Ultrasensitive Assay Development toward the Early Detection of Breast Cancer

A dissertation submitted by

Shazia Baig

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Chemistry

Tufts University

May 2017

Advisor: Professor David R. Walt

Abstract

Breast cancer is one of the most commonly diagnosed cancers in women today, with an estimated 250,000 new cases in the United States this year. The advent of new therapeutic strategies has vastly improved disease-free survival rates, but screening mammography can miss up to 20% of cancer cases. This lack of clinical sensitivity prevents early detection and in turn, timely treatment. Moreover, current diagnostic processes used to determine treatment plans are invasive and subjective. A quantitative, cancer biology-focused blood test for breast cancer would provide a minimally invasive method for screening and characterizing the disease. This blood test could also serve the purpose of real-time monitoring of therapeutic efficacy and disease recurrence.

This thesis focuses on the development of ultrasensitive protein assays toward a blood test for the detection of breast cancer. The technology behind the single molecule array (Simoa) assays used to achieve ultralow limits of detection, as well as the advantages of using such a sensitive method are described in Chapter 2. Chapter 3 details the development and preliminary testing of singleplex cancer biomarker assays in serum. Chapter 4 describes the incorporation of these markers and others into several multiplex assays for the simultaneous detection of proteins in small volumes of serum. The multiplexed serum data was then used with a multivariate classification method to evaluate the utility of an eight-protein fingerprint for the detection of breast cancer.

To my parents, Sayeed and Shakira Baig, whose unconditional love and support have made me everything I am. Thank you for everything.

Acknowledgements

I would like to first thank Dr. Walt, whose unending support and constant challenges to succeed made me an independent scientist during my time at Tufts. Thank you to my committee members, Dr. Samuel Thomas and Dr. Joshua Kritzer, who have been an unwavering source of valuable insight and support for the last several years. I also would like to thank Dr. Marsha Moses of Harvard Medical School and Boston Children's Hospital, who took the time to serve on my thesis committee and provide her expertise to the breast cancer biomarker project.

There are several other individuals who helped bring this project to its current state. I would like to thank Dr. Rachel Buchsbaum, who contributed her clinical experience to this project, and coordinated the collection and characterization of the clinical samples from Tufts Medical Center. The biomarker signature would not be what it is without the valuable contributions from Dr. Charlotte Kuperwasser and Dr. Gail Sonenshein of Tufts Medical Center, Dr. Marsha Moses of Boston Children's Hospital, and Dr. Akhilesh Pandey of John's Hopkins University. Dr. Stephanie Schubert and Dr. Stephanie Walter were also critical to the development and progress in this project, as well as Dr. Liangxia Xie who contributed to the final set of data. I was also fortunate enough to mentor Kudret Usmani and Julie Xia during my time at Tufts, and I thank them for their help with this project. Thanks to Dr. Manuel Palacios, who guided me during the multivariate analysis of the clinical data, and remained patient with me throughout. I want to acknowledge the Department of Defense for funding all of my graduate work.

I would also like to thank the Walt lab members, past and present, who have taught me to be a better scientist, colleague, leader, and friend. Specifically, thank you to Dr. Sarah Brunker, Dr. Stephanie Schubert, Dr. Stephanie Walter, Limor Cohen, Soyoon Hwang, Dr. Liangxia Xie, and many others. I also deeply appreciate the other faculty such as Dr. Charles Mace for fostering a supportive environment within the department, as well as the support staff for holding it all together.

Thank you to the team at Quanterix, especially Purvish Patel, Melissa Gardel, Tomasz Piech, Ray Meyer, Jennifer Geldart-Flashman, Derek DuPont, Dr. Cheuk Kan, Dr. David Rissin, and others who have supported my use of the Simoa technology for the last several years.

Thank you to Dr. Mary Carroll and Dr. Kristin Fox, who fostered my love for chemistry and research at Union College. Your mentorship and leadership roles in the department inspired me to go to graduate school. There are a number of people who have supported me throughout this journey, turning colleagues into friends, and friends into family. Thank you to Sarah Brunker, Amanda Aldous, Cristy Zamora, Steph Schubert, Steph Walter, Limor Cohen, and Lauren Gibson for making me look forward to going into lab everyday, and encouraging me to see this through to the end. To my friends outside of graduate school: Marti Gabriella, Maha Mian, Madiha Rasool, Sabrina Ghaus, Rifat Mamun, Hyder and Diya Kazmi, the Gabriella family, and other friends, old and new, thank you for being and endless source of love and support.

Lastly, I would like to thank my parents, Sayeed and Shakira Baig. I will never have the words to express how grateful I am for their love, encouragement,

strength, and sacrifice. Everything they have done has brought me to this point, and I am lost without them. Thank you to my sister, Sadaf Baig, who has been a parent, sibling, friend, mentor, ally, and whatever else she has needed to be to help me succeed. It takes a lot of love and support to create the environment this family has, and I am forever grateful. Thank you all. I wouldn't be where I am without you.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Introduction

Breast cancer is the second most common cancer in women, constituting 29% of newly diagnosed cancer cases.[1] In 2016, there will be an estimated 247,000 newly diagnosed cases of breast cancer, with a predicted 40,000 deaths due to the disease.[1, 2] A woman who lives in the United States has a 12.3% (or one in eight) chance of developing the disease in her lifetime.[1] Given the prevalence of breast cancer in women, there are efforts to improve the screening process to detect cancer in its early stages in a robust, minimally invasive manner, and in turn improve the outcome of the disease.[3] Current screening methods rely on imaging techniques in which detection sensitivity and specificity are affected by variables such as tissue density and the size of the tumor. Physical exams are also limited by similar factors, where small tumors and a lack of symptoms allow the disease to escape detection. Furthermore, false positives lead to inaccurate diagnoses and often require invasive procedures such as biopsies to confirm the initial findings. These screening procedures can miss small tumors or a rapidly progressing disease, which can lead to poor prognosis due to delayed diagnosis and more aggressive treatment. Thus, an alternative method for breast cancer detection would have to bypass the typical limitations of the current techniques and provide biological information about the disease.

In this chapter, breast cancer characteristics and biomarkers are described, as well as clinically accepted screening and diagnostic practices. These practices are evaluated, and the need for a more sensitive, specific and minimally invasive screening method is established. Single molecule arrays (Simoa) assays are

introduced as a novel method to address the problems with current breast cancer screening practices.

Breast Cancer

Symptoms and Risk Factors

Breast cancer does not typically display any physically noticeable symptoms in early stages, which is often when it is most treatable. Aside from a palpable tumor, there may be breast soreness or redness, as well as nipple discharge, but there may not be any early warning signs of the disease. The lack of early indication of disease is the reason for current screening guidelines for early detection.[1]

There are a number of known risk factors, including genetic predisposition, lifestyle, age, and pre-existing carcinomas that inform the screening process. One of the larger risk factors is having one or more family members, primary or otherwise, with a history of breast cancer. Among the most widely known genetic factors in breast cancer risk are the BRCA1 and BRCA2 genes. Breast cancer due to hereditary factors accounts for only 5-10% of cases, but having either of these sets of mutations confers a 40-80% lifetime risk of getting breast cancer. Hereditary factors combined with age (70 years of age or older) increases the likelihood of breast cancer development to 80-87%. Despite the wide variation in risk based on age, family, and ethnicity, a personal history or family history of breast cancer is still one of the strongest known predictors of the disease.[4] Other risk factors of equal importance include previous breast disease or noninvasive carcinomas in situ,

either in the duct or the lobule (DCIS and LCIS). These in situ carcinomas are often considered precursors to invasive breast cancer.[1]

Other potential factors have also been identified, including but not limited to age at menarche, parity, BMI, oral contraceptives, and the use of hormonal therapy for menopause. These conditions or lifestyle choices do not have a clear link to the risk of breast cancer, as multiple studies have reported different results.[5] Furthermore, the same risk factor may decrease the risk of one subtype of breast cancer, but increase the risk of another.[5] Generally, healthy lifestyle choices have a stronger link to a better prognosis for breast cancer survivors.[6]

Disease Classification

Though breast cancer describes a tumor of the breast, it is actually a diverse set of diseases that must be further stratified and characterized before it can be appropriately treated. Breast tumors can vary in a number of ways, including size, malignance, aggressiveness, and therapeutic response, which can greatly influence a patient's prognosis and treatment plans. Current techniques that identify and characterize such features will be described and evaluated further.

Staging and Grading

One of the most well-known ways that breast cancer is stratified is by staging. The American Joint Committee on Cancer (AJCC) created a standardized method for assigning a stage to cancer at the time of diagnosis. This system, known as the TNM staging system, is based on T (tumor size and spread within local area), N (lymph node involvement), and M (distant metastases), shown in Table 1.1. Each part of the system has a degree of severity associated with it, denoted by a number

(0 to 4 for T, 0 to 3 for N, and 0 or 1 for M), and the stage is determined by the combination of severity of each factor. A more detailed description of the TNM system is included in Appendix Table A1.

Stage	Tumor	Node	Metastasis
Stage 0	Tis	N0	M0
Stage IA	T1	N0	M0
Stage IB	T0	N1mi	M0
	T1	N1mi	M0
Stage IIA	T0	N1	M0
	T1	N1	M0
	T2	N0	M0
Stage IIB	T2	N1	M0
	T3	N0	M0
Stage IIIA	T0	N2	M0
	T1	N2	M0
	T2	N2	M0
	T3	N1	M0
	T3	N2	M0
Stage IIIB	T4	N0	M0
	T4	N1	M0
	T4	N2	M0
Stage IIIC	Any T	N3	M0
Stage IV	Any T	Any N	M1

Table 1.1. American Joint Committee on Cancer staging guidelines for breast cancer (used with permission from Reference 7.)

Early stages (Stages I and II) generally include T1-3, N0-2, and M0, which describe a relatively small primary tumor with little to no lymph node involvement. The assigned stage increases as the mass spreads beyond the immediate area of the breast. Stages I-III have no distant metastases, but any degree of T or N (including zero) combined with M1 automatically classifies the disease as Stage IV. Examples of distant metastases include anything far from the primary tumor, like the lungs or bones. The earlier the stage of the disease at diagnosis, the higher the five-year relative survival rate (Figure 1.1).

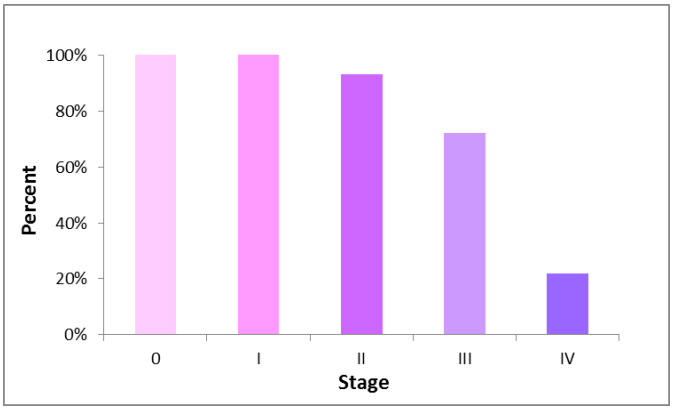


Figure 1.1. Five year relative survival rate by stage at diagnosis based on National Cancer Institute SEER database (adapted from References 3 and 8)

The Nottingham grading system also helps characterize breast cancer by semi-quantitatively describing the aggressiveness of the disease. Stained tumor tissue is analyzed by a pathologist and the morphology of the cells is scored based on three criteria: (1) differentiation, the similarity of tumor cells to normal glands, (2) cellular pleomorphism, the variability in size and shape of the nucleus in tumor cells, and (3) mitotic activity, based on the presence of dividing cells. Each category is assigned a value of 1 to 3, and is then combined into an overall score. A lower value indicates a less aggressive disease. This information can provide insight for treatment, as more aggressive disease would require more aggressive treatment options.[9]

Subtype

Breast cancer is a heterogeneous disease; each subtype exhibits unique characteristics that can aid in formulating an appropriate therapeutic strategy. It is important to identify the subtype of the disease in order to avoid ineffective treatments to which the tumor would not respond, or overtreatment, which can be unnecessarily difficult for the patient due to side effects. There are four major

molecular subtypes of breast cancer, listed in Table 1.2. The main criteria for each subtype classification are based on the expression of hormone receptors and growth factors, including estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor (HER2).

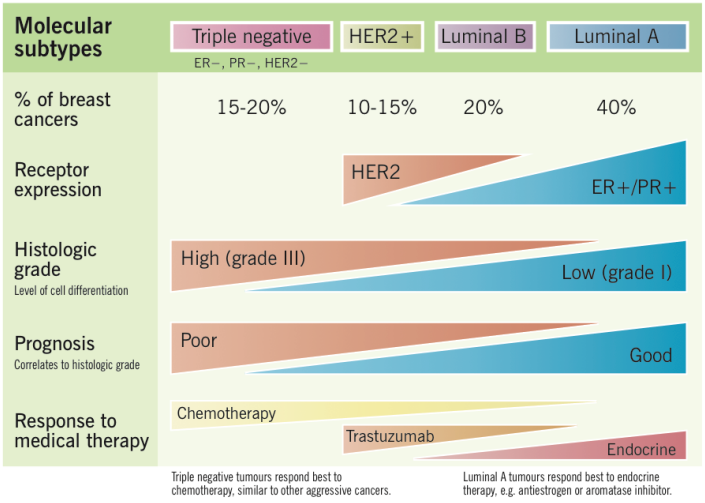


Table 1.2 Four major subtypes of breast cancer, description, and prevalence of cases[10]

Estrogen receptor alpha (ERα, also referred to as ER) is a hormone receptor involved in normal development and function of the mammary gland, but is found to be overexpressed or "ER-positive" in 70% of invasive breast cancer cases.[11, 12] ERα is a ligand responsive transcription factor that typically has low expression levels in normal cells, though basal levels vary by age. It has potential value as a risk factor, as it is found in normal tissue of breast cancer patients and populations with higher breast cancer incidence.[13, 14] The major role of ERα in formulating a therapeutic strategy is important, but 50% of ER+ tumors are resistant to certain kinds of therapy, and previously susceptible tumors can develop resistance due to ER mutation.[12] The adaptability of the disease necessitates the use of other biomarkers to create a signature for a more detailed molecular profile of the disease.

Progesterone receptor (PR), like ER, is also necessary for the normal development and function of the mammary gland. Expression of PR does not fluctuate widely like ER, but it generally has low expression levels in normal cells, and drops further after menopause. PR positive status is found in 50-70% of invasive breast cancer, and when the disease is both ER+ and PR+, there is a better chance of responsiveness to endocrine therapy.[12] PR negative status is associated with more aggressive, metastatic disease.[15]

Human epidermal growth factor 2 (HER2) is a transmembrane glycoprotein that interacts with a number of growth factor ligands. HER2 is the only HER family receptor without a known natural ligand; instead, it dimerizes with itself and other HER receptors to activate a number of signal transduction pathways.[16] Like ER and PR, it is essential for normal breast growth and development.[17] In breast cancer, HER2 does not properly regulate cell proliferation, survival, differentiation, angiogenesis, invasion, and metastasis.[18] It is overexpressed in 20-30% of breast tumors, and the presence of such overexpression is associated with a more aggressive disease, higher recurrence rate, and increased mortality. In addition to being a diagnostic and prognostic marker, it is a therapeutic target for clinically approved drugs, such as Trastuzumab (Herceptin), an anti-HER2 antibody. [19]

The four subtypes of breast cancer are based on histological characterization of ER, PR, and HER2 expression. Luminal A, luminal B, HER2 positive, and triple negative/basal-like (TNBC) disease exhibit unique characteristics and as such, these subtypes have different therapeutic options and prognostic outcomes.

Luminal A is the most commonly diagnosed subtype of breast cancer, accounting for 50-60% of all breast cancer. It is characterized as ER positive, with lower expression levels of proliferation genes. Due to the low histological grade, low mitotic activity, and relatively low relapse rate, it is considered a highly treatable disease with good prognosis. Luminal B is less prevalent, constituting 15-20% of breast cancer cases. This subtype is histologically defined as ER and HER2 positive—30% of HER2+ tumors belong to the Luminal B classification. HER2 negative tumor tissue can also be classified as Luminal B disease if paired with high Ki67 expression, which is a cellular marker for proliferation. The ER/HER2+ classification is not always consistent; up to 6% of Luminal B cancers are classified as ER-/HER2- by immunohistochemistry. Unlike Luminal A, Luminal B is less responsive to hormone therapy, but more responsive to neoadjuvant therapy. Overall, it is considered a more aggressive disease, with a higher histological grade and a worse prognosis. Luminal B also has a higher recurrence rate and lower survival rate than Luminal A.[18] These two subtypes are difficult to separate initially, as their hormone receptor expression levels can be similar. The way they behave and their responsiveness to treatment however, are very different, so it important to distinguish the two subtypes. This differentiation is where Ki67 expression can be informative—a combination of microarray analysis and IHC data has shown that a 14% Ki67 index in is enough to distinguish the two subtypes[20], as Luminal B is highly proliferative and Luminal A is not.

The third subtype is HER2 positive, which is characterized by ER-/PR-/HER2+ expression. Nearly half of all HER2 positive tumors display some ER

expression, but it is generally low. HER2+ tumors are aggressive, with an increased propensity for metastasis to visceral organs and low responsiveness to hormone therapy. HER2 positive tumors have increased sensitivity to cytotoxic therapy however, and HER2-specific treatment is available for clinical use.[18]

The last subtype of breast cancer is basal-like or "triple negative", accounting for 8-37% of all breast cancers and ~75% of BRCA1 gene related cancers.[18]. The basal-like subgroup is characterized by the expression of basal myoepithelial markers, and a lack of ER, PR, and HER2 expression. The term "basal-like" is not a clinical classification, as it is only available as a research term and it is determined by gene expression microarray. Triple negative status is determined by immunohistochemistry, which is similar to the way the other subtypes are classified. This type of cancer is very aggressive, highly proliferative, associated with invasive ductal tumors, and has a high rate of metastasis to the lungs and brain, with poor outcome.

Current Methods: Breast Cancer Detection and Characterization

Image-Based Techniques for Breast Cancer Detection

The current methods used for breast cancer detection are image-based techniques, including mammography, ultrasonography, and MRI. These methods have traditionally been useful to screen patients and guide biopsies and other surgical procedures for breast cancer therapeutics. The current standard method for breast cancer screening is mammography, which is a radiographic examination of the breast. In this process, the breast is compressed between a plastic paddle and an

X-ray detecting plate while an image of the breast is generated. Mammographic screening was more widely adopted in the 1990s, with full-field digital mammography largely replacing traditional film-screen mammography after FDA approval in 2000. Current recommendations by the U.S. Preventive Services Task Force suggest biennial screening mammograms for asymptomatic low-risk women over the age of 50.[21]

Additional methods include ultrasonography, which uses high frequency sound waves to generate an image, and is used as a method to image dense breast tissue.[22] Ultrasound has been useful in detecting occult breast cancer and early stage breast cancer that may not be detectable by mammography.[23] MRI is an another method that can be used to image breast cancer, using a combination of radio waves and a magnetic field to image tissue. Additional features can be enhanced using contrast agents. MRI evaluation of patients can be utilized for newly diagnosed patients to assess the extent of disease at the time of diagnosis, or to clarify inconclusive results from another imaging method. MRI is also used to guide biopsy or evaluate therapeutic efficacy by monitoring tumor shrinkage after treatment. Generally, MRI is not used as a screening tool unless a patient has a higher (> 20%) risk of developing breast cancer.[22, 24, 25] More recently, breast tomosynthesis has been introduced in the clinic as an improvement upon current methods. Breast tomosynthesis takes multiple images and reconstructs a three-dimensional image; this method is less prone to obstruction by dense tissue, as is the case with traditional mammography. This method has been shown to decrease recall rates and false positive rates compared to mammography alone. [26-28]

The reported clinical sensitivity and specificity of mammography is 79% and 90%, respectively, but these values are lower in women with dense breast tissue.[22] The efficacy of mammography is highly dependent on a number of factors, including tissue density, tumor size, and tumor location. The size at which a tumor is detectable by mammography can be as small as 1-3mm in diameter with < 5% efficiency, whereas detecting tumors between 4mm and 11mm can range from 15%-75% efficiency, with a median of 7.5mm tumor size for a limit of detection.[29] MRI has its own limitations, as it may not detect ductal carcinoma in situ (DCIS) as effectively as mammography. Additionally, recall rates for high-risk patients after MRI have ranged from 10% to 24%.[30] As is the case with any imaging technique, discerning between healthy and malignant masses is a challenge.

Image-based techniques are used to detect masses, calcifications, and other abnormalities in the tissue that would indicate cancer growth. Though these methods provide information for TNM staging, they have their limitations. Imaging techniques provide very little molecular information about the biology of the disease (i.e. grading and subtype). Additionally, an actual mass or evidence of one must be detectable by the imaging modality in order to further investigate the disease. All of these imaging techniques require one or more specialists in order to obtain and interpret test results, so the skill of the technician acquiring the image, as well as the expertise of the specialist interpreting the results can lead to subjectivity in the detection process.

Molecular Characterization of Breast Cancer

Molecular characterization of breast cancer requires a tumor tissue sample, which is obtained by biopsy. There are multiple biopsy methods with varying degrees of invasiveness, including fine needle biopsy, core needle biopsy, and open or surgical biopsy. These procedures may use some of the previously mentioned imaging techniques to pinpoint the location of the tumor.[31] Once the tissue has been extracted, it is usually prepared by formalin fixation and embedded in paraffin. The sample is then sliced and mounted on glass sides for the pathologist to analyze by microscopy.[32] The pathologist can use a number of histological stains to highlight the appropriate cellular features. For example, the standard H&E protocol uses hematoxylin to bind DNA, which stains the nuclei purple, and eosin binds to protein and DNA, staining other cellular structures pink.[32] This stain allows the pathologist to grade the tumor tissue by observing cellular morphology.

