.7 cm with 24 (35.3%) patients.

The average amount of time in months for the post-operative plasma collection was 6.43 months (SD=1.56). Subsequently, the cases were separated again into two groups with less than 6 months post-operative plasma collection being only 18 (26.1%) patients and greater than or equal to 6 months having a large proportion with 51 (73.9%) patients.

Lastly, patients were separated into two groups depending on if recurrence of cancer had occurred at the time of the post-operative plasma sample collection. Only 8 (11.6%) of patients had recurrence, with 2 having recurrence in the lung and the other 6 in different regions ranging from the brain, liver and colon.

3.1.2: Demographic Comparison of Pre-operative and Post-operative Cases versus Controls

In total, 64 cases (Table 3.2) and 110 controls (Table 3.3) were included in the statistical analysis when comparing the cases pre-operative plasma samples to controls. 5 cases were excluded in the pre-operative samples group due to the cel-39 spiked-in control not being detectable after

performing qRT-PCR, thus miRNA levels were not able to be normalized. These pre-operative cases were W0044, V0231, V0275, V0170, and V0166 (Table 3.2). An additional 4 cases were excluded in the post-operative samples group due to the same cel-39 issue. These post-operative cases were S0093, V0197, V0158, and S0172 (Table 3.2). In total, 110 controls, 64 pre-operative cases, and 60 post-operative cases were included in the analysis.

There was an even gender ratio in the control group with 55 females and 55 males, but due to the limitations of early-stage NSCLC patient supplies, the tissue banks that provided the cancer samples sent an uneven gender ratio. For the pre-operative cases group there was 25 females (39.1%) and 39 males (60.9%), and for the post-operative cases group there was 22 females (36.7%) and 38 males (63.3%) Included in the analysis (Table 3.4).

Regarding age, the controls had a mean age of 61.4 years (SD=7.95), the pre-operative cases had a mean age of 62 years (SD=7.76) and the post-operative cases had a mean age of 62.1 years (SD=7.9). For smoking history the 110 controls had a lower mean of 30.98 pack-years (SD=10.84) when compared to the 62 pre-operative cases with a mean of 51.5 pack-years (SD=29.46) and the 58 post-operative cases with a mean of 51.69 pack-years (SD=27.91), as two patients' smoking history was not known (Table 3.4).

3.1.3: Demographic Comparison of Post-operative Sample Recurrence versus No Recurrence

Out of the 60 post-operative samples that were used in the analysis, 52 (86.7%) did not have recurrence of cancer while 8 (13.3%) did have recurrence of cancer (loco-regional or distant metastasis). When comparing gender ratios according to recurrence, there was a much larger disparity in gender for the recurrence group, with 1 female (12.5%) and 7 males (87.5%)

compared to the non-recurrence group with 21 females (40.4%) and 31 males (59.6%) (Table 3.4).

Post-operative cases without recurrence had a mean age of 62.42 years (SD=7.76) while cases with recurrence had a mean age of 59.88 years (SD=8.42). Smoking history calculations produced an average pack-years of 48.61 (SD=26.74) for cases without recurrence and 74.14 (SD=25.86) for cases with recurrence (smoking history was available for 51 non-recurrence cases and 7 recurrence cases) (Table 3.4).

3.2: Binary Logistic Regression Risk Score Analysis

The spiked-in control cel-39 was quantified through qRT-PCR and measured as a C_T value (Table 3.5). The mean C_T measurements for the control group were 28.89 (SD=3.31), 29.21 (SD=5.68) for the pre-operative group, and 28.28 (SD=4.8) for the post-operative group.

After the RQ for all miRNAs for all samples were calculated, logistic regression was not successful in finding statistical significance using all different combinations of miRNAs. With extreme values affecting the data (see Tables 3.5 and 3.6), RQ was log transformed to assist in analysis without exclusion of any miRNA RQ measurements. Following this, logistic regression was successfully performed with statistically significant results using all 4 miRNAs in one profile.

3.2.1: Pre-operative Cases versus Controls

When comparing pre-operative cases with controls, the ROC analysis gave an area under the curve (AUC) of 72.3% (95% confidence interval (C.I.)=(0.641, 0.805)) and was subsequently used to create an RQ cut-off point of -0.4169 (Figure 3.1). Binary logistic regression was

performed and yielded a statistically significant OR of 3.000 (p-value=0.003, 95% C.I.=(1.440, 6.249)) (Table 3.7). Thus, samples with a combined miRNA RQ above the -0.4169 cut-off value are 3.000 times more likely to be a case than a control. Sensitivity and specificity were also calculated comparing pre-operative cases to controls resulting in 81% and 41%, respectively.

Binary logistic regression analysis was conducted comparing gender, smoking status, and age between the pre-operative cases and controls using the same cut-off point as above to look for confounding factors or bias in these parameters. Age was dichotomized into two groups, the first being less than 62 years of age and the other being greater than or equal to 62 years of age. This was done using the whole sample size (110 controls versus 69 cases) as well as the sample size included in the analysis (110 controls versus 64 cases).

Using the entire sample size, comparing gender between the two groups yielded an OR of 1.654 (p-value=0.108, 95% C.I.=(0.895, 3.055)). The same was done comparing ex-smokers to current smokers giving an OR of 1.288 (p-value=0.436, 95% C.I=(0.681, 2.437)) as well as age giving an OR of 1.430 (p-value=0.248, 95% C.I.=(0.780, 2.624)) (Table 3.8).

Comparatively, using only the cases and controls included in the final analysis, comparison of the pre-operative cases and controls gender resulted in an OR of 1.553 (p-value=0.184, 95% C.I.=(0.811, 2.972)). For smoking status and age, the analysis gave an OR of 1.424 (p-value=0.331, 95% C.I.=(0.699, 2.902)) and an OR of 1.386 (p-value=0.341, 95% C.I.=(0.707, 2.718)), respectively (Table 3.9).

3.2.2: Post-operative Cases versus Controls

An ROC analysis between post-operative cases and controls gave an AUC of 67.0% (95% C.I.=(0.577, 0.763)), yielding a cut-off point of -0.3255 (Figure 3.1). This cut-off point was

subsequently used in the binary logistic regression analysis giving a statistically significant OR of 2.275 (p-value=0.023, 95% C.I.=(1.120, 4.621)) (Table 3.7). Therefore, samples with a combined miRNA RQ above the -0.3255 cut-off value are 2.275 times more likely to be a case than a control. Further calculations were done yielding a sensitivity and specificity of 77% and 41%, respectively.

Binary logistic regression was again performed comparing gender, smoking status, and age, similarly as the previous pre-operative case versus control analysis. Tests were done comparing these possible confounding factors using the whole sample size (110 controls versus 69 cases) (Table 3.8) as well as the samples included in the analysis (110 controls versus 60 cases) (Table 3.9).

With the entire sample size, binary logistic regression yielded an OR of 1.654 (p-value=0.108, 95% C.I.=(0.895, 3.055)), 1.288 (p-value=0.436, 95% C.I.=(0.681, 2.437)), and 1.430 (p-value=0.248, 95% C.I.=(0.780, 2.624)) comparing gender, smoking status, and age, respectively (Table 3.8).

In comparison, using binary logistic regression on only the patients in the analysis gave an OR of 1.824 regarding gender (p-value=0.078, 95% C.I.=(0.936, 3.555), an OR of 1.925 regarding smoking status (p-value=0.084, 95% C.I.=(0.916, 4.047), and an OR of 1.532 regarding age (p-value=0.225, 95% C.I.=0.769, 3.050) (Table 3.9).

3.2.3: Pre-operative Cases versus Post-operative Cases

Separate analyses were done to compare pre-operative to post-operative cases. The first uses all 64 pre-operative cases and all 60 post-operative cases, regardless of cancer recurrence. The cut-off point of -0.7277 was determined through ROC analysis with an AUC value of 52.4% (95%

C.I.=(0.421, 0.627)) (Figure 3.1). This cut-off value was used in the binary logistic regression analysis yielding a statistically insignificant OR of 1.204 (p-value=0.718, 95% C.I.=(0.441, 3.287)) (Table 3.7).

The second set of analyses takes into account specific patients that did or did not have cancer recurrence. 8 patients had cancer recurrence and collectively the pre-operative and post-operative combined miRNA RQ values were compared using binary logistic regression. The 7 pre-operative and 8 post-operative samples gave a cut-off point of -1.2541 through ROC analysis and an AUC value of 53.6% (95% C.I.=(0.228, 0.843)) (Figure 3.2) and consequently a statistically insignificant OR of 2.500 (p-value=0.403, 95% C.I.=(0.292, 21.399)) (Table 3.7). When looking at the 57 pre-operative cases and 52 post-operative cases of patients that did not have cancer recurrence, a cut-off point of -0.6350 was determined via ROC analysis and an AUC value of 53.1% (95% C.I.=(0.421, 0.641)) (Figure 3.2). Thus, giving a statistically insignificant OR of 1.410 (p-value=0.624, 95% C.I.=(0.357, 5.559)) (Table 3.7).

3.3: Hierarchical Clustering Analysis

Both before and after the log transformation of all miRNA RQ values, hierarchical clustering analysis did not yield any significant results. All different combinations of miR-21, 155, 210, and 223 were tried, as well as each miRNA on its own.

3.4: Tumour Tissue Analysis

The 20 NSCLC cases that had accompanying tumour tissue samples were not compared to their partnered blood plasma samples, as previously planned. Due to the difference in miRNA isolation protocols, miRNA RQ's could not be compared between the two tissue types. The plasma miRNA isolation protocol uses cel-39 as its spiked-in control, while the tumour miRNA

isolation protocol does not. Thus, it was decided to not perform qRT-PCR on the tumour tissue and subsequent statistical analysis.

Characteristics	Cases, n=69	Controls, n=110
Age, years		
Median [range]	61 [36]	63 [34]
Gender, n [%]		
Female	26 [37.7]	55 [50]
Male	43 [62.3]	55 [50]
Smoking status, n [%]		
Yes	22 [31.9]	42 [38.2]
No	47 [68.1]	68 [61.8]
Smoking history[1], pack-years		
Mean [SD]	53.5 [31.9]	30.8 [10.8]
Tumour histology[2], n [%]		
Adenocarcinoma	34 [49.3]	
Squamous cell carcinoma	31 [44.9]	
Other	4 [5.8]	
Stage (I-IV), n [%]		
IA	25 [36.2]	
IB	13 [18.8]	
IIA	19 [27.5]	
IIB	12 [17.4]	
Stage (TNM)[3], n [%]		
T1aN0M0	17 [24.6]	
T1aN1M0	1 [1.4]	
T1bN0M0	8 [11.6]	
T1bN1M0	4 [5.8]	
T2aN1M0	24 [34.9]	
T2bN0M0	5 [7.2]	
T2bN1M0	1 [1.4]	
T3N0M0	9 [13.1]	
Tumour size[4] (cm), mean [SD]	3.67 [2.58]	
< 3.7, n [%]	44 [64.7]	
≥ 3.7, n [%]	24 [35.3]	
Post-operative collection (months), mean [SD]	6.43 [1.56]	
< 6, n [%]	18 [26.1]	
≥ 6, n [%]	51 [73.9]	
Cancer recurrence, n [%]		
No	61 [88.4]	
Yes	8 [11.6]	

Recurrence location[5] (n=8), n [%]	
Lung	2 [25]
Other	6 [75]

Table 3.1: Patient Demographics and Other Baseline Characteristics

Information included provided by the Lung Cancer Biospecimen Resource Network, Conservant Bio, and Alberta Tomorrow Project. Tumour size is represented as the greatest dimension of the mass.

[1] Two cases were smokers but did not have smoking history provided.

[2] "Other" category for tumour histology contains one case of adenocarcinoma mixed features, adenosquamous carcinoma, bronchioalveolar carcinoma, and non-small cell carcinoma.

[3] One case was categorized as T1N1M0 and two other cases were categorized as T2N0M0, with no specification on T subtypes a or b, thus they were added to T1bN1M0 and T2bN0M0, respectively.

[4] One case did not have a given greatest dimension of tumour mass value.

[5] Other cancer recurrence locations present with metastasis to the brain, liver, or colon.

(SD, standard deviation; TNM, tumour, lymph node, and metastasis classification of malignant tumours).

Sample ID	Gender	Age (years)	Smoking Status	Smoking History (pys)	Tumour Histology	Overall Stage
S0014	Male	66	Current	20	Adenocarcinoma	IIB
S0025	Male	70	Ex	50	Squamous Cell Carcinoma	IIA
S0027	Male	74	Ex	105	Adenocarcinoma	IA
S0030	Female	58	Ex	66	Adenocarcinoma	IA
S0031	Female	74	Ex	32	Adenocarcinoma	IIA
S0036	Male	74	Ex	60	Squamous Cell Carcinoma	IB
S0043	Male	54	Current	40	Adenocarcinoma	IIA
S0084	Male	49	Current	32	Adenocarcinoma	IA
S0090	Female	60	Ex	20.5	Squamous Cell Carcinoma	IB
S0092	Female	62	Ex	30	Squamous Cell Carcinoma	IIB
S0093	Male	55	Current	8	Adenocarcinoma	IIB
S0097	Female	64	Ex	10.5	Adenocarcinoma	IA
S0112	Female	74	Ex	40	Adenocarcinoma	IIA
S0135	Male	61	Current	22	Squamous Cell Carcinoma	IIA
S0168	Male	65	Current	47	Squamous Cell Carcinoma	IB
S0172	Female	68	Ex	7	Adenocarcinoma	IA
S0177	Male	75	Ex	90	Squamous Cell Carcinoma	IA
S0186	Female	67	Ex	40	Adenocarcinoma - Mixed features	IB
S0195	Male	69	Ex	40	Squamous Cell Carcinoma	IIB
V0015	Male	62	Ex	40	Adenocarcinoma	IA
V0021	Female	61	Current	102	Squamous Cell Carcinoma	IA
V0031	Female	56	Ex	40	Squamous Cell Carcinoma	IIA
V0041	Male	61	Current	46	Adenocarcinoma	IIB
V0056	Male	59	Ex	100	Squamous Cell Carcinoma	IIA
V0070	Male	54	Current	Unknown	Squamous Cell Carcinoma	IB
V0077	Male	75	Ex	100	Squamous Cell Carcinoma	IIB
V0096	Male	60	Current	88	Squamous Cell	IIA

					Carcinoma	
V0097	Male	61	Ex	60	Squamous Cell Carcinoma	IB
V0101	Female	72	Ex	50	Squamous Cell Carcinoma	IA
V0116	Male	69	Ex	70	Adenocarcinoma	IIA
V0128	Female	72	Ex	40	Adenocarcinoma	IA
V0129	Male	65	Ex	40	Adenocarcinoma	IA
V0141	Male	67	Ex	40	Adenocarcinoma	IIA
V0151	Male	62	Current	90	Squamous Cell Carcinoma	IIA
V0155	Male	62	Ex	42	Adenocarcinoma	IB
V0158	Female	60	Current	120	Adenocarcinoma	IA
V0166	Male	55	Current	60	Squamous Cell Carcinoma	IIA
V0170	Male	68	Ex	168	Squamous Cell Carcinoma	IIB
V0173	Female	51	Current	6	Adenocarcinoma	IA
V0180	Female	66	Current	55	Adenocarcinoma	IA
V0195	Male	54	Ex	40	Adenocarcinoma	IB
V0197	Female	58	Ex	60	Squamous Cell Carcinoma	IA
V0206	Male	52	Ex	72	Adenocarcinoma	IA
V0211	Male	68	Ex	30	Squamous Cell Carcinoma	IA
V0213	Male	56	Current	40	Squamous Cell Carcinoma	IB
V0229	Male	69	Ex	50	Squamous Cell Carcinoma	IB
V0231	Female	55	Current	28	Adenocarcinoma	IIA
V0233	Female	60	Current	10	Squamous Cell Carcinoma	IA
V0241	Male	75	Ex	80	Adenocarcinoma	IIB
V0245	Male	58	Ex	57	Adenocarcinoma	IIB
V0253	Male	58	Ex	66	Squamous Cell Carcinoma	IA
V0263	Female	57	Ex	60	Adenocarcinoma	IA
V0269	Female	74	Ex	50	Adenocarcinoma	IIA
V0275	Male	67	Ex	45	Squamous Cell Carcinoma	IIA
W0002	Female	59	Ex	35	Bronchioalveolar Carcinoma	IA
W0030	Male	54	Ex	35	Non-small Cell	IIA

					Carcinoma	
W0039	Male	58	Ex	30	Squamous Cell Carcinoma	IB
W0042	Female	65	Ex	Unknown	Adenocarcinoma	IIB
W0044	Female	56	Current	80	Adenocarcinoma	IA
W0047	Female	57	Ex	120	Squamous Cell Carcinoma	IIA
W0052	Male	63	Ex	100	Adenocarcinoma	IB
W0093	Female	68	Ex	20	Adenocarcinoma	IIA
W0109	Male	39	Ex	20	Adenosquamous Carcinoma	IIB
W0137	Male	47	Current	111	Adenocarcinoma	IIA
W0164	Female	52	Current	30	Adenocarcinoma	IB
W0189	Male	61	Ex	67.5	Squamous Cell Carcinoma	IA
W0200	Male	64	Current	40	Squamous Cell Carcinoma	IIB
W0230	Male	49	Ex	14.5	Squamous Cell Carcinoma	IA
W0270	Male	58	Ex	78	Adenocarcinoma	IA
Sample ID	**TNM Stage**	**Tumour Size (cm)**	**Post-op Sample Collection (Months)**	**Recurrence**	**Recurrence Site**	**Included in Analysis[1]**
S0014	T3N0M0	7.8	7	No		Yes
S0025	T1N1M0	Unknown	7	No		Yes
S0027	T1aN0M0	1.5	7	No		Yes
S0030	T1aN0M0	2	5	Yes	Brain	Yes
S0031	T2aN1M0	1.4	5	No		Yes
S0036	T2aN0M0	1.2	6	No		Yes
S0043	T2aN1M0	6	5	No		Yes
S0084	T1aN0M0	7.9	4	No		Yes
S0090	T2aN0M0	3	7	No		Yes
S0092	T3N0M0	1.7	8	No		Yes
S0093	T3N0M0	6.4	7	No		Pre-op only
S0097	T1aN0M0	1.1	7	No		Yes
S0112	T2aN1M0	3.7	6	No		Yes
S0135	T2aN1M0	3.4	7	No		Yes
S0168	T2aN0M0	10.6	7	No		Yes
S0172	T1aN0M0	2.6	7	No		Pre-op only
S0177	T1bN0M0	2.1	6	No		Yes

S0186	T2aN0M0	2.3	7	No		Yes
S0195	T3N0M0	6.9	6	No		Yes
V0015	T1bN0M0	2.6	7	No		Yes
V0021	T1aN0M0	1.5	9	No		Yes
V0031	T2aN1M0	3.8	7	No		Yes
V0041	T2bN1M0	3.5	8	No		Yes
V0056	T1bN1M0	2.5	7	No		Yes
V0070	T2aN0M0	3.2	10	Yes	Lung, mediastinum	Yes
V0077	T3N0M0	6.1	8	Yes	Lung, chest wall	Yes
V0096	T2aN1M0	2.4	7	No		Yes
V0097	T2aN0M0	4	9	No		Yes
V0101	T1aN0M0	1.2	7	No		Yes
V0116	T1aN1M0	2.6	8	No		Yes
V0128	T1aN0M0	2.6	7	No		Yes
V0129	T1bN0M0	2.6	7	No		Yes
V0141	T2bN0M0	7	7	Yes	Colon	Yes
V0151	T1bN1M0	8.2	7	Yes	Brain	Yes
V0155	T2aN0M0	1.2	9	No		Yes
V0158	T1aN0M0	1.8	5	No		Pre-op only
V0166	T2aN1M0	3.1	6	No		No
V0170	T2N0M0	8.8	1	No		No
V0173	T1aN0M0	0.8	7	No		Yes
V0180	T1aN0M0	0.9	7	No		Yes
V0195	T2aN0M0	3.2	6	No		Yes
V0197	T1aN0M0	2	8	No		Pre-op only
V0206	T1bN0M0	3.2	7	Yes	Liver	Yes
V0211	T1aN0M0	2	8	No		Yes
V0213	T2aN0M0	3	6	No		Yes
V0229	T2aN0M0	9.6	6	No		Yes
V0231	T2aN1M0	4.8	8	No		No
V0233	T1bN0M0	2	7	No		Yes
V0241	T3N0M0	0.8	6	No		Yes
V0245	T3N0M0	1.3	4	No		Yes
V0253	T1aN0M0	1.5	7	No		Yes
V0263	T1bN0M0	2	7	No		Yes
V0269	T2aN1M0	3.2	8	No		Yes
V0275	T1bN1M0	3	7	No		No
W0002	T1bN0M0	2.4	5	No		Yes

W0030	T2bN0M0	5.2	4	No		Yes
W0039	T2aN0M0	6	5	No		Yes
W0042	T2N0M0	4.2	6	No		Yes
W0044	T1bN0M0	2.8	4	No		No
W0047	T2aN1M0	4.6	4	No		Yes
W0052	T2aN0M0	4.4	5	No		Yes
W0093	T2aN1M0	1.9	10	No		Yes
W0109	T3N0M0	7	6	No		Yes
W0137	T2bN0M0	3.9	4	Yes	Brain	Yes
W0164	T2aN0M0	1.7	5	No		Yes
W0189	T1aN0M0	0.7	5	No		Yes
W0200	T3N0M0	11.8	6	Yes	Liver	Yes
W0230	T1aN0M0	8.4	4	No		Yes
W0270	T1aN0M0	1.2	5	No		Yes

Table 3.2: Patient Specific Demographics for Cases

Information included provided by the Lung Cancer Biospecimen Resource Network, Conservant Bio, and Alberta Tomorrow Project. Tumour size is represented as the greatest dimension of the mass. Each patient gave a blood plasma sample pre-operatively and post-operatively. Note that underlined sample ID's are patients not included or partially included in the analysis.