Immunohistochemistry (IHC) is a more advanced staining technique that utilizes tagged antibodies to bind to protein biomarkers in a tissue sample—this technique is particularly useful when assigning hormone receptor status (ER, PR, HER2). The percentage of positively stained cells (nuclei for ER and PR, cellular membranes for HER2) determines the status in a semi-quantitative manner.[32] IHC methodology varies from lab to lab, as antibodies, fixatives, and staining methods are not standardized. Because of this variation, there are different interpretations of the same image, as was shown in a study that compared 240 observations from three pathologists. They unanimously agreed on 75% of cases, but agreement scored much lower for DCIS and atypia cases.[33] Though there is a large degree of

agreement between individual pathologists, an objective quantification of these markers could mitigate some of the ambiguity associated with disease characterization and allow for a more objective, quantitative view of the disease.

Genomic Characterization of Breast Cancer

Genomic characterization of tissue samples can also be performed to gain insight into the disease. Multiple gene signatures are available for clinical use as prognostic indicators; additionally, they are useful for making treatment decisions and assessing the risk of recurrence. Three such genomic tests are Mammaprint, Oncotype Dx, and the PAM50-based Prosigna assay. The Mammaprint assay is a 70-gene fingerprint used to predict tumor recurrence. This microarray test was approved by the FDA in 2007 for use in freshly frozen tissue, and is most useful for Stage I or II disease in cases with no spread to the lymph node (node-negative). Though it does hold value in risk assessment, it also requires a 3mm diameter tissue sample obtained by surgical biopsy, and it is not currently recommended in the American Society of Clinical Oncology (ASCO) guidelines to use this array to guide adjuvant systemic therapy.[34, 35] Oncotype Dx is a 21-gene assay that utilizes qRT-PCR and microarray to predict recurrence of invasive breast cancer.[36] It has become a standard test for early stage, node-negative, ER+ breast cancer. The ASCO recommended guidelines found clinical utility in this assay to help clinicians decide whether a patient could avoid adjuvant systemic therapy. The Prosigna assay uses a 50-gene signature to classify tumor tissue into a subtype, and uses other proliferative markers to calculate the risk of recurrence. The FDA approved this assay for clinical use in 2013, and it is most useful for characterizing untreated and

tamoxifen-treated cases. Like the other gene assays, RNA is extracted from tissue, and the nCounter instrument (Nanostring) is used to directly quantify specific gene expression. When combined with other pathological factors such as tumor stage and node status, this robust assay has proven to be useful in breast cancer prognosis to guide adjuvant therapy.[37, 38]

Circulating Biomarkers

While the prevalent screening methods are imaging techniques useful for finding a potential mass, they do not provide any biological or molecular information about the disease. Biopsy and tissue staining help clinicians classify the disease, which subsequently allows for a treatment plan to be formed, but the sampling process is invasive. Circulating biomarkers can bridge the gap between noninvasive sampling and obtaining biological information about the disease. Blood tests are a common and straightforward process; additional tests could easily be incorporated into routine healthcare appointments instead of specialized procedures like mammography and biopsy. Furthermore, circulating biomarkers could potentially be used for initial diagnosis, but could also prove useful for monitoring disease progression to test for therapeutic efficacy and recurrence. The few clinically approved blood biomarkers for breast cancer are CA 15-3, CA 27.29, and CEA.[5]

CA 15-3 and CA 27.29 are both fragments of the mucin protein MUC1, a transmembrane glycoprotein. As with other mucins, MUC1 is associated with cell hydration and lubrication, inhibiting cell-cell interactions, and playing a protective role against microorganisms. MUC1 also has a cytoplasmic tail known to regulate

signal transduction pathways in the cell.[39] MUC1 is often overexpressed in breast cancer and aberrantly located fully in the cell membrane. In this case, parts of the protein are proteolytically cleaved and shed into the bloodstream.[40]

The CA 15-3 portion of MUC1 is extracellularly located and is known to be cleaved and released into the bloodstream, but the level of this shed ectodomain is associated with the progression of breast cancer in patients. Typically, expression levels of >35 U/mL are considered to be indicative of breast cancer progression or recurrence.[40] Elevated levels of CA 15-3 have been found in 60% of preoperative breast cancer patients and 80% of advanced metastatic patients. Despite the prevalence of overexpression in breast cancer patients, it is not considered a strong indicator of early stage breast cancer due to its low clinical sensitivity for stages I and II. It is better used as a recurrence indicator, and in conjunction with other biomarkers.[41, 42] CA 27.29 is another portion of MUC1 that may be more specific for cancerous conditions than CA 15-3, but not as sensitive. It is more stably expressed in healthy patients, but is not seen as any more valuable than CA 15-3 as a diagnostic or prognostic marker.[41]

Carcinoembryonic antigen (CEA) is a glycoprotein thought to be associated with adhesion, and is also used as a marker for colorectal cancer.[43] Though CEA expression can be indicative of tumor size and nodal involvement in breast cancer, it is not as valuable of a predictive biomarker when compared to CA 15-3.[41, 43, 44] Overall, there are clinically accepted circulating biomarkers are indicative of multiple cancers including breast cancer. Their individual sensitivity and

specificity however, are not sufficient for early detection on their own. A protein signature is necessary to fully utilize the value of several biomarkers.

Early Detection

Early stage breast cancer is defined as a cancer that has not spread beyond the breast and the axillary lymph nodes according to the National Cancer Institute, and there is a higher chance of survival if the disease is diagnosed at an early stage. Stage is directly related to the presence of node involvement and metastasis, so size is not the only factor to consider with prognosis.[45] The other conditions of staging are indications of how aggressive and advanced the disease has become, which is used with molecular information about the disease to make treatment decisions and formulate a prognosis. The impact of early detection on breast cancer survival has been a point of contention in the field of breast cancer research, as the benefit of early detection is debated due to overdiagnosis and subsequent overtreatment for breast cancer that was not terminal or aggressive in the first place. Improved survival rates have further been explained by the advent of newer, more effective treatments upon detection.[45-48] Early detection of breast cancer and timely biological information about the disease can help inform and better utilize available treatment options. If these aspects of detection could be applied to a relatively noninvasive sampling method, such as a blood test, the disease could be detected, characterized, and appropriately treated without unnecessary invasive methods or unnecessarily aggressive treatments.

Ultrasensitive Protein Detection

Biomarkers are defined as measurable indicators of a biological process such as a disease state. These indicators can include biomolecules such as DNA, RNA, proteins, carbohydrates, etc.[49] Using proteins as biomarkers is advantageous for several reasons, as proteins are the biomolecules that act as cellular effectors that closely reflect the biology of the disease. Differential DNA or RNA expression levels may be indicative of a risk, but may not correlate with actual protein expression or account for post-translation modifications relevant to a disease state. Proteins are also more likely to remain stable while freely circulating in serum, which allows for easier sampling and handling with little risk of degradation.

As previously mentioned, circulating biomarkers are valuable due to the ease of sampling and the ability to incorporate tests into existing protocols. The advantage of using serum over other biofluids such as urine and saliva for biomarker discovery and detection is the availability of many more proteins at a predicted range of concentrations spanning nine orders of magnitude.[50] Serum and plasma are more likely to contain proteins and other biomarkers that have been shed from a tumor, whether it was a result of tissue damage, aberrant secretion, or signaling.[51] Urine and saliva are generally less complex biofluids, but may contain only a fraction of the analytes present in blood.

Clinical screening and diagnostic processes previously discussed are dependent upon the presence of a mass and the extraction of that tissue for characterization—all of which are invasive and specialized processes. Tumors are known to secrete markers, and the smaller the mass, the fewer molecules it will secrete.[52] When diluted into the entire volume of blood in the body, aberrant protein

expression or changes in expression would be extremely low and undetectable by standard methods. Current blood testing for circulating biomarkers approved for clinical use have very little clinical sensitivity and diagnostic value in early stage breast cancer. The detection of novel low abundance biomarker proteins requires a method that can overcome the analytical sensitivity barrier encountered by current methods if it is to be useful for the early detection of breast cancer.

The single molecule array (Simoa) based assay is an ultrasensitive ELISA-based method for the detection and quantification of proteins, with limits of detection 10-1000X lower than standard ELISA. This enhanced sensitivity has provided an advantage in detecting disease state and recurrence in multiple cases, including monitoring TNF-α and IL-6 levels in patients with Crohn's disease,[53] and predicting the recurrence of prostate cancer based on ultrasensitive measurement of prostate specific antigen in patients who had undergone radical prostatectomy.[54] Simoa was also demonstrated to be a method for early detection of cancer in mice; a cohort of mice was inoculated with a low number of LNCaP cells and PSA was measured in the serum to monitor the progression of tumor growth. The levels of PSA increased as the disease progressed, but were detectable before a palpable tumor was formed, and at much lower concentrations than standard "ultrasensitive" PSA assays.[55] This proof-of-concept study demonstrated that with the appropriate biomarkers, a tumor could be detected at earlier stages than by standard methods. The Simoa platform also has the potential to detect and track biomarkers to establish healthy baselines and monitor changes.

Scope of Thesis

The aims of this project are based on the hypothesis that breast cancer tumors secrete biomarkers that can be measured by Simoa and used to detect and track the disease as it progresses from an early stage. This body of work describes efforts at creating a novel biomarker fingerprint for the presence of breast cancer based on protein concentrations in serum. The current methods for breast cancer detection and characterization are largely subjective, with very little absolute quantification. Furthermore, clinical diagnostics require multiple specialists and invasive tests, hence the appeal of a simple blood test with highly sensitive detection limits. Chapter 2 describes the theory and methodology of the single molecule array (Simoa) assay, the platform upon which ultrasensitive protein measurements have been performed in the described work. Chapter 3 outlines the biomarker selection process, as well as preliminary measurements in commercially available patient serum. Baseline protein concentrations are established, and possible correlation of each marker's expression level to stage or subtype of breast cancer is examined. Chapter 4 takes this work a step further by describing the development of multiplex assays and the measurement of eight different biomarkers in clinical serum samples. The analysis in this chapter evaluates the impact of combining all eight markers as a signature for detecting breast cancer, and further attempts to classify samples in a stage-specific or subtype-specific manner. The appendix contains cancer-staging guidelines for breast cancer, relevant information regarding patient information and additional data for Chapters 3 and 4, as well as results of cross-reactivity and recovery experiments.

References

1. American Cancer Society (American Cancer Society, Inc., Atlanta; 2015).
2. Siegel, R.L., Miller, K.D. & Jemal, A. Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians* **66**, 7-30 (2016).
3. Howlader, N. et al. based on November 2014 SEER data submission (National Cancer Institute, 2015).
4. Fackenthal, J.D. & Olopade, O.I. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat Rev Cancer* **7**, 937-948 (2007).
5. Barnard, M.E., Boeke, C.E. & Tamimi, R.M. Established breast cancer risk factors and risk of intrinsic tumor subtypes. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1856**, 73-85 (2015).
6. Irwin, M.L. et al. Influence of pre- and postdiagnosis physical activity on mortality in breast cancer survivors: the health, eating, activity, and lifestyle study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **26**, 3958-3964 (2008).
7. Edge, S.B. & American Joint Committee on Cancer. AJCC cancer staging manual, Edn. 7th. (Springer, New York; 2010).
8. , Vol. 2016 (American Cancer Society, 2016).
9. Rakha, E.A. et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res* **12**, 207 (2010).
10. Wong, E. & Rebelo, J. (McMaster Pathophysiology Review; 2012).
11. Sommer, S. & Fuqua, S.A.W. Estrogen receptor and breast cancer. *Seminars in Cancer Biology* **11**, 339-352 (2001).
12. Murphy, L.C. & Watson, P. Steroid receptors in human breast tumorigenesis and breast cancer progression. *Biomedicine & Pharmacotherapy* **56**, 65-77 (2002).
13. Khan, S.A., Rogers, M.A.M., Obando, J.A. & Tamsen, A. Estrogen receptor expression of benign breast epithelium and its association with breast cancer. *Cancer research* **54**, 993-997 (1994).
14. Lawson, J.S. et al. Low oestrogen receptor α expression in normal breast tissue underlies low breast cancer incidence in Japan. *The Lancet* **354**, 1787-1788 (1999).
15. Mote, P.A., Balleine, R.L., McGowan, E.M. & Clarke, C.L. Heterogeneity of progesterone receptors A and B expression in human endometrial glands and stroma. *Human reproduction* **15 Suppl 3**, 48-56 (2000).
16. Tai, W., Mahato, R. & Cheng, K. The role of HER2 in cancer therapy and targeted drug delivery. *Journal of Controlled Release* **146**, 264-275 (2010).
17. Yarden, Y. Biology of HER2 and Its Importance in Breast Cancer. *Oncology* **61(suppl 2)**, 1-13 (2001).
18. Yersal, O. & Barutca, S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology* **5**, 412-424 (2014).

19. Mitri, Z., Constantine, T. & O'Regan, R. The HER2 Receptor in Breast Cancer: Pathophysiology, Clinical Use, and New Advances in Therapy. *Chemotherapy Research and Practice* **2012**, 743193 (2012).
20. Cheang, M.C.U. et al. Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. *JNCI Journal of the National Cancer Institute* **101**, 736-750 (2009).
21. Siu, A.L. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation StatementScreening for Breast Cancer. *Annals of internal medicine* **164**, 279-296 (2016).
22. Smetherman, D.H. Screening, imaging, and image-guided biopsy techniques for breast cancer. *Surg Clin North Am* **93**, 309-327 (2013).
23. Brem, R.F., Lenihan, M.J., Lieberman, J. & Torrente, J. Screening Breast Ultrasound: Past, Present, and Future. *American Journal of Roentgenology* **204**, 234-240 (2015).
24. Drukteinis, J.S., Mooney, B.P., Flowers, C.I. & Gatenby, R.A. Beyond mammography: new frontiers in breast cancer screening. *Am J Med* **126**, 472-479 (2013).
25. Menezes, G.L., Knuttel, F.M., Stehouwer, B.L., Pijnappel, R.M. & van den Bosch, M.A. Magnetic resonance imaging in breast cancer: A literature review and future perspectives. *World J Clin Oncol* **5**, 61-70 (2014).
26. Friedewald, S.M., Rafferty, E.A., Rose, S.L. & et al. Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA* **311**, 2499-2507 (2014).
27. Svahn, T.M. et al. Breast tomosynthesis and digital mammography: a comparison of diagnostic accuracy. *The British Journal of Radiology* **85**, e1074-e1082 (2012).
28. Sechopoulos, I. A review of breast tomosynthesis. Part I. The image acquisition process. *Medical Physics* **40**, 014301-n/a (2013).
29. Michaelson, J. et al. Estimates of the sizes at which breast cancers become detectable on mammographic and clinical grounds. (2003).
30. Gundry, K.R. The application of breast MRI in staging and screening for breast cancer. *Oncology (Williston Park)* **19**, 159-169; discussion 170, 173-154, 177 (2005).
31. , Vol. 2016 (American Cancer Society, 2014).
32. Veta, M., Pluim, J.P., van Diest, P.J. & Viergever, M.A. Breast cancer histopathology image analysis: a review. *IEEE Trans Biomed Eng* **61**, 1400-1411 (2014).
33. Elmore, J.G. et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* **313**, 1122-1132 (2015).
34. Zoon, C.K. et al. Current molecular diagnostics of breast cancer and the potential incorporation of microRNA. *Expert Rev Mol Diagn* **9**, 455-467 (2009).
35. Harris, L.N. et al. Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline.

Journal of clinical oncology : official journal of the American Society of Clinical Oncology (2016).
36. Sparano, J.A. et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med* **373**, 2005-2014 (2015).
37. Nielsen, T.O. et al. A Comparison of PAM50 Intrinsic Subtyping with Immunohistochemistry and Clinical Prognostic Factors in Tamoxifen-Treated Estrogen Receptor–Positive Breast Cancer. *Clinical Cancer Research* **16**, 5222-5232 (2010).
38. Nielsen, T. et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. *BMC cancer* **14**, 177 (2014).
39. Brayman, M., Thathiah, A. & Carson, D.D. MUC1: A multifunctional cell surface component of reproductive tissue epithelia. *Reproductive biology and endocrinology : RB&E* **2**, 4-4 (2004).
40. Chourb, S., Mackness, B. , Farris, L. and McDonald, M. Improved detection Of the MUC1 cancer antigen CA 15-3 by ALYGNSA fluorimmunoassay. *Health* **3**, 524-528 (2011).
41. Sunwoo, H.H. & Suresh, M.R. in The Immunoassay Handbook (Fourth Edition). (ed. D. Wild) 833-856 (Elsevier, Oxford; 2013).
42. Duffy, M.J., Evoy, D. & McDermott, E.W. CA 15-3: Uses and limitation as a biomarker for breast cancer. *Clinica Chimica Acta* **411**, 1869-1874 (2010).
43. Banin Hirata, B.K. et al. Molecular Markers for Breast Cancer: Prediction on Tumor Behavior. *Disease Markers* **2014**, 12 (2014).
44. Ebeling, F.G. et al. Serum CEA and CA 15-3 as prognostic factors in primary breast cancer. *British Journal of Cancer* **86**, 1217-1222 (2002).
45. Saadatmand, S., Bretveld, R., Siesling, S. & Tilanus-Linthorst, M.M.A. Influence of tumour stage at breast cancer detection on survival in modern times: population based study in 173 797 patients. *BMJ* **351** (2015).
46. Miller, A.B. et al. Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. *BMJ* **348** (2014).
47. Jørgensen, K.J. & Gøtzsche, P.C. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* **339** (2009).
48. Welch, H.G. & Black, W.C. Overdiagnosis in Cancer. *Journal of the National Cancer Institute* (2010).
49. Strimbu, K. & Tavel, J.A. What are Biomarkers? *Current opinion in HIV and AIDS* **5**, 463-466 (2010).
50. Pieper, R. et al. The human serum proteome: Display of nearly 3700 chromatographically separated protein spots on two-dimensional electrophoresis gels and identification of 325 distinct proteins. *PROTEOMICS* **3**, 1345-1364 (2003).

51. Anderson, N.L. & Anderson, N.G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Molecular & Cellular Proteomics* **1**, 845-867 (2002).
52. Zhang, S., Garcia-D'Angeli, A., Brennan, J.P. & Huo, Q. Predicting detection limits of enzyme-linked immunosorbent assay (ELISA) and bioanalytical techniques in general. *Analyst* **139**, 439-445 (2014).
53. Song, L. et al. Single molecule measurements of tumor necrosis factor alpha and interleukin-6 in the plasma of patients with Crohn's disease. *Journal of immunological methods* **372**, 177-186 (2011).
54. Lepor, H. et al. Clinical evaluation of a novel method for the measurement of prostate-specific antigen, AccuPSATM, as a predictor of 5-year biochemical recurrence-free survival after radical prostatectomy: results of a pilot study. *BJU international* **109**, 1770-1775 (2012).
55. Schubert, S.M. et al. Ultra-sensitive protein detection via Single Molecule Arrays towards early stage cancer monitoring. *Scientific reports* **5**, 11034 (2015).

Chapter 2

Methodology

Introduction

The need and rationale for highly sensitive assays in cancer screening and diagnostics were previously described in Chapter 1. Simoa overcomes the current limitations in assay sensitivity and improves upon conventional protein detection methods, and this chapter explains the methodology used to achieve higher sensitivity. The standard method for protein detection, enzyme linked immunosorbent assay (ELISA) is explained, and this chapter describes how Simoa uses the principles of ELISA to achieve digital detection of proteins. The governing principles, experimental details, and materials of Simoa singleplex and multiplex assays are outlined. A description of the instrument used to run the automated Simoa assays is also provided.

Enzyme-Linked Immunosorbent Assays

ELISA is the current gold standard method for protein detection. The method was first introduced in the 1960's and has since become one of the most commonly used methods for protein detection and quantification.[1, 2] ELISA has been a powerful diagnostic tool, used to evaluate a variety of conditions including HIV, pregnancy, allergies, and infectious diseases. ELISA has a number of direct and indirect configurations that can be used to detect both antigens and antibodies, making it a versatile method.[3] ELISA is often combined with noninvasive sampling methods, as it has been used to analyze a variety of bodily fluids, including saliva, blood, and urine[4]

ELISA utilizes antibody-antigen interactions to capture and detect analytes of interest in a sample. Though there are various configurations available for ELISA, such as direct, indirect, and competitive formats,[5] the focus will be on the direct sandwich ELISA, pictured in Figure 2.1. In a sandwich ELISA, a capture surface is prepared by adsorbing antibodies to well surfaces in a standard polystyrene microtiter plate. This passive adsorption process allows noncovalent interactions between the antibody molecules and the surface to bind the antibodies to the wells. Unbound antibodies are then washed away, and a blocking buffer is added to occupy any vacant surface sites on the well surface, thus preventing non-specific binding of molecules in the sample to the well.

Once the capture surface has been prepared, the sample is incubated in the antibody-coated wells of the plate, shown in Figure 2.1a. The wells are then washed with buffer multiple times, removing unbound molecules and leaving only the antigen of interest capture by the antibodies. A biotin-labeled secondary detector antibody is then incubated with the capture antibody and antigen molecules in the plate; this second antibody binds to a different epitope on the antigen, forming a sandwich immunocomplex.