[1] Specifications on which patients samples were included in study. Yes means both pre-operative and post-operative plasma samples were kept in the final analysis. Pre-op only does not have the patients corresponding post-operative sample included. No has neither pre nor post-operative included.

(pys, pack-years; TNM, tumour node metastasis).

Sample ID	Gender	Age (years)	Smoking Status	Smoking History (pys)	Included in Analysis
C1	Female	64	Ex	24	Yes
C10	Female	60	Current	24	Yes
C100	Male	57	Current	24.3	Yes
C101	Female	58	Ex	28	Yes
C102	Male	70	Ex	40	Yes
C103	Female	59	Ex	31	Yes
C104	Male	72	Current	42.6	Yes
C105	Male	56	Current	30.4	Yes
C106	Male	59	Ex	24	Yes
C107	Male	58	Ex	22.8	Yes
C108	Male	49	Current	31	Yes
C109	Female	61	Ex	25	Yes
C11	Male	65	Ex	44	Yes
C110	Male	72	Ex	57	Yes
C12	Female	74	Current	15	Yes
C13	Female	51	Ex	17	Yes
C14	Male	51	Ex	25	Yes
C15	Male	65	Ex	22.5	Yes
C16	Male	68	Ex	20	Yes
C17	Female	52	Ex	30	Yes
C18	Female	52	Ex	54	Yes
C19	Male	57	Ex	20	Yes
C2	Female	65	Ex	25	Yes
C20	Male	63	Ex	30	Yes
C21	Male	73	Ex	30.4	Yes
C22	Female	68	Ex	27.6	Yes
C23	Male	62	Current	27.6	Yes
C24	Female	71	Ex	25.1	Yes
C25	Male	70	Ex	23	Yes
C26	Female	69	Ex	37.2	Yes
C27	Female	62	Current	31.1	Yes
C28	Female	68	Ex	25.8	Yes
C29	Male	70	Ex	28	Yes
C3	Female	52	Ex	54	Yes
C30	Male	70	Ex	32	Yes
C31	Male	61	Ex	26.6	Yes
C32	Male	68	Ex	26.6	Yes
C33	Male	52	Current	30	Yes
C34	Female	58	Current	39	Yes

C35	Male	68	Ex	28.2	Yes
C36	Male	60	Current	24	Yes
C37	Female	71	Ex	23	Yes
C38	Male	59	Current	30.4	Yes
C39	Male	68	Ex	35.2	Yes
C4	Male	53	Current	30	Yes
C40	Female	48	Current	16.8	Yes
C41	Female	65	Ex	24	Yes
C42	Female	66	Current	45	Yes
C43	Male	64	Current	16	Yes
C44	Male	71	Current	33	Yes
C45	Male	66	Ex	31.9	Yes
C46	Female	74	Current	41	Yes
C47	Male	61	Ex	45	Yes
C48	Female	51	Current	18	Yes
C49	Male	50	Current	22.8	Yes
C5	Female	69	Current	20	Yes
C50	Female	73	Ex	24	Yes
C51	Male	53	Current	21	Yes
C52	Female	47	Current	35	Yes
C53	Male	65	Current	30	Yes
C54	Female	63	Ex	51	Yes
C55	Female	67	Ex	22.8	Yes
C56	Male	53	Current	36	Yes
C57	Female	64	Ex	24	Yes
C58	Female	61	Current	45	Yes
C59	Female	48	Ex	28	Yes
C6	Female	66	Ex	50	Yes
C60	Male	74	Ex	16	Yes
C61	Female	62	Current	16.8	Yes
C62	Male	53	Ex	35	Yes
C63	Female	67	Ex	49	Yes
C64	Male	70	Ex	33	Yes
C65	Male	61	Ex	40	Yes
C66	Male	65	Ex	35	Yes
C67	Female	51	Ex	21.3	Yes
C68	Male	56	Ex	26	Yes
C69	Female	64	Ex	28	Yes
C7	Male	67	Current	20	Yes
C70	Female	65	Ex	30	Yes
C71	Male	63	Ex	64	Yes

C72	Female	70	Current	19.6	Yes
C73	Male	51	Current	49	Yes
C74	Female	63	Current	43	Yes
C75	Female	60	Current	31.9	Yes
C76	Male	46	Current	25	Yes
C77	Female	62	Ex	29	Yes
C78	Female	42	Ex	23	Yes
C79	Female	52	Current	28	Yes
C8	Male	76	Ex	15	Yes
C80	Female	49	Ex	25	Yes
C81	Male	51	Ex	29	Yes
C82	Female	46	Ex	30	Yes
C83	Female	63	Ex	35	Yes
C84	Female	67	Current	35.7	Yes
C85	Female	60	Current	17.6	Yes
C86	Female	57	Current	42	Yes
C87	Female	48	Current	22.8	Yes
C88	Male	65	Ex	38	Yes
C89	Male	74	Ex	29.4	Yes
C9	Male	67	Current	49	Yes
C90	Female	52	Current	27	Yes
C91	Male	57	Current	21.4	Yes
C92	Male	62	Ex	44	Yes
C93	Female	64	Ex	25.1	Yes
C94	Female	63	Ex	30	Yes
C95	Male	71	Ex	74.2	Yes
C96	Female	56	Ex	25.8	Yes
C97	Female	63	Ex	50.4	Yes
C98	Male	48	Current	32	Yes
C99	Male	73	Ex	25	Yes

Table 3.3: Patient Specific Demographics for Controls
Information included provided by Conservant Bio, and Alberta Tomorrow Project.
(pys, pack-years)

	# Samples Included	Female, n [%]	Male, n [%]	Age, mean [SD]	Smoking history in pys, mean [SD][1]
Pre-op	64	25 [39.1]	39 [60.9]	62 [7.76]	51.5 [29.46]
Post-op	60	22 [36.7]	38 [63.3]	62.08 [7.9]	51.69 [27.91]
Controls	110	55 [50]	55 [50]	61.4 [7.95]	30.98 [10.84]
Post-op without recurrence	52	21 [40.4]	31 [59.6]	62.42 [7.76]	48.61[26.74]
Post-op with recurrence	8	1 [12.5]	7 [87.5]	8.42 [59.88]	74.14 [25.86]

Table 3.4: Demographic Comparison of Pre-operative, Post-operative, Controls, and Recurrence State Included in Analysis

Information included provided by the Lung Cancer Biospecimen Resource Network, Conservant Bio, and Alberta Tomorrow Project.

[1] Two cases are not known for the pre-operative and post-operative groups, and one for the control, recurrence, and no recurrence groups for smoking history.

(SD, standard deviation; pys, pack-years; pre-op, pre-operative; post-op, post-operative).

Sample ID	Pre-op/ Post-op	Control C_T (cel-39)	RQ miR-155	RQ miR-21	RQ miR-210	RQ miR-223
S0014	Post	3.71E+01	9.18E+00	4.20E+01	1.08E+01	3.24E+02
S0014	Pre	3.93E+01		3.98E+00	1.09E+01	2.81E+02
S0025	Post	2.55E+01	8.96E-04	2.10E-02	1.32E-03	9.29E-02
S0025	Pre	3.58E+01	5.81E-01	1.42E+01	5.68E-01	1.01E+02
S0027	Post	3.61E+01		9.78E-01	1.52E+00	4.19E+01
S0027	Pre	3.60E+01	5.39E+00	7.22E+01	3.14E+00	1.09E+03
S0030	Pre	3.41E+01	8.41E-01	1.20E+01	3.30E-01	1.09E+02
S0030	Post	2.61E+01	1.47E-03	3.47E-02	1.13E-03	1.11E-01
S0031	Pre	3.60E+01	2.16E+00	3.70E+01	1.62E+00	2.81E+02
S0031	Post	2.96E+01		1.27E-02		3.13E-01
S0036	Post	2.82E+01	1.14E-02	2.10E-02	2.27E-03	1.87E+00
S0036	Pre	3.51E+01	1.12E+00	4.70E+00	5.11E-01	3.80E+01
S0043	Post	2.59E+01	1.83E-03	3.84E-02	5.25E-04	1.24E-01
S0043	Pre	3.49E+01	6.54E-01	8.43E+00	2.30E-01	6.05E+01
S0084	Post	2.52E+01	5.98E-03	2.38E-02	3.36E-03	4.76E-01
S0084	Pre	3.54E+01	1.39E+00	5.55E+00	8.78E-01	3.30E+01
S0090	Pre	3.95E+01	1.25E+01	1.13E+01	4.44E+00	9.74E+03
S0090	Post	3.67E+01	2.38E+00	1.52E+00		9.38E+00
S0092	Post	3.49E+01	2.34E-01	2.42E-01		3.96E+01
S0092	Pre	3.84E+01	1.47E+00	1.68E+00		1.24E+03
S0093	Pre	3.63E+01	9.01E+00	9.78E+01	5.10E+00	1.09E+03
S0093	Post					
S0097	Post	2.85E+01	8.61E-03	2.70E-03		4.06E+00
S0097	Pre	2.69E+01	8.47E-03	9.90E-03	1.76E-04	2.36E+00
S0112	Pre	3.71E+01		8.31E+00	7.11E-01	2.30E+02
S0112	Post	3.44E+01	8.20E-01	1.82E+00	1.29E+00	4.53E+01
S0135	Post	2.39E+01	1.16E-04	5.36E-04	1.03E-04	3.00E-02
S0135	Pre	2.36E+01	1.83E-04	4.50E-04	3.07E-04	2.65E-02
S0168	Pre	2.36E+01	4.31E-04	5.85E-03	6.13E-04	6.59E-02
S0168	Post	2.56E+01	3.60E-04	9.80E-03		1.52E-01
S0172	Pre	2.72E+01	2.19E-02	2.85E-02	2.80E-03	1.07E-03
S0172	Post					
S0177	Pre	2.66E+01	9.44E-03	1.93E-02	2.36E-03	7.71E+00
S0177	Post	2.58E+01	3.63E-03	2.82E-02	1.03E-03	3.51E+00
S0186	Pre	2.69E+01	2.68E-03	2.21E-02	4.55E-03	7.36E+00
S0186	Post	2.66E+01	6.42E-03	7.04E-03	3.37E-03	3.56E+00
S0195	Pre	2.40E+01	6.86E-04	1.34E-02	1.14E-04	1.14E-01
S0195	Post	2.47E+01	2.48E-03	4.51E-02		3.86E-01
V0015	Pre	3.86E+01	4.70E+00	1.04E+01		3.68E+03
V0015	Post	3.69E+01	3.58E+00	2.10E+01	3.23E+00	4.99E+02
V0021	Post	2.36E+01	1.86E-04	3.59E-03	1.51E-04	1.47E-02
V0021	Pre	2.09E+01		2.91E-04	5.94E-06	1.24E-03
V0031	Post	3.85E+01	6.93E+00	3.90E+01	1.85E+00	7.69E+02

V0031	Pre	3.79E+01	5.09E+00	2.54E+00	1.01E+00	1.47E+03
V0041	Pre	2.56E+01	3.04E-03	7.26E-03	1.91E-03	1.91E-01
V0041	Post	2.61E+01	2.94E-03	1.59E-02	1.04E-03	1.88E-01
V0056	Post	3.70E+01	2.95E+00	2.47E+01	1.72E+00	2.50E+02
V0056	Pre	3.78E+01	2.38E+00	1.55E+00	6.66E-01	2.47E+03
V0070	Post	2.03E+01	2.38E-05	7.19E-04	2.15E-05	3.87E-03
V0070	Pre	2.12E+01	3.64E-05	6.84E-04	3.19E-05	5.46E-03
V0077	Post	2.58E+01	1.49E-03	3.81E-03	1.02E-03	5.65E-02
V0077	Pre	2.60E+01	1.00E-03	2.16E-02	6.80E-04	1.48E-01
V0096	Pre	2.47E+01	5.90E-04	3.16E-03	3.35E-04	3.80E-02
V0096	Post	2.64E+01	1.15E-03	2.89E-02	1.22E-03	3.24E-01
V0097	Post	3.05E+01	7.84E-02	1.11E-02		2.50E+00
V0097	Pre	2.27E+01	4.12E-04	2.57E-03		2.00E-02
V0101	Post	2.18E+01	8.48E-05	9.20E-04		4.67E-03
V0101	Pre	2.42E+01	8.36E-04	9.20E-03		1.28E-02
V0116	Pre	2.57E+01	4.10E-04	1.40E-02	4.05E-04	4.04E-02
V0116	Post	2.61E+01	8.13E-04	1.72E-02	1.27E-03	1.44E-01
V0128	Pre	2.20E+01	3.49E-04	1.29E-03	3.30E-04	9.14E-02
V0128	Post	2.32E+01	4.70E-04	1.68E-03	6.99E-04	3.61E-02
V0129	Post	2.36E+01	3.04E-04	8.00E-04		9.45E-03
V0129	Pre	2.23E+01		1.89E-04		1.32E-02
V0141	Post	3.70E+01	6.27E+00	5.35E+01	2.23E+01	1.11E+03
V0141	Pre	3.06E+01	1.75E-02	3.81E-01	3.62E-02	3.05E+00
V0151	Post	2.46E+01	2.22E-04	1.18E-02		1.35E-02
V0151	Pre	2.25E+01	1.34E-04	2.65E-03	5.09E-05	5.18E-03
V0155	Pre	2.52E+01	5.81E-04	7.23E-04		4.69E-02
V0155	Post	2.56E+01	5.89E-04	5.17E-04		4.87E-02
V0158	Pre	3.63E+01	5.86E-01		6.94E+00	
V0158	Post					
V0166	Pre					
V0166	Post					
V0170	Post					
V0170	Pre					
V0173	Pre	2.52E+01	4.18E-03	5.92E-02	1.33E-03	6.41E-01
V0173	Post	2.60E+01	1.81E-03	3.94E-02	1.04E-03	4.26E-01
V0180	Post	3.67E+01	2.80E+00	1.60E+01	1.73E+00	2.35E+02
V0180	Pre	3.71E+01	1.08E+00	2.35E+01	3.53E+00	2.23E+02
V0195	Post	2.68E+01	4.44E-03	4.59E-02	1.50E-03	3.83E-01
V0195	Pre	2.72E+01	1.09E-02	6.30E-02	1.25E-03	4.50E-01
V0197	Pre	3.71E+01	1.23E+00	3.12E+00	1.69E+00	2.52E+02
V0197	Post					
V0206	Pre	2.67E+01	1.37E-03	3.65E-02	2.60E-03	3.61E-01
V0206	Post	2.50E+01	4.53E-04	9.93E-03	2.96E-04	6.83E-02
V0211	Pre	2.68E+01	3.62E-03	9.69E-03	2.21E-03	2.31E+00
V0211	Post	2.68E+01	2.03E-02	2.28E-02		5.84E+00