After another wash step, unbound detector antibody is washed away, and a streptavidin-labeled enzyme such as horseradish peroxidase or beta-galactosidase is incubated with the immunocomplex. The streptavidin on the enzyme molecule strongly binds to the biotin on the secondary antibody, labeling the immunocomplex with an enzyme molecule. Following another series of washes to remove the unbound enzyme molecules, the enzyme's substrate is added to the

plate. Signal amplification occurs as enzyme molecules turn over molecules of substrate to generate a measureable signal over time such as a chemiluminescense, fluorescence, or a color change measured by absorbance, shown in Figure 2.1b.[3] The resulting change in signal intensity can then be quantified and correlated to a concentration using a standard curve.[5] The analytical sensitivity and specificity of the sandwich ELISA has been well characterized, and reported limits of detection are typically in picomolar concentrations of protein.[6]

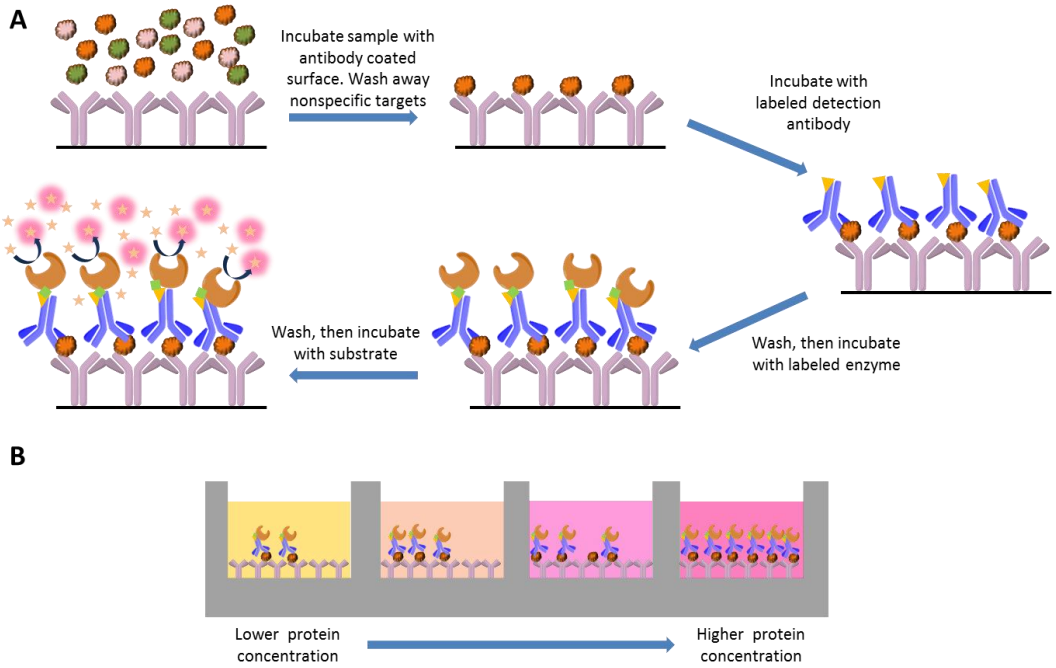


Figure 2.1. Traditional sandwich ELISA scheme: (a) depicts the capture, labeling, and detection steps of a typical ELISA with an enzymatic readout, (b) shows the color/fluorescence change as the concentration of target protein increases

Though ELISA has proven to be useful for protein detection in the past, there are two major disadvantages to this current gold standard. One disadvantage is limited sensitivity, as a 1 pM detection limit in 100 µL of sample is still 60 million molecules of analyte required in order to be measurable, and proteins present at lower concentrations would not be detectable by ELISA. Secondly, traditional ELISA cannot multiplex, or measure multiple different proteins simultaneously in

the same sample; measuring several proteins would require more sample volume to be assayed in separate wells. In both scenarios, information about the sample is lost due to either a barrier in sensitivity or a lack of multiplexing capability.

The readout in ELISA limits the sensitivity of the method because a large amount of fluorescent product must be generated in order to be detectable. The fluorescent molecules generated by a reporter enzyme at low concentrations are quickly diluted into a typical bulk reaction volume of 100 µL, and the resulting signal becomes indistinguishable from the background of the assay. Thus, millions of enzyme molecules are needed to generate enough fluorescence to overcome this dilution barrier. ELISA can also be limiting if sample volume is scarce and multiple proteins need to be assayed, since separate measurements may not be possible. An improved method would be able to measure multiple proteins at low abundance, maximizing the utility of a single sample, and allowing access to biological information that may not have been available due to limitations in sensitivity.

The Luminex platform has improved upon traditional ELISA in terms of multiplexing capabilities. This technology uses antibody-coated beads as a capture surface instead of antibody-coated microtiter plate wells, which allows for more efficient capture. Additionally, the beads are encoded with dyes as unique identifiers, and each uniquely encoded bead is used to capture a different analyte, which facilitates multiplexed protein detection. The assay process is very similar to traditional ELISA: the capture beads are incubated with sample, labeled with a biotinylated secondary antibody that is also specific to the antigen, then tagged with streptavidin-conjugated phycoerythrin (SAPE). The beads are then processed

through a flow cytometry setup where two lasers measure the fluorescence of each identifying dye in the beads, and the third measures the fluorescence intensity of the phycoerythrin fluorophores bound to the bead. The first set of measurements corresponds to the unique bead identifier, and the second measurement corresponds to the amount of analyte bound to the bead. This system of multiplexed protein detection has proven to be a valuable alternative to ELISA, as it can be more sensitive within an order of magnitude, and is less sample-depleting than standard methods. However, this technology does not address the issue of analytical sensitivity, and this is a concern that Simoa assays can address further.[7, 8]

Single Molecule Array Assays

Single Molecule Arrays (Simoa)

The basis of the Simoa technology is the microarrays used for protein detection. The microarrays in Simoa assays are composed of ~216,000 wells contained in a 3mm x 4mm area. Each well is 4.25 μm in diameter and 3.25 μm deep, with 8 μm spacing from center to center. Each well is a total of ~ 46 fL in volume, which is large enough to fit a single 2.7 μm microsphere (the capture suface in the assay) and a small volume of substrate. There are 24 arrays positioned radially in a 120 mm disc format (Sony DADC).[9] The disc is comprised of two layers: one contains the arrays, while the other contains the fluidic channels. Both parts utilize injection molding based on DVD manufacturing—the array is composed of cyclic olefin polymer (COP) and the fluidic channels are made of the same material doped with 3% carbon black, giving the disc a black appearance, shown in Figure 2.2. The two parts are then laser bonded to create a single disc.

The fluidic portion of the disc contains the inlet port, channel, and outlet, which allows the pipet to load and flow the beads, substrate, and fluorocarbon oil into the disc. The clear portion of the disc contains the arrays and is positioned closer to the imaging module below, such that the disc is loaded from the top and imaged from the bottom.

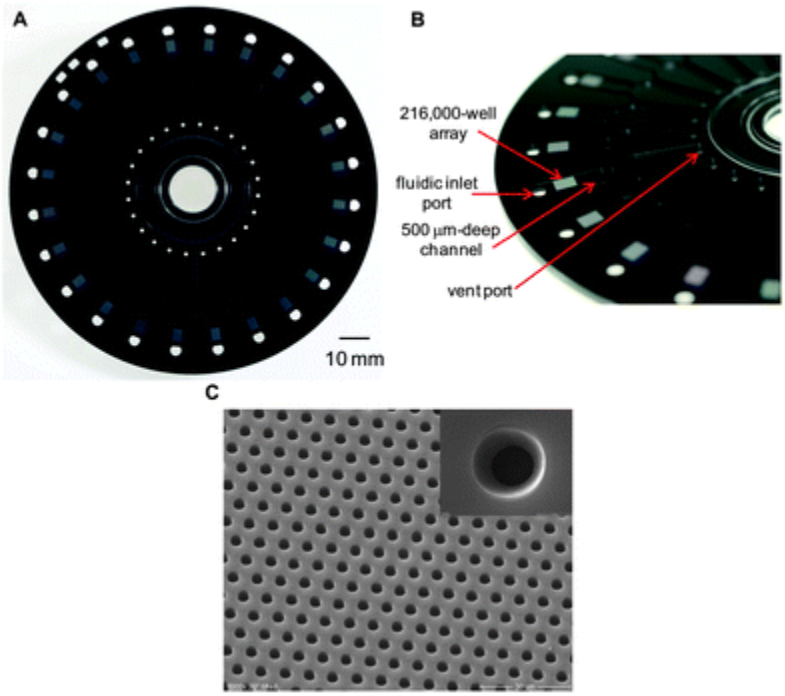


Figure 2.2 Simoa HD-1 Discs: (a) Sony DADC disc with 24 microwell arrays radially arranged, (b) a photographic close-up with the different liquid channel features labeled, and (c) is a scanning electron microscopy (SEM) image of a single microwell array on a disc, with a single well pictured on the inset. (Reprinted with permission from Reference 9)

Simoa Assays

Single Molecule Array (Simoa) assays are similar to the traditional sandwich ELISA in that antibodies are used to capture and label proteins for subsequent detection via formation of an immunocomplex and production of a

measurable signal. In Simoa assays, capture antibodies are covalently coupled to paramagnetic 2.7 μm beads and incubated with target protein in solution. There are several advantages to the bead-based platform: (1) the antibodies are covalently bound to the capture surface instead of physically adsorbed, so the antibody coating is stable during the assay, (2) the beads are stable for several months, so a single batch can be conjugated to antibody and stored for later usage, and (3) the bead suspension in solution allows for a more accessible capture surface for antigens, whereas a traditional plate ELISA is limited by the kinetics of the antigen traveling to a fixed planar surface. A biotinylated secondary detection antibody, which recognizes a different epitope on the target analyte than that of the capture antibody, is added to the solution and binds to the target analyte. After several washes, the beads are incubated with streptavidin-conjugated beta-galactosidase (SβG). The streptavidin on the enzyme binds to the biotin conjugated to the detection antibody during this incubation, and the beads are washed several times before they are re-suspended in fluorogenic substrate, resorufin-β-D-galactopyranoside. (RGP). These assay steps are illustrated in Figure 2.3.

The Simoa assay procedure diverges from standard methods when the bead and substrate suspension is taken from the reaction cuvette and loaded into the disc microarray via fluidics. Once the array is loaded, oil is used to remove excess beads that did not settle into wells, and it seals the array. Wells containing a bead with an enzyme-labeled immunocomplex build a high local concentration of the fluorescent product over time as the enzymatic reaction progresses. A series of images are then taken of the array, and the wells that contain a bead with an enzyme-labeled

immunocomplex will display fluorescence, while empty wells or beaded wells without an enzyme molecule will appear to be dark. The product generated from a single enzyme is easily detectable because the fluorescence product is contained within a 50 fL volume, which is 2 billion times smaller than the working volume of a standard ELISA.

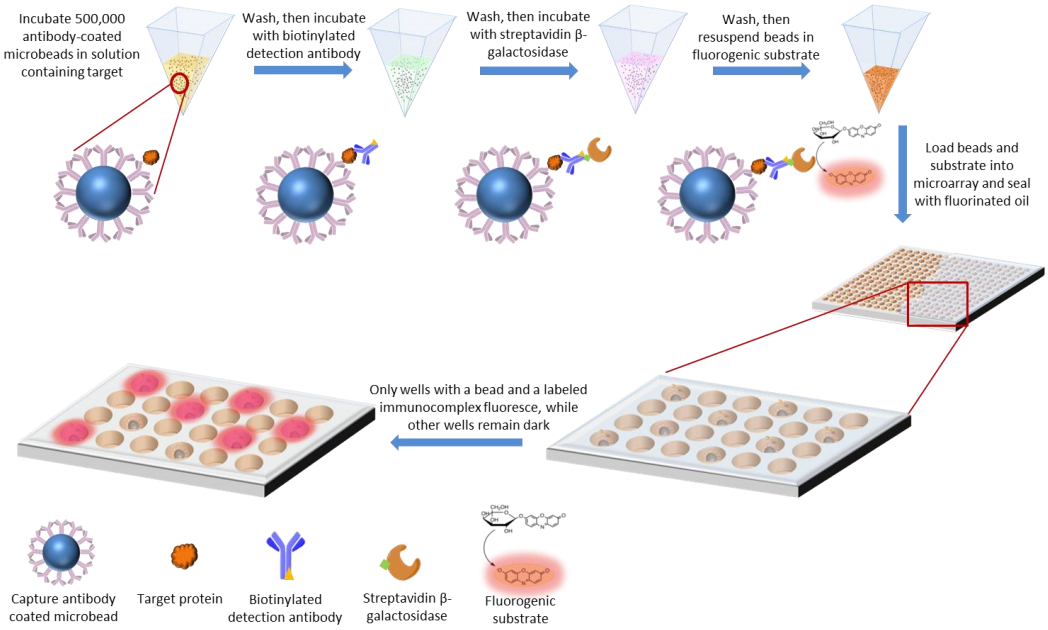


Figure 2.3. Simoa Scheme: 500,000 capture antibody coated beads are incubated with the sample, washed, and then incubated with biotinylated detection antibody. The beads are washed, then incubated with streptavidin β-galactosidase, washed, and re-suspended in RGP. This solution is loaded into the microarray, sealed with oil, and imaged. The wells that contain a bead with labeled immunocomplexes fluoresce, while other wells remain dark.

Multiplex Assays

Multiplex assays work similarly to standard Simoa assays, but dye-encoded beads are utilized to detect multiple proteins simultaneously. Each bead type or "plex" has one of four fluorescent dyes coupled to its surface. The intensity at which each bead type fluoresces under certain wavelengths becomes a unique identifier. Each encoded bead type is coated with capture antibodies to different proteins, and combined to perform a standard Simoa assay with pooled detection antibodies. The

multiplex Simoa assay is then performed the same way as a standard singleplex assay, with the same fluorescence readout in the results. The beads in the array are then decoded in the image analysis process based on the wavelength and fluorescence intensity of each bead subpopulation or plex, while the enzymatic readout provides the signal values for each individual protein assay.

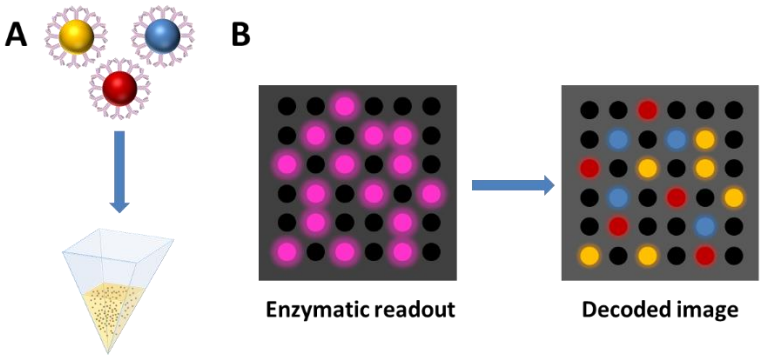


Figure 2.4. Multiplex Simoa assay scheme: (a) Beads with different antibodies are combined into a single reaction vessel to undergo a Simoa assay. Each bead represents a different dye and protein antibody conjugated to the surface, or a different "plex". (b) The illustration of the enzymatic readout shows no difference between the plexes until it is decoded, identifying which wells contained each plex.

Simoa HD-1 Automation

The Simoa assay process has been automated using the Simoa HD-1 Analyzer (Quanterix), pictured in Figure 2.5. The instrument contains separate bays for loading assay reagents and 96-well plates with samples. Once the reagents and samples are programmed and loaded, two automated pipettors in the instrument are used to distribute user-programmed volumes of reagents and samples from their respective bays into individual cuvettes, where the binding steps of the assay take place. The incubation and wash steps are performed in two rings inside the instrument. The rotating incubation ring shakes the cuvettes to keep the beads suspended in solution, allowing the capture and detection reagents to interact with

the sample. The rotating wash ring contains four wash stations and magnets that pellet the beads to the side of the cuvette, aspirates the solution, and re-disperses the beads by pipetting wash buffer directly toward the bead pellet. The wash buffer steps are all pre-programmed, and the instrument transfers sample cuvettes between the two rings based on the process needed. Following the incubation steps, the pipettor loads the beads into the disc arrays to imaged.

A major advantage provided by the automation is the high-throughput capability, with a steady-state usage capacity of 66 samples per hour.[10] Additionally, the instrument schedules sequential sample processing in 45 second "cadences" such that each sample is treated identically. Variation between replicate measurements are often below 15% CV as a result of the instrument's precise pipetting and consistent timing.

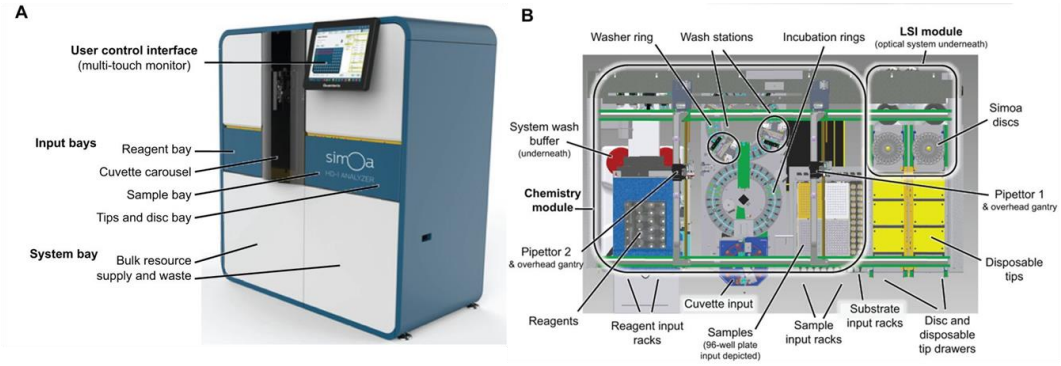


Figure 2.5 (a) is the outside of the Simoa HD-1 Analyzer and (b) shows where the assay reagents and consumables are located inside the instrument. The wash and incubation rings, as well as the imaging module are shown. (Reprinted with permission from Reference 7)

Poisson Statistics

The ability to achieve single molecule detection is explained by Poisson statistics. Poisson statistics describe the probability of a discrete number of events

occurring within a given time frame, based on a known or expected average, described in the equation below:

$$\boldsymbol{P_{\mu}(v) = e^{-\mu}\left(\frac{\mu^{v}}{v!}\right)} \quad \textbf{Eq. 1}$$

In the case of Simoa assays, μ is the expected average, or the ratio of bound immunocomplexes to beads in the assay, which is kept in excess at 500,000 beads per assay. The variable ν is a non-negative integer (e.g. 0, 1, 2) that represents the number of bound immunocomplexes or "events" in the population of beads. Therefore, the P(0) represents the subpopulation of beads with nothing bound, or "off" beads. When ν=1, 2, 3, etc., the population is known as fraction "on" (f_{on}). Another way to express the fraction of "off" beads is $1 - f_{on}$. [11]

In the case of a 1 fM protein solution in 100 µL, there are 60,000 molecules. Assuming 500,000 beads in the same solution, the µ is 0.12 and the f_{off} = $P_{(0.12)}(0)$ = 0.88. The f_{on}= (1.0-0.88)= 0.12. Using the same equation to account for only one immunocomplex bound to the bead, $P_{(1.12)}(1)$ = 0.106. The remaining 1.4% of the beads would have two or more enzyme molecules bound, theoretically. At low concentrations of protein, μ is low, and P(0) or f_{off} is high, so most of the beads in the assay would have nothing bound. Any beads in the f_{on} population would have only one immunocomplex bound, while the probability of having two or more enzymes bound to a bead is unlikely. In this scenario, one can assume each "on" bead represents a single protein molecule bound, and counting the number of "on" beads in a population equates to counting single protein molecules. The f_{on} value is then related to the measure of signal output in Simoa, the "average enzyme per bead" or AEB.[12]

Average Enzyme per Bead

The $AEB_{digital}$ is calculated by using the fraction of active beads (f_{on}) in the array according to the following equation:

$$\boldsymbol{AEB_{digital} = -ln[1 - f_{on}]} \quad \textbf{Eq. 2}$$

This is based on Poisson counting of molecules in the array, which is used to evaluate arrays with a f_{on} below 0.7. This approach is not reliable at higher f_{on} values, where Poisson statistics predict multiple enzyme molecules per bead; assuming a single molecule per bead is no longer an accurate representation of the bead occupancy. The AEB_{analog} value is used to analyze the data when the assay is not in the digital range. AEB_{analog} is calculated using Eq. 3 below:

$$\boldsymbol{AEB_{analog} = \frac{f_{on} \times \bar{I}_{bead}}{\bar{I}_{single}}} \quad \textbf{Eq. 3}$$

The f_{on} value is related back to AEB using the average fluorescence intensity of the active beads ($\bar{I}_{bead}$) in the array in conjunction with the average fluorescence intensity value generated by a single enzyme molecule ($\bar{I}_{single}$) to calculate the average number of enzymes per bead. $\bar{I}_{single}$ is calculated from arrays where f_{on}< 0.2, shown below. [11]

$$\boldsymbol{\bar{I}_{single} = \frac{f_{on} \times \bar{I}_{bead}}{-ln[1 - f_{on}]}} \quad \textbf{Eq. 4}$$

These AEB calculations can be combined into one plot, thus extending the dynamic range of an assay from digital to the analog range, spanning five logs of protein concentration.[12]

Simoa Assay Efficiency

Since the Simoa platform is relatively new, there are several steps to consider when evaluating the efficiency of an assay. Among these steps is the efficiency of the capture, labeling, and detection steps. The capture step is highly efficient because there 500,000 beads in solution coated in capture antibodies. With an estimated 274,000 antibodies per bead, the capture antibody concentration in solution is 2.3 nM.[13] This excess amount of capture antibody and its distribution throughout the sample solution provide favorable conditions for target proteins to bind to a capture surface. Unlike ELISA, where there is a planar surface to which the protein has to travel to bind, the bead-based approach is not diffusion-limited. If a protein were to break away from the capture surface, it could easily rebind to the same surface, or collide with a number of other microbeads in solution, based on the estimated 2-20 collisions per 31 seconds.[13]

The detection antibody is also added to the reaction in excess, with a final concentration of at least 0.1 μg/mL or 1 nM. The labeling efficiency is independent of the number of captured proteins, but is highly dependent upon the affinity of the antibody towards the protein. Lower affinity antibodies may require a higher working concentration to achieve a high labeling efficiency, but this can also contribute to a higher background due to nonspecific binding. Detection antibody labeling is a less stable interaction because dissociation can occur and a re-binding

event is less likely, since there is not an excess of antigen molecules available for a re-binding event. If a detection antibody dissociates, it will most likely be removed from the solution in the following wash step.

Though this step is relatively less efficient than the capture step, the automated Simoa HD-1 Analyzer process requires every sample to go through the exact timing and treatment process. The impact of the dissociation of the detection antibody would then be similar across all samples, thus removing another factor in sample-to-sample variability. Enzyme labeling of the immunocomplex is considered to be very reliable based on the strong affinity between the streptavidin enzyme label and the detection antibody's multiple biotin labels. Once bound, the enzyme is unlikely to be washed away, given the 4×10^{-14} M K_d of the biotin-streptavidin interaction.[14] The same balance between efficient labeling and minimizing background must be achieved in this step as well, as SβG is also a source of nonspecific binding.

Though Simoa has proven to be a highly sensitive method, there are still some limitations. One of these limitations is the loading efficiency of the beads onto the array. 500,000 beads are incubated in solution initially, and resuspended in 25 µL of RGP substrate—15 µL of this bead suspension are flowed into the array (~300,000 beads assuming no loss), and 20,000-25,000 beads are loaded and imaged. The sensitivity of the assay could be fully realized if more beads are interrogated. This increased bead sampling could be accomplished with lower dead volumes and improved fluidics for improved bead loading.[13]

Overall, Simoa overcomes the sensitivity limitation of traditional ELISA. The beads used in the assay contribute to favorable binding kinetics, while the microarrays allow for beads to be interrogated individually for single molecule sensitivity. Bead encoding enables the detection of multiple proteins in a single sample, and the combination of Poisson statistics with traditional analog measurements make data analysis possible over a large dynamic range.

Materials and Methods

Bead Coupling

Approximately 100 μg of capture antibody is buffer exchanged into 50 mM MES, pH 6.2 using a 50 kDa, 0.5 mL Amicon Ultra centrifugal filter unit (Millipore), per the manufacturer's instructions. The concentration of the antibody is then measured using a Nanodrop ND-1000 Spectrophotometer (Nanodrop) and the volume is adjusted to 200 μL by adding MES. The final concentration of the capture antibody in this volume typically varies between 0.3 mg/mL and 0.5 mg/mL. Beads are prepared by transferring 2.8 x 10^8 paramagnetic carboxylated beads (Quanterix) into a conical 1.7 mL microcentrifuge tube. In the case of multiplex beads, the encoded beads are purchased with various dyes already conjugated to the surface (dyes listed below in Table 2.1). The beads are washed by placing the tube on a magnetic separator, waiting for the beads to pellet, taking the supernatant out, resuspending the beads in buffer, vortexing the tube for five seconds, and centrifuging briefly. The beads are washed three times with 200 μL 1X PBS/1% Tween 20, then twice with cold MES. The final volume of the beads in MES is 190 μL. For multiplex beads, the final volume is 195 μL. The

carboxylated beads are activated with EDC, or 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (Pierce Biotechnology). After dissolving 10 mg EDC in one mL of MES (10 mg/mL final concentration), 10 µL (5 µL for multiplex beads) of the EDC solution is added to the 190 or 195 µL bead solution. The bead solution is immediately placed on a microplate shaker (IKA) at 1000 rpm for 30 minutes in order to activate the bead surface for conjugation.

Plex	Dye
488	Alexa Fluor 488 (AF488)
647	Cyanine 5 Mono Hydrazide (Cy5)
700	Cyanine 5.5 (Cy5.5)
750	HiLyte Fluor 750 Hydrazide

Table 2.1. Dyes coupled to each bead plex in Simoa multiplex assays

After the 30 minute incubation, the beads are washed with 200 µL cold MES buffer and 200 µL of antibody solution, and then added to the beads and vortexed for 10 seconds. The bead solution is then placed back on the microplate shaker at 1000 rpm for two hours. After incubation, the supernatant is aspirated and placed into a separate microcentrifuge tube. The beads are washed twice with 1X PBS/ 1% Tween 20. The first of these washes is also saved in a separate microcentrifuge tube. 200 µL blocking buffer (1X PBS/ 1% BSA) is added to the beads, vortexed for five seconds, and placed on the microplate shaker for 30 minutes at 1000 rpm. After the blocking incubation, the beads are washed three times with 1X PBS/ 1% Tween 20, and twice with Bead Diluent Buffer (50 mM Tris buffer with Proclin, Quanterix). The beads are then transferred to a clean microcentrifuge tube and stored at 4°C in bead diluent.

Bead Characterization

The antibody coupling efficiency is evaluated by measuring the antibody concentration in the saved supernatant and wash performed after the two hour coupling step. The Nanodrop ND-2000 (Nanodrop) was used to measure absorbance at 280 nm and calculate the amount of antibody in each solution. The total amount of antibody coupled to the beads is calculated by subtracting the amount of antibody in the washes from the original amount of antibody recovered from the buffer exchange.

The concentration and aggregation of the beads are characterized by using a Coulter Counter Z2 (Beckman Coulter). 10 µL of the coupled bead stock solution is pipetted into 10 mL of Zpak electrolyte buffer (Beckman Coulter) in a 15 mL Falcon tube. The tube is vortexed for 10 seconds and the solution is placed in a 20 mL cuvette (Accuvette) and placed into the instrument. Parameters are set to count particles between two and six µm. Results include particle concentration and size distribution of the particle population. The beads must be at least 80% monomeric to be of adequate quality for use in SiMoA assays.

Detector Antibody Biotinylation

Approximately 100 µg of antibody is buffer exchanged into 1X PBS using 0.5 mL Amicon Ultra centrifugal filter unit (Millipore). The concentration of antibody is measured using a Nanodrop ND-1000 instrument (Nanodrop). A single, two mg vial of EZ-Link NHS-PEG_4-Biotin, no-weigh format (Thermo Scientific) is reconstituted in water. A working dilution was made in water and added to the tube of antibody at a 20X molar excess, 2.5% by volume. The antibody and biotin

mixture is pipet mixed and allowed to incubate at room temperature for 30 minutes. After this incubation another 0.5 mL Amicon Ultra centrifugal filter unit (Millipore) is used to remove the excess unreacted biotin and buffer exchange the remaining antibody into fresh 1X PBS. The final concentration of the antibody is measured using a Nanodrop ND-1000 instrument (Nanodrop). The final product is stored at 4°C for short term storage, and -20°C for long term storage.

Reagent Preparation for HD-1 Assays

Capture antibody conjugated beads is diluted in Bead Diluent Buffer (Quanterix) to a concentration of 5 x 10^6 beads/mL in a 15 mL bottle (Quanterix). For multiplex assays (with at least three plexes), the total number of beads is increased to 6 x 10^6, but is split evenly between the number of plexes (e.g. 6 x 10^6 beads/mL divided by four plexes is 1.5 x 10^6 beads/mL). The appropriate volume was determined by multiplying the number of samples by 110 µL and adding 0.6 mL to account for dead volume in the bottle.

Biotinylated detection antibody is diluted to a working concentration in Detector & Sample Diluent (Quanterix). Two-step assays typically require a working concentration of 1 µg/mL. The appropriate volume for a two-step assay is determined by multiplying the number of samples by 35 µL and adding the 0.6 mL dead volume. Streptavidin-β-galactosidase enzyme (SBG, Quanterix) is diluted to a concentration of 100-200 pM in SBG Diluent (Quanterix), depending on the individual assay. The appropriate volume is calculated in the same manner as the capture beads. The beads are placed in a Hulamixer rotator (Thermo Scientific) at 35 rpm for 10-15 minutes to prevent the beads from settling, and all reagents bottles

are loaded into the reagent bay of the Simoa HD-1 Analyzer (Quanterix). The RGP substrate (Quanterix) is supplied by the manufacturer and used at a concentration of 100 μM.

Calibration and Sample Preparation

Calibrators are prepared by diluting protein stock into a standard diluent (1 XPBS/ 1% BSA or 25% newborn calf serum in PBS, 5mM EDTA, 0.01% Tween 20, and ProClin 300, depending on the assay) to appropriate concentrations. Serum samples are prepared by pipetting into a 96-well round-bottom plate (VWR). If the HD-1 Analyzer has been programmed to dilute the samples by a factor of four, the serum would be pipetted neat, with a volume of 25 μL per replicate plus the dead volume (75 μL previously, 30 μL after the v1.5 software upgrade). If diluting offline, 100 μL of serum would be diluted in 300 μL of appropriate sample diluent in the plate. The plate wells have a maximum volume capacity of 450 μL.

HD-1 Procedure

For a two-step assay, 100 μL of 5 X 10^6 beads/mL bead solution is pipetted into a cuvette. The cuvette is held against a magnet to pellet the beads while the bead diluent is aspirated from the cuvette. 100 μL of sample is deposited into the cuvette, as well as 20 μL of detection antibody. This first shaking incubation lasts 35 minutes, followed by three washes with System Wash Buffer 1. 100 μL of SBG enzyme solution is then added to the cuvette and allowed to incubate for five minutes, followed by five washes with System Wash Buffer 1 and one wash with System Wash Buffer 2. After this last wash, 25 μL of the RGP substrate solution is

added to the beads to resuspend them and 15 μL of this solution is loaded onto the Simoa HD-1 disc array and sealed with oil (Krytox ®, Dupont).

Image Acquisition

The HD-1 Analyzer (Quanterix) is equipped with a CCD camera that takes images of the arrays over a 45-second time period. The images are taken in different excitation/emission fluorescence channels in the following order: (1) 622/615 nm "dark field image", (2) 574/615 nm (resorufin), (3) 740/800 nm, (4) 680/720 nm, (5) 622/667 nm, (6) 574/615 nm, and (7) 490/530 nm. (1) establishes the position of the array to create a "well mask" for the other images, (2) and (4) image the fluorescence intensity of the product of the enzymatic reaction, (3)-(5) decode the identity of any dye-encoded beads, and (7) is used to decode the level of fluorescence in AF488-encoded beads and locate the position of all beads in the array. [9]

Data Analysis

The array images are analyzed and decoded based on activity and bead type. A bead is considered "on" or "active" if fluorescence intensity of the well increases above a known threshold (40 counts of fluorescence) in the 30 seconds between the first and second resorufin channel images.[9] The bead type or plex is determined by the presence and fluorescence intensity of the bead in a particular channel, which confirms bead identity.

The calibration AEB values extracted from the imaging data are fit to a four-parameter logistic curve with a $1/y^2$ weighting. The protein concentration of each sample is calculated using the fit equation of the generated curve. The LOD of the

assay is determined by adding three standard deviations to the average signal generated by the blank, and using this signal value in the curve fit equation to calculate the concentration limit.

References

1. Lequin, R.M. Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA). *Clinical chemistry* **51**, 2415-2418 (2005).
2. Zhang, S., Garcia-D'Angeli, A., Brennan, J.P. & Huo, Q. Predicting detection limits of enzyme-linked immunosorbent assay (ELISA) and bioanalytical techniques in general. *Analyst* **139**, 439-445 (2014).
3. Crowther, J. in The Elisa Guidebook, Vol. 149 9-44 (Humana Press, 2000).
4. Tighe, P.J., Ryder, R.R., Todd, I. & Fairclough, L.C. ELISA in the multiplex era: Potentials and pitfalls. *PROTEOMICS – Clinical Applications* **9**, 406-422 (2015).
5. Crowther, J. in ELISA, Vol. 42. (ed. J. Crowther) 35-61 (Humana Press, 1995).
6. Davies, C. in The Immunoassay Handbook (Fourth Edition). (ed. D. Wild) 29-59 (Elsevier, Oxford; 2013).
7. duPont, N.C., Wang, K., Wadhwa, P.D., Culhane, J.F. & Nelson, E.L. Validation and comparison of luminex multiplex cytokine analysis kits with ELISA: Determinations of a panel of nine cytokines in clinical sample culture supernatants. *Journal of Reproductive Immunology* **66**, 175-191 (2005).
8. Elshal, M.F. & McCoy, J.P. Multiplex bead array assays: Performance evaluation and comparison of sensitivity to ELISA. *Methods* **38**, 317-323 (2006).
9. Kan, C.W. et al. Isolation and detection of single molecules on paramagnetic beads using sequential fluid flows in microfabricated polymer array assemblies. *Lab on a chip* **12**, 977-985 (2012).
10. Wilson, D.H. et al. The Simoa HD-1 Analyzer: A Novel Fully Automated Digital Immunoassay Analyzer with Single-Molecule Sensitivity and Multiplexing. *J Lab Autom* (2015).
11. Rissin, D.M. et al. Simultaneous Detection of Single Molecules and Singulated Ensembles of Molecules Enables Immunoassays with Broad Dynamic Range. *Analytical Chemistry* **83**, 2279-2285 (2011).
12. Rissin, D.M. et al. Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nat Biotechnol* **28**, 595-599 (2010).
13. Chang, L. et al. Single molecule enzyme-linked immunosorbent assays: theoretical considerations. *Journal of immunological methods* **378**, 102-115 (2012).
14. Green, N.M. Avidin and streptavidin. *Methods in enzymology* **184**, 51-67 (1990).

Chapter 3

Simoa Assay Development and Preliminary Serum Tests

Introduction

The shortcomings of current breast cancer detection and characterization methods and the need for a more sensitive and less invasive diagnostic procedure were introduced in Chapter 1. Using the methodology described in Chapter 2, this work focuses on the preliminary efforts to develop ultrasensitive protein assays towards establishing a breast cancer biomarker signature. The biomarker selection process, the candidate biomarkers, and the assay development process are described below. Commercially sourced serum samples were purchased and tested in order to (1) confirm the presence of the markers in blood serum, (2) establish baseline concentrations for these markers and assess the feasibility of measuring clinical samples, and (3) evaluate any observed differences between healthy and breast cancer patients.

Biomarker Selection

Initial steps to develop a diagnostic test focused on finding potential biomarkers indicative of breast cancer. Protein biomarkers were selected using several criteria, including differential expression of the marker in early-stage breast cancer, specificity to the disease, presence and stability in serum, and the availability of biomarker-specific antibodies. Markers that were not detectable in serum by standard methods such as ELISA were of special interest, given that Simoa assays are highly sensitive. In addition to searching the literature, several collaborators and breast cancer researchers were consulted to develop a list of potential biomarkers. Of the developed assays for the protein biomarkers described

below, Dr. Marsha Moses (Harvard Medical School/ Boston Children's Hospital) suggested LCN2 and CYR61, while CA 19-9 and HIF1α were found in the literature to be indicative of breast cancer. Each marker's function and connection to breast cancer is described below.

LCN2

LCN2 is part of the lipocalin protein family, which is a group of extracellular proteins secreted in order to bind small lipophilic ligands and transport them into the cell. LCN2 was found to form a complex with matrix metalloproteinase-9 (MMP-9) in a protective capacity. It has also been linked to roles in bacteriostatic function, iron trafficking, apoptosis, and survival.[1] In breast cancer, LCN2 was found to upregulate VEGF in breast cancer cells, and was correlated with increased angiogenic activity. Urinary levels of LCN2 were also found to correlate with patients who had a poor prognosis, aggressive disease, and metastasis. This secreted marker may be useful for assessing the stage and grade of disease.[2, 3]

CYR61

CYR61 is a cysteine rich member of the CCN family of proteins, known to regulate a number of cellular functions, including cell adhesion, motility, cell division, and apoptosis.[4, 5] This secreted extracellular matrix associated protein is a ligand for integrin proteins, and a known angiogenic inducer. CYR61 has also been implicated in tumorigenesis and growth factor-driven proliferation. In breast cancer, elevated levels of CYR61 protein expression are indicative of advanced

disease. Moreover, CYR61 is thought to be associated with survival, playing a protective role in chemotherapy-induced apoptosis. In breast cancer, 70% of breast cancer patient tissue samples presenting with Stage II invasive ductal carcinoma expressed CYR61 protein levels 3.5 to 6X higher than healthy mammary controls.[4, 6] Overall, the secretion of CYR61 and its known association with more aggressive and advanced stages of disease make this protein a potentially valuable circulating biomarker for both staging and grading of breast cancer. Although CYR61 has been detected in urine and serum, the current indications for invasive breast cancer makes this marker potentially valuable in a diagnostic signature.

CA 19-9

CA 19-9 is a carbohydrate antigen related to the Lewis A blood group;[7] in cancer, it is expressed as an O-linked glycoprotein on the surface of cells, and is thought to be associated with invasion and metastasis.[8] CA 19-9 is already a prognostic biomarker for pancreatic and gastrointestinal cancers, with a normal expression level in the blood of < 37 U/mL.[9] Increased expression of CA 19-9 has also been observed in cases of colorectal cancer where lymph node invasion and metastasis have occurred. Though it is not a traditional breast cancer marker, one case study describes a patient who presented with a potential relapse with a ten-fold increase of circulating CA19-9, but had normal levels of CA 15-3 and CEA, which are clinically accepted biomarkers for breast cancer.[10] The presence of CA 19-9 in several other invasive cancers described above make it a potential candidate for evaluating the aggressiveness of the disease.

HIF1α

Hypoxia Inducible Factor-1 alpha (HIF1α) is a subunit of a heterodimeric transcription factor that regulates gene expression in multiple pathways, including glucose transport, growth factors, iron transport, and heme metabolism.[11] HIF1α levels in cells are normally in flux, as it goes through constant degradation via the ubiquitination pathway, but this process is oxygen-dependent. In hypoxic conditions, ubiquitination does not occur and HIF1α levels increase. As tumor environments are often hypoxic due to increased usage of oxygen by rapidly dividing cancer cells and limited oxygen availability from the formation of abnormal blood vessels,[12] it can lead to an increased presence of HIF1α as a tumor develops. Additionally, hypoxic tumors have been associated with increased risk of metastasis and mortality, thus HIF1α protein levels in tissue have been associated with early relapse, more aggressive disease, and metastasis.[13] This protein could serve as a valuable marker for aggressive, invasive disease if measured in serum.

Assay Development

The first step in assay development is to screen antibody pairs in a bulk bead-based assay format, using a fluorescence plate reader (Tecan) to quantify fluorescence. The bead-based platform has a two-fold purpose: (1) assess the compatibility between the antibodies and the bioconjugation chemistry, and (2) transfer beads directly to the Simoa platform if an antibody pair performs well. Once it is established that an antibody pair yields a responsive signal dependent upon target protein concentration, the assay conditions are optimized to lower the

limit of detection and maximize the signal-to-noise ratio**.** This optimization can usually be achieved by changing the duration or combining the protein capture and detection steps, changing the sample and/or calibration diluent, and adjusting detector antibody or enzyme concentrations. Figure 3.1 shows representative calibration curves for the LCN2, CYR61, CA 19-9, and HIF1α assays.

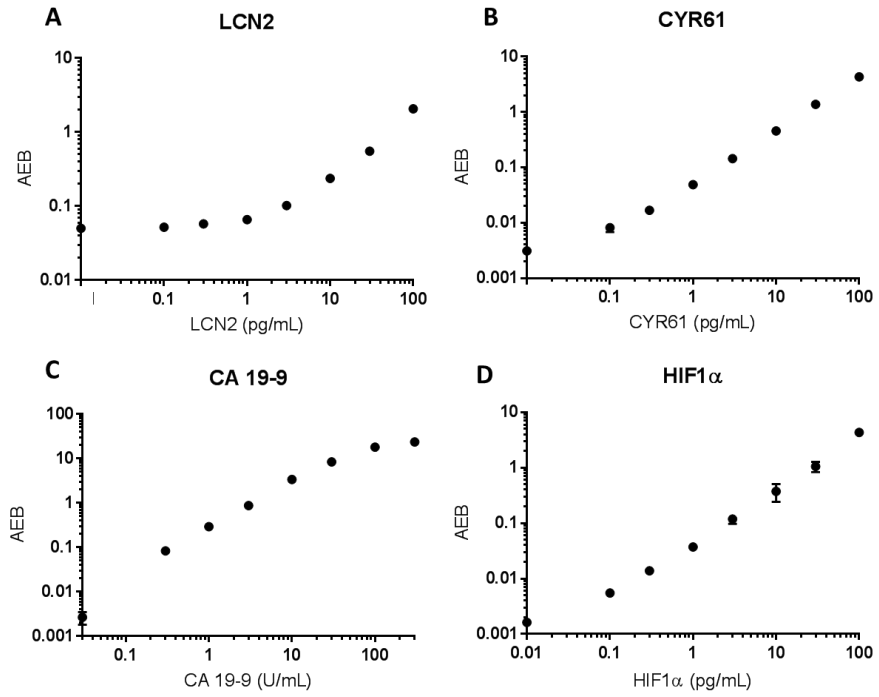


Figure 3.1. Representative calibration curves for (a) LCN2, LOD = 0.958 pg/mL (b) CYR61, LOD= 0.0191 pg/mL (c) CA 19-9, LOD = 0.0126 U/mL, and (d) HIF1α, LOD = 0.015 pg/mL. Error bars represent standard deviations of triplicate measurements.

The Simoa assays were 40-400 times more sensitive than their ELISA counterparts, shown in Table 3.1. The ELISA LOD values for LCN2 and CYR61 were provided by the manufacturer (R&D Systems). A CA 19-9 ELISA LOD was not provided by the manufacturer (Fitzgerald), since the antibody pair did not have an optimized assay procedure associated with it. The results of an assay using this pair would depend entirely on the conditions optimized by the product consumer. A kit from Thermo Fisher Scientific listed a different CA 19-9 ELISA antibody

pair's LOD as 5 U/mL, which was used for comparison. The HIFα antibody pair did not list an LOD by the manufacturer (R&D Systems), but a study that utilized the same manufacturer's ELISA kit listed the lower detection limit as 3 pg/mL.[14]

Protein Target	ELISA LOD (pg/mL)	Simoa LOD (pg/mL)	Fold Difference
LCN2	40	0.96	42X
CYR61	3.8	0.019	200X
CA 19-9	5 U/mL	0.013 U/mL	400X
HIF1α	3	0.015	200X

Table 3.1. Developed assays with ELISA LODs, Simoa LODs, and the fold difference.

Protein Concentration in Serum

After the assays were developed, they were tested in commercially available serum purchased from BioreclamationIVT. There was a total of 62 breast cancer samples and 38 healthy samples tested, although not every sample was tested by every assay due to limited sample volume. Serum sample information, including gender, age, cancer stage, and hormone receptor status, are listed in Appendix Table A1.