V0213	Post	2.73E+01	2.25E-03	2.96E-03		8.94E-01
V0213	Pre	2.84E+01		5.87E-03		3.02E-01
V0229	Post	2.92E+01	2.20E-02	5.23E-03		4.01E-01
V0229	Pre	3.05E+01	3.79E-02	1.54E-02		2.18E+00
V0231	Pre					
V0231	Post					
V0233	Post	2.60E+01	3.92E-03	3.54E-03		4.22E-01
V0233	Pre	2.75E+01	1.77E-03	4.74E-03		1.24E+00
V0241	Pre	2.69E+01	5.52E-03	9.97E-03	3.08E-03	4.51E+00
V0241	Post	3.05E+01		9.54E-03		1.53E-01
V0245	Pre	3.14E+01	3.10E-02			5.43E+00
V0245	Post	3.19E+01	7.43E-01	4.03E-03	4.18E-02	2.12E+02
V0253	Post	2.60E+01	7.25E-04	4.22E-03	9.23E-04	1.20E-01
V0253	Pre	2.37E+01	2.65E-04	3.31E-04		2.14E-02
V0263	Post	3.42E+01		1.74E-01		2.31E+00
V0263	Pre	3.71E+01		2.11E+00		1.53E+02
V0269	Post	2.35E+01	3.81E-04	7.07E-04	4.36E-04	4.34E-02
V0269	Pre	2.60E+01	2.09E-03	3.02E-03	2.01E-04	9.33E-02
V0275	Post					
V0275	Pre					
W0002	Pre	3.37E+01	3.46E-01	4.06E+00	9.00E-02	6.75E+01
W0002	Post	3.71E+01	1.76E+00	1.22E+01	2.62E+00	1.48E+02
W0030	Pre	3.67E+01		7.89E+00	3.52E+00	5.05E+00
W0030	Post	3.67E+01	6.88E+00	5.71E+01	1.41E+00	2.78E+02
W0039	Post	2.66E+01	2.28E-03	4.96E-02	1.80E-03	1.36E-01
W0039	Pre	2.79E+01	1.10E-02	1.20E-01	4.62E-03	5.89E-01
W0042	Post	3.08E+01	5.61E-02	1.85E-01	2.85E+03	3.85E+00
W0042	Pre	3.07E+01	1.35E-01	1.89E-02		7.10E-01
W0044	Pre					
W0044	Post					
W0047	Pre	2.28E+01	6.35E-05	1.18E-02	5.50E-04	5.77E-02
W0047	Post	2.76E+01	7.93E-03	5.98E-02	3.47E-03	7.32E+00
W0052	Pre	2.25E+01	1.72E-04	3.02E-03	2.34E-04	1.72E-02
W0052	Post	2.30E+01	8.92E-04	1.11E-02	5.77E-04	7.64E-02
W0093	Pre	2.37E+01	2.73E-04	3.65E-03	6.56E-04	1.11E-01
W0093	Post	2.70E+01	1.00E-02	1.68E-02	2.73E-03	4.73E+00
W0109	Post	2.46E+01	3.43E-03	6.43E-03	1.43E-03	2.14E+00
W0109	Pre	2.73E+01	7.01E-03	1.77E-02	1.21E-03	1.50E+00
W0137	Post	2.80E+01	1.13E-02	4.34E-02	8.45E-03	1.75E+01
W0137	Pre	2.78E+01	1.67E-03	3.16E-02		2.47E+00
W0164	Pre	2.27E+01	2.67E-04	3.13E-03	2.24E-04	6.34E-02
W0164	Post	2.12E+01	1.13E-03	1.53E-02	3.24E-04	1.74E-01
W0189	Post	2.54E+01	1.38E-03	1.59E-02	1.24E-03	2.57E-01
W0189	Pre	2.50E+01	1.15E-03	1.21E-01	3.31E-04	1.42E-01
W0200	Post	2.78E+01	2.92E-03	8.57E-03	2.66E-04	3.90E-01

W0200	Pre	2.38E+01	5.27E-04	8.74E-04	1.05E-04	1.20E-02
W0230	Pre	2.73E+01	1.03E-03	2.85E-02	3.08E-03	3.13E-01
W0230	Post	2.47E+01	1.94E-03	1.02E-02	2.47E-04	1.57E-01
W0270	Pre	2.56E+01	3.41E-03	2.79E-03		1.57E-01
W0270	Post	2.50E+01	1.64E-03	7.73E-03	5.33E-04	2.55E-01
Sample ID	Pre-op/ Post-op	Log RQ miR-155	Log RQ miR-21	Log RQ miR-210	Log RQ miR-223	Combined RQ
S0014	Post	3.20E+00	5.39E+00	3.44E+00	8.34E+00	0.65
S0014	Pre		1.99E+00	3.45E+00	8.13E+00	0.59
S0025	Post	-1.01E+01	-5.57E+00	-9.56E+00	-3.43E+00	-0.41
S0025	Pre	-7.83E-01	3.83E+00	-8.15E-01	6.66E+00	0.47
S0027	Post		-3.15E-02	6.03E-01	5.39E+00	0.39
S0027	Pre	2.43E+00	6.17E+00	1.65E+00	1.01E+01	0.77
S0030	Pre	-2.50E-01	3.59E+00	-1.60E+00	6.76E+00	0.48
S0030	Post	-9.41E+00	-4.85E+00	-9.79E+00	-3.18E+00	-0.38
S0031	Pre	1.11E+00	5.21E+00	6.94E-01	8.14E+00	0.60
S0031	Post		-6.29E+00		-1.67E+00	-0.12
S0036	Post	-6.46E+00	-5.57E+00	-8.78E+00	9.04E-01	-0.04
S0036	Pre	1.64E-01	2.23E+00	-9.67E-01	5.25E+00	0.38
S0043	Post	-9.10E+00	-4.70E+00	-1.09E+01	-3.01E+00	-0.36
S0043	Pre	-6.12E-01	3.08E+00	-2.12E+00	5.92E+00	0.42
S0084	Post	-7.39E+00	-5.39E+00	-8.22E+00	-1.07E+00	-0.20
S0084	Pre	4.74E-01	2.47E+00	-1.88E-01	5.04E+00	0.37
S0090	Pre	3.65E+00	3.50E+00	2.15E+00	1.32E+01	1.01
S0090	Post	1.25E+00	6.06E-01		3.23E+00	0.25
S0092	Post	-2.09E+00	-2.05E+00		5.31E+00	0.35
S0092	Pre	5.58E-01	7.52E-01		1.03E+01	0.75
S0093	Pre	3.17E+00	6.61E+00	2.35E+00	1.01E+01	0.78
S0093	Post					
S0097	Post	-6.86E+00	-8.53E+00		2.02E+00	0.04
S0097	Pre	-6.88E+00	-6.66E+00	-1.25E+01	1.24E+00	-0.02
S0112	Pre		3.05E+00	-4.92E-01	7.84E+00	0.56
S0112	Post	-2.86E-01	8.67E-01	3.67E-01	5.50E+00	0.39
S0135	Post	-1.31E+01	-1.09E+01	-1.32E+01	-5.06E+00	-0.57
S0135	Pre	-1.24E+01	-1.11E+01	-1.17E+01	-5.24E+00	-0.58
S0168	Pre	-1.12E+01	-7.42E+00	-1.07E+01	-3.92E+00	-0.46
S0168	Post	-1.14E+01	-6.67E+00		-2.72E+00	-0.38
S0172	Pre	-5.51E+00	-5.13E+00	-8.48E+00	-9.87E+00	-0.80
S0172	Post					
S0177	Pre	-6.73E+00	-5.70E+00	-8.72E+00	2.95E+00	0.10
S0177	Post	-8.10E+00	-5.15E+00	-9.92E+00	1.81E+00	0.00
S0186	Pre	-8.54E+00	-5.50E+00	-7.78E+00	2.88E+00	0.07
S0186	Post	-7.28E+00	-7.15E+00	-8.21E+00	1.83E+00	0.02
S0195	Pre	-1.05E+01	-6.22E+00	-1.31E+01	-3.13E+00	-0.39
S0195	Post	-8.65E+00	-4.47E+00		-1.37E+00	-0.24

V0015	Pre	2.23E+00	3.37E+00		1.18E+01	0.89
V0015	Post	1.84E+00	4.40E+00	1.69E+00	8.96E+00	0.67
V0021	Post	-1.24E+01	-8.12E+00	-1.27E+01	-6.09E+00	-0.64
V0021	Pre		-1.17E+01	-1.74E+01	-9.66E+00	-0.70
V0031	Post	2.79E+00	5.29E+00	8.88E-01	9.59E+00	0.73
V0031	Pre	2.35E+00	1.35E+00	1.63E-02	1.05E+01	0.79
V0041	Pre	-8.36E+00	-7.11E+00	-9.03E+00	-2.39E+00	-0.31
V0041	Post	-8.41E+00	-5.97E+00	-9.91E+00	-2.41E+00	-0.31
V0056	Post	1.56E+00	4.63E+00	7.82E-01	7.96E+00	0.60
V0056	Pre	1.25E+00	6.28E-01	-5.85E-01	1.13E+01	0.83
V0070	Post	-1.54E+01	-1.04E+01	-1.55E+01	-8.01E+00	-0.82
V0070	Pre	-1.47E+01	-1.05E+01	-1.49E+01	-7.52E+00	-0.78
V0077	Post	-9.39E+00	-8.03E+00	-9.93E+00	-4.14E+00	-0.45
V0077	Pre	-9.96E+00	-5.53E+00	-1.05E+01	-2.75E+00	-0.36
V0096	Pre	-1.07E+01	-8.30E+00	-1.15E+01	-4.72E+00	-0.51
V0096	Post	-9.77E+00	-5.11E+00	-9.68E+00	-1.62E+00	-0.27
V0097	Post	-3.67E+00	-6.49E+00		1.32E+00	0.04
V0097	Pre	-1.12E+01	-8.61E+00		-5.64E+00	-0.59
V0101	Post	-1.35E+01	-1.01E+01		-7.74E+00	-0.77
V0101	Pre	-1.02E+01	-6.76E+00		-6.29E+00	-0.62
V0116	Pre	-1.13E+01	-6.16E+00	-1.13E+01	-4.63E+00	-0.51
V0116	Post	-1.03E+01	-5.86E+00	-9.62E+00	-2.80E+00	-0.37
V0128	Pre	-1.15E+01	-9.60E+00	-1.16E+01	-3.45E+00	-0.43
V0128	Post	-1.11E+01	-9.22E+00	-1.05E+01	-4.79E+00	-0.52
V0129	Post	-1.17E+01	-1.03E+01		-6.72E+00	-0.67
V0129	Pre		-1.24E+01		-6.24E+00	-0.45
V0141	Post	2.65E+00	5.74E+00	4.48E+00	1.01E+01	0.77
V0141	Pre	-5.84E+00	-1.39E+00	-4.79E+00	1.61E+00	0.02
V0151	Post	-1.21E+01	-6.41E+00		-6.21E+00	-0.64
V0151	Pre	-1.29E+01	-8.56E+00	-1.43E+01	-7.59E+00	-0.75
V0155	Pre	-1.07E+01	-1.04E+01		-4.41E+00	-0.49
V0155	Post	-1.07E+01	-1.09E+01		-4.36E+00	-0.49
V0158	Pre	-7.71E-01		2.79E+00		-0.01
V0158	Post					
V0166	Pre					
V0166	Post					
V0170	Post					
V0170	Pre					
V0173	Pre	-7.90E+00	-4.08E+00	-9.56E+00	-6.42E-01	-0.17
V0173	Post	-9.11E+00	-4.67E+00	-9.91E+00	-1.23E+00	-0.23
V0180	Post	1.49E+00	4.00E+00	7.93E-01	7.88E+00	0.59
V0180	Pre	1.14E-01	4.55E+00	1.82E+00	7.80E+00	0.56
V0195	Post	-7.82E+00	-4.45E+00	-9.38E+00	-1.39E+00	-0.22
V0195	Pre	-6.52E+00	-3.99E+00	-9.65E+00	-1.15E+00	-0.19
V0197	Pre	2.95E-01	1.64E+00	7.55E-01	7.97E+00	0.58

V0197	Post					
V0206	Pre	-9.51E+00	-4.77E+00	-8.59E+00	-1.47E+00	-0.26
V0206	Post	-1.11E+01	-6.65E+00	-1.17E+01	-3.87E+00	-0.46
V0211	Pre	-8.11E+00	-6.69E+00	-8.82E+00	1.21E+00	-0.04
V0211	Post	-5.62E+00	-5.46E+00		2.55E+00	0.09
V0213	Post	-8.79E+00	-8.40E+00		-1.61E-01	-0.15
V0213	Pre		-7.41E+00		-1.73E+00	-0.12
V0229	Post	-5.51E+00	-7.58E+00		-1.32E+00	-0.18
V0229	Pre	-4.72E+00	-6.02E+00		1.12E+00	0.01
V0231	Pre					
V0231	Post					
V0233	Post	-8.00E+00	-8.14E+00		-1.24E+00	-0.22
V0233	Pre	-9.15E+00	-7.72E+00		3.07E-01	-0.12
V0241	Pre	-7.50E+00	-6.65E+00	-8.34E+00	2.17E+00	0.04
V0241	Post		-6.71E+00		-2.71E+00	-0.20
V0245	Pre	-5.01E+00			2.44E+00	0.10
V0245	Post	-4.29E-01	-7.95E+00	-4.58E+00	7.72E+00	0.55
V0253	Post	-1.04E+01	-7.89E+00	-1.01E+01	-3.06E+00	-0.39
V0253	Pre	-1.19E+01	-1.16E+01		-5.55E+00	-0.59
V0263	Post		-2.52E+00		1.21E+00	0.09
V0263	Pre		1.08E+00		7.26E+00	0.52
V0269	Post	-1.14E+01	-1.05E+01	-1.12E+01	-4.53E+00	-0.51
V0269	Pre	-8.90E+00	-8.37E+00	-1.23E+01	-3.42E+00	-0.39
V0275	Post					
V0275	Pre					
W0002	Pre	-1.53E+00	2.02E+00	-3.47E+00	6.08E+00	0.41
W0002	Post	8.20E-01	3.61E+00	1.39E+00	7.21E+00	0.53
W0030	Pre		2.98E+00	1.82E+00	2.34E+00	0.17
W0030	Post	2.78E+00	5.84E+00	4.98E-01	8.12E+00	0.63
W0039	Post	-8.77E+00	-4.33E+00	-9.11E+00	-2.88E+00	-0.35
W0039	Pre	-6.50E+00	-3.06E+00	-7.76E+00	-7.64E-01	-0.16
W0042	Post	-4.16E+00	-2.44E+00	1.15E+01	1.94E+00	0.07
W0042	Pre	-2.89E+00	-5.73E+00		-4.93E-01	-0.08
W0044	Pre					
W0044	Post					
W0047	Pre	-1.39E+01	-6.41E+00	-1.08E+01	-4.12E+00	-0.52
W0047	Post	-6.98E+00	-4.06E+00	-8.17E+00	2.87E+00	0.10
W0052	Pre	-1.25E+01	-8.37E+00	-1.21E+01	-5.86E+00	-0.62
W0052	Post	-1.01E+01	-6.50E+00	-1.08E+01	-3.71E+00	-0.43
W0093	Pre	-1.18E+01	-8.10E+00	-1.06E+01	-3.17E+00	-0.42
W0093	Post	-6.64E+00	-5.89E+00	-8.52E+00	2.24E+00	0.06
W0109	Post	-8.19E+00	-7.28E+00	-9.45E+00	1.10E+00	-0.05
W0109	Pre	-7.16E+00	-5.82E+00	-9.69E+00	5.87E-01	-0.07
W0137	Post	-6.47E+00	-4.53E+00	-6.89E+00	4.13E+00	0.19
W0137	Pre	-9.22E+00	-4.99E+00		1.31E+00	-0.05

W0164	Pre	-1.19E+01	-8.32E+00	-1.21E+01	-3.98E+00	-0.48
W0164	Post	-9.79E+00	-6.03E+00	-1.16E+01	-2.52E+00	-0.34
W0189	Post	-9.50E+00	-5.97E+00	-9.65E+00	-1.96E+00	-0.29
W0189	Pre	-9.76E+00	-3.05E+00	-1.16E+01	-2.82E+00	-0.36
W0200	Post	-8.42E+00	-6.87E+00	-1.19E+01	-1.36E+00	-0.23
W0200	Pre	-1.09E+01	-1.02E+01	-1.32E+01	-6.38E+00	-0.63
W0230	Pre	-9.92E+00	-5.13E+00	-8.34E+00	-1.68E+00	-0.28
W0230	Post	-9.01E+00	-6.62E+00	-1.20E+01	-2.67E+00	-0.34
W0270	Pre	-8.20E+00	-8.49E+00		-2.67E+00	-0.32
W0270	Post	-9.25E+00	-7.02E+00	-1.09E+01	-1.97E+00	-0.29

Table 3.5: microRNA Relative Expression Values and Combined microRNA Relative Expression Values for Pre-operative and Post-operative Blood Plasma Samples

Pre-operative and post-operative blood plasma samples were first analyzed through miRNA isolation, then quantified using qRT-PCR. C_T values for cel-39, miR-155, miR-21, miR-210, and miR-223 were measured, then normalized via cel-39 as a control to obtain the RQ. Due to some extreme values in the dataset (underlined) all miRNA RQ levels were log transformed. Then, the log RQ values were combined to be further analyzed through binary logistic regression. Note that blank spaces are due to samples which did not have measurable cel-39, thus RQ for each miRNA could not be calculated.

(Pre-op, pre-operative; post-op, post-operative; C_T, cycle threshold; RQ, relative expression).

Sample ID	Control C_T (cel-39)	RQ miR-155	RQ miR-21	RQ miR-210	RQ miR-223
C1	3.75E+01	7.71E-01	3.05E+00	1.43E+00	1.05E+03
C10	3.43E+01	8.08E-01	2.09E+01	4.95E-01	2.42E+02
C100	2.55E+01	1.18E-03	3.75E-04	5.43E-04	3.28E-01
C101	2.75E+01		3.31E-03		3.47E+00
C102	3.10E+01		4.15E-01	1.45E-02	8.17E+00
C103	3.00E+01		1.74E-01		4.31E+00
C104	2.62E+01	2.19E-03	6.15E-02	5.09E-03	3.89E-01
C105	2.62E+01	1.63E-03	3.12E-03	1.16E-03	1.32E+00
C106	2.61E+01	1.06E-03	1.96E-02	7.88E-04	2.00E-01
C107	2.45E+01	1.85E-03	1.54E-02		2.18E-01
C108	2.55E+01	9.20E-04	6.53E-03		5.44E-01
C109	2.73E+01	1.08E-03	8.14E-03		3.86E+00
C11	3.56E+01	2.92E+00	6.57E+01	5.39E-01	6.89E+02
C110	2.59E+01	1.13E-02	3.52E-03	5.76E-04	1.25E-01
C12	3.48E+01	1.80E+00	2.61E+01	4.36E-01	4.30E+02
C13	3.50E+01	4.34E+00	2.70E+01	7.88E-01	1.37E+02
C14	3.53E+01	1.26E+00	6.13E+00	4.36E+10	6.02E+01
C15	3.60E+01	4.46E+00	1.21E+01	5.03E-01	2.16E+02
C16	2.66E+01	1.89E-03	6.46E-02	2.16E-03	4.98E-02
C17	2.66E+01	2.00E-03	6.86E-02	2.29E-03	5.26E-02
C18	2.59E+01	3.46E-03	7.61E-03	5.76E-04	9.57E-02
C19	3.52E+01	1.21E+00	1.69E+01	9.42E-01	2.02E+02
C2	2.92E+01		7.17E-01	4.32E-03	4.76E-01
C20	3.50E+01	2.48E-01	9.92E-01	2.84E-01	3.21E+01
C21	3.16E+01	2.89E-02	6.96E-02	2.36E-02	1.43E+01
C22	3.69E+01		2.75E+00		2.56E+03
C23	2.85E+01	9.62E-03	1.28E-02		5.19E+00
C24	2.67E+01	2.50E-03	1.17E-02		4.90E+00
C25	2.94E+01	1.76E+06	1.40E-02		2.22E+01
C26	2.76E+01	2.66E+04	1.43E-03	1.66E-03	5.79E+00
C27	2.77E+01	3.57E-03	8.52E-03	2.56E-03	6.68E+00
C28	3.10E+01	1.53E-02	7.15E-02		8.86E+00
C29	2.85E+01	8.70E-03	9.74E-03		1.09E+01
C3	3.47E+01	3.65E-01	4.79E+00		5.87E+00
C30	3.41E+01	2.22E-01	6.40E-01		7.66E+01
C31	3.34E+01	3.44E-01	2.51E-01		3.29E+01
C32	3.30E+01		2.91E-01		2.59E+01
C33	3.32E+01	3.53E-01	9.71E-02		3.56E+01
C34	3.15E+01	5.99E-02	7.19E-02		1.52E+01
C35	2.97E+01		8.82E-03		3.38E+00
C36	3.56E+01	2.75E-01	2.64E+00		3.37E+02
C37	2.92E+01		1.80E-02	4.19E-03	1.34E+01
C38	3.22E+01		1.90E-01		3.20E+01