Serum Samples

Serum samples were diluted by a factor of four in 1X PBS/ 1% BSA to mitigate serum matrix effects, described below in the methods. All of the serum samples tested with the LCN2 assay were at concentrations above the range of the calibration curve or at high concentrations that saturated the detector in the instrument. HIF1α assays had the opposite trend, where the signal recovered from serum was equivalent to that of background seen in the calibration curve. Typically,

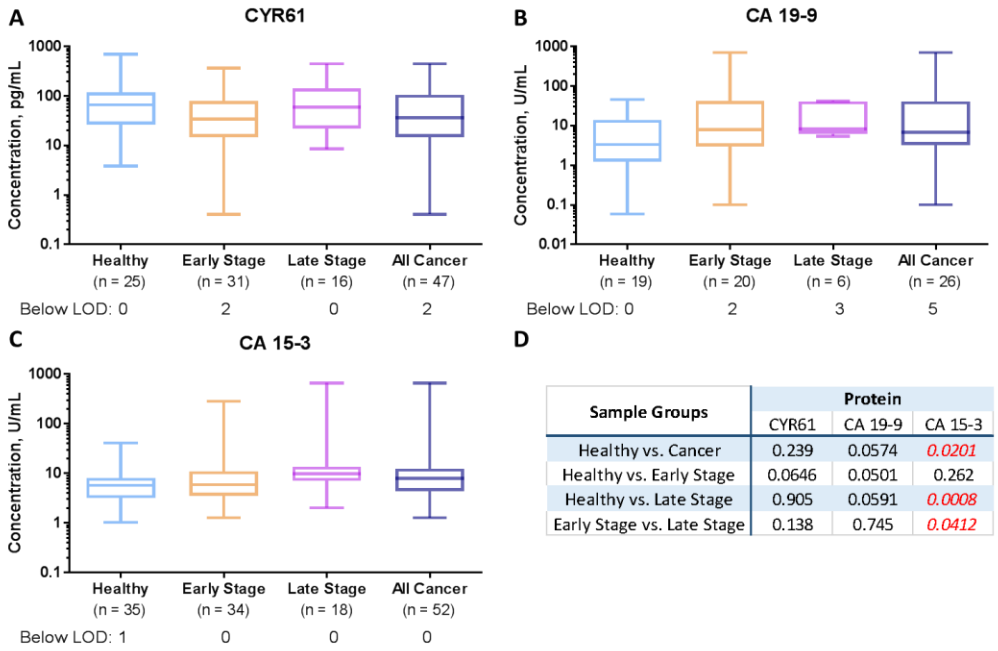
this lack of signal would indicate protein concentrations below the limit of detection, but this finding does not agree with values found in the literature, where normal, healthy patients display a detectable baseline level of HIF1α.[14, 15] The sample buffer (1X PBS/ 1% BSA) used in this assay will have to be adjusted in order to mitigate matrix effects. Adjustments to these diluents would include changing the protein content, adding animal serum, adding surfactants, and changing the salt concentration. Serum sample volumes were limited, so the same samples could not be re-tested under optimized conditions. Alternatively, the antibodies and protein standard used in this assay may not accurately represent the endogenous protein present in the serum (i.e. the protein could be truncated or folded in such a way that the antibodies cannot bind their epitopes). In this case, the assay reagents would have to be changed completely. The serum data for the HIF1α assay are shown in Appendix Table A2.

Protein Expression by Breast Cancer Stage

CYR61, CA 19-9, and CA 15-3 were all measured in healthy and breast cancer serum at detectable levels. CA 15-3, previously described in Chapter 1, was developed and assayed by Dr. Stephanie Walter. Figure 3.2 shows box plots of the serum data obtained for CYR61, CA19-9, and CA 15-3. Assays were performed in 4X diluted serum as previously described; the data in the figure accounts for this dilution. Early stage samples (Stages I and II) were grouped, while Stage III and IV samples were considered "late stage" patients. This grouping was primarily based whether the tumor was still localized to the breast and surrounding area; Stage III and beyond implies the tumor has spread beyond the breast. Sample size for each

stage and disease state, as well as the number of samples that measured below the detection limit are also denoted below the plots.

Mann-Whitney statistical analysis was used to assess the difference between the two patient groups based on the expression of a single marker.[16] Results showed only CA 15-3 was differentially expressed between healthy and breast cancer patients, healthy and late stage, and early stage and late stage groups; the p values calculated were 0.0201, 0.0008, and 0.0412 respectively, where p <0.05 denotes a significant difference between groups (Figure 3.2.d).



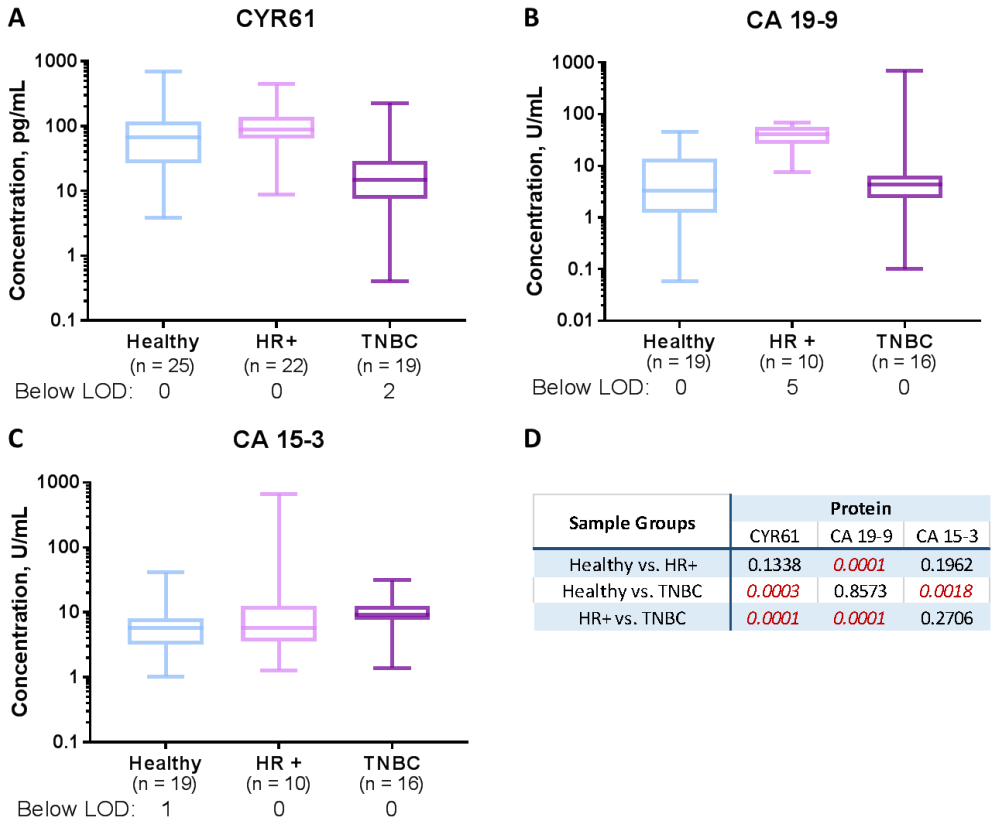
Sample Groups	Protein		
	CYR61	CA 19-9	CA 15-3
Healthy vs. Cancer	0.239	0.0574	*0.0201*
Healthy vs. Early Stage	0.0646	0.0501	0.262
Healthy vs. Late Stage	0.905	0.0591	*0.0008*
Early Stage vs. Late Stage	0.138	0.745	*0.0412*

Figure 3.2. Box plots of serum data for (a) CYR61, (b) CA 19-9, and (c) CA 15-3 by stage and disease status. Sample size n is listed below each cohort label, and the numbers of samples that were not detectable by the assay are listed below the sample size. Serum measurements were made using 4X dilutions; the concentrations listed account for this dilution factor. The table in (d) lists the Mann-Whitney statistical tests performed for each marker, comparing different stages and disease states. A p value below 0.05 indicates a significant difference between the two populations at a 95% confidence level (highlighted in red).

Protein Expression by Breast Cancer Subtype

Though stage is an important metric by which to assess the severity of the disease, subtype also has an impact on a patient's diagnosis and prognosis. As outlined in Chapter 1, the expression of hormone receptors or lack thereof helps the clinician make treatment decisions. Thus, the same data were analyzed to assess potential expression differences between different subtypes of breast cancer. ER+, PR+, and HER2+ patients (or a combination of the three), were grouped into the hormone receptor positive (HR+) cohort, while triple negative breast cancer patients were grouped together under the TNBC label. Figure 3.3 shows the serum measurements grouped by hormone receptor status as compared to a healthy cohort. Assays were performed in 4X diluted serum; the data in the figure accounts for this dilution. Mann-Whitney statistics are shown in Figure 3.3.d comparing the different groups, with a $p < 0.05$ denoting a significant difference between the two groups.

CYR61 expression appears to be differentially expressed in triple negative serum samples from both the healthy cohort and the hormone receptor positive samples with p values of 0.0003 and 0.0001, respectively. CA 19-9 serum measurements showed no significant difference between healthy and TNBC samples, but had a p value of 0.0001 when comparing the healthy and HR+ groups, as well as the TNBC and HR+ groups. CA 15-3 in contrast only showed a significant difference ($p = 0.0018$) when comparing healthy and TNBC samples.



Sample Groups	Protein		
	CYR61	CA 19-9	CA 15-3
Healthy vs. HR+	0.1338	*0.0001*	0.1962
Healthy vs. TNBC	*0.0003*	0.8573	*0.0018*
HR+ vs. TNBC	*0.0001*	*0.0001*	0.2706

Figure 3.3. Box plots of (a) CYR61, (b) CA 19-9, and (c) CA 15-3 serum measurements grouped by healthy, hormone receptor positive samples, and triple negative breast cancer. Sample size n is listed below each cohort label, and the numbers of samples that were not detectable by the assay are listed below the sample size. Serum measurements were made in 4X dilutions; the concentrations listed account for this dilution factor. The table in (d) lists the Mann-Whitney statistical tests performed for each marker, comparing different molecular subtypes. A p value below 0.05 indicates a significant difference between the two populations at a 95% confidence level (highlighted in red).

Discussion

This work outlines the development of ultrasensitive assays for LCN2, CYR61, CA 19-9, and HIF1α, which yielded LODs up to 400X lower than their ELISA counterparts. These protein markers, along with CA 15-3, were then tested in commercially available serum to (1) confirm the presence of the markers in blood serum, (2) establish baseline concentrations for these markers and assess the likelihood of measuring clinical samples, and (3) evaluate any observed differences

between healthy and breast cancer patients. By developing ultrasensitive assays, protein expression could be measured at baseline levels and early stage disease, where concentrations fall below current ELISA detection techniques, as they did with CYR61, CA 19-9, and CA 15-3. This analytical sensitivity could allow for new thresholds for disease to be established on a more sensitive basis.

Assay development was a largely successful endeavor, as the standard curves had much lower LODs than those of their established ELISAs, but not all of the assays were successfully applied to serum samples. Since many of the LCN2 measurements made in serum were above the upper limit of the assay and even saturated the detector, it was clear that a LCN2 Simoa assay was not compatible with the serum samples, and did not fit the original criteria for candidate Simoa biomarkers. The HIF1α assay was also sensitive in buffer, but it failed to detect the protein in serum. This assay will have to be optimized for serum in the future. CYR61, CA 15-3, and CA 19-9 were successfully established assays with detectable amounts of protein in serum, displaying concentrations that were below established ELISA LOD.

Each protein marker was individually evaluated in different stages and subtypes of breast cancer, with telling results. Overall, no markers outside of CA 15-3 showed any stage-specific expression. CA 15-3 expression is correlated with metastasis, a trait of late-stage cancer, so the findings agree with CA 15-3 known function in breast cancer. The other markers displayed differential expression between disease subtypes. Studies associate higher levels of CYR61 expression with more aggressive, triple negative disease subtypes, with luminal subtypes

(HR+) showing lower expression of CYR61 in tissue, but this study's findings show the opposite trend, likely due to an external factor such as patient age, as the different populations were not age-matched (Figure 3.3.a). [17] This work tested a relatively small group with unpaired samples, so data does not reflect expression changes in the general population. CA 19-9 showed higher expression in HR+ cancers, but 5 of the 15 samples were below the LOD, which would change the degree of separation between the HR+ samples and the other subtypes.

Overall, it appears that each of these markers could potentially contribute to a diagnostic fingerprint for breast cancer, based on their individual expression patterns. Additional sample volume would be required in order to evaluate the contribution of the three markers as a standalone molecular signature. Nonetheless, this work successfully establishes the need for ultrasensitive assays for CYR61, CA 15-3, and CA 19-9 in serum.

Future Directions

Though initial progress was made to identify potential biomarkers, develop ultrasensitive assays to measure them in serum, and assess their potential for identifying breast cancer, there are a number of ways to further improve the work. On the assay development, HIF1α is a potentially valuable biomarker that is not currently validated for serum measurements due to poor recovery. This recovery will have to be modified in order to accurately determine serum concentrations. Additionally, all of the measurements made in this work were done in singleplex

assays. Multiplex assays will be developed to detect multiple proteins simultaneously, saving time and conserving precious sample volume.

The serum that was used in this work was commercially obtained and based entirely on donors who self-reported their condition at the time of diagnosis, not at the time of donation. The sample information provided indicated that many of these patients had already undergone some form of therapeutic intervention (i.e. chemotherapy, radiation, or surgery), so it is likely that the disease had receded to some degree at the time the samples were collected. For some of these patients, certain biomarkers could have returned to their individual baseline levels, so the self-reported diagnosis would not have been an accurate representation of the patient's disease state at the time of sampling.

Clinical samples that have not yet undergone treatment are a more suitable sample set to measure in order to accurately report biomarker levels throughout the progression of breast disease. Furthermore, it would also be useful to obtain time point samples from individuals in order to determine baseline levels of different biomarkers and observe the changes in expression as an individual patient undergoes treatment. The measurements made in these samples better represent current biomarker levels throughout disease progression on a personalized basis. Clinical serum samples collected at Tufts Medical Center will account for these conditions, and the preliminary results from commercial serum show that Simoa is a valuable tool for measuring these biomarkers.

Future data analysis will be done by accounting for all biomarker concentrations simultaneously to evaluate a protein signature. Multivariate analysis

can build a predictive model for identifying and stratifying breast cancer. This type of analysis could not be performed on the collected data because there was not enough sample overlap between the biomarkers to build any models. Sample data collected in the future will be a more complete dataset to determine whether the combination of CYR61, CA 19-9, CA 15-3, and other markers correlates with disease state.

Materials and Methods

All serum samples were purchased from BioreclamationIVT and diluted by a factor of four in PBS (Thermo Fisher AM9625) and 1% BSA (Millipore 820451). All assays were performed in a two-step format as previously described unless otherwise indicated.

LCN2 antibodies and standard were purchased from R&D Systems (MAB17571, AF1757, 1757-LC-050). The standard was reconstituted in 1 X PBS/ 1% BSA. The capture antibody was coupled to the magnetic beads and the detection antibody was biotinylated as previously described in Chapter 2. The detection antibody was kept at a concentration of 1 μg/mL and SβG at a concentration of 100 pM.

CYR61 antibodies and standard were purchased from R&D Systems (MAB4055, AF6009, 4055-CR). The standard was reconstituted and diluted in 1 X PBS/ 1% BSA for the standard curve. The capture antibody was coupled to the magnetic beads and the detection antibody was biotinylated as previously

described. The detection antibody was kept at a concentration of 1 μg/mL and SβG at a concentration of 100 pM.

CA 19-9 antibodies and standard were purchased from Fitzgerald, Inc (10-CA9B, 10-CA19A, 30-AC14S). The standard was diluted in 1 X PBS/ 1% BSA for the calibration curve. The capture antibody was coupled to the magnetic beads and the detection antibody was biotinylated as previously described. The detection antibody was kept at a concentration of 2 μg/mL and SβG at a concentration of 100 pM.

HIF1α antibodies were purchased from R&D Systems (MAB1935 and AF1935), and the standard was purchased from Abcam (ab48734). The standard was reconstituted in 1 X PBS/ 1% BSA for the calibration curve. The capture antibody was coupled to the magnetic beads and the detection antibody was biotinylated as previously described. The detection antibody was kept at a concentration of 1 μg/mL and SβG at a concentration of 100 pM.

CA 15-3 antibodies and standard were purchased from Fitzgerald, Inc. (10-CA15C, 10-CA15B, and 30-AC17). The standard was reconstituted in 1X PBS/ 1% BSA for the calibration curve. The beads were coupled with capture antibody, and the detection antibody was biotinylated as previously described. The assay was run in a three-step format with a 15 minute capture step, a five minute detection step with 0.5 μg/mL detection antibody, and a five minute SβG incubation with 150 pM enzyme.

Assay LOD was determined by taking three standards of deviation above the average blank signal and fitting the value to the calculated calibration curve. All calibration curves were fit with four parameter logistic (4PL) equations by the Simoa HD-1 software (Quanterix). The Mann-Whitney statistical tests were performed using Prism 7 (Graphpad).

References

1. Yang, J. & Moses, M.A. Lipocalin 2: a multifaceted modulator of human cancer. *Cell cycle* **8**, 2347-2352 (2009).
2. Yang, J. et al. Lipocalin 2 promotes breast cancer progression. *Proceedings of the National Academy of Sciences* **106**, 3913-3918 (2009).
3. Yang, J., McNeish, B., Butterfield, C. & Moses, M.A. Lipocalin 2 is a novel regulator of angiogenesis in human breast cancer. *FASEB J* **27**, 45-50 (2013).
4. Sampath, D., Winneker, R.C. & Zhang, Z. Cyr61, a member of the CCN family, is required for MCF-7 cell proliferation: regulation by 17beta-estradiol and overexpression in human breast cancer. *Endocrinology* **142**, 2540-2548 (2001).
5. Tan, T.W. et al. Cyr61 increases migration and MMP-13 expression via alphavbeta3 integrin, FAK, ERK and AP-1-dependent pathway in human chondrosarcoma cells. *Carcinogenesis* **30**, 258-268 (2009).
6. Menendez, J.A., Mehmi, I., Griggs, D.W. & Lupu, R. The angiogenic factor CYR61 in breast cancer: molecular pathology and therapeutic perspectives. *Endocrine-Related Cancer* **10**, 141-152 (2003).
7. Pall, M., Iqbal, J., Singh, S.K. & Rana, S.V. CA 19-9 as a serum marker in urothelial carcinoma. *Urology annals* **4**, 98 (2012).
8. Ballehaninna, U.K. & Chamberlain, R.S. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *Journal of gastrointestinal oncology* **3**, 105-119 (2012).
9. Galli, C., Basso, D. & Plebani, M. CA 19-9: handle with care. *Clinical chemistry and laboratory medicine : CCLM / FESCC* **51**, 1369-1383 (2013).
10. Papantoniou, V. et al. Increased serum carbohydrate antigen 19–9 in relapsed ductal breast carcinoma. *Hell J Nucl Med* **9**, 36-38 (2006).
11. Bos, R. et al. Levels of hypoxia-inducible factor-1 alpha during breast carcinogenesis. *Journal of the National Cancer Institute* **93**, 309-314 (2001).
12. Gilkes, D.M. & Semenza, G.L. Role of hypoxia-inducible factors in breast cancer metastasis. *Future oncology (London, England)* **9**, 1623-1636 (2013).
13. Semenza, G.L. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* **3**, 721-732 (2003).
14. Hamed, E.A., El-Abaseri, T.B., Mohamed, A.O., Ahmed, A.R. & El-Metwally, T.H. Hypoxia and oxidative stress markers in pediatric patients undergoing hemodialysis: cross section study. *BMC Nephrology* **13**, 1-7 (2012).
15. Jia, Z.Z., Jiang, G.M. & Feng, Y.L. Serum HIF-1alpha and VEGF levels pre- and post-TACE in patients with primary liver cancer. *Chin Med Sci J* **26**, 158-162 (2011).
16. Elliott, A. in Statistical analysis quick reference guidebook. (ed. W. Woodward) 191-207 (SAGE Publications, Inc., 2007).
17. Sánchez-Bailón, M.P. et al. Cyr61 as mediator of Src signaling in triple negative breast cancer cells. *Oncotarget* **6**, 13520-13538 (2015).

Chapter 4

Multiplex Assay Development and Clinical Sample Testing

Introduction

Current approaches to breast cancer screening, as previously discussed in Chapter 1, outlined how mammography and other imaging techniques have proven to be limited in clinical sensitivity and specificity.[1] Furthermore, invasive biopsy techniques are required to characterize the disease biologically and formulate a treatment plan. Circulating biomarkers were introduced as a relatively noninvasive way to screen and track disease progression in breast cancer, but current clinically approved biomarkers, CA 15-3, CEA, and CA 27.29, are not sufficiently sensitive or specific individually to be used for this purpose.[2-5]

Early stage tumors are known to secrete small amounts of biomarkers into the bloodstream,[6] and high analytical sensitivity is necessary to detect these markers. Simoa assays were demonstrated to have high analytical sensitivity for low abundance proteins. This sensitivity was shown to be advantageous in studies measuring PSA to track tumor growth and recurrence in prostate cancer, as these early changes were only detectable at levels below the LOD of current "ultrasensitive" tests.[7, 8] The approach to the prostate cancer work was extended to breast cancer biomarkers in Chapter 3, where ultrasensitive singleplex protein assays were developed and tested in commercially available serum. These assays were 40-400X more sensitive than their bulk ELISA counterparts, and this sensitivity was found to be necessary, as some samples read below the LOD of standard ELISA. Individual biomarkers were then statistically evaluated to assess differences between healthy and breast cancer populations.

This chapter describes the development of several multiplex assays from previously established singleplex assays. These assays are intended to detecting multiple proteins simultaneously in limited volumes of sample. The proteins used for this signature include ER, PR, CYR61, CDKN2D, CA 15-3, CA 19-9, HER2, and VEGF, which are associated with various aspects of breast cancer, including proliferation, invasiveness, and disease progression.[3, 9-12] Clinical serum samples obtained from Tufts Medical Center were tested with these assays and the resulting measurements were compared to commercially available healthy controls. Each marker was individually evaluated for significant differences between subgroups within healthy and breast cancer cohorts. Protein concentrations and patient age were then used as inputs for a supervised classification technique to assess all of the markers as a signature to discriminate between healthy samples and different groups of breast cancer patients. The development of these predictive models, the results of their classifications, and implications for their potential diagnostic utility are described in this chapter.

Multiplex Assay Development

Biomarkers

The biomarker signature tested in this body of work includes two additional markers not previously discussed: VEGF and CDKN2D. VEGF, or vascular endothelial growth factor, is an angiogenic cytokine that is overexpressed in breast cancer tissue. This marker is associated with the presence of solid tumors in other cancers such as brain, lung, and ovarian cancer.[13] Elevated levels of circulating VEGF have been found metastastic breast cancer, making this marker a potentially

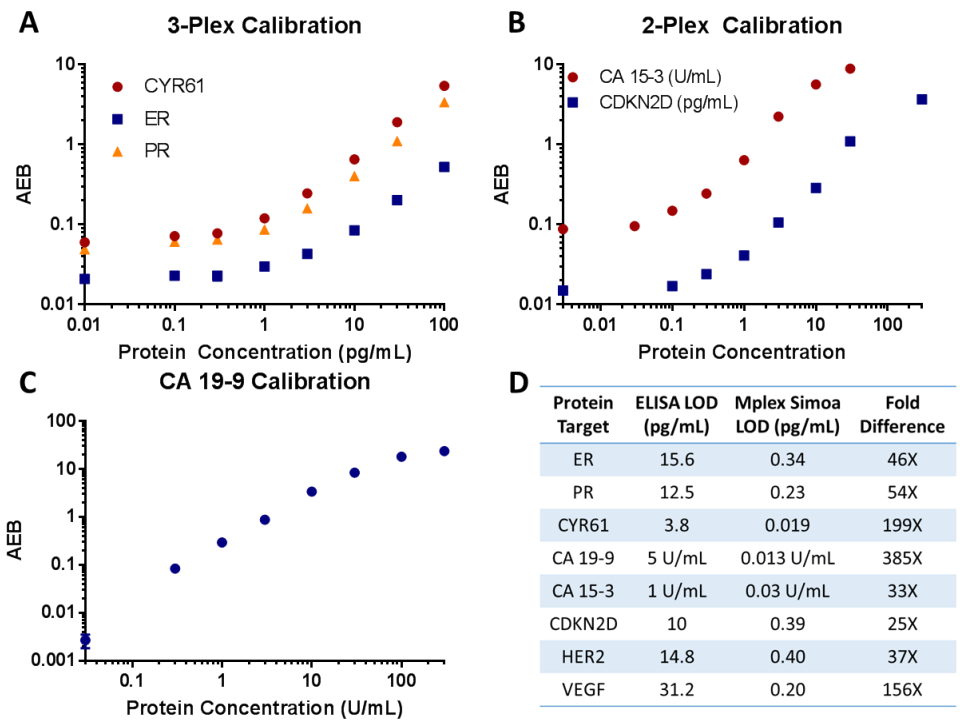
valuable indicator of aggressive disease. Cyclin-dependent kinase inhibitor 2D (CDKN2D) is a cell cycle inhibitor that acts on CDK 4 and 6 kinases to prevent them interacting with Cyclin D. Cyclin D drives the transition between G1 and S phase in the cell cycle, which is when DNA replication occurs. Elevated tissue levels of this protein have been associated with ovarian cancer, as well as other cancers.[14, 15] In addition to cell cycle regulation, this protein has been found to respond to genotoxic stress, facilitating DNA repair. Though it is unclear what role CDKN2D plays in tumorigenesis, it was overexpressed in commercially available breast cancer serum, and was an influential marker in a three-protein signature indicative of breast cancer.[16]

Multiplex assays were developed and optimized for ER, PR, CDKN2D, CA 15-3, and CYR61 using their singleplex assays as a guide. The singleplex Simoa assays for ER, PR, and CDKN2D were developed by Dr. Stephanie Schubert, and the CA 15-3 assay was developed by Dr. Stephanie Walter. There were three multiplex assays developed: one was a 3-plex assay measuring ER, PR, and CYR61 simultaneously, another was a 2-plex assay for CA 15-3 and CDKN2D, and the last was for HER2 and VEGF (developed and validated by Dr. Liangxia Xie).

Multiplex Assay Results

The multiplex assays were developed using established singleplex assays that had been tested in commercially available serum. These assays were combined and initially tested for cross-reactivity to assess whether the proteins, detection antibodies, or dye-encoded beads resulted in any biochemical crosstalk between

assays. The cross-reactivity experiments are described in Appendix A4. The calibration curves were then evaluated by evaluating the background of the assay, the signal-to-noise ratio, limit of detection, and the dynamic range of the assay as compared to expected values in serum found in previously obtained singleplex results. The multiplex calibration curves and LODs are shown in Figure 4.1. The LODs for the three-plex proteins ER, PR, and CYR61 were 0.34 pg/mL, 0.23 pg/mL, and 0.019 pg/mL respectively. The two-plex assay LODs were 0.39 pg/mL for CDKN2D and 0.03 U/mL for CA 15-3, and the singleplex CA 19-9 assay LOD was 0.013 U/mL. The remaining two-plex assay for HER2 and VEGF had LODs of 0.40 pg/mL and 0.20 pg/mL, respectively. There was some loss of sensitivity and an increase in background for these assays when transitioning from singleplex to multiplex format, but the sensitivity appeared to be sufficient for these assays based on previous measurements in commercially available serum. The reagents and conditions used in these assays are described in the Materials and Methods section.



Protein Target	ELISA LOD (pg/mL)	Mplex Simoa LOD (pg/mL)	Fold Difference
ER	15.6	0.34	46X
PR	12.5	0.23	54X
CYR61	3.8	0.019	199X
CA 19-9	5 U/mL	0.013 U/mL	385X
CA 15-3	1 U/mL	0.03 U/mL	33X
CDKN2D	10	0.39	25X
HER2	14.8	0.40	37X
VEGF	31.2	0.20	156X

Figure 4.1. Calibration curves for (a) three-plex assay for ER, PR, and CYR61, (b) two-plex assays for CA 15-3 and CDKN2D, and (c) singleplex CA 19-9. Error bars depict standard deviation of triplicate measurements. Error bars smaller than the marker are not visible. The table in (d) lists the ELISA and multiplex Simoa LODs for each measured protein and the fold difference between the two.

Once the calibration curves were established, several experiments were performed to optimize the assay for protein detection in serum samples. Spike and recovery experiments were performed in commercially available healthy serum (BioreclamationIVT) to determine several factors for serum sample testing: (1) determine which sample and calibration diluents were compatible, (2) choose the appropriate dilution to maximize protein recovery and mitigate matrix effects, and (3) ensure that measured concentrations would fall within the assay's dynamic range. These experiments and their results are described in Appendix A5. The assay reagents did not cross-react significantly in multiplex format, and the spike and recovery experiments supported the use of a 25% newborn calf serum, PBS-based diluent for both calibrators and serum samples.

Serum Measurement Results

There were two sets of serum samples tested, one healthy, and one breast cancer. The healthy serum samples tested in this study were obtained from BioreclamationIVT from women who self-reported as healthy. Subject age information is provided in Appendix A6. For breast cancer serum samples, a prospective study was initiated at Tufts Medical Center. Patients were screened and diagnosed with breast cancer via standard methods, including mammography followed by biopsy. Patients who had not undergone surgical and/or therapeutic intervention were eligible. Eligible patients consented to blood donation for the study upon a positive diagnosis, and the serum was stored at -80°C for later analysis (IRB Study #1103017). AJCC staging, 7th edition was used for categorizing stages of breast cancer. Information provided at the time of serum testing included patient age, date of sampling, and information regarding biopsy results where available. Complete cancer stage and subtype information were provided after the serum testing was complete in order to perform a supervised classification analysis. The donors were all female, with an age range of 32-53 years. The breast cancer group (n=66), comprised of women ages 37-90 years old who had been diagnosed with breast cancer, but had not undergone any therapeutic intervention at the time of blood donation. The majority of these patients were diagnosed at Stage I or II with hormone receptor-positive cancer, with eight patients in Stage III-IV disease and eight triple negative breast cancer (TNBC) patients. Details about both sample cohorts are provided in Appendix A6.

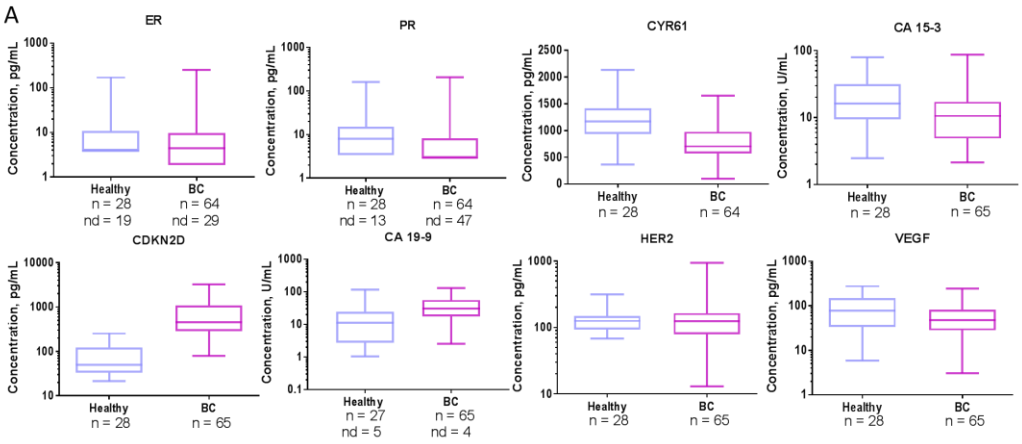
The breast cancer and healthy serum samples were tested for eight different markers, with results for both cohorts shown in Figure 4.2. The box plots reflect the samples above the calculated limit of detection for the assay, and the concentrations listed account for the sample dilution factor. ER and PR had the lowest number of detectable samples out of all the markers, suggesting that optimization can be performed to improve the assay LOD. The other six assays were sufficiently sensitive to measure serum levels of these proteins.

Although these markers were expected to be overexpressed in the serum of breast cancer patients,[9, 13] CYR61, CA 15-3, and VEGF were shown to be significantly lower in breast cancer patients (Figure 4.2.i.), likely due the age-dependent expression of these markers and the significant age difference between the healthy and breast cancer cohorts (Appendix A6). CA 15-3 was already described to be a poor individual marker of early stage breast cancer; given that 58 out of the 66 samples are Stage II or earlier, it is not unexpected that only 25% of the detectable samples were above the healthy 35 U/mL threshold.[2, 3] These markers may prove to be more informative as part of a larger biomarker signature, rather than individual markers.

The table in Figure 4.2 compares individual marker expression in further stratified groups of breast cancer patients using Mann-Whitney statistics.[17] Healthy expression levels were compared to "early stage" or Stage 0 through Stage II patients, where ER, CYR61, CDKN2D, CA 15-3, and CA 19-9 were shown to be differentially expressed. This appears to be promising for a predictive signature for early-stage cancer. Subtype-specific expression was also examined to compare

hormone receptor (HR+) cancer and TNBC samples. CDKN2D appeared to be the only marker differentially expressed, but there were only a total of eight TNBC samples in the breast cancer group. A larger sample size in that subgroup may provide more insight on subtypes, as there is not enough information to draw conclusions about indications of TNBC.

A



B

Sample Groups	Protein							
	ER	PR	CYR61	CA 15-3	CDKN2D	CA 19-9	HER2	VEGF
Healthy vs. All Cancer	0.213	***<0.0001***	***<0.0001***	***0.0068***	***<0.0001***	***< 0.0001***	0.820	***0.0418***
Healthy vs. Early Stage	0.0759	***<0.0001***	***<0.0001***	***0.0018***	***<0.0001***	***0.0002***	0.968	0.0580
HR+ vs. TNBC	0.925	0.978	0.307	0.8323	***0.0181***	0.436	0.593	0.394

Figure 4.2. Concentrations of different protein biomarkers in healthy and breast cancer serum. Each graph reflects measurements above the limit of detection, with the sample size listed below each group. Samples below detection limit are listed as "nd". Listed concentrations account for the assay's dilution factor. The markers shown are (a) ER, (b) PR, (c) CYR61, (d) CA 15-3, (e) CDKN2D, (f) CA 19-9, (g) HER2, and (h) VEGF. The table in (i) shows the p value results from Mann-Whitney statistical analysis of serum protein concentrations in healthy, all breast cancer, early stage (Stage 0, I, II), hormone receptor positive (HR+), and triple negative breast cancer samples. Bolded values indicate a significant difference between the two groups, with a p value <0.05.

Based on the protein concentrations of individual biomarkers, ER and PR are generally at relatively low levels in serum, but there was not enough sample information provided to determine whether the level of expression would be indicative of disease state. HER2 did not display any statistical difference between subgroups of patients; it is not anticipated to contribute to a biomarker signature for

this patient population. Furthermore, there were only three HER2 positive patients in the breast cancer population, so these patients were not taken into account when examining subtype groups. Individual statistical analyses of CYR61, CA 15-3, CDKN2D, and CA 19-9 show these markers to be the most promising candidates for early breast cancer detection.

Multivariate Classification of Serum Samples

Multivariate analysis was used to evaluate all eight markers simultaneously as a signature for the identification and stratification of breast cancer serum samples. Several multivariate methods are available for analyzing complicated data, with the aim of either clustering or classifying data based on the given variables. Clustering is an unsupervised technique, which is an unbiased approach to grouping data without including corresponding class group for each data point. Classification is a supervised technique that uses the sample class as an input to train the algorithm, which can then be validated and utilized on unknown samples.[18]

Partial Least Squares- Discriminant Analysis (PLS-DA)

Principal Component Analysis (PCA) is an unsupervised technique that was used to reduce the dimensionality of the dataset to facilitate a simpler analysis without losing important information. In this analysis, principal components of the data are identified in which the data has maximum variance, and allows for the data to be visualized differently and identify any underlying structure.[19,20] Data imputation was also implemented by the software (PLS Toolbox, Eigenvector, Inc.) during the PCA process, where missing data points were replaced using the current

model as a template. Seven data points were imputed in a 94 X 9 matrix consisting of 94 samples, breast cancer and healthy combined, and input variables (eight protein marker concentrations and patient age). Serum measurements that fell below the LOD for an assay were assigned a value at half the LOD, and accounted for the dilution factor of the assay. These assigned values allowed for serum samples below the LOD to be included in the model with minimal bias.

Once the principal components were identified, PLS-DA (Partial Least Squares-Discriminant Analysis) was used to find latent variables, which are linear combinations of the original variables, to maximize covariance between y-variables. This analysis method is common in chemometrics,[21] and has been used for the interpretation of complex microarray data[22] and metabolomics data[23, 24] toward predicting diagnosis and clinical outcome of a disease state. PLS-DA in this case was used to assign samples to a class (such as "healthy" or "breast cancer") based on the input variables.

Before PCA and PLS analysis were performed, the data was treated in order to account for a large spread in data; for this dataset, the data was autoscaled. The cross-validation method used 80% of the sample set for calibration, while the remaining 20% was used for validation. The process was repeated until the entire sample set had been used for validation. The autoscaling, data imputation, and cross validation procedures were executed for each of the models tested, which are described below.

Predictive Modeling Results

The first model tested compared healthy samples and all breast cancer samples. A graphical representation of the PCA is shown in Figure 4.3, showing scores along PC1 and PC2. The healthy samples are shown in green and the red samples are all of the breast cancer samples. The 95% confidence level is shown as a dotted circle. As this method is unsupervised, the group classification was not taken into account creating this model—the clustering was entirely based upon protein expression and ages of the patients. Though most of the variance is contained in PC1 by definition (23%), the separation between the two groups primarily occurs along the PC2 axis. There is some overlap between the breast cancer and healthy groups, but each group clusters differently. PCA was also used to examine outliers, but eliminating these samples did not make a significant difference in downstream analysis, thus all samples were retained for the prediction model.

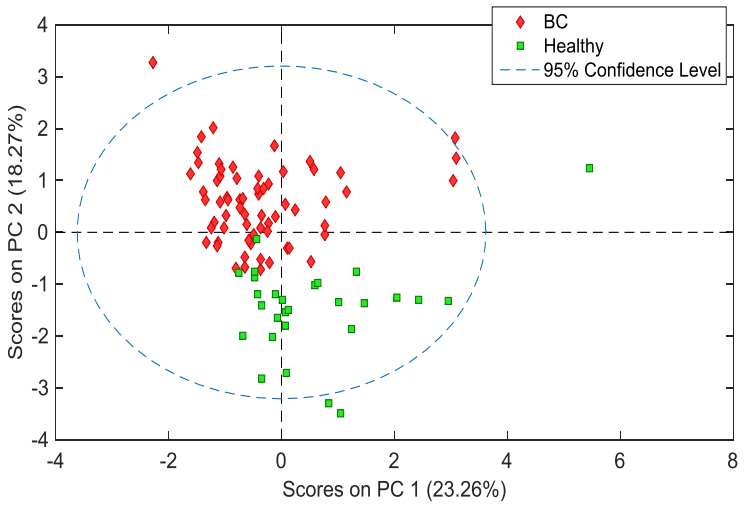


Figure 4.3. Principal Component Analysis biplot of scores along PC1 and PC2. All breast cancer samples are shown in red, and healthy samples are shown in green. The 95% confidence level is shown as a dotted line.

There are several ways to describe a classification model. Sensitivity, specificity, and precision (also known as positive predictive value) are metrics by which models can be assessed. Sensitivity is defined as the number of true positives divided by the number of true positives and false negatives, describing the ability of the model to detect or classify all of the true positive samples in a group.[25] In the case of healthy vs breast cancer samples, breast cancer is considered the "positive". Sensitivity is a similar metric that evaluates the model's performance in identifying all of the "negative" (or in this case healthy) samples. Precision, or positive predictive value takes true positives and false positives into account by evaluating what percentage of positive classifications predicted by the model were accurate.[26] Overall accuracy is defined by the number of correct classifications (true positives and true negatives) divided by the total number of samples.[26]

PLS-DA was performed on all serum samples, with seven imputed data points and 122 undetectable measurements replaced with new values. Figure 4.4 shows the classification results from the first model, comparing all breast cancer and healthy samples. Breast cancer samples are shown in red, healthy samples are shown in green, and the line of discrimination between the two classes based on the model is shown as a red dotted line. Three of the 66 breast cancer samples were misclassified as healthy in the cross validation, and one of the 28 healthy samples was classified as breast cancer.

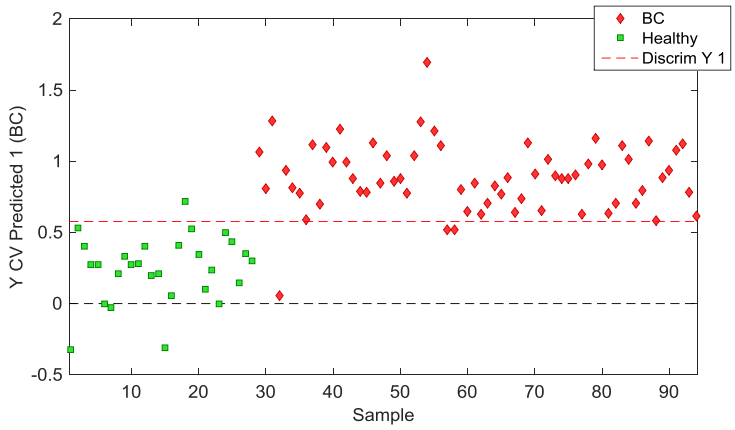


Figure 4.4. Partial Least Squares Discriminant Analysis of all breast cancer and healthy samples. Known healthy samples are green, while the breast cancer samples are shown in red. The y-axis shows the y values returned by the cross-validation prediction of sample. The dotted red line shows the line of discrimination between predicted class groups. Samples on the other side of the discrimination line were misclassified in the cross validation.

A common way of visualizing the sensitivity and specificity of a predictive model is a Receiver Operating Characteristic (ROC) curve.[25] This curve plots sensitivity of the model against 1-specificity at different discrimination threshold settings. As the curve shifts to the top left quadrant of the plot or the (0,1) coordinate, the model is considered a better classifier. The diagonal line that spans from the origin to the (1,1) coordinate represents the threshold of no discrimination, where any points below this line would represents poor classifications. The area under the ROC curve, known as AUC, is used to quantify how proficient the model would be at discriminating positive and negative samples. An AUC value of 0.5 is a poor classifier, as it is no better than a random guess, while a score of 1 represents a perfect classifier. The PLS-DA models for the breast cancer serum samples will be described and evaluated using these classification terms.

The first model compares all breast cancer and healthy samples, illustrated previously in Figure 4.4. The ROC curves in Figure 4.5 show a high degree of sensitivity for classifying both breast cancer and healthy samples, and breast cancer

detection sensitivity at 95% and healthy sample detection sensitivity at 96%. The blue curve in the plot displays the results from the calibrated data, and the green line shows the data from the cross validation. The red circles mark the threshold for discrimination in each model. The overall accuracy of the cross validated model is 96%, with an AUC value of 0.98.

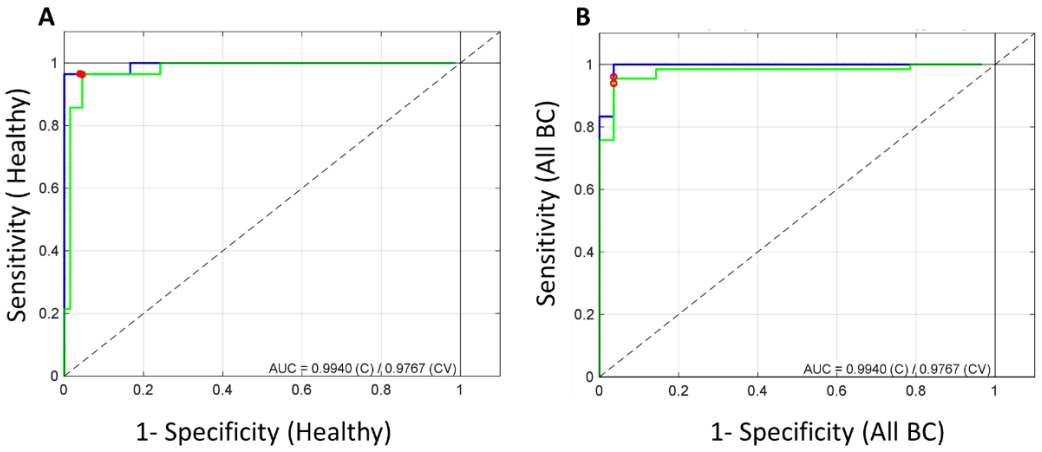


Figure 4.5. ROC curve for Model 1 showing specificity and sensitivity for (a) all healthy (n=28) and (b) all breast cancer (n=66) serum samples based on patient age and an eight-protein signature. The red circles on the blue and green curves mark the sensitivity and specificity at discrimination threshold values for the calibration and cross validation models, respectively The AUC for the cross validated model is 0.98.