C39	2.70E+01	8.53E-03	9.98E-03		4.19E+00
C4	2.39E+01	2.37E-04	2.30E-02		1.32E-02
C40	2.97E+01	1.89E+07	3.85E-02		1.35E+01
C41	2.57E+01	2.11E-03	1.56E-03	2.06E-04	2.10E+00
C42	2.86E+01		1.46E-02	3.09E-03	1.69E+01
C43	2.88E+01		1.31E-02		1.76E+01
C44	2.89E+01		1.52E-02		2.54E+00
C45	2.72E+01	5.83E-03	3.58E-03		5.30E+00
C46	2.94E+01		4.94E-03		3.77E+00
C47	2.99E+01	3.08E-02	6.08E-02		7.06E+00
C48	2.64E+01	2.96E-03	6.28E-03		3.04E+00
C49	2.59E+01	3.04E-03	1.82E-03	6.37E-04	2.23E+00
C5	3.68E+01	6.17E-01	9.57E-01	4.21E-01	2.20E+03
C50	3.23E+01		1.31E+00		4.24E+01
C51	2.97E+01		5.42E-02		4.60E+01
C52	3.01E+01	8.46E-03	7.31E-03		8.52E+01
C53	2.90E+01	1.09E-02	9.01E-03		2.03E+00
C54	2.83E+01	7.43E-03	1.53E-02		2.32E+01
C55	2.91E+01		4.63E-02		3.72E+01
C56	2.59E+01	5.80E-03	6.60E-03	6.02E-04	2.23E+00
C57	2.67E+01		2.74E-02	5.81E-04	8.80E+00
C58	2.90E+01	7.57E-04	3.45E-02		3.27E+01
C59	2.83E+01	4.11E-03	1.69E-02		2.49E+01
C6	2.56E+01		1.57E-02	9.81E-03	3.20E-02
C60	2.78E+01	1.71E-03	1.05E-02		1.74E+01
C61	2.70E+01	5.04E-03	1.95E-02	2.08E-03	1.68E+01
C62	2.75E+01		2.29E-01	2.18E-03	1.56E+01
C63	2.68E+01	2.51E-03	5.76E-02	2.65E-03	1.10E+01
C64	2.62E+01		5.16E-02	1.77E-03	7.73E+00
C65	2.56E+01	1.25E-03	3.06E-02	1.54E-03	4.21E+00
C66	2.57E+01	8.43E-04	2.77E-02	1.16E-04	4.40E+00
C67	2.52E+01	3.55E-04	4.02E-02		3.53E+00
C68	2.39E+01	1.13E-03	1.85E-02	1.22E-03	1.58E+00
C69	2.88E+01		8.80E-03		6.36E+00
C7	2.60E+01	1.54E-03	4.21E-02	3.24E-04	2.20E-02
C70	2.66E+01		3.64E-03		4.41E-01
C71	3.21E+01		5.96E-01		8.84E+00
C72	2.69E+01		3.47E-03		6.86E-01
C73	2.57E+01	2.15E-03	2.80E-03		3.77E-01
C74	2.66E+01	3.09E-03	9.57E-03	9.96E-04	5.47E-01
C75	2.61E+01		3.57E-03	5.09E-04	3.75E-01
C76	2.59E+01	5.28E-03	6.68E-03	4.40E-04	5.14E-01
C77	2.59E+01	5.66E-04	4.50E-03		3.99E-01
C78	2.93E+01	3.14E-02	4.09E-02	3.24E-03	1.22E+01
C79	2.70E+01	5.90E-03	3.78E-03		8.80E-01

C8	2.56E+01	9.12E-04	2.86E-04	1.18E-03	2.68E-03
C80	2.80E+01	1.74E-03	6.76E-03	7.84E-03	1.48E+00
C81	2.64E+01	8.84E-04	5.38E-03		7.60E-01
C82	2.82E+01	2.43E-03	3.17E-03	1.14E-02	3.69E+00
C83	2.82E+01	9.64E-03	8.84E-03		3.36E+00
C84	2.77E+01	1.77E-03	6.37E-03		3.21E+00
C85	2.60E+01		3.20E-03	3.14E-04	5.15E-01
C86	2.52E+01	4.18E-04	3.70E-03		2.90E-01
C87	2.83E+01	2.89E-03	8.03E-03		4.62E+00
C88	2.95E+01	1.95E-02	5.53E-03	4.58E-03	1.16E+01
C89	2.67E+01	1.01E-03	5.32E-03		1.84E+00
C9	3.48E+01	2.98E-01	4.84E+00	5.84E-01	2.81E+01
C90	2.56E+01	9.30E-04	3.23E-03		4.07E-01
C91	2.60E+01		2.30E-03	1.36E-03	1.57E+00
C92	2.55E+01	9.32E-04	1.88E-03	3.21E-04	3.07E-01
C93	3.02E+01	1.55E-03	1.97E-02	1.44E-02	5.29E+00
C94	3.24E+01	3.99E-02	3.69E-02		1.63E+01
C95	2.93E+01		3.18E-02	2.70E-03	2.13E+00
C96	2.72E+01	2.26E-03	5.22E-03		5.49E+00
C97	2.83E+01	4.04E-03	4.57E-02		1.16E+00
C98	2.72E+01	1.47E-03	3.35E-03		2.70E+00
C99	2.97E+01	3.27E-02	3.47E-02		1.92E+01
Sample ID	Log RQ miR-155	Log RQ miR-21	Log RQ miR-210	Log RQ miR-223	Combined RQ
C1	-3.75E-01	1.61E+00	5.18E-01	1.00E+01	0.72
C10	-3.07E-01	4.38E+00	-1.01E+00	7.92E+00	0.57
C100	-9.73E+00	-1.14E+01	-1.08E+01	-1.61E+00	-0.27
C101		-8.24E+00		1.80E+00	0.13
C102		-1.27E+00	-6.11E+00	3.03E+00	0.22
C103		-2.52E+00		2.11E+00	0.15
C104	-8.83E+00	-4.02E+00	-7.62E+00	-1.36E+00	-0.24
C105	-9.26E+00	-8.32E+00	-9.76E+00	4.02E-01	-0.12
C106	-9.88E+00	-5.67E+00	-1.03E+01	-2.32E+00	-0.33
C107	-9.08E+00	-6.02E+00		-2.20E+00	-0.30
C108	-1.01E+01	-7.26E+00		-8.78E-01	-0.22
C109	-9.85E+00	-6.94E+00		1.95E+00	-0.02
C11	1.55E+00	6.04E+00	-8.91E-01	9.43E+00	0.70
C110	-6.46E+00	-8.15E+00	-1.08E+01	-3.00E+00	-0.32
C12	8.51E-01	4.71E+00	-1.20E+00	8.75E+00	0.64
C13	2.12E+00	4.76E+00	-3.44E-01	7.10E+00	0.55
C14	3.31E-01	2.62E+00	3.53E+01	5.91E+00	0.43
C15	2.16E+00	3.59E+00	-9.92E-01	7.75E+00	0.59
C16	-9.05E+00	-3.95E+00	-8.86E+00	-4.33E+00	-0.46
C17	-8.96E+00	-3.86E+00	-8.77E+00	-4.25E+00	-0.45
C18	-8.17E+00	-7.04E+00	-1.08E+01	-3.39E+00	-0.37

C19	2.76E-01	4.08E+00	-8.59E-02	7.66E+00	0.56
C2		-4.80E-01	-7.85E+00	-1.07E+00	-0.08
C20	-2.01E+00	-1.12E-02	-1.81E+00	5.00E+00	0.33
C21	-5.11E+00	-3.85E+00	-5.41E+00	3.83E+00	0.19
C22		1.46E+00		1.13E+01	0.82
C23	-6.70E+00	-6.29E+00		2.38E+00	0.06
C24	-8.65E+00	-6.41E+00		2.29E+00	0.03
C25	2.08E+01	-6.16E+00		4.48E+00	0.65
C26	1.47E+01	-9.45E+00	-9.23E+00	2.53E+00	0.42
C27	-8.13E+00	-6.87E+00	-8.61E+00	2.74E+00	0.07
C28	-6.03E+00	-3.81E+00		3.15E+00	0.13
C29	-6.85E+00	-6.68E+00		3.45E+00	0.14
C3	-1.45E+00	2.26E+00		2.55E+00	0.16
C30	-2.17E+00	-6.43E-01		6.26E+00	0.42
C31	-1.54E+00	-2.00E+00		5.04E+00	0.34
C32		-1.78E+00		4.70E+00	0.34
C33	-1.50E+00	-3.36E+00		5.16E+00	0.35
C34	-4.06E+00	-3.80E+00		3.93E+00	0.22
C35		-6.83E+00		1.76E+00	0.13
C36	-1.86E+00	1.40E+00		8.39E+00	0.57
C37		-5.79E+00	-7.90E+00	3.75E+00	0.27
C38		-2.40E+00		5.00E+00	0.36
C39	-6.87E+00	-6.65E+00		2.07E+00	0.04
C4	-1.20E+01	-5.44E+00		-6.24E+00	-0.64
C40	2.42E+01	-4.70E+00		3.76E+00	0.66
C41	-8.89E+00	-9.33E+00	-1.22E+01	1.07E+00	-0.07
C42		-6.10E+00	-8.34E+00	4.08E+00	0.29
C43		-6.25E+00		4.13E+00	0.30
C44		-6.04E+00		1.34E+00	0.10
C45	-7.42E+00	-8.13E+00		2.41E+00	0.05
C46		-7.66E+00		1.91E+00	0.14
C47	-5.02E+00	-4.04E+00		2.82E+00	0.12
C48	-8.40E+00	-7.32E+00		1.61E+00	-0.02
C49	-8.36E+00	-9.10E+00	-1.06E+01	1.16E+00	-0.05
C5	-6.96E-01	-6.40E-02	-1.25E+00	1.11E+01	0.79
C50		3.91E-01		5.41E+00	0.39
C51		-4.21E+00		5.52E+00	0.40
C52	-6.89E+00	-7.10E+00		6.41E+00	0.35
C53	-6.52E+00	-6.79E+00		1.02E+00	-0.03
C54	-7.07E+00	-6.03E+00		4.53E+00	0.21
C55		-4.43E+00		5.22E+00	0.38
C56	-7.43E+00	-7.24E+00	-1.07E+01	1.16E+00	-0.04
C57		-5.19E+00	-1.07E+01	3.14E+00	0.23
C58	-1.04E+01	-4.86E+00		5.03E+00	0.20
C59	-7.93E+00	-5.89E+00		4.64E+00	0.21

C6		-5.99E+00	-6.67E+00	-4.97E+00	-0.36
C60	-9.20E+00	-6.57E+00		4.12E+00	0.15
C61	-7.63E+00	-5.68E+00	-8.91E+00	4.07E+00	0.17
C62		-2.13E+00	-8.84E+00	3.97E+00	0.29
C63	-8.64E+00	-4.12E+00	-8.56E+00	3.45E+00	0.11
C64		-4.28E+00	-9.14E+00	2.95E+00	0.21
C65	-9.64E+00	-5.03E+00	-9.34E+00	2.07E+00	-0.01
C66	-1.02E+01	-5.17E+00	-1.31E+01	2.14E+00	-0.01
C67	-1.15E+01	-4.64E+00		1.82E+00	-0.05
C68	-9.78E+00	-5.76E+00	-9.68E+00	6.64E-01	-0.11
C69		-6.83E+00		2.67E+00	0.19
C7	-9.34E+00	-4.57E+00	-1.16E+01	-5.50E+00	-0.55
C70		-8.10E+00		-1.18E+00	-0.08
C71		-7.46E-01		3.14E+00	0.23
C72		-8.17E+00		-5.44E-01	-0.04
C73	-8.86E+00	-8.48E+00		-1.41E+00	-0.24
C74	-8.34E+00	-6.71E+00	-9.97E+00	-8.71E-01	-0.20
C75		-8.13E+00	-1.09E+01	-1.42E+00	-0.10
C76	-7.57E+00	-7.23E+00	-1.11E+01	-9.60E-01	-0.19
C77	-1.08E+01	-7.79E+00		-1.32E+00	-0.27
C78	-4.99E+00	-4.61E+00	-8.27E+00	3.61E+00	0.18
C79	-7.41E+00	-8.05E+00		-1.84E-01	-0.13
C8	-1.01E+01	-1.18E+01	-9.72E+00	-8.54E+00	-0.78
C80	-9.17E+00	-7.21E+00	-7.00E+00	5.70E-01	-0.11
C81	-1.01E+01	-7.54E+00		-3.96E-01	-0.19
C82	-8.69E+00	-8.30E+00	-6.46E+00	1.88E+00	0.00
C83	-6.70E+00	-6.82E+00		1.75E+00	0.02
C84	-9.14E+00	-7.29E+00		1.68E+00	-0.03
C85		-8.29E+00	-1.16E+01	-9.56E-01	-0.07
C86	-1.12E+01	-8.08E+00		-1.79E+00	-0.31
C87	-8.43E+00	-6.96E+00		2.21E+00	0.02
C88	-5.68E+00	-7.50E+00	-7.77E+00	3.53E+00	0.16
C89	-9.95E+00	-7.55E+00		8.77E-01	-0.10
C9	-1.75E+00	2.27E+00	-7.75E-01	4.81E+00	0.32
C90	-1.01E+01	-8.27E+00		-1.30E+00	-0.25
C91		-8.76E+00	-9.52E+00	6.52E-01	0.05
C92	-1.01E+01	-9.05E+00	-1.16E+01	-1.70E+00	-0.28
C93	-9.33E+00	-5.66E+00	-6.12E+00	2.40E+00	0.02
C94	-4.65E+00	-4.76E+00		4.03E+00	0.22
C95		-4.98E+00	-8.53E+00	1.09E+00	0.08
C96	-8.79E+00	-7.58E+00		2.46E+00	0.04
C97	-7.95E+00	-4.45E+00		2.12E-01	-0.11
C98	-9.41E+00	-8.22E+00		1.43E+00	-0.05
C99	-4.94E+00	-4.85E+00		4.26E+00	0.23

Table 3.6: microRNA Relative Expression Values and Combined microRNA Relative Expression Values for Pre-operative and Post-operative Blood Plasma Controls

Blood plasma control samples were first analyzed through miRNA isolation, then quantified using qRT-PCR. C_T values for cel-39, miR-155, miR-21, miR-210, and miR-223 were measured, then normalized via cel-39 as a control to obtain the RQ. Due to some extreme values in the dataset (underlined) all miRNA RQ levels were log transformed. Then, the log RQ values were combined to be further analyzed through binary logistic regression. Note that blank spaces are due to samples which did not have measurable cel-39, thus RQ for each miRNA could not be calculated.

(C_T, cycle threshold; RQ, relative expression).

Comparing	AUC, % (95% C.I.)	Cut-off value	Odds Ratio	P-value	95% C.I.
Pre-op vs controls[1]	72.3 (0.641, 0.805)	-0.4169	3	0.003	(1.440, 6.249)
Post-op vs controls[2]	67.0 (0.577, 0.763)	-0.3255	2.275	0.023	(1.120, 4.621)
Pre-op vs post-op[3]	52.4 (0.421, 0.627)	-0.7277	1.204	0.718	(0.441, 3.287)
Recurrence[4]	53.6 (0.228, 0.843)	-1.2541	2.5	0.403	(0.292, 21.399)
No recurrence[5]	53.1 (0.421, 0.641)	-0.635	1.41	0.624	(0.357, 5.559)

Table 3.7: Binary Logistic Regression Risk Score Analysis Comparing Pre-operative, Post-operative, and Controls

ROC analysis was performed to calculate the AUC and obtain the cut-off point used in the binary logistic regression analysis. This calculated an OR showing that if a samples combined miRNA RQ is higher than the cut-off point, the sample is that much more probable to be a case than a control.

[1] When comparing the pre-operative cases (n=64) to the controls (n=110), ROC analysis gave an AUC of 72.3% (95% C.I.=(0.641, 0.805)). This created a cut-off point of -0.4169. Binary logistic regression analysis yielded an OR of 3.000 (p-value=0.003, 95% C.I.= (1.440, 6.249)). Thus, samples with a combined miRNA RQ above -0.4169 are 3 times more likely to be a pre-operative case than a control.

[2] When comparing the post-operative cases (n=60) to the controls (n=110), ROC analysis gave an AUC of 67.0% (95% C.I.=(0.577, 0.763)). This created a cut-off point of -0.3255. Binary logistic regression analysis yielded an OR of 2.275 (p-value=0.023, 95% C.I.= (1.120, 4.621)). Thus, samples with a combined miRNA RQ above -0.3255 are 2.275 times more likely to be a post-operative case than a control.

[3] When comparing the pre-operative cases (n=64) to the post-operative cases (n=60), ROC analysis gave an AUC of 52.4% (95% C.I.=(0.421, 0.627)). This created a cut-off point of -0.7277. Binary logistic regression analysis yielded a statistically insignificant OR, thus samples could not be categorized into risk groups according to the cut-off point.

[4] When comparing the pre-operative cases (n=7) to the post-operative cases (n=8) that had cancer recurrence, ROC analysis gave an AUC of 53.6% (95% C.I.=(0.228, 0.843)). This created a cut-off point of -1.2541. Binary logistic regression analysis yielded a statistically insignificant OR, thus samples could not be categorized into risk groups according to the cut-off point.

[5] When comparing the pre-operative cases (n=57) to the post-operative cases (n=52) that did not have cancer recurrence, ROC analysis gave an AUC of 53.1% (95% C.I.=(0.421, 0.641)). This created a cut-off point of -0.635. Binary logistic regression analysis yielded a statistically insignificant OR, thus samples could not be categorized into risk groups according to the cut-off point.

(AUC, area under the curve; C.I., confidence interval; ROC, receiver operative curve; RQ, relative expression).

Sample size type	Comparing	AUC, % (95% C.I.)	Cut-off value	Odds Ratio	P-value	95% C.I.
	Pre-op vs controls[1]	72.3 (0.641, 0.805)	-0.4169	3	0.003	(1.440, 6.249)
All samples	Gender		-0.4169	1.654	0.108	(0.895, 3.055)
All samples	Smoking status		-0.4169	1.288	0.436	(0.681, 2.437)
All samples	Age		-0.4169	1.43	0.248	(0.780, 2.624)
	Post-op vs controls[2]	67.0 (0.577, 0.763)	-0.3255	2.275	0.023	(1.120, 4.621)
All samples	Gender		-0.3255	1.654	0.108	(0.895, 3.055)
All samples	Smoking status		-0.3255	1.288	0.436	(0.681, 2.437)
All samples	Age		-0.3255	1.43	0.248	(0.780, 2.624)

Table 3.8: Binary Logistic Regression Risk Score Analysis Comparing Pre-operative, Post-operative, and Controls with Followed Analysis of the Parameters Gender, Smoking Status, and Age for Entire Sample Size

ROC analysis was performed to calculate the AUC and obtain the cut-off point used in binary logistic regression analysis. This created an OR showing that if a samples combined miRNA RQ is higher than the cut-off point, the sample is that much more likely to be a case than a control. Using the cut-off points from comparing either the pre-operative, post-operative, or controls groups, binary logistic regression was repeated to look at gender, smoking status, or age affecting the results. When comparing pre-operative cases and controls, as well as post-operative cases and controls, none of the parameters showed statistically significant results, inferring that the results are not due to gender, smoking status, or age. This was done including the entire sample size (pre-operative (n-69), post-operative (n=69), and controls (n=110).

[1] When comparing the pre-operative cases (n=64) to the controls (n=110), ROC analysis gave an AUC of 72.3% (95% C.I.=(0.641, 0.805)). This created a cut-off point of -0.4169. Binary logistic regression analysis yielded an OR of 3.000 (p-value=0.003, 95% C.I.= (1.440, 6.249)). Thus, samples with a combined miRNA RQ above -0.4169 are 3 times more likely to be a pre-operative case than a control.

[2] When comparing the post-operative cases (n=60) to the controls (n=110), ROC analysis gave an AUC of 67.0% (95% C.I.=(0.577, 0.763)). This created a cut-off point of -0.3255. Binary logistic regression analysis yielded an OR of 2.275 (p-value=0.023, 95% C.I.= (1.120, 4.621)). Thus, samples with a combined miRNA RQ above -0.3255 are 2.275 times more likely to be a post-operative case than a control.

(AUC, area under the curve; C.I., confidence interval; ROC, receiver operative curve; RQ, relative expression).

Sample size type	Comparing	AUC, % (95% C.I.)	Cut-off value	Odds Ratio	P-value	95% C.I.
	Pre-op vs controls[1]	72.3 (0.641, 0.805)	-0.4169	3	0.003	(1.440, 6.249)
Samples in analysis	Gender		-0.4169	1.553	0.184	(0.811, 2.972)
Samples in analysis	Smoking status		-0.4169	1.424	0.331	(0.699, 2.902)
Samples in analysis	Age		-0.4169	1.386	0.341	(0.707, 2.718)
	Post-op vs controls[2]	67.0 (0.577, 0.763)	-0.3255	2.275	0.023	(1.120, 4.621)
Samples in analysis	Gender		-0.3255	1.824	0.078	(0.936, 3.555)
Samples in analysis	Smoking status		-0.3255	1.925	0.084	(0.916, 4.047)
Samples in analysis	Age		-0.3255	1.532	0.225	(0.769, 3.050)

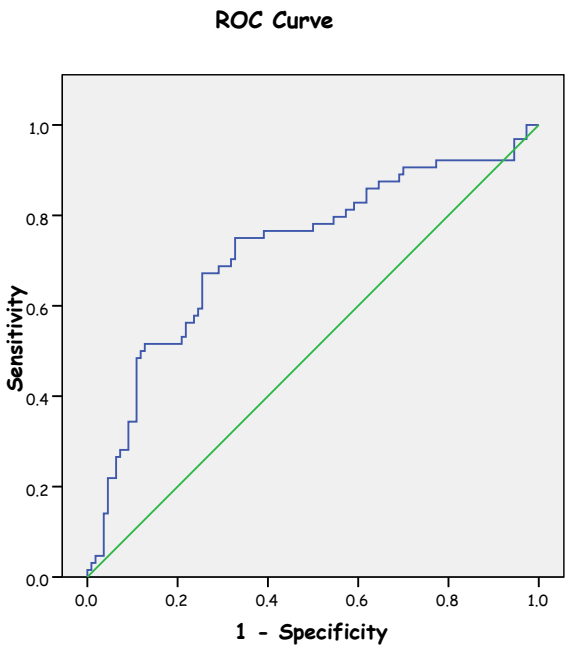
Table 3.9: Binary Logistic Regression Risk Score Analysis Comparing Pre-operative, Post-operative, and Controls with Followed Analysis of the Parameters Gender, Smoking Status, and Age in Samples Only Included in the Analysis

ROC analysis was performed to calculate the AUC and obtain the cut-off point used in binary logistic regression analysis. This created an OR showing that if a samples combined miRNA RQ is higher than the cut-off point, the sample is that much more likely to be a case than a control. Using the cut-off points from comparing either the pre-operative, post-operative, or controls groups, binary logistic regression was repeated to look at gender, smoking status, or age affecting the results. When comparing pre-operative cases and controls, as well as post-operative cases and controls, none of the parameters showed statistically significant results, inferring that the results are not due to gender, smoking status, or age. This was done in the sample size included in the final analysis (pre-operative (n=64), post-operative (n=60), and controls (n=110)).

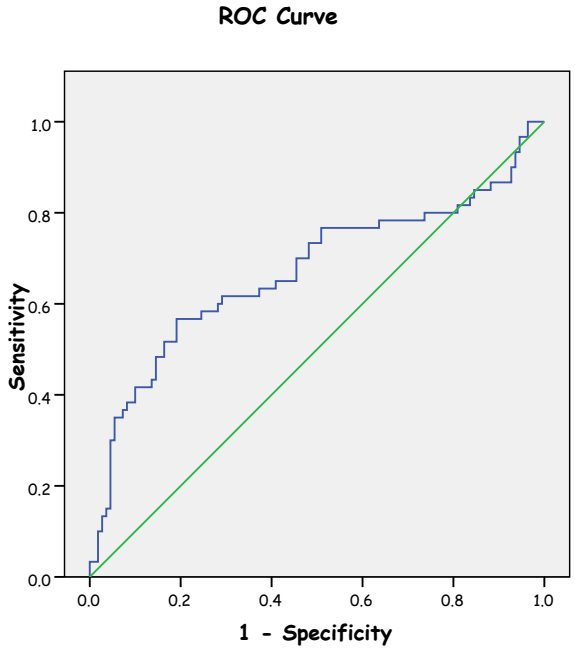
[1] When comparing the pre-operative cases (n=64) to the controls (n=110), ROC analysis gave a. AUC of 72.3% (95% C.I.=(0.641, 0.805)). This created a cut-off point of -0.4169. Binary logistic regression analysis yielded an OR of 3 (p-value=0.003, 95% C.I.= (1.440, 6.249)). Thus, samples with a combined miRNA RQ above -0.4169 are 3 times more likely to be a pre-operative case than a control.

[2] When comparing the post-operative cases (n=60) to the controls (n=110), ROC analysis gave an AUC of 67.0% (95% C.I.=(0.577, 0.763)). This created a cut-off point of -0.3255. Binary logistic regression analysis yielded an OR of 2.275 (p-value=0.023, 95% C.I.= (1.120, 4.621)). Thus, samples with a combined miRNA RQ above -0.3255 are 2.275 times more likely to be a post-operative case than a control.

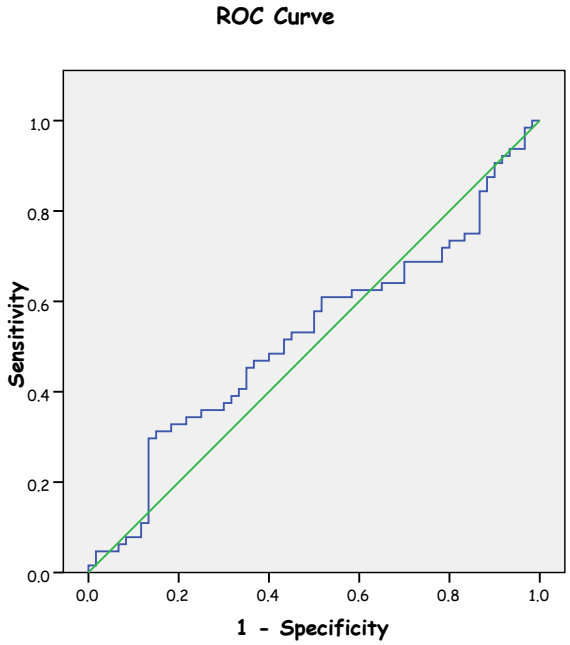
(AUC, area under the curve; C.I., confidence interval; ROC, receiver operative curve; RQ, relative expression).



A)



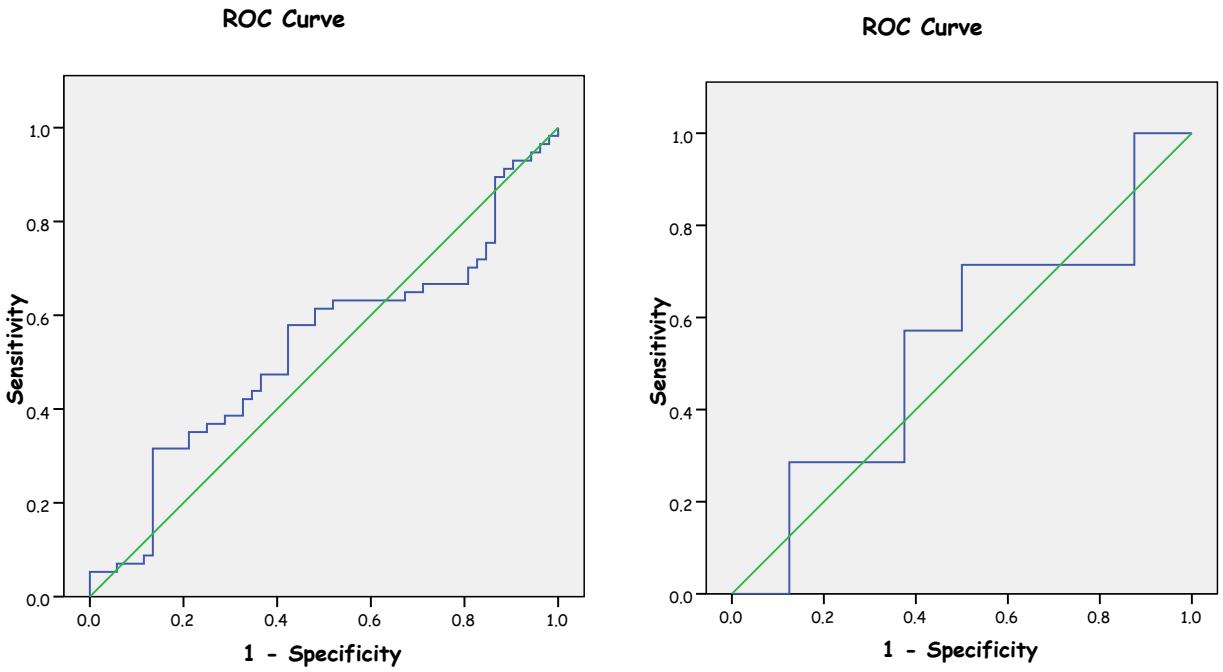
B)



C)

Figure 3.1: Receiver Operating Curve (ROC) Analysis Comparing Pre-operative, Post-operative, and Controls, Disregarding Recurrence

A) Resulting ROC analysis from comparison of pre-operative and controls. AUC was then calculated resulting in an AUC of 72.3% (95% C.I.=(0.641, 0.805)).
B) Resulting ROC analysis from comparison of pre-operative and controls. AUC was then calculated resulting in an AUC of 67.0% (95% C.I.=(0.577, 0.763)).
C) Resulting ROC analysis from comparison of pre-operative and post-operative. All post-operative samples were included regardless of recurrence or not. AUC was then calculated resulting in an AUC of 52.4% (95% C.I.= (0.421, 0.627).
(AUC, area under the curve; C.I., confidence interval; ROC, receiver operating curve).



A) B)

Figure 3.2: Receiver Operating Curve (ROC) Analysis Comparing Recurrence and No Recurrence of Cancer Between Pre-operative and Post-operative Plasma Samples

A) Resulting ROC analysis from comparison of pre-operative and post-operative samples that did not have cancer recurrence at the time of post-operative plasma sample collection. AUC was then calculated resulting in an AUC of 53.1% (95% C.I.=(0.421, 0.641)).
B) Resulting ROC analysis from comparison of pre-operative and post-operative samples that have cancer recurrence at time of post-operative plasma sample collection. AUC was then calculated resulting in an AUC of 53.6% (95% C.I.=(0.228, 0.843)).
(AUC, area under the curve; C.I., confidence interval; ROC, receiver operating curve).

Chapter 4: Discussion

4.1: Demographic Analysis

4.1.1: Population Demographics

When looking in depth at the population demographics, there are issues that could affect the results seen in Chapter 3. The first of which are the higher proportion of male to female cases than controls. Even gender representation was sought for during experimental planning, but ultimately not possible due to the availability of early stage NSCLC cases in tissue banks.

Another contrast in the populations' demographics is smoking history. Cases have a much higher smoking history with 53.5 pack years compared to the controls 30.8 pack years. This was another parameter that the study had little control over, due to tissue bank availabilities. It is difficult to determine if smoking history has an effect on miRNA levels, as this has yet to be investigated thoroughly. Despite this, the NLST criteria for participants eligible for annual screening must have at least a 30 pack-year smoking history, which both the case and control groups meet (23). It would be beneficial to investigate miRNA levels of miR-21, 155, 210, and 223 in order to determine if the amount a person has previously smoked does affect these parameters.

The cases population was comprised of NSCLC, but there are many different histological subtypes. AC comprised 49.3% of cases, and SCC with 44.9% of cases. It is important to note that it is not known if there are different dysregulation levels of the miRNAs used in this study when comparing different subtypes of NSCLC, as well as if other miRNAs are greatly affected in different subtypes. Also, our population contained 5.8% of "other" types in NSCLC (Table 3.1). In order to maintain statistical power, we were not able to separate these histological groups and do separate analyses to look for differences in miRNA expression.

All patient demographic information provided from the tissue banks was blinded, therefore there is no possibility of patient follow up. Although all cases provided in the study are in stage I-II, a select few had a greatest tumour dimension > 10 cm being reported. The information given from the tissue banks measured tumour dimensions from CT imaging before tumour resection, resulting in the sizes being approximate rather than definite. True tumour dimension is not given post-resection, and with patient follow up not being possible, we cannot get an accurate tumour measurement. It is possible that some samples received had tumour sizes large enough to be classified as stage III NSCLC, but without being able to verify this information, and the need of samples for proper sample size, these samples were kept in the study.

Overall, smoking history, tumour histology, and tumour size accuracy are factors that could contribute to the results seen in this study, which will be discussed further.

4.1.2: Demographics of Pre-operative and Post-operative versus Controls and Patient Recurrence Status

Although some samples were excluded in analysis, this did not affect the overall demographics of the study.

Gender in the 64 pre-operative samples and 60 post-operative samples were predominantly male, with 60.9% and 63.3% being male, respectively. Also, average age was similar between the pre and post-operative groups with a mean age of 62 and 62.08 years, respectively. Comparing smoking history, the pre-operative and post-operative group have similar measurements of 51.5 and 51.69 pack-years, respectively. But, it is not known what the smoking status differences were between these two groups. As previously stated, smoking status was similar between cases and controls overall, but it is not known if a proportion of patients quit smoking during the 5-8

months after tumour resection and post-operative blood sample retrieval. It is possible that with the serious diagnosis of lung cancer, this could motivate the patients' decision to quit smoking. Unfortunately, this information was not provided by the tissue banks.

When looking at patient recurrence, only 11.6% of samples were affected.

4.2: Quantitative Reverse Transcriptase Polymerase Chain Reaction (qRT-PCR) Expression

4.2.1: microRNA Extreme Value Expression

When measuring miRNA relative expression, there were values that were considerably higher than the majority of other measurements in the qRT-PCR analysis. To assist with this, the values underwent a logarithmic transformation. Despite this, there were still extreme values present. Rather than remove these extreme values, they ultimately were kept in the final analysis to ensure accuracy of the overall study. These extreme values could be due to the issue of possible inaccurate end point quantification, caused by the difficulty of maintaining linearity during the multiple cycles required to measure miRNA (177-179).

4.2.2: Extreme microRNA Relative Expression Values in Pre-operative and Post-operative Blood Plasma Samples

As depicted in Table 3.5, there are several underlined values that are considered to be extreme. Overall, miR-223 was the only miRNA with occurrences of this, with 7 pre-operative samples and 1 post-operative sample having extreme values. One possibility for these samples having extreme miR-223 expression could be due to sample contamination. It is uncertain at which step during sample analysis this could have occurred. The first person in contact with the sample is

the medical professional collecting the blood, followed with another professional isolating the plasma, and putting it into storage. All samples are stored in small aliquots to ensure no to minimal freeze-thaw cycles can occur, thus these samples are not opened again until they are received at the University of Alberta and analyzed with strict protocols to ensure contamination does not occur, as described in Chapter 2. When samples arrived, all packaging and sample aliquots were inspected to ensure no damage, thawing, or leakage was present, and this did not occur. Also, all samples were inspected, handled, and analyzed by one person to ensure consistency in the experimental conditions. With more variation, and no ability to look at personnel records, occurring before samples were received at the University of Alberta, it is more likely that contamination of samples, if any, ensued during this time.

Another possibility for the presented extreme miR-223 values is the possibility that miR-223 may not be a good indicator for dysregulation due to NSCLC. This is possible, as all other miRNAs did not present extreme values in the pre-operative and post-operative groups.

The extreme values for miR-223 are peculiar, as they are seen predominantly in the pre-operative samples than post-operative. Currently, there is not great understanding in the pathways and proteins targeted by miR-223, thus it is unclear on why these extreme values are occurring only for miR-223 in the pre-operative and post-operative samples and not the other miRNAs measured.

4.2.3: Extreme microRNA Relative Expression Values in Control Blood Plasma Samples

In control samples, there are several extreme miRNA relative expression values as well. There are 3 samples with miR-155 extreme values, 2 for miR-223, and 1 sample for miR-210 (Table 3.6). As expressed above, this is likely due to sample contamination.

4.3: Pre-operative Blood Plasma Samples Can Be Differentiated from Controls Using Binary Logistic Regression

When comparing pre-operative blood plasma samples to control blood plasma samples, a significant binary logistic regression analysis occurred. But, there are factors of the analysis that must be highlighted and challenged regardless of this result.

When the AUC was calculated, a result of 72.3% was found. Although this is acceptable for the statistical analysis to continue, it is not optimal. For a test to be considered "good", an AUC should be above 85%. This is to ensure accuracy of the test being performed, ensuring that biases and confounding factors seen commonly in screening studies using biomarkers (180). This further reflects on the sensitivity and specificity calculated, which will be discussed further below. Using the AUC, a cut-off value of -0.4169 was found. This cut-off point is the combined miRNA relative expression profile, and patients with a miRNA profile above this cut-off value are 3 times more likely to be a case than a control.

Although these results are significant ($p=0.003$), interpretation of the results must be done with caution. As previously stated, the less than optimal AUC value resulted in a sensitivity of 81% and specificity of 41%. With specificity below 95%, in order for the risk score analysis to be considered a proper screening test, the sensitivity would have to be above 95%, as recommended for effective screening of different cancers (12). Unfortunately, this is not the case.

To further ensure that the significant results seen are due to NSCLC and not another factor, more binary logistic regression analyses were done to rule out that age, smoking status, and gender were not factors that produced these results. It is important to note that due to the available information given by the tissue banks, patients were separated into current or ex-

smoker groups, but it is not known how long a patient had to have quit smoking for in order to be considered an ex-smoker. None of the factors showed significant binary logistic regression results using the same miRNA profile cut-off of value -0.4169. Therefore, this cut-off value cannot show risk of being a case when looking at these factors. This allows us to eliminate the factors from confounding our results, as well as the exclusion of samples from affecting the overall trend seen in our analysis.