The second model compares all of the healthy samples (n =28) to Stage 0-II breast cancer, which includes eight DCIS samples (Stage 0), 34 Stage I samples, and 16 Stage II samples. The ROC curves in Figure 4.6 describe the results of this model; healthy samples were classified with a sensitivity of 96%, and breast cancer classification sensitivity was 97% these values are denoted by the red circles on the green curves in each plot. Three samples were misclassified in total, with one false positive and two false negative samples. The AUC for this model is 0.99, with an overall accuracy of 97%. This model performed similarly to Model 1, which is

expected given the overlap in sample class; the only data points excluded from this model were from eight Stage III-IV samples.

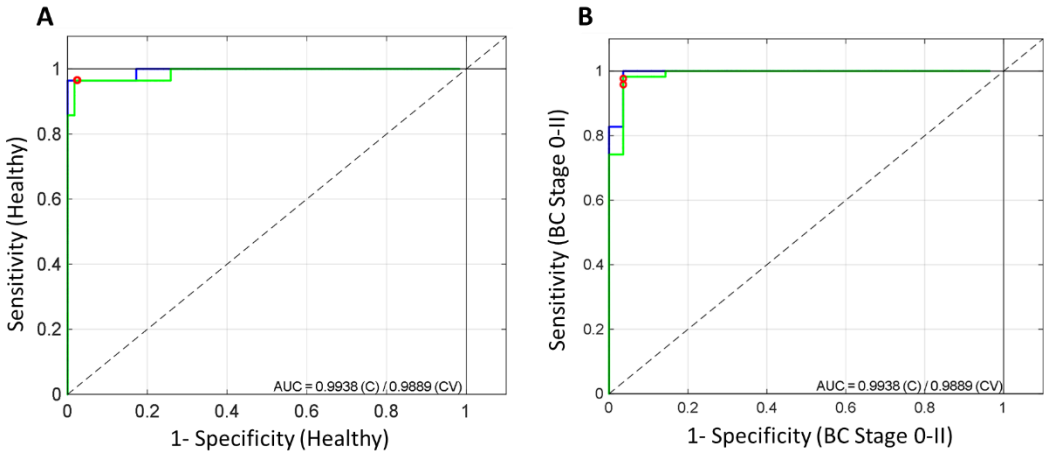


Figure 4.6 ROC curve for Model 2 showing sensitivity and specificity for (a) all healthy (n=28) and (b) Stage 0-II breast cancer (n=58) serum samples based on patient age and an eight protein signature. The AUC for this model is 0.99. The red circles on the blue and green curves mark the sensitivity and specificity at discrimination threshold values for the calibration and cross validation models, respectively

Model 3, illustrated in Figure 4.7 compares Stage 0-II to Stage III-IV samples; Stage I-II samples were classified with a sensitivity of 88%, and Stage III-IV samples were classified with a sensitivity of 38%. The AUC for this model is 0.78, which does not perform as well as Model 1 or 2, but is still higher than the line of no discrimination. Seven of the 58 Stage I-II samples were misclassified, while five of the eight Stage III-IV samples were misclassified. Model 3 clearly classifies Stage 0-II samples with much higher sensitivity and precision than Stage III-IV, and there is less agreement between the calibration and cross validation than in either Model 1 or 2.

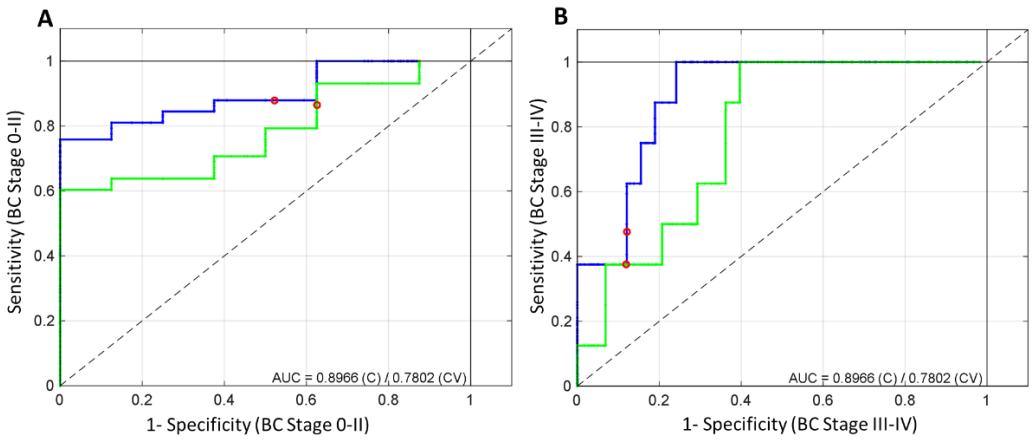


Figure 4.7 ROC curve for Model 3 showing sensitivity and specificity for (a) Stage 0-II (n=58) breast cancer and (b) Stage III-IV breast cancer (n=8) serum samples based on patient age and an eight-protein signature. The red circles on the blue and green curves mark the sensitivity and specificity at discrimination threshold values for the calibration and cross validation models, respectively. The AUC for this model is 0.78.

Model 4 compares ER and/or PR positive (hormone receptor positive) samples (n=54) and triple negative breast cancer samples (n=8), shown in Figure 4.8. This model is relevant to the diagnostic process, as treatment plans can be formulated to target one subtype of the disease, but can be ineffective for another subtype. The sensitivity for HR+ detection was 70%, with 38 of the 54 samples being classified correctly. The sensitivity for triple negative samples was 38%, with three of the eight TNBC samples correctly assigned. The AUC for this model is 0.56, which does not describe this model as a proficient classifier. The metrics of all four models are summarized in Table 4.1, including precision, sensitivity, AUC, and accuracy.

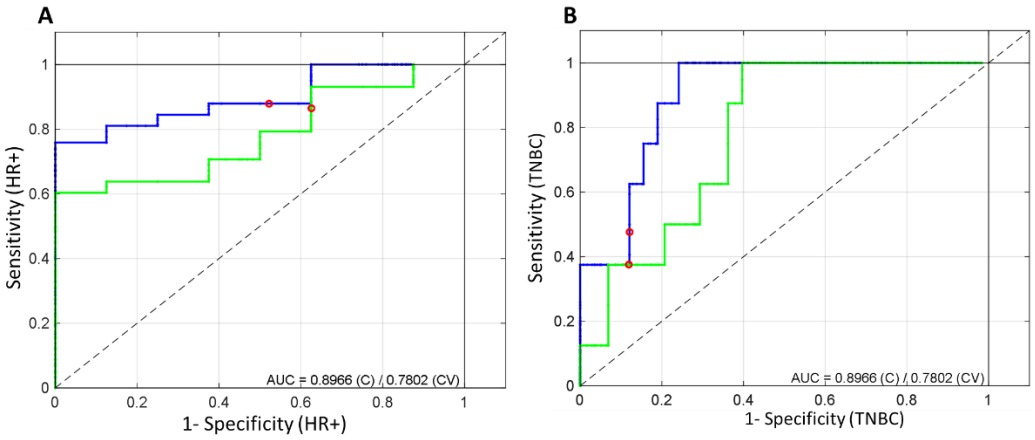


Figure 4.8. ROC curve for Model 4 showing sensitivity and specificity for (a) HR+ (n=54) breast cancer and (b) triple negative breast cancer (n=8) serum samples based on patient age and an eight protein signature. The red circles on the blue and green curves mark the sensitivity and specificity at discrimination threshold values for the calibration and cross validation models, respectively The AUC for this model is 0.56.

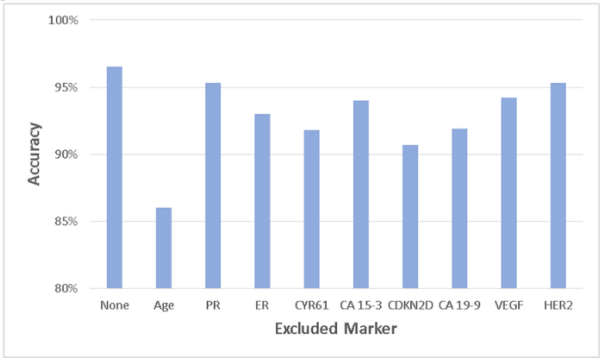
	Class Groups	Precision	True Positives	AUC	# Samples	Accuracy
Model 1	Healthy	90%	96%	0.98	28	96%
	All BC	98%	95%		66	
Model 2	Healthy	93%	96%	0.99	28	97%
	BC Stage 0-II	98%	97%		58	
Model 3	BC Stage 0-II	91%	88%	0.78	58	82%
	BC Stage III-IV	30%	38%		8	
Model 4	HR+ BC	88%	70%	0.56	54	66%
	TNBC	84%	38%		8	

Table 4.1. Description of each PLS-DA model, with precision, true positive rate, AUC value, the number of samples in each group, and the overall accuracy.

Individual contributions of each marker to the model were also evaluated by excluding one variable at a time and re-running the model. Each model had the same preprocessing and cross validation method, and the latent variable number with the best accuracy was chosen to represent that instance of the PLS-DA model. Model 2 (Healthy vs Stage 0-II) was chosen for this evaluation, since it displayed the highest accuracy and the largest AUC. The model assigned classes to samples with an overall 97% accuracy when all nine variables were included, and this accuracy declined to varying degrees when a single variable was excluded from the

analysis. The largest change in accuracy occurred upon the exclusion of patient age, with a resulting accuracy of 86%. Age has been previously identified as a risk factor for breast cancer, so the impact on the model is plausible, though this impact is more likely due to the significant age difference between the two patient cohorts. The protein markers with the most notable impact on the model were CDKN2D, CYR61, and CA 19-9. The accuracy of Model 2 decreased to 91% for CDKN2D, and 92% for CYR61 and CA 19-9. The other marker exclusions resulted in accuracy values ranging 93-95%. Detailed results of individual marker contributions to Model 2 are displayed in Figure 4.9, summarizing the accuracy for each instance of the model.

A



B

Excluded Marker	Total Sample Number	Correct Assignments	Accuracy
None	86	83	97%
Age	86	74	86%
PR	86	82	95%
ER	86	80	93%
CYR61	86	79	92%
CA 15-3	86	81	94%
CDKN2D	86	78	91%
CA 19-9	86	79	92%
VEGF	86	81	94%
HER2	86	82	95%

Figure 4.9. (a) Overall accuracy of Model 2 (Healthy vs Stage 0-II) is plotted against the marker excluded from the model. (b) The number of correctly assigned samples, the total number of samples, and the resulting accuracy is listed for each marker exclusion scenario.

Discussion

This work describes the development and validation of three multiplex assays, for (a) CYR61, ER, and PR, (b) CA 15-3 and CDKN2D, (c) HER2 and VEGF. The previously developed CA 19-9 singleplex assay was used as originally designed. These assays were all validated through cross-reactivity and spike and recovery experiments for serum measurements. Every Simoa assay displayed

LODs lower than their ELISA counterparts. The serum samples used for assay testing consisted of serum taken from women who had positive mammogram and a positive diagnosis from a clinician. These newly diagnosed women had not yet undergone any therapeutic intervention at the time of sampling (n=66). Healthy serum controls from 28 females were purchased from a commercial source. The four assays were then used to test all 94 serum samples.

Initial investigation of the resulting data showed that the sensitivity achieved by Simoa was necessary in order to detect several biomarkers at low levels. Even with this heightened sensitivity, a number of samples were below the LOD for ER and PR. The dilution factor contributed to this need for sensitivity, but this also allowed for low volumes of serum (45 – 60 µL) to be consumed per assay. A total of 200 µL of serum were used to collect data on eight different protein markers. Statistical analysis of individual marker concentrations in serum showed CYR61, CA 19-9, CA 15-3, ER, and CDKN2D levels were correlated with the presence of early stage breast cancer compared to a healthy cohort. This finding suggested potential value in a multivariate approach for discriminating between healthy samples and breast cancer samples. The other four markers showed no significant difference between healthy and breast cancer samples, nor did they distinguish between early and late stages in disease progression or subtype.

All of the data obtained for these serum samples were used as inputs to evaluate the combined utility of eight protein markers and patient age toward a breast cancer diagnostic signature. Four models were tested comparing healthy

samples to all breast cancer samples, healthy and Stage 0-II samples, Stage 0-II and Stage III-IV samples, and HR+ and TNBC samples.

Model 1 compared all healthy cohort to the entirety of the breast cancer cohort, a sensitivity and specificity of 95% and 96%, respectively. Model 1 displayed an overall accuracy of 96% and an AUC of 0.98, which describes a successful classifier. Model 2, which compared healthy samples to Stage 0-II breast cancer, performed slightly better with an overall of 97%, with sensitivity and specificity values at 97% and 96%, respectively. The metrics for both Model 1 and Model 2 show an improvement compared to the current gold standard in screening, mammography, which has 79% sensitivity and 90% specificity.[1] These preliminary results are promising for the development of a noninvasive screening test for breast cancer. Additionally, the variables with the highest impact on Model 2 were age, CDKN2D, CYR61, and CA 19-9 based on the exclusion of these markers from the models. These findings agree with the Mann-Whitney statistical evaluation of individual markers, which supports the use of such univariate statistics to help assess the utility of individual markers. The impact of some of these markers may be due to the statistical difference in age between the healthy and breast cancer cohorts—some markers have age-dependent expression, so the markers in the signature may not be completely independent of one another.

Model 3 evaluated the same biomarker signature to differentiate Stage 0-II and Stage III-IV breast cancer samples—this model demonstrated proficiency in classifying Stage 0-II samples with a sensitivity of 88%, but was less successful in identifying Stage III-IV, which only had a sensitivity of 38%. The AUC for this

model was 0.78, with 82% accuracy. Although Model 3 is not as successful as Model 1 and 2, the signature shows some promise for use in tracking disease progression. The most influential variables in this model were CA 15-3 and CA 19-9, which supports the idea that different markers in a signature could be used for different purposes (i.e. screening, therapeutic efficacy, recurrence monitoring). Model 4 compared the hormone positive (HR+) population to the TNBC cohort, which performed poorly. Is model had an accuracy of 66% and an AUC of 0.56, which places this model slightly above the discrimination of a random guess. This signature does not appear to be appropriate for differentiating breast cancer subtypes. Out of the four models, the protein signature combined with patient age displayed the most discrimination between Stage 0-II breast cancer and healthy samples.

Future Work

The high sensitivity and accuracy demonstrated by Models 1 and 2 are both promising and encouraging, but additional work is necessary to further validate these models and improve discrimination within breast cancer cohorts (i.e. stage and subtype). There are several ways to improve upon this preliminary data, which focus on two aspects of the model: the biomarker assays and the tested samples. The ER and PR biomarker assays were not sensitive enough to detect half or more of the serum samples, and having these values may contribute to better discrimination in a stage-specific or disease-specific manner. The ER and PR assays should be further optimized to be more sensitive in order to assess whether they can contribute to a diagnostic signature. Another way to improve upon the

biomarker assays is to add more markers to the protein signature. HIF1α for example, is a protein that has been associated with aggressive breast cancer in cells, and has potential to be a discriminating factor in aggressive cancer as a circulating marker.

The models can also be improved by diversifying the sample pool further. The four models had an overwhelming majority of HR+, Stage I-II samples, so it is not surprising that, when trained on this set of data, the model identifies these samples with higher sensitivity and accuracy. Giving the model more data from aggressive, late-stage, and triple negative breast cancer cases may help train the model better, and thus allow it to identify these types of samples more accurately. The healthy cohort may also be improved upon by getting serum samples from women who have had negative mammograms. This type of cohort provides a better control for newly diagnosed women, and does not rely on self-reporting healthy patients, as was the case with the current healthy cohort. Additionally, the healthy cohort was significantly younger than the breast cancer cohort, which directly factored into the influential variables of the model. Collecting more healthy samples to age-match the breast cancer cohorts will likely address this bias.

Another interesting aspect of the work worth investigating is the other samples received from Tufts Medical Center. The 66 samples tested in this work were chosen for the lack of therapeutic intervention, which avoided marker expression bias due to treatment. Samples outside of this group were classified "NED", or no evidence of disease detected. These should be tested with the eight-marker signature to find whether these samples would be classified as healthy.

Additionally, there was a small group of serial timepoint samples that were taken from women who were diagnosed with breast cancer and underwent treatment. The serum taken from these patients at the time of diagnosis, treatment, and afterwards, contributes to the effort to create a test that tracks therapeutic efficacy. Overall, the current work has produced a biomarker signature that classified early breast cancer with an overall accuracy of 97%, and through this we have identified important predictive markers for breast cancer. Though further validation is needed, the creation of a cancer biology-focused signature for breast cancer could change the future of screening and facilitate timely, effective treatment of breast cancer.

Materials and Methods

The three-plex assay was comprised of the ER, PR, and CYR61 reagents, with 700, 750, and 488 dye encoding, respectively. The antibodies and standards for all three assays were obtained from R&D Systems Duoset products (DYC5715, DYC5415, and DY4055). The assay format was a standard two-step procedure with a working concentration of 2 x 10^6 beads/mL per assay plex, for a total bead concentration of 6 x 10^6 beads/mL. The detector antibody working stock concentration was kept at 1 µg/mL per plex, and the enzyme concentration was 240 pM.

The CA 15-3 and CDKN2D two-plex assay was a standard two-step assay that was 488 and 750 dye encoded. The CA 15-3 assay reagents were obtained from Fitzgerald (10-CA15C, 10-CA15B, 30-AC17), with a bead concentration of 2.5 x 10^6 bead/mL, 3 µg/mL working stock detector concentration, and 200 pM enzyme.

The CDKN2D capture antibody was obtained from Lifespan Biosciences (LS-C37972), the detector was from Abnova (H00001032-D01P), and the standard was from Origene (TB14065). The bead concentration was 2.5 x 10^6 bead/mL, 1 µg/mL working stock detector concentration, and 200 pM enzyme concentration.

The HER2 and VEGF 2-plex assay was a standard two-step format with 488 and 700 dye encoding. HER2 reagents were purchased from R&D Systems, with the capture antibody and standard from a kit (DYC1129), and the detector antibody was purchased separately (BAF1129). The bead concentration was 2.5 x 10^6 beads/mL, with a detector antibody working stock concentration of 1 µg/mL, and an enzyme concentration of 25 pM. The VEGF assay capture antibody was obtained from Life Technologies (M808), and the standard and detector antibody were from R&D Systems (DY293B-05 and BAF293). The bead concentration was 2.5 x 10^6 beads/mL, with a detector antibody working stock concentration of 2 µg/mL, and an enzyme concentration of 25 pM.

The CA 19-9 assay was run as a two-step singleplex assay on unencoded beads, with a bead concentration of 5 x 10^6/mL. CA 19-9 antibodies and standard were purchased from Fitzgerald, Inc (10-CA9B, 10-CA19A, 30-AC14S). The capture antibody was coupled to the magnetic beads and the detection antibody was biotinylated as previously described. The detection antibody was kept at a concentration of 2 µg/mL and SβG at a concentration of 100 pM.

The protein standards and serum samples tested in this chapter were all diluted in 25% newborn calf serum PBS-based buffer. The serum samples tested by the HER2 and VEGF assays were manually diluted by a factor of six, while all

other assay dilutions were by a factor of 8 before being placed in the HD-1 Analyzer (Quanterix).

Calibration curve fitting was performed the HD-1 Analyzer software (Stratec) using a 4PL equation, which was then used to calculate the protein concentrations in samples. Sample concentrations were corrected for the assay's dilution factor. The assay limit of detection was determined by adding three standard deviations of the blank measurement to the average blank signal, and using this value in the curve fitting formula to interpolate the LOD concentration. Statistical analysis of the single markers was performed using Prism 7 (Graphpad), and all multivariate analysis of the healthy and breast cancer serum data was done using a Matlab add-on software called PLS Toolbox v8.0.2 (Eigenvector).

Data was treated by first by replacing all undetectable samples with a value equal to half of the LOD of the assay, and multiplied by the dilution factor to minimally bias the model. Any missing data points were imputed by the software using the model as a template. The data was then autoscaled by the software and into the model. Cross validation was performed by splitting the data into five equal portions by way of random subsets, using 80% of the data for the calibration model, and the remaining 20% was used for validation. This was repeated until every portion of the data had been used for validation, then repeated five times. The sensitivity, specificity, precision, and accuracy values reported for each model were calculated from the confusion matrix generated from the PLS-DA classification.