Overall, a risk score analysis could be useful in screening for lung cancer in the high risk population, but not on its own due to the low AUC measurements, and subsequent low sensitivity and specificity measurements.

4.4: Post-operative Blood Plasma Samples Can Be Differentiated from Controls Using Binary Logistic Regression

When using binary logistic regression, post-operative blood plasma samples were found to be differentiable ($p=0.023$) from controls 5-8 months post tumour resection. Similarly to the pre-operative versus controls analysis, a less than optimal AUC of 67% was calculated giving a miRNA profile cut-off point of -0.3255. Therefore, patients with miRNA profile values above this point were 2.275 times more likely to be a case than control. A sensitivity of 77% and specificity of 41% were calculated. Also, analysis looking into the factors age, smoking status and gender were not significant.

This finding was surprising, as tumours were resected several months before blood withdrawal, and miRNA levels did not return to “normal”, or similar levels as the controls. One reason for this could be due to cell-to-cell communication via exosomal miRNAs.

miRNAs are spilled extracellularly via exosomes from within the cytoplasm, allowing the proteins to explore into the blood stream (where they are further measurable in blood plasma) and can be re-uptaken into neighbouring cells, regardless of malignancy (5-7, 181). Due to miRNAs affecting targeted mRNA expression, they can affect the healthy cells next to the tumour tissue (5-7, 181). What is unknown is how long miRNAs can affect these healthy cells for. Previously, we hypothesized that through waiting long enough these levels would return to normal in the neighbouring cells, allowing for miRNA profiling to also look at cancer recurrence. There is very little research regarding miRNA levels post-operatively in cancer. One study performed in China investigated lung carcinoma measuring miR-21, 24, 30d, and 205 in blood serum using 82 cases and 50 controls pre and post-operatively. Pre-operatively, these 4 miRNAs were statistically significant when comparing cases and controls, but 10 days post-operatively only miR-21 and 24 were statistically significantly decreased (182). Another study investigating miRNAs in cancer pre and post-operatively was a study done in Japan looking at gastric carcinoma. miR-451 and 486 were measured in blood plasma pre-operatively and 1-2 months post-operatively. They found that there was a statistically significant decrease in these two miRNAs expression (183). Unfortunately, there are not any studies similar to ours that investigates NSCLC in blood plasma pre and post-operatively. These studies do show that miRNAs decrease after tumour resection, but these measurements are taken much earlier post-operatively than our study. Further experimentation is required to understand these mechanisms, as they are not well described.

To assist with these findings, experimentation that thoroughly analyzes miRNA levels over time would need to occur. Future research into this could characterize the changing miRNA levels post tumour resection, as with the limited information given from the tissue banks, we are unable

to determine if the miRNA levels changed at some point between blood withdrawal pre-operatively and post-operatively, then perhaps returned to their pre-operative levels.

4.5: Pre-operative and Post-operative Blood Plasma Samples Cannot Be Differentiated Using Binary Logistic Regression Regardless of Cancer Recurrence

Through the use of binary logistic regression, pre-operative and post-operative blood plasma samples are not significantly differentiable from one another, regardless of cancer recurrence. This finding follows the trends previously shown, and supports the finding that miRNA levels return to similar pre-operative levels 5-8 months post tumour resection or do not change.

4.6: Use of CT Scanning in Conjunction with microRNA Profiling Could Have Better Sensitivity and Specificity in Screening for Non-Small Cell Lung Cancers (NSCLC)

Despite the statistically significant results found in this study, our binary logistic regression analysis does not yield high enough sensitivity and specificity values (12) to be a screening protocol on its own. But, our findings do show promise that miRNA profiling could be the right direction for screening the high risk population in NSCLC. Out of all the other studies investigating this, the only study that achieves a high enough sensitivity and specificity is one conducted in the USA with a sensitivity of 86% and specificity of 96% (145). But, this study does not have a large enough sample size (cases=58, controls=29), and investigates miR-21, 126, 21-, and 486-5p making it difficult to compare to one another.

Another important point is that there are 12 miRNAs investigated in NSCLC in blood plasma (miR-20a, 21, 210, 145, 155, 182, 197, 221, 223, 486-5p, 944, and 3662) (143-151). With further investigation into these miRNAs, it would be beneficial to find the key players that would be useful for a NSCLC screening test. Through the use of the miRNAs that show the highest

disparity of dysregulation between cases and controls, a more accurate miRNA profile could be formulated.

Another important issue to consider is the current screening protocols used for detecting lung cancer. As discussed in chapter 1, CT scanning is the current standard for this, but has issues with low specificity and high false positive rates (10, 24, 25). But, the criteria used for selecting the high risk population does not include important parameters such as family history of cancer, socioeconomic status, smoking intensity and exposure to carcinogens other than smoking tobacco (23). It would be beneficial for this to be changed to ensure that the high risk population includes persons with parameters other than just smoking and age.

Overall, through the optimization of both miRNA profiling and CT scanning, through finding the key miRNAs that are dysregulated in lung cancer, as well as including other important parameters into the high risk population, the accuracy of screening for lung cancer could be improved by combining these techniques. This would maximize the efficacy of the test and minimize the impact on the patient.

Chapter 5: General Conclusions

In cancer, it is well known that lung cancer has the highest mortality rates of all cancers worldwide due to general late stage diagnosis. Late stage diagnosis occurs due to the disease being asymptomatic in early stages (1, 2). But, in the late stages of lung cancer, the disease is not easily treatable with less than 14% of cases having a 5-year survival rate (8). The use of screening methodologies has been used to attempt to counteract this issue through screening the high risk population.

As previously discussed, the current screening method used to screen for lung cancer is low-dose CT scanning, which has low specificity and results in a high rate of false positive lung cancer screenings (10). This is leading to unneeded biopsies, hospital stays, costs to the economy, and patient anxiety, showing a great need for more efficacious methods to screen for lung cancer, and increase early stage diagnosis.

miRNAs have shown to be useful candidates as biomarkers for lung cancer screening, as they are tissue specific, detectable in small quantities, highly stable, and show dysregulation in different cancers. They are detectable in biological fluids, such as blood plasma, and can be combined together to create a "profile", allowing researchers to use these to attempt to screen for lung cancer in the high risk population (5, 6, 7).

The aims for this thesis are to use miRNA profiling and test its ability to distinguish cases from controls in early stage NSCLC pre-operatively and post-operatively through blood plasma measurement.

Experimentation started with obtaining 70 early stage NSCLC blood plasma samples from tissue banks, receiving two samples per patient with the first being pre-operatively and the second 5-8

months post tumour resection. 110 controls with similar smoking history and no past history of malignancy were used. Through qRT-PCR and binary logistic regression risk score analysis, we were able to differentiate cases from controls and characterize miRNA profiles in order to create a risk score for the probability that a sample is a case.

We found that when comparing pre-operative blood plasma samples to controls, binary logistic regression differentiated the two groups with relatively good sensitivity, but not great specificity. Similar results were found when comparing post-operative blood samples to controls. The results showed that, despite our hypotheses, miRNA levels either did not decrease over the 5-8 months between the two sample retrievals, or they lowered, then increased again similar to their pre-operative levels.

As previously discussed, these post-operative levels could be due to the cell-to-cell communication via miRNAs once they leave the malignant cell. But, it is not yet known how long miRNAs can communicate to their neighbouring cells, or how strong of an effect they have in general to miRNA levels in neighbouring, non-malignant lung tissue.

Although these results have shown promise, there are limitations. The overall findings in this study of miRNA profiling are not strong enough to be an independent screening test, and would need to be perfected and improved, or combined with other screening methodology, such as CT screening. It is possible that with the combination of CT scanning and miRNA profiling, that we could improve the current issue of high false positive screening currently seen in common protocols. Studies combining these two methodologies would be very beneficial.

Another limitation is the use of qRT-PCR for measurement of miRNAs. Overall, qRT-PCR is expensive and requires different levels of professional personnel to obtain a result. Medical

professionals are needed to retrieve blood samples, lab technicians are required to isolate and analyze the miRNA samples, as well as statisticians to perform the binary logistic regression analysis. The need for many levels of professionals and cost would make it difficult to implement this protocol to the bedside. Future investigation into other methods of miRNA measurement would be useful to the implementation of miRNA profiling into medical practice.

Despite these limitations to miRNA profiling, there are advantages when compared to low-dose CT scanning as a lung cancer screening method. CT scans have long waits, and are not available to persons living in rural areas. Also, there is some concern with repeated exposures to radiation through CT scans. Receiving a blood test could be a minimally invasive alternative to these issues.

Our study also only investigates miRNA profiling in NSCLC, leaving out 10-15% of SCLC cases. Due to the different characteristics, prognostics, and malignant nature of SCLC, future investigation would be required to characterize the miRNA levels of these lung cancer cases.

Overall, this analysis shows that miRNA profiling could be a useful screening method in early stage NSCLC. We propose that through more research and refinement of the protocols presented in this study, miRNA profiling could be the first step in screening protocols. By giving a high risk patient a simple blood test, this could eliminate the long waits and limited availability of CT scans. But, with miRNAs not being strong enough as an independent screening method, patients with a positive blood test results could go on to receive a low-dose CT scan, to check for a false positive result and locate the proposed lung cancer for biopsy. With an additional step in screening protocols, this could decrease the amount of unneeded biopsies that are currently occurring with the use of CT scanning alone.

This study demonstrates the promise in miRNA profiling as a screening method in the high risk population, and warrants future research in the field of lung cancer screening.

References

1. Ferlay, J.; Soerjomataram, O.; Ervik, M.; Dikshit, R.; Eser, S.; Mathers, S.; Rebelo, M.; Parkin, D. M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer* **2015**, 136(5), E359-86.

2. Canadian Cancer Society's Advisory Committee on Cancer Statistics. Canadian Cancer Statistics 2016. Toronto, ON: *Canadian Cancer Society* **2016**.

3. Non-small cell lung cancer risk factors. *American Cancer Society* **2016**. Retrieved from: http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-risk-factors.

4. Lung cancer (Small Cell). Atlanta, GA: *American Cancer Society* **2012**.

5. Billeter, A. T.; Barnett, R. E.; Druen, D.; Polk, Jr H. C.; van Berkel, V. H. MicroRNA as a new factor in lung and esophageal cancer. *Seminars in Thoracic and Cardiovascular Surgery* **2012**, 24(3), 155-165.

6. Ulivi, P.; Zoli, W. miRNAs as non-invasive biomarkers for lung cancer diagnosis. *Molecules* **2014**, 19, 8220-8237.

7. Fanini, F.; Vannini, I.; Amadori, D.; Fabbri, M. Clinical implications of microRNAs in lung cancer. *Seminars in Oncology* **2011**, 38(6), 776-780.

8. Lung cancer (Non-Small Cell). Atlanta GA: *American Cancer Society* **2012**.

9. Cole P, Morrison AS. Basic issues in population screening for cancer. *J Natl Caner Inst* **1980**; 64:1263-1272.

10. Shen, J.; Liao, J.; Guarnera, M.; Fang, H.; Cai, L.; Stass, S.; Jiang, F. Analysis of microRNAs in Sputum to Improve Computer Tomography for Lung Cancer Diagnosis. *Journal of Thoracic Oncology* **2014**; 9(1): 33-40.

11. Anjuman, N.; Li, N.; Guarnera, M.; Stass, S.; Jiang, F. Evaluation of Lung Flute in Sputum Samples for Molecular Analysis of Lung Cancer. *Clinical and Translational Medicine* **2013**; 2: 15-20

12. Obuchowski, N. Ten Criteria for Effective Screening. *AJR* **2001**; 176.

13. Black, W.; Welch H. Screening for Disease. *AJR* **1997**; 168: 3-11

14. Faulkner, K.; Moores, B. Radiation Dose and Smokatic Risk From CT. *Acta Radiol.* **1987**.

15. Mossman, K. Analysis of Risk in CT and Other Diagnostic Radiology Procedures. *Comput Radiol.* **1982**.

16. Renston, J. Survey of Physicians' Attitudes About Risks and Benefits of Chest CT. *South Med Journal* **1996**.

17. Bacchus, B.; Feizue, R. Waiting Your Turn: Wait Times for Health Care in Canada. *Fraser Institute* **2016**.

18. Lung Disease Imposes Major Costs on Canada's Economy **2015**. Retrieved from: http://www.conferenceboard.ca/press/newsrelease/12-03-15/lung_disease_imposes_major_costs_on_canada_s_economy.aspx

19. Flehinger, B.; Melamed, M.; Zamam, M.; Heelan, R.; Perchick, W.; Martini, N. Early Lung Cancer Detection: Results of the Initial (Prevalence) Radiologic and Cytologic Screening in the Memorial Sloan-Kettering Study. *Am Rev Respir Dis* **1984**; 130(4): 555-560.

20. Melamed, M. Lung Cancer Screening Results in the National Cancer Institute New York Study. *Cancer* **2000**; 89(11): 2356-2362.

21. Frost, J.; Ball, W.; Levin, M.; Tockman, M.; Baker, M.; Carter, D.; Eggleston, J.; Erozan, Y.; Gupta, P.; Khouri, N. Early Lung Cancer Detection: Results of the Initial (Prevalence) Radiologic and Cytologic Screening in the Johns Hopkins Study. *Am Rec Respir Dis* **1984**; 130(4): 549-554.

22. Tockman, M. Survival and Mortality from Lung Cancer in a Screened Population: The Johns Hopkins Study. *Chest* **1986**; 89: 324S-325S.

23. National Lung Screening Trial research team. *Radiology* **2011**; 258(1): 243-253.

24. Patz, E.; Pinsky, P.; Gatsonis, C.; Sicks, J.; Kramer, B.; Tammemagi, M.; Chiles, C.; Black, W.; Aberle, D. Overdiagnosis in Low-Dose Computed Tomography Screening for Lung Cancer. *JAMA Internal Medicine* **2013**; 174(2): 269-274.

25. Bach, P.; Mirkin, J.; Oliver, T.; Azzoli, C.; Berry, D.; Brawley, O.; Byers, T.; Colditz, G.; Gould, M.; Jett, J.; et al. Benefits and Harms of CT Screening for Lung Cancer: A Systematic Review. *JAMA* **2012**; 307(22): 2418-2429.

26. Mozzoni, P.; Banda, I.; Goldoni, M.; Corradi, M.; Tiseo, M.; Acampa, O.; Balestra, V.; Ampollini, L.; Casalini, A.; Carbognani, P.; et al. Plasma and EBC microRNAs as Early Biomarkers of Non-Small-Cell Lung Cancer. *Biomarkers* **2013**; 18(8): 679-686

27. Wiley, H.; Shyartsman, S.; % Lauffenburger, D. Computational modeling of the EGF-receptor system: a paradigm for systems biology. *Trends Cell Biol.* **2003**, 13(1), 43-50.

28. Kholodenko, B.; Demin, O.; Moehren, G.; & Hoek, J. Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem*. **1999**, 274(42), 30169-81.

29. Schoeberl, B.; Eichler-Jonsson, C.; Gilles, E.; & Müller, G. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol.* **2002**, 20(4), 370-5.

30. Oda, K.; Matsuoka, Y.; Funahashi, A.; & Kitano, H. A comprehensive pathway map of epidermal growth factor receptor signaling. *Molecular Systems Biology* **2005**, 1(1).

31. Lynch, T.; Bell, D.; Sordella, R.; Gurubhgavatula, S.; Okimoto, R.; Brannigan, B.; Harris, P.; Haserlat, S.; Supko, J.; Haluska, F et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small cell lung cancer to gefitinib. *New England Journal of Medicine* **2004**, 350(21), 2129-39.

32. Walker, F.; Abramowitz, L.; Benabderrahmane, D.; Descatoire, V.; Henin, D.; Lehy, T.; & Aparicio, T. Growth factor receptor expression in anal squamous lesions: modifications associated with oncogenic human papillomavirus and human immunodeficiency virus. *Human Pathology* **2009**, 40(11), 1517-27.

33. Kuan, C.; Wikstrand, C.; & Bigner, D. EGF mutant receptor vIII as a molecular target in cancer therapy. *Endocrine-Related Cancer* **2011**, 8(2), 83–96.

34. Kumar, V.; Abbas, A.; & Aster, J. Robbins basic pathology. *Philadelphia: Elsevier/Saunders* **2013,** 179.

35. Paez, J.; Jänne, P.; Lee, J.; Tracy, S.; Greulich, H.; Gabriel, S.; Herman, P.; Kaye, F.; Lindeman, N.; Boggon, T. et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **2004**, 304(5676), 1497–500.

36. Liang, W.; Wu, X.; Fang, W.; Zhao, Y.; Yang, Y.; Hu, Z.; Xue, C,; Zhang, J.; Zhang, J.; Ma, Y. et al. Network meta-analysis of erlotinib, gefitinib, afatinib, and icotinib in patients with advanced non-small-cell lung cancer harboring EGFR mutations. *PLoS ONE* **2014**, 9(2), e85245.

37. Rodríguez, P.; Rodríguez, G.; González, G.; & Lage, A. Clinical development and perspectives of CIMAvax EGF, Cuban vaccine for non-small-cell lung cancer therapy. *MEDICC Review* **2010**, 12(1), 17–23.

38. Patel, N. Cuba has a lung cancer vaccine – And America Wants It. *Wired* **2015**.

39. Jackman, D.; Miller, V.; Cioffredi, L.; Yeap, B.; Jänne, P.; Riely, G.; Ruiz, M.; Giaccone, G.; Sequist. LV.; & Johnson, B. Impact of epidermal growth factor receptor and KRAS mutations on clinical outcomes in previously untreated non-small cell lung cancer patients: results of an online tumor registry of clinical trials. *Clinical Cancer Research* **2009**, 15(16), 5267–73.

40. Tsuchida, N.; Ryder, T.; & Ohtsubo, E. Nucleotide sequence of the oncogene encoding p21 transforming protein of Kirsten murine sarcoma virus. *Science* **1982**, 217, 937–939.

41. Yun, J.; Rago, C.; Cheong, I.; Pagliarini, R.; Angenendt, P.; Rajagopalan, H.; Schmidt, K.; Willson, J.; Markowitz, S.; Zhou, S. et al. Glucose deprivation contributes to the development of KRAS pathway mutations in tumor cells. *Science* **2009**, 325(5947), 1555–9.

42. Chiosea, S.; Sherer, C.; Jelic, T.; & Dacic, S. KRAS mutant allele-specific imbalance in lung adenocarcinoma. *Modern Pathology* **2001**, 24(12), 1571–7.

43. Hartman, D.; Davison, J.; Foxwell, T.; Nikiforova, M.; & Chiosea, S. Mutant allele-specific imbalance modulates prognostic impact of KRAS mutations in colorectal adenocarcinoma and is associated with worse overall survival. *International Journal of Cancer* **2012**, 131(8), 1810–7.

44. Krasinskas, A.; Moser, A.; Saka, B.; Adsay, N.; & Chiosea, S. KRAS mutant allele-specific imbalance is associated with worse prognosis in pancreatic cancer and progression to undifferentiated carcinoma of the pancreas. *Modern Pathology* **2013**, 26(10), 1346–54.

45. Burmer, G.; & Loeb, L. Mutations in the KRAS2 oncogene during progressive stages of human colon carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **1989**, 86(7), 2403–7.

46. Almoguera, C.; Shibata, D.; Forrester, K.; Martin, J.; Arnheim, N.; & Perucho, M. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. *Cell* **1988**, 53(4), 549–54.

47. Tam, I.; Chung, L.; Suen, W.; Wang, E.; Wong, M.; Ho, K.; Lam, W.; Chiu, S.; Girard, L.; Minna, J. et al. Distinct epidermal growth factor receptor and KRAS mutation patterns in non-small cell lung cancer patients with different tobacco exposure and clinicopathologic features. *Clinical Cancer Research* **2006**, 12(5), 1647–53.

48. Suda, K.; Tomizawa, K.; & Mitsudomi, T. Biological and clinical significance of KRAS mutations in lung cancer: an oncogenic driver that contrasts with EGFR mutation. *Cancer Metastasis Reviews* **2010**, 29(1), 49–60.

49. Riely, G.; Marks, J.; & Pao, W. KRAS mutations in non-small cell lung cancer. *Proceedings of the American Thoracic Society* **2009**, 6(2), 201–5.

50. Pao, W.; Wang, T.; Riely, G.; Miller, V.; Pan, Q.; Ladanyi, M.; Zakowski, M.; Heelan, R.; Kris, M.; & Varmus, H. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLOS Medicine* **2005**, 2(1), e17.

51. Iwahara, T.; Fujimoto, J.; Wen, D.; Cupples, R.; Bucay, N.; Arakawa, T.; Mori, S.; Ratzkin, B.; Yamamoto, T. Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. *Oncogene* **1997**, 14(4), 439–49.

52. Mossé, Y.; Laudenslager, M.; Longo, L.; Cole, K.; Wood, A.; Attiyeh, E.; Laquaglia, M.; Sennett, R.; Lynch, J.; Perri, P. et al. Identification of ALK as a major familial neuroblastoma predisposition gene. Nature 2008, 455(7215), 930–5.

53. Cools, J.; Wlodarska, I.; Somers, R.; Mentens, N.; Pedeutour, F.; Maes, B.; De Wolf-Peeters, C.; Pauwels, P.; Hagemeijer, A. et al. Identification of novel fusion partners of ALK, the anaplastic lymphoma kinase, in anaplastic large-cell lymphoma and inflammatory myofibroblastic tumor. *Genes, Chromosomes & Cancer* **2002**, 34(4), 354–62.

54. Lawrence, B.; Perez-Atayde, A.; Hibbard, M.; Rubin, B.; Dal Cin, P.; Pinkus, J.; Pinkus, G.; Xiao, S.; Yi, E. et al. TPM3-ALK and TPM4-ALK oncogenes in inflammatory myofibroblastic tumors. *The American Journal of Pathology* **2000**, 157(2), 377–84.

55. Sukov, W.; Hodge, J.; Lohse, C.; Akre, M.; Leibovich, B.; Thompson, R.; & Cheville, J. ALK alterations in adult renal cell carcinoma: frequency, clinicopathologic features and outcome in a large series of consecutively treated patients. *Modern Pathology* **2012**, 25(11), 1516–25.

56. Sugawara, E.; Togashi, Y.; Kuroda, N.; Sakata, S.; Hatano, S.; Asaka, R.; Yuasa, T.; Yonese, J.; Kitagawa, M.; Mano, H. et al. Identification of anaplastic lymphoma kinase fusions in renal cancer: large-scale immunohistochemical screening by the intercalated antibody-enhanced polymer method. *Cancer* **2012**, 118(18), 4427–36.

57. Debelenko, L.; Raimondi, S.; Daw, N.; Shivakumar, B.; Huang, D.; Nelson, M.; & Bridge, J. Renal cell carcinoma with novel VCL-ALK fusion: new representative of ALK-associated tumor spectrum. *Modern Pathology* **2011**, 24(3), 430–42.

58. Mariño-Enríquez, A.; Ou, W.; Weldon, C.; Fletcher, J.; Pérez-Atayde, A. ALK rearrangement in sickle cell trait-associated renal medullary carcinoma. *Genes, Chromosomes & Cancer* **2011**, 50(3), 146–53.

59. Jazii, F.; Najafi, Z.; Malekzadeh, R.; Conrads, T.; Ziaee, A.; Abnet, C.; Yazdznbod, M.; Karkhane, A.; & Salekdeh, G. Identification of squamous cell carcinoma associated proteins by proteomics and loss of beta tropomyosin expression in esophageal cancer. *World Journal of Gastroenterology* **2006**, 12(44), 7104–12.

60. Yaakup, H.; Sagap, I.; & Fadilah, S. Primary oesophageal Ki (CD30)-positive ALK+ anaplastic large cell lymphoma of T-cell phenotype. *Singapore Medical Journal* **2008**, 49(10), e289–92.

61. Lin, E.; Li, L.; Guan, Y.; Soriano, R.; Rivers, C.; Mohan, S.; Pandita, A.; Tang, J.; & Modrusan, Z. Exon array profiling detects EML4-ALK fusion in breast, colorectal, and non-small cell lung cancers. *Molecular Cancer Research* **2009**, 7(9), 1466–76.

62. Tuma, R. ALK gene amplified in most inflammatory breast cancers. *Journal of the National Cancer Institute* **2012**, 104(2), 87–8.

63. Powers, C.; Aigner, A.; Stoica, G.; McDonnell, K.; & Wellstein, A. Pleiotrophin signaling through anaplastic lymphoma kinase is rate-limiting for glioblastoma growth. *The Journal of Biological Chemistry* **2002**, 277(16), 14153–8.

64. Stoica, G.; Kuo, A.; Aigner, A.; Sunitha, I.; Souttou, B.; Malerczyk, C.; Caughey, D.; Wen, D.; Karavanov, A. et al. Identification of anaplastic lymphoma kinase as a receptor for the growth factor pleiotrophin. *The Journal of Biological Chemistry* **2011**, 276(20), 16772–9.

65. Murugan, A.; & Xing, M. Anaplastic thyroid cancers harbor novel oncogenic mutations of the ALK gene. *Cancer Research* **2011**, 71(13), 4403–11.

66. Travis, W.; Brambilla, E.; Noguchi, M.; Nicholson, A.; Geisinger, K.; Yatabe, Y.; Beer, D.; Powell, C.; Riely, G.; Van Schil, P. et al. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. *Journal of Thoracic Oncology* **2011**, 6(2), 244–85.

67. Xalkori Approved for Lung Cancer. *FDA* **2011**.

68. ZYKADIA (certinib) capsules, for oral use Initial U.S. Approval. United States Food and Drug Administration **2014**.

69. Chapman, C.; Murray, A.; McElveen, J.; Sahin, U.; Luxemburger, U.; Tureci, O.; Wiewrodt, R.; Barnes, A.; & Robertson, J. Autoantibodies in lung cancer: possibilities for early detection and subsequent cure. *Thorax* **2008**, 63, 228-33.

70. Chapman, C.; Healey, G.; Murray, A.; Boyle, P.; Robertson, C.; Peek, L.; Allen, J.; Thorpe, A.; Hamilton-Fairley, G.; Parsy-Kowalska, C. et al. EarlyCDT®-Lung test: improved clinical utility through additional autoantibody assays. *Tumour Biol* **2012**, 33, 1319-26.

71. Lam, S.; Boyle, P.; Healey, G.; Maddison, P.; Peek, L.; Murray, A.; Chapman, C.; Allen, J.; Wood, W.; Sewell, H. et al. EarlyCDT-Lung: an immunobiomarker test as an aid to early detection of lung cancer. *Cancer Prev Res* **2011**, 4, 1126-34.

72. Boyle, P.; Chapman, C.; Holdenrieder, S.; Murray, A.; Robertson, C.; Wood, W.; Maddison, P.; Healey, G.; Fairley, G.; Barnes, A. et al. Clinical validation of an autoantibody test for lung cancer. *Ann Oncol* **2011**, 22, 383-9.

73. Healey, G.; Lam, S.; Boyle, P.; Hamilton-Fairley, G.; Peek, L.; & Robertson, J. Signal stratification of autoantibody levels in serum samples and its application to the early detection of lung cancer. *J Thorac Dis* **2013**, 5, 618-25.

74. Lee, R.; Feinbaum, R.; Ambros, V.; Feinbaum; Ambros. The C. elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to lin-14: *Cell* **1993**; 75 (5): 843–54.

75. Reinhart, B.; Slack, F.; Basson, M.; Pasquinelli, A.; Bettinger, J.; Rougvie, A.; Horvitz, H.; Ruvkun, G.; Slack; et al. The 21-nucleotide let-7 RNA Regulates Developmental Timing in *Caenorhabditis elegans*. *Nature* **2000**; 403 (6772): 901–6.

76. Pasquinelli, A.; Reinhart, B.; Slack, F.; Martindale, M.; Kuroda, M.; Maller, B.; Hayward, D.; Ball, E.; Degnan, B.; Müller, P.; Conservation of the Sequence and Temporal Expression of let-7 Heterochronic Regulatory RNA. *Nature* **2000**; 408 (6808): 86–9.

77. Ambros, V.; Bartel, B.; Bartel, D.; Burge, C.; Carrington, J.; Chen, X.; Dreyfuss, G.; Eddy, S.; Griffiths-Jones, S.; Marshall, M.; et al. A Uniform System for microRNA Annotation. *RNA* **2003**; 9(3): 277-279.

78. Lee, Y.; Kim, M.; Han, J.; Yeom, K.; Lee, S.; Baek, S.; Kim, V.; Kim; Han; Yeom; et al. microRNA Genes are Transcribed by RNA polymerase II. *EMBO J* **2004**; 23(20): 4051-4060.

79. Winter, J.; Jung, S.; Keller, S.; Gregory, R.; Diederichs, S. Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nature Cell Biology* **2009**; 11: 228-234.

80. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Rådmark, O.; Kim, S.; et al. The Nuclear RNase III Drosha Initiates microRNA Processing. *Nature* **2003**; 425(6956): 415–9.

81. Gregory, R.; Chendrimada, T.; Shiekhattar, R.; Chendrimada; Shiekhattar. microRNA Biogenesis: Isolation and Characterization of the Microprocessor Complex". *Methods Mol. Biol*. **2006**; 342: 33–47.

82. Murchison, E.; Hannon, G.; Hannon. miRNAs on the Move: miRNA Biogenesis and the RNAi Machinery. *Curr. Opin. Cell Biol*. **2004**; 16 (3): 223–9.

83. Lund, E.; Dahlberg, J.; Dahlberg. Substrate Selectivity of Exportin 5 and Dicer in the Biogenesis of microRNAs". *Cold Spring Harb. Symp. Quant. Biol*. **2006**; 71: 59–66.

84. Pratt, A.; MacRae, I.; MacRae. The RNA-Induced Silencing Complex: A Versatile Gene-Silencing Machine. *J. Biol. Chem.* **2009**; 284(27): 17897-17901.

85. Huntzinger, E.; Izaurralde, E. Gene Silencing by microRNAs: Contributions of Translational Repression and mRNA Decay. *Nat. Rev. Genet.* **2011**; 12: 99–110.

86. Subtelny, A.; Eichhorn, S.; Chen, G.; Sive, H.; Bartel, D. Poly(A)-tail Profiling Reveals an Embryonic Switch in Translational Control. *Nature*. **2014**; 508: 66–71.

87. Bazzini, A.; Lee, M.; Giraldez, A. Ribosome Profiling Shows That miR-430 Reduces Translation Before Causing mRNA Decay in Zebrafish. *Science*. **2012**; 336: 233–237.

88. Lewis, B.; Burge, C.; Bartel, D. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates That Thousands of Human Genes are microRNA targets. *Cell* **2005**; 120(1): 15–20.

89. Lewis, B.; Shih, I.; Jones-Rhoades, M.; Bartel, D.; Burge, C. Prediction of Mammalian MicroRNA Targets. *Cell* **2003**; 115(7): 787–798.

90. Mazière, P.; Enright, A.; Enright. Prediction of microRNA Targets. *Drug Discov. Today* **2007**; 12(11–12): 452–8.

91. Rhoades, M.; Reinhart, B.; Lim, L.; Burge, C.; Bartel, B.; Bartel, D. Prediction of Plant microRNA Targets. *Cell* **2002**; 110: 513–520.

92. Schwab, R.; Palatnik, J.; Riester, M.; Schommer, C.; Schmid, M.; Weigel, D. Specific Effects of MicroRNAs on the Plant Transcriptome. *Dev. Cell*. **2005**; 8: 517–527.

93. Llave, C.; Xie, Z.; Kasschau, K.; Carrington, J. Cleavage of Scarecrow-like mRNA Targets Directed by a Class of Arabidopsis miRNA. *Science*. **2002**; 297: 2053–2056.

94. Bohmert, K.; Camus, I.; Bellini, C.; Bouchez, D.; Caboche, M.; Benning, C. AGO1 Defines a Novel Locus of Arabidopsis Controlling Leaf Development. *EMBO J.* **1998**; 17:170–180.

95. Chen, C.; Ridzon, D.; Broomer, A.; Zhou, Z.; Lee, D.; Nguyen, J.; Barbisin, M.; Xu, N.; Mahuvakar, V.; Andersen, M. Real-time Quantification of microRNAs by Stem-loop RT-PCR. *Nucleic Acids Res* **2005**; 33(20): e179.

96. Mraz, M.; Malinova, K.; Mayer, J.; Pospisilova, S. MicroRNA isolation and stability in stored RNA samples. *Biochem. Biophys. Res. Commun.* **2009**; 390(1): 1–4.

97. Liu, C.; Calin, G.; Volinia, S.; Croce, C. MicroRNA expression profiling using microarrays. *Nat Protoc.* **2008**; 3(4): 563–78.

98. Shingara, J.; Keiger, K.; Shelton, J.; Laosinchai-Wolf, W.; Powers, P.; Conrad, R.; Brown, D.; Labourier, E. An Optimized Isolation and Labeling Platform for Accurate microRNA Expression Profiling. *RNA* **2005**; 11(9): 1461-1470.

99. Mencía, A.; Modamio-Høybjør, S.; Redshaw, N.; Morín, M.; Mayo-Merino, F.; Olavarrieta, L.; Aguirre, L.; del Castillo I, Steel, K.; Dalmay, T.; Moreno, F. Mutations in the Seed Region of Human miR-96 are Responsible for Nonsyndromic Progressive Hearing Loss". *Nat. Genet.* **2009**; 41(5): 609–13.

100. Hughes, A.; Bradley, D.; Campbell, M.; Lechner, J.; Dash, D.; Simpson, D.; Willoughby, C. Mutation Altering the miR-184 Seed Region Causes Familial Keratoconus with Cataract. *The American Journal of Human Genetics* **2011**; 89(5): 628-633.

101. de Pontual, L.; Yao, E.; Callier, P.; Faivre, L.; Drouin, V.; Cariou, S.; Van Haeringen, A.; Geneviève, D.; Goldenberg, A.; Oufadem, M. Germline Deletion of the miR-17~92 Cluster Causes Skeletal and Growth Defects in Humans. *Nat. Genet.* **2011**; 43(10): 1026-1030.

102. Thum, T.; Galuppo, P.; Wolf, C.; Fiedler, J.; Kneitz, S.; van Laake, L.; Doevendans, P.; Mummery, C.; Borlak, J.; Haverich, A. MicroRNAs in the Human Heart: A Clue to Fetal Gene Reprogramming in Heart Failure". *Circulation* **2007**; 116(3): 258–67.

103. van Rooij, E.; Sutherland, L.; Liu, N.; Williams, A.; McAnally, J.; Gerard, R.; Richardson, J.; Olson, E. A Signature Pattern of Stress-Responsive microRNAs that Can Evoke Cardiac Hypertrophy and Heart Failure. *Proc. Natl. Acad. Sci. U.S.A* **2006**; 103(48): 18255–60.

104. Tatsuguchi, M.; Seok, H.; Callis, T.; Thomson, J.; Chen, J.; Newman, M.; Rojas, M.; Hammond, S.; Wang, D. Expression of microRNAs is Dynamically Regulated During Cardiomyocyte Hypertrophy. *J. Mol. Cell. Cardiol.* **2007**; 42(6): 1137–41.

105. Zhao, Y.; Ransom, J.; Li, A.; Vedantham, V.; von Drehle, M.; Muth, A.; Tsuchihashi, T.; McManus, M.; Schwartz, R.; Srivastava, D. Dysregulation of Cardiogenesis, Cardiac Conduction, and Cell Cycle in Mice Lacking miRNA-1-2. *Cell* **2007**; **129** (2): 303–17.

106. Zhao, Y.; Samal, E.; Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **2005**; 436(7048): 214–20.

107. Carè, A.; Catalucci, D.; Felicetti, F.; Bonci, D.; Addario, A.; Gallo, P.; Bang, M.; Segnalini, P.; Gu, Y.; Dalton, N. MicroRNA-133 Controls Cardiac Hypertrophy". *Nat. Med.* **2007**; 13(5): 613–8.

108. van Rooij, E.; Sutherland, L.; Qi, X.; Richardson, J.; Hill, J.; Olson, E. Control of Stress-dependent Cardiac Growth and Gene Expression by a microRNA. *Science* **2007**; 316(5824): 575–9.

109. Insull, W. The Pathology of Atherosclerosis: Plaque Development and Plaque Responses to Medical Treatment. *The American Journal of Medicine* **2009**; 122(1): S3–S14.

110. Son, D.; Kumar, S.; Takabe, W.; Kim, C.; Ni, C.; Alberts-Grill, N.; Jang, I.; Kim, S.; Kim, W.; Won Kang, S. The Atypical Mechanosensitive microRNA-712 Derived from pre-ribosomal RNA Induces Endothelial Inflammation and Atherosclerosis. *Nature Communications* **2013**; 4:3000.

111. Phua, Y.; Chu, J.; Marrone, A.; Bodnar, A.; Sims-Lucas, S.; Ho, J. Renal Stromal miRNAs are Required for Normal Nephrogenesis and Glomerular Mesangial Survival. *Physiological reports* **2015**; 3(10): e12537.

112. Musilova, K.; Mraz, M. MicroRNAs in B Cell Lymphomas: How a Complex Biology Gets More Complex". *Leukemia* **2014**; 29:1004–17.

113. Screening Tool Can Detect Colorectal Cancer from a Small Blood Sample (Press Release). *American Association for Cancer Research* **2010**.

114. Nielsen, B.; Jørgensen, S.; Fog, J.; Søkilde, R.; Christensen, I,; Hansen, U.; Brünner, N.; Baker, A.; Møller, S.; Nielsen, H. High levels of microRNA-21 in the Stroma of Colorecta Cancers Predict Short Disease-free Survival in Stage II Colon Cancer Patients. **2010**; 28(1): 27–38.

115. Eyking, A.; Reis, H.; Frank, M.; Gerken, G.; Schmid, K.; Cario, E. miR-205 and miR-373 Are Associated with Aggressive Human Mucinous Colorectal Cancer. *PloS One* **2016**; 11(6): e0156871.