References

1. Smetherman, D.H. Screening, imaging, and image-guided biopsy techniques for breast cancer. *Surg Clin North Am* **93**, 309-327 (2013).
2. Sunwoo, H.H. & Suresh, M.R. in The Immunoassay Handbook (Fourth Edition). (ed. D. Wild) 833-856 (Elsevier, Oxford; 2013).
3. Duffy, M.J., Evoy, D. & McDermott, E.W. CA 15-3: Uses and limitation as a biomarker for breast cancer. *Clinica Chimica Acta* **411**, 1869-1874 (2010).
4. Banin Hirata, B.K. et al. Molecular Markers for Breast Cancer: Prediction on Tumor Behavior. *Disease Markers* **2014**, 12 (2014).
5. Ebeling, F.G. et al. Serum CEA and CA 15-3 as prognostic factors in primary breast cancer. *British Journal of Cancer* **86**, 1217-1222 (2002).
6. Zhang, S., Garcia-D'Angeli, A., Brennan, J.P. & Huo, Q. Predicting detection limits of enzyme-linked immunosorbent assay (ELISA) and bioanalytical techniques in general. *Analyst* **139**, 439-445 (2014).
7. Schubert, S.M. et al. Ultra-sensitive protein detection via Single Molecule Arrays towards early stage cancer monitoring. *Scientific reports* **5**, 11034 (2015).
8. Lepor, H. et al. Clinical evaluation of a novel method for the measurement of prostate-specific antigen, AccuPSATM, as a predictor of 5-year biochemical recurrence-free survival after radical prostatectomy: results of a pilot study. *BJU international* **109**, 1770-1775 (2012).
9. Menendez, J.A., Mehmi, I., Griggs, D.W. & Lupu, R. The angiogenic factor CYR61 in breast cancer: molecular pathology and therapeutic perspectives. *Endocrine-Related Cancer* **10**, 141-152 (2003).
10. Galli, C., Basso, D. & Plebani, M. CA 19-9: handle with care. *Clinical chemistry and laboratory medicine : CCLM / FESCC* **51**, 1369-1383 (2013).
11. Sommer, S. & Fuqua, S.A.W. Estrogen receptor and breast cancer. *Seminars in Cancer Biology* **11**, 339-352 (2001).
12. Cheang, M.C.U. et al. Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer. *JNCI Journal of the National Cancer Institute* **101**, 736-750 (2009).
13. Byrne, G.J. et al. Serum Vascular Endothelial Growth Factor in Breast Cancer. *Anticancer Research* **27**, 3481-3487 (2007).
14. Felisiak-Golabek, A. et al. p19(INK4d) mRNA and protein expression as new prognostic factors in ovarian cancer patients. *Cancer Biology & Therapy* **14**, 973-981 (2013).
15. Cánepa, E.T. et al. INK4 proteins, a family of mammalian CDK inhibitors with novel biological functions. *IUBMB Life* **59**, 419-426 (2007).
16. Schubert, S.M. The Development and Implementation of Single Molecule Protein Assays for Application in Early Cancer Diagnostics and Single Cell Studies. (Tufts University, 2015).
17. Elliott, A. in Statistical analysis quick reference guidebook. (ed. W. Woodward) 191-207 (SAGE Publications, Inc., 2007).

18. James, G., Witten, D., Hastie, T. & Tibshirani, R. An Introduction to Statistical Learning with Applications in R. (Springer, New York; 2013).
19. Ringner, M. What is principal component analysis? *Nature biotechnology* **26**, 303-304 (2008).
20. Jolliffe, I.T. & Cadima, J. Principal component analysis: a review and recent developments. *Philos T R Soc A* **374** (2016).
21. Wold, S., Sjostrom, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometr Intell Lab* **58**, 109-130 (2001).
22. Perez-Enciso, M. & Tenenhaus, M. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Hum Genet* **112**, 581-592 (2003).
23. Worley, B. & Powers, R. Multivariate Analysis in Metabolomics. *Current Metabolomics* **1**, 92-107 (2013).
24. Zeng, J. et al. Metabolomics Identifies Biomarker Pattern for Early Diagnosis of Hepatocellular Carcinoma: from Diethylnitrosamine Treated Rats to Patients. *Scientific reports* **5**, 16101 (2015).
25. Hanley, J.A. & McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29-36 (1982).
26. Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G. & Thomas, R. Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology* **56**, 45-50 (2008).

Appendix

The table below lists detailed criteria for the AJCC's TNM staging system described in Chapter 1.

Primary tumor (T)		Regional lymph nodes (N)	
TX	Primary tumor cannot be assessed	NX	Regional lymph nodes cannot be assessed (eg, previously removed)
T0	No evidence of primary tumor	N0	No regional lymph node metastasis
Tis	Carcinoma in situ	N1	Metastasis to movable ipsilateral level I, II axillary lymph node(s)
Tis (DCIS)	Ductal carcinoma in situ	N2	Metastases in ipsilateral level I, II axillary lymph nodes that are clinically fixed or matted or in clinically detected* ipsilateral internal mammary nodes in the absence of clinically evident axillary lymph node metastasis
Tis (LCIS)	Lobular carcinoma in situ	N2a	Metastases in ipsilateral level I, II axillary lymph nodes fixed to one another (matted) or to other structures
T1	Tumor ≤ 20 mm in greatest dimension	N2b	Metastases only in clinically detected* ipsilateral internal mammary nodes and in the absence of clinically evident level I, II axillary lymph node metastases
T1mi	Tumor ≤ 1 mm in greatest dimension	N3	Metastases in ipsilateral infraclavicular (level III axillary) lymph node(s), with or without level I, II axillary node involvement, or in clinically detected * ipsilateral internal mammary lymph node(s) and in the presence of clinically evident level I, II axillary lymph node metastasis; or metastasis in ipsilateral supraclavicular lymph node(s), with or without axillary or internal mammary lymph node involvement
T1a	Tumor > 1 mm but ≤ 5 mm in greatest dimension		
T1b	Tumor > 5 mm but ≤ 10 mm in greatest dimension	N3a	Metastasis in ipsilateral infraclavicular lymph node(s)
T1c	Tumor > 10 mm but ≤ 20 mm in greatest dimension	N3b	Metastasis in ipsilateral internal mammary lymph node(s) and axillary lymph node(s)
T2	Tumor > 20 mm but ≤ 50 mm in greatest dimension	N3c	Metastasis in ipsilateral supraclavicular lymph node(s)
T3	Tumor > 50 mm in greatest dimension	**Distant Metastasis (M)**	
T4	Tumor of any size with direct extension to the chest wall and/or to the skin (ulceration or skin nodules)	M0	No clinical or radiographic evidence of distant metastasis
T4a	Extension to chest wall, not including only pectoralis muscle adherence/invasion	cM0(i+)	No clinical or radiographic evidence of distant metastases, but deposits of molecularly or microscopically detected tumor cells in circulating blood, bone marrow, or other nonregional nodal tissue that are no larger than 0.2 mm in a patient without symptoms or signs of metastases
T4b	Ulceration and/or ipsilateral satellite nodules and/or edema (including peau d'orange) of the skin, which do not meet the criteria for inflammatory carcinoma		
T4c	Both T4a and T4b	M1	Distant detectable metastases as determined by classic clinical and radiographic means and/or histologically proven > 0.2 mm
T4d	Inflammatory carcinoma		

Figure A 1. Detailed AJCC criteria for TNM Staging

Sample ID	Stage	Age
BRH813984	Healthy	22
BRH813985	Healthy	33
BRH813986	Healthy	36
BRH813987	Healthy	21
BRH813989	Healthy	52
BRH813990	Healthy	46
BRH813991	Healthy	25
BRH813992	Healthy	32
BRH813993	Healthy	42
BRH813994	Healthy	45
BRH813995	Healthy	45
BRH813996	Healthy	38
BRH813997	Healthy	47
BRH813998	Healthy	50
BRH888412	Healthy	72
BRH888413	Healthy	73
BRH888415	Healthy	54
BRH888416	Healthy	52
BRH888417	Healthy	57
BRH943383	Healthy	54
BRH943384	Healthy	72
BRH943385	Healthy	58
BRH943386	Healthy	53
BRH943387	Healthy	58
BRH943388	Healthy	57
BRH813988	Healthy	43
BRH888414	Healthy	70
BRH943389	Healthy	52
BRH943390	Healthy	58
BRH943391	Healthy	52
BRH943392	Healthy	58
BRH943393	Healthy	47
BRH943394	Healthy	50
BRH943395	Healthy	56
BRH943396	Healthy	51
BRH943397	Healthy	52

Figure A 2. Information provided for all serum samples tested in Chapter 3.

Sample ID	Stage	Subtype	Age
BRH824412	1	HR+	52
BRH824413	1	HR+	77
BRH824414	2	HR+	84
BRH824415	1	HR+	71
BRH824416	2	HR+	79
BRH824417	1	HR+	65
BRH824418	2	HR+	56
BRH824419	2	HR+	44
BRH824420	1	HR+	36
BRH845591	4	HR+	79
BRH845592	4	HR+	75
BRH845593	4	HR+	71
BRH845594	4	HR+	49
BRH845596	4	HR+	39
BRH845597	4	HR+	75
BRH845598	4	HR+	75
BRH845599	4	HR+	72
BRH853584	3	TNBC	70
BRH885843	2	unknown	40
BRH885844	4	HER2	40
BRH885845	2	TNBC	81
BRH885846	4	HER2	56
BRH885847	3	TNBC	53
BRH885848	3	TNBC	53
BRH885849	2	unknown	69
BRH885851	1	HR+	69
BRH885852	1	HR+	70
BRH885853	4	HR+	75
BRH885854	4	unknown	71
BRH885855	1	HR+	71
BRH885856	1	HR+	74
BRH885857	3	HR+	40
BRH885858	0	HR+	61
BRH921357	2	HR+	61
BRH921358	3	TNBC	73
BRH921359	2	TNBC	74
BRH921360	2	TNBC	58
BRH921361	2	TNBC	74
BRH921362	2	TNBC	79
BRH921363	2	TNBC	59
BRH921364	2	TNBC	49
BRH921365	4	TNBC	68
BRH921366	1	TNBC	82
BRH921367	2	TNBC	62
BRH921368	2	TNBC	67
BRH921369	2	TNBC	69
BRH921370	2	TNBC	63
BRH921371	2	TNBC	54
BRH921372	3	TNBC	54
BRH922986	2	HR+	51
BRH922987	4	HR+	68
BRH922988	3	HR+	55
BRH922989	2	HR+	51
BRH922990	3	HR+	68
BRH922991	3	HR+	47
BRH922992	3	HR+	64
BRH922993	1	HR+	71
BRH922994	2	HR+	43
BRH922995	1	HR+	38
BRH922996	unknown	HR+	68
BRH922997	3	HR+	47
BRH922998	2	HR+	69

Healthy	
Sample ID	HIF1α, pg/mL
BRH774363	Below LOD
BRH774364	Below LOD
BRH813984	0.0245
BRH813985	Below LOD
BRH813986	Below LOD
BRH813987	Below LOD
BRH813988	Below LOD
BRH813989	0.603
BRH813990	Below LOD
BRH813991	Below LOD
BRH813992	Below LOD
BRH813993	Below LOD
BRH943383	Below LOD
BRH943384	167.8
BRH943385	3.26
BRH943386	Below LOD
BRH943387	Below LOD
BRH943388	Below LOD
BRH943389	Below LOD
BRH943390	Below LOD

Breast cancer	
Sample ID	HIF1α, pg/mL
BRH922987	Below LOD
BRH922988	Below LOD
BRH922989	0.0949
BRH922990	Below LOD
BRH922991	Below LOD
BRH922992	Below LOD
BRH922993	Below LOD
BRH922994	Below LOD
BRH922995	Below LOD
BRH922996	Below LOD
BRH922997	Below LOD
BRH922998	0.133
BRH921360	1.38
BRH921361	Below LOD
BRH921362	0.277
BRH921363	Below LOD
BRH921364	0.165
BRH921365	Below LOD
BRH921366	0.268
BRH921367	13.6

Figure A 3. Serum samples tested against the HIF1α assay. All concentrations shown account for serum dilution factor. Most of the samples tested were not detectable by the assay. Patient information for these samples can be found in Figure A2.

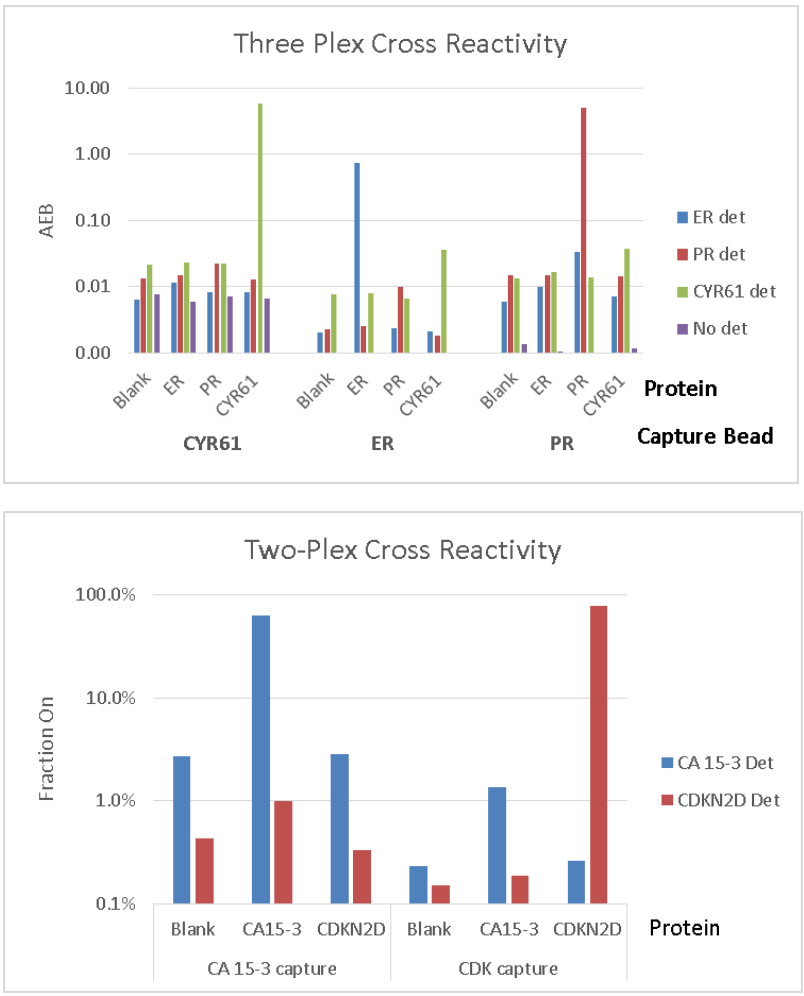


Figure A 4. Cross reactivity results for the ER, PR, and CYR61 three-plex assay and the two-plex CA 15-3 and CDKN2D assay.

The cross-reactivity testing was performed with all of the encoded beads combined as they would be in typical assay conditions. Detector and enzyme concentrations were also kept at standard assay conditions. The conditions were then separated such that each detector antibody and protein standard were individually combined to test specific interactions. The negative controls for each bead type were blank buffer samples tested with no detector antibody, while the positive controls were the correct detector antibody combined with a high concentration of the correct protein. For the three-plex assay, these protein concentrations were 500 pg/mL for ER and PR, and 200 pg/mL for CYR61. The

concentrations of CA 15-3 and CDKN2D were 100 U/mL and 500 pg/mL, respectively. There did not appear to be a significant amount of cross-reactivity between assay reagents, even in the presence of high concentrations of protein standard.

The following tables shown in Figure A5 summarize the results of the spike/recovery experiments to determine the appropriate dilutions and sample buffers ito use for the assay.

		Serum Dilution Factor			
Analyte and Spike Level	CYR61	Buffer only	1:04	1:08	1:10
	None (pg/mL)	0.32	29.88	42.09	22.88
	2	68.9%	3430.6%	535.2%	1020.3%
	50	109.4%	189.4%	86.3%	101.2%
	100	109.6%	144.9%	103.8%	NaN
	ER				
	None (pg/mL)	not detected	7.75	3.71	7.05
	2	114.1%	30.3%	288.4%	241.2%
	50	98.6%	68.1%	75.1%	62.2%
	500	97.4%	67.8%	78.7%	NaN
	PR				
	None (pg/mL)	0.1573	0.8379	0.9382	0.5215
	2	67.4%	88.4%	97.3%	123.0%
	50	99.9%	49.8%	87.3%	71.6%
	100	109.9%	81.1%	91.0%	NaN

		Serum Dilution Factor			
Analyte and Spike Level	CA 15-3	Neat serum	1:04	1:08	Buffer only
	None (U/mL)	4.37	5.24	0.96	0
	0.5	-22%	-70%	144%	46%
	5	14%	133%	147%	80%
	50	5%	NaN	119%	50%
	CDKN2D				
	None (pg/mL)	25.97	9.09	3.34	0.3
	1	-708%	-54%	243%	32%
	10	-39%	66%	91%	86%
	100	19%	50%	68%	73%

CA19-9				
Spike, U/mL	Neat	1:04	1:08	Control
0	507.33	9.43	4.06	0.60
2	-177.02	84%	189%	76%
10	-30.47	139%	108%	77%
50	-2.28	152%	127%	95%

Figure A 5. Spike and recovery results for the multiplex assays and the CA 19-9 singleplex assay.

The columns on the left describe the spiked protein and the spiked amount of standard. The values in the neat rows display the unspiked serum concentrationused as a reference for recovery. Recovery percentages were calculated by subtracting

the unspiked value from the spiked value and dividing by the known spike value. The control for this experiment was a buffer-only set of samples that were spiked with the same solution as the serum dilutions. Serum and calibrators were diluted in a PBS-based buffer with 25% newborn calf serum. Serum samples were diluted by a factor of eight based on the results of these experiments.

Sample Name	Age	Stage	Grade	ER	PR	HER2	Subytpe
BCS-1019	68	IIA	II	Y	Y	1+	HR+
BCS-1021	61	IIA	I	Y	Y	0	HR+
BCS-1025	71	IA	II	Y	Y	0	HR+
BCS-1026	55	IV	III	Y	N	1+	HR+
BCS-1029	72	(II)	III	N	N	0	TNBC
BCS-1030	48	IIIA	III	N	N	0	TNBC
BCS-1047	61	IB	II	Y	Y	0	HR+
BCS-1049	42	0	III	Y	Y	3+	HR+
BCS-1053	90	IIIB	II	Y	Y	1+	HR+
BCS-1054	47	IIA	II	Y	Y	2+	HR+
BCS-1058	64	IA	III	Y	Y	0	HR+
BCS-1061	57	IA	II	Y	Y	1+	HR+
BCS-1062	54	IA	II	Y	Y	1+	HR+
BCS-1065	54	IA	I	Y	Y	1+	HR+
BCS-1075	47	0	III	Y	Y		HR+
BCS-1091	47	IIA	II	Y	Y	1+	HR+
BCS-1098	53	IIB	II	Y	Y	0	HR+
BCS-1099	39	IIA	I	Y	Y	1+	HR+
BCS-1108	71	0	III	N	N	3+	HER2+
BCS-1117	57	IA	II	Y	Y	0	HR+
BCS-1121	53	IA	II	Y	Y	0	HR+
BCS-1122	63	I	II	Y	N	0	HR+
BCS-1128	44	IIIA	II	Y	Y	0	HR+
BCS-1134	57	0	III	Y	Y		HR+
BCS-1144	70	IA	II	Y	Y	1+	HR+
BCS-1145	60	IA	I	Y	N	1+	HR+
BCS-1156	76	IA	III	N	N	1+	TNBC
BCS-1160	45	IIIB	II	Y	N	1+	HR+
BCS-1165	53	IIIA	III	N	N	0	TNBC
BCS-1173	42	IB	II	Y	Y	1+	HR+
BCS-1177	62	IB	II	Y	Y	0	HR+
BCS-1182	63	IA	II	N	N	0	TNBC

Sample Name	Age	Stage	Grade	ER	PR	HER2	Subytpe
BCS-1185	68	IA	II	Y	Y	1+	HR+
BCS-1194	37	IIA	III	Y	Y	3+	HR+
BCS-1195	66	IA	II	Y	N	0	HR+
BCS-1204	71	IA	I	Y	Y	1+	HR+
BCS-1205	51	IIA	I	Y	Y	0	HR+
BCS-1208	47	IIA	II	N	N	1+	TNBC
BCS-1217	47	IA	I	Y	Y	0	HR+
BCS-1218	56	II	III	N	N	3+	HER2+
BCS-1219	70	IA	I	Y	Y	1+	HR+
BCS-1224	51	IA	II	Y	Y	0	HR+
BCS-1229	66	0	III	Y	Y		HR+
BCS-1230	70	0	III	Y	Y		HR+
BCS-1231	74	IB	II	Y	Y	0	HR+
BCS-1239	60	IA	II	N	N	1+	TNBC
BCS-1240	64	0	II	Y	Y		HR+
BCS-1241	64	I		?	?		?
BCS-1246	61	IIB	II	Y	Y	0	HR+
BCS-1247	58	IA	I	Y	Y	1+	HR+
BCS-1271	68	I		Y	Y	0	HR+
BCS-1276	68	IA	I	Y	N	1+	HR+
BCS-1278	56	IV		Y	Y	0	HR+
BCS-1288	49	0	III	N	N	3+	HER2+
BCS-1290	70	IIIA	II	Y	Y	0	HR+
BCS-1295	67	IA	II	Y	Y	N	HR+
BCS-1296	41	IA	I	Y	Y	1+	HR+
BCS-1298	82	IIA	II	Y	Y	0	HR+
BCS-1300	58	IA	I	Y	Y	0	HR+
BCS-1301	47	IA	II	Y	Y	1+	HR+
BCS-1302	69	IIA	I	Y	Y	0	HR+
BCS-1303	65	IIA	II	Y	Y	0	HR+
BCS-1309	82	IA	I	Y	N	1+	HR+
BCS-1312	52	IIB	II	Y	Y	1+	HR+
BCS-1315	51	IB	I	Y	Y	1+	HR+
BCS-1318	54	IB	III	N	N	0	TNBC

Sample Name	Age
BRH1166211	44
BRH1166212	40
BRH1166213	43
BRH1166214	56
BRH1166215	46
BRH1166216	51
BRH1166217	37
BRH1166218	38
BRH1166219	53
BRH1166220	53
BRH1166221	47
BRH1166222	39
BRH1166223	32
BRH1166224	51
BRH1166225	33
BRH1166226	45
BRH1166227	47
BRH1166228	53
BRH1166229	35
BRH1166230	42
BRH1166231	47
BRH1166232	53
BRH1166233	36
BRH1166234	43
BRH1166235	35
BRH1166236	34
BRH1166237	43
BRH1166238	42

Population	Mean	Median	IQR	Range	N
Healthy	44	43	38-48	32-56	28
All BC	59	59	51-68	37-90	66
HR+	59	59	51-68	37-90	54
Stage 0-II	59	61	53-68	37-82	58
Stage III-IV	58	54	47-54	44-90	8
TNBC	59	57	52-65	53-76	8

Figure A 6.The tables above list clinical information about the breast cancer patient serum tested for eight protein biomarkers (n=66). The healthy cohort is described in the table to the left with sample ID and donor ages. All donors were female. A summary table lists sample sizes and age statistics for each cohort.