116. Xu, G.; Wei, J.; Jia, W.; Ge, Z.; Zhang, Y.; Zhang, Z.; Liu, X. MicroRNA-21 Promotes Hepatocellular Carcinoma HepG2 Cell Proliferation Through Repression of Mitogen-Activated Protein Kinase-kinase 3. *BMC Cancer* **2013**; 13(469).

117. Jones, K.; Nourse, J.; Keane, C.; Bhatnagar, A.; Gandhi, M. Plasma MicroRNA Are Disease Response Biomarkers in Classical Hodgkin Lymphoma". *Clin Can Res* **2014**; 20(1): 253–64.

118. Wu, H.; Mo, Y. Targeting miR-205 in Breast Cancer". *Expert Opin. Ther. Targets* **2009**; 13(12): 1439–48.

119. Gregory, P.; Bert, A.; Paterson, E.; Barry, S.; Tsykin, A.; Farshid, G.; Vadas, M.; Khew-Goodall, Y.; Goodall, G. The miR-200 Family and miR-205 Regulate Epithelial to Mesenchymal Transition by Targeting ZEB1 and SIP1. *Nat. Cell Biol* **2008**; 10(5): 593–601.

120. Volinia, S.; Cailin, G.; Liu, C.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; et al. A microRNA Expression Signature of Human Solid Tumors Defines Cancer Gene Targets. *Proc. Natl. Acad. Sci. USA* **2006**, 103, 2257–2261.

121. Xie, Y.; Todd, N.; Liu, Z.; Zhan, M.; Fang, H.; Peng, H.; Alattar, M.; Deepak, J.; Stass, S.; Jiang, F. Altered miRNA Expression in Sputum for Diagnosis of Non-small Cell Lung Cancer. *Lung Cancer* **2010**, 67, 170–176.

122. Li, N.; Ma, J.; Guarnera, M.; Fang, H.; Cai, L.; Jiang, F. Digital PCR Quantification of miRNAs in Sputum for Diagnosis of Lung Cancer. *J. Clin. Res. Surg. Oncol.* **2014**, 140, 145–150.

123. Xing, L.; Todd, N.; Yu, L.; Fang, H.; Jiang, F. Early Detection of Squamous Cell Lung Cancer in Sputum by a Panel of microRNA Markers. *Mod. Pathol*. **2010**, 23, 1157–1164.

124. Yu, L.; Todd, N.; Xing, L.; Xie, Y.; Zhang, H.; Liu, Z.; Fang, H.; Zhang, J.; Katz, R.; Jiang, F. Early Detection of Lung Adenocarcinoma in Sputum by a Panel of microRNA Markers. *Int. J. Cancer* **2010**, 127, 2870–2878.

125. Su, J.; Anjuman, N.; Guarnera, M.; Zhang, H.; Stass, S.; Jiang, F. Analysis of Lung Flute-collected Sputum for Lung Cancer Diagnosis. *Biomark. Insights* **2015**, 10, 55–61.

126. Roa, W.; Kim, J.; Razzak, R.; Du, H.; Guo, L.; Singh, R.; Gazala, S.; Ghosh, S.; Wong, E.; Joy, A.; et al. Sputum microRNA Profiling: A Novel Approach for the Early Detection of Non-small Cell Lung Cancer. *Clin. Investig. Med.* **2012**, 35, E271–E281.

127. Kim, O.; Gazala, S.; Razzak, R.; Guo, L.; Ghosh, S.; Roa, W.; Bédard, E. Non-small Cell Lung Cancer Detection Using microRNA Expression Profiling of Bronchoalveolar Lavage Fluid and Sputum. *Anticancer Res.* **2015**, 35, 1873–1880.

128. Li, Y.; Li, W.; Ouyang, Q.; Hu, S.; Tang, J. Detection of Lung Cancer with Blood microRNA-21 Expression Levels in Chinese Population. *Oncol. Lett.* **2011**, 2, 991–994.

129. Yang, Y.; Xu, L.; Zhou, F.; Wang, T. Prognostic Value of microRNA-10b Overexpression in Peripheral Blood Mononuclear Cells of Non-small Cell Lung Cancer Patients. *Tumor Biol.* **2015.**

130. Patnaik, S.; Yendamuri, S.; Kannisto, E.; Kucharczuk, J.; Singhal, S.; Vachani, A.; Corvalan, A. A microRNA Expression Profiles of Whole Blood in Lung Adenocarcinoma. *PLoS ONE* **2012**, 7, e46045.

131. Ulivi, P.; Foschi, G.; Mengozzi, M.; Scarpi, E. Peripheral Blood miR-328 Expression as a Potential Biomarker for the Early Diagnosis of NSCLC. *Int. J. Mol. Sci.* **2013**, 14, 10332–10342.

132. Chen, X.; Hu, Z.; Wang, W.; Ba, Y.; Ma, L.; Zhang, C.; Wang, C.; Ren, Z.; Zhao, Y.; Wu, S.; et al. Identification of Ten Serum microRNAs from a Genome-wide Serum microRNA Expression Profile as Novel Noninvasive Biomarkers for Nonsmall Cell Lung Cancer Diagnosis. *Int. J. Cancer* **2012**, 130, 1620–1628.

133. Ma, Y.; Tian, Z.; Zhang, W. Circulating miR-125b is a Novel Biomarkers for Screening Non-small-cell Lung Cancer and Predicts Poor Prognosis. *J. Cancer Res. Clin. Oncol*. **2012**, 138, 2045–2050.

134. Li, Z.; Zhang, H.; Yang, Z.; Wen, G.; Cui, Y.; Shao, G. Prognostic Significance of Serum microRNA-210 Levels in Nonsmall-cell Lung Cancer. *J. Int. Med. Res*. **2013**, 41, 1437–1444.

135. Jiang, M.; Zhang, P.; Hu, G.; Xiao, Z.; Xu, F.; Zhong, T.; Huang, F.; Kuang, H.; Zhang, W. Relative Expressions of miR-205-5p, miR-205-3p and miR-21 in Tissues and Serum of Non-small Cell Lung Cancer Patients. *Mol. Cell. Biochem*. **2013**, 383, 67–75.

136. Wang, P.; Yang, D.; Zhang, H.; Wei, X.; Ma, T.; Cheng, Z.; Hong, Q.; Hu, J.; Zhuo, H.; Song, Y.; et al. Early Detection of Lung Cancer in Serum by a Panel of MicroRNA Biomarkers. *Clin. Lung Cancer* **2015.**

137. Zhou, C.; Chen, Z.; Dong, J.; Li, J.; Shi, X.; Sun, N.; Luo, M.; Zhou, F.; Tan, F.; He, J. Combination of Serum miRNAs with Cyfra21-1 for the Diagnosis of Non-small Cell Lung Cancer. *Cancer Lett.* **2015**, 367, 138–146.

138. Foss, K.; Sima, C.; Ugolini, D.; Neri, M.; Allen, K.; Weiss, G. miR-1254 and miR-574-5p: Serum-based microRNA Biomarkers for Early-stage Non-small Cell Lung Cancer. *J. Thorac. Oncol.* **2011**, 6, 482–488.

139. Hennessey, P.; Sanford, T.; Choudhary, A.; Mydlarz, W.; Brown, D.; Adai, A.; Ochs, M.; Ahrendt, S.; Mambo, E.; Califano, J. Serum microRNA Biomarkers for Detection of Non-small Cell Lung Cancer. *PLoS ONE* **2012**, 7, e32307.

140. Bianchi, F.; Nicassio, F.; Marzi, M.; Belloni, E. A Serum Circulating miRNA Diagnostic Test to Identify Asymptomatic High-risk Individuals with Early Stage Lung Cancer. *EMBO Mol. Med.* **2011**, 3, 495–503

141. Wang, C.; Ding, M.; Xia, M.; Chen, S.; Le, A.; Soto-Gil, R.; Shen, Y.; Wang, N.; Wang, J.; Gu, W.; et al. A Five-miRNA Panel Identified From a Multicentric Case-control Study Serves as a Diagnostic Tool for Ethnically Diverse Non-small Cell Lung Cancer Patients. *EBioMedicine* **2015**, 2, 1377–1385.

142. Shen, J.; Liu, Z.; Todd, N.; Zhang, H.; Liao, J.; Yu, L.; Guarnera, M.; Li, R.; Cai, L.; Zhan, M.; et al. Diagnosis of Lung Cancer in Individuals with Solitary Pulmonary Nodules by Plasma microRNA Biomarkers. *BMC Cancer* **2011**, 11, 374.

143. Zheng, D.; Haddadin, S.; Wang, Y.; Gu, L.; Perry, M.; Freter, C.; Wang, M. Plasma microRNAs as Novel Biomarkers for Early Detection of Lung Cancer. *Int. J. Clin. Exp. Pathol.* **2011**, 4, 575–586.

144. Shen, J.; Todd, N.; Zhang, H.; Yu, L.; Lingxiao, X.; Mie, Y.; Guarnera, M.; Liao, J.; Chou, A.; Lu, C.; et al. Plasma microRNAs as Potential Biomarkers for Non-small Cell Lung Cancer. *Lab. Investig.* **2011**, 91, 579–587.

145. Wei, J.; Gao, W.; Zhu, C.; Liu, Y.; Mei, Z.; Cheng, T.; Shu, Y. Identification of Plasma microRNA-21 as a Biomarker for Early Detection and Chemosensitivity of Non-small Cell Lung Cancer. *Chin. J. Cancer* **2011**, 30, 407–414.

146. Sanfiorenzo, C.; Illie, M.; Belaid, A.; Barlesi, F.; Mouroux, J.; Marquette, C.; Brest, P.; Hofman, P. Two Panels of Plasma microRNAs as Non-invasive Biomarkers for Prediction of Recurrence in Resectable NSCLC. *PLoS ONE* **2013**, 8, e54596.

147. Tang, D.; Shen, Y.; Wang, M.; Yang, R.; Wang, Z.; Sui, A.; Jiao, W.; Wang, Y. Identification of Plasma microRNAs as Novel Noninvasive Biomarkers for Early Detection of Lung Cancer. *Eur. J. Cancer Prev*. **2013**, 22, 540–548.

148. Sozzi, G.; Boeri, M.; Rossi, M.; Verri, C.; Suatoni, P.; Bravi, F.; Roz, L.; Conte, D.; Grassi, M.; Sverzellati, N.; et al. Clinical Utility of a Plasma-based miRNA Signature Classifier Within Computed Tomography Lung Cancer Screening: A Correlative Mild Trial Study. *J. Clin. Oncol.* **2014**, 32, 768–773.

149. Geng, Q.; Fan, T.; Zhang, B.; Wang, W.; Xu, Y.; Hu, H. Five miRNAs in Plasma as Novel Biomarkers for Screening of Early-stage Non-small Cell Lung Cancer. *Respir. Res*. **2014**, 15.

150. Powrozek, T.; Krawczyk, P.; Kowalski, D.; Winiarczyk, K.; Olszyna-Serementa, M.; Milanowski, J. Plasma Circulating microRNA-944 and microRNA-3662 as Potential Histologic Type-specific Early Lung Cancer Biomarkers. *Transl. Res.* **2015**, 166, 315–322.

151. Baker, S.; Kramer, B.; Srivastava, S. Markers for Early Detection of Cancer: Statistical Guidelines for Nested Case-control Studies. *BMC Medical Research Methodology* **2002**; 2:4.

152. Osaki, M.; Oshimura, M.; Ito, H. The PI3K-Akt Pathway: Its Functions and Alterations in Human Cancer". *Apoptosis* **2004**; 9(6): 667–676.

153. Manning, B.; Cantley, L. AKT/PKB Signaling: Navigating Downstream. *Cell* **2007**; 29(5): 1004-1017.

154. Musilova, K.; Mraz, M. MicroRNAs in B-cell Lymphomas: How a Complex Biology Gets More Complex. *Leukemia* **2015**; 29(5): 1004–17.

155. Orton, R.; Sturm, O.; Vyshemirsky, V.; Calder, M.; Gilbert, D.; Kolch, W. Computational Modelling of the Receptor-tyrosine-kinase-activated MAPK Pathway. *The Biochemical Journal* **2005**; 392: 249-261.

156. Asangani, I.; Rasheed, S.; Nikolova, D.; Leupold, J.; Colburn, N.; Post, S.; Allgayer, H. MicroRNA-21 (miR-21) Post-transcriptionally Downregulates Tumor Suppressor Pdcd4 and Stimulates Invasion, Intravasation and Metastasis in Colorectal Cancer". *Oncogene* **2008**; 27(15): 2128–36.

157. Meng, F.; Henson, R.; Wehbe-Janek, H.; Ghoshal, K.; Jacob, S.; Patel, T. microRNA-21 Regulates Expression of the PTEN Tumor Suppressor Gene in Human Hepatocellular Cancer. *Gastroenterology* **2007**; 133(2): 647–58.

158. Iorio, M.; Ferracin, M.; Liu, C.; Veronese, A.; Spizzo, R.; Sabbioni, S.; Magri, E.; Pedriali, M.; Fabbri, M.; Campiglio, M.; et al. MicroRNA Gene Expression Deregulation in Human Breast Cancer". *Cancer Research* **2005**; 65(16): 7065–70.

159. Iorio, M.; Visone, R.; Di Leva, G.; Donati, V.; Petrocca, F.; Casalini, P.; Taccioli, C.; Volinia, S.; Liu, CG.; Alder, H.; et al. MicroRNA signatures in human ovarian cancer. *Cancer Research* **2007**; 67(18): 8699–707.

160. Lui, W.; Pourmand, N.; Patterson, B.; Fire, A. Patterns of known and novel small RNAs in human cervical cancer. *Cancer Research* **2007**; 67(13): 6031–43.

161. Volinia, S.; Calin, G.; Liu, C.; Ambs, S.; Cimmino, A.; Petrocca, F.; Visone, R.; Iorio, M.; Roldo, C.; Ferracin, M.; et al. A microRNA Expression Signature of Human Solid Tumors Defines Cancer Gene Targets. *Proceedings of the National Academy of Sciences of the United States of America* **2006**; 103(7): 2257–61.

162. Chan, J.; Krichevsky, A.; Kosik, K. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Research* **2005**; 65(14): 6029–33.

163. Hu, Y.; Correa, A.; Hoque, A.; Guan, B.; Ye, F.; Huang, J.; Swisher, S.; Wu, T.; Ajani, J.; Xu, X. Prognostic Significanc of Differentially Expressed miRNAs in Esophageal Cancer. *International Journal of Cancer. Journal International Du Cancer* **2011**; 128(1): 132–43.

164. Tetzlaff, M.; Liu, A.; Xu, X.; Master, S.; Baldwin, D.; Tobias, J.; Livolsi, V.; Baloch, Z. Differential expression of miRNAs in papillary thyroid carcinoma compared to multinodular goiter using formalin fixed paraffin embedded tissues. *Endocrine Pathology* **2007**; 18(3): 163–73.

165. Teng, G.; Papavasiliou, F. Shhh! Silencing by microRNA-155. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **2009**; 364(1517): 631–637.

166. Faraoni, I.; Antonetti, F.; Cardone, J.; Bonmassar, E. miR-155 Gene: A Typical Multifunctional microRNA". *Biochimica et Biophysica Acta* **2009**; 1792(6): 497–505.

167. Tsuchiya, S.; Fujiwara, T.; Sato, F.; Shimada, Y.; Tanaka, E.; Sakai, Y.; Shimizu, K.; Tsujimoto, G. microRNA-210 Regulates Cancer Cell Proliferation Through Targeting Fibroblast Growth Factor Receptor-like 1 (FGFRL1). *Journal of Biological Chemistry* **2011**; 286(1): 420-428.

168. Eyholzer, M.; Schmid, S.; Schardt, J.; Haefliger, S.; Mueller, B.; Pabst, T. Complexity of miR-223 regulation by CEBPA in human AML. *Leuk Res* **2010**; 34(5): 672–6.

169. Stamatopoulos, B.; Meuleman, N.; Haibe-Kains, B.; Saussoy, P.; Van Den Neste, E.; Michaux, L.; Heimann, P.; Martiat, P.; Bron, D.; Lagneaux, L. microRNA-29c and microRNA-223 Down-regulation Has in vivo Significance in Chronic Lymphocytic Leukemia and Improves Disease Risk Stratification". *Blood* **2009**; 113(21): 5237–45.

170. Chiaretti, S.; Messina, M.; Tavolaro, S.; Zardo, G.; Elia, L.; Vitale, A.; Fatica, A.; Gorello, P.; Piciocchi, A.; Scappucci, G.; et al. Gene Expression Profiling Identifies a Subset of Adult T-cell Acute Lymphoblastic Leukemia with Myeloid-like Gene Features and Over-expression of miR-223. *Haematologica* **2010**; 95(7): 1114–21.

171. Pulikkan, J.; Dengler, V.; Peramangalam, P.; Peer Zada, A.; Müller-Tidow, C.; Bohlander, S.; Tenen, D.; Behre, G. Cell-cycle Regulator E2F1 and microRNA-223 Comprise an Autoregulatory Negative Feedback Loop in Acute Myeloid Leukemia. *Blood* **2010**; 115(9): 1768–78.

172. Liu, T.; Chen, S.; Kuo, S.; Cheng, A.; Lin, C. E2A-positive Gastric MALT Lymphoma Has Weaker Plasmacytoid Infiltrates and Stronger Expression of the Memory B-cell-associated miR-223: Possible Correlation with Stage and Treatment Response". *Mod Pathol* **2010**; 23(11): 1507–17.

173. Laios, A.; O'Toole, S.; Flavin, R.; Martin, C.; Kelly, L.; Ring, M.; Finn, SP.; Barrett, C.; Loda, M.; Gleeson, N.; et al. Potential Role of miR-9 and miR-223 in Recurrent Ovarian Cancer. *Mol Cancer* **2008**; 7:35.

174. Drake S. *Clinical Chemistry* **2004**; 50: 2398-2401.

175. WHO Guidelines for Drawing Blood: Best Practices in Phlebotomy. *WHO Press, Geneva* **2010.**

176. Shiao, Y. A new reverse transcription-polymerase chain reaction method for accurate quantification. *BMC Biotechnol* **2003**, 3(22).

177. Gettemy, J.; Ma, B.; Alic, M.; & Gold, M. Reverse transcription-PCR analysis of the regulation of the manganese peroxidase gene family. *Appl. Environ. Microbiol.* **1998**, 64(2), 569-74.

178. Ramakers, C.; Ruijter, J.; Deprez, R.; & Moorman, A. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters* **2003**, 339(1), 62-66.

179. Hajian-Tilaki, K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med* **2013**, 4(2), 627-635.

180. Vanni, I.; Alama, A.; Grossi, F.; Dal Bello, M.; & Coco, S. Exosomes: a new horizon in lung cancer. *In Drug Discovery Today* **2017**, 22(6), 927-936.

181. Le, H.; Zhang, Y.; Zhu, W.; Chen, D.; He, J.; Huang, Y.; & Liu, X. Evaluation of dynamic change of serum miR-21 and miR-24 in pre- and post-operative lung carcinoma patients. *Medical Oncology* **2012**, 29(5), 3190-97.

182. Ren, C.; Chen, H.; Han, C.; Fu, D.; Wang, D.; &Shen, M. High expression of miR-16 and miR-451 predicating better prognosis in patients with gastric cancer. *J Cancer Res Clin Oncol* **2016**, 142(12), 2489-96.