Ultrasound Spectroscopy of Biofluid Properties for Health Assessment

Andrien A. Rackov

Doctor of Philosophy

Department of Chemistry

McGill University, Montreal

Quebec, Canada

December 2014

A dissertation submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Abstract

The development of portable diagnostic platforms for rapid measurement of complex biological samples has become a target of great interest in analytical chemistry. Ideal protocols provide accurate and precise results with minimal sample handling and make use of cost-effective instrumentation. However, the high concentration of absorbing and scattering species present in biological samples such as blood, serum, and milk can make conventional optical measurements very challenging. Dilution, separation, and the addition of reagents are usually necessary to reduce the complexity of the sample matrix prior to analysis. This thesis presents a measurement methodology for monitoring markers of health in biological fluids by the analysis of ultrasound wave frequencies.

Initially, ultrasound was explored for monitoring pH in homogeneous, albumin-containing media. During propagation, ultrasound waves are subject to distortion processes that are characteristic of the chemicals present in solution. The nonlinear distortion of the waves can be measured in the ultrasound frequency profile. Nonlinear distortion and attenuation of ultrasound waves were shown to be sensitive to changes in the molecular conformation of albumin with pH. Multivariate analysis of ultrasound frequency profiles identified frequencies associated with conformational changes of albumin. Although a large number of

other constituents are present in serum, the pH of human serum samples from multiple donors could be estimated with an error of 0.077 pH units from a single ultrasound measurement.

The application of nonlinear distortion was extended to estimation of total cell counts in a heterogeneous system, milk. Total cell counts were associated with the presence or absence of subclinical mastitis and were grouped as low or high for classification. Despite the presence of fat globules, samples from women with low (≤150K cells/mL) and high (≥600K cells/mL) somatic cell counts were classified with a sensitivity and specificity of 90% and 79% respectively. Estimates by ultrasound were comparable to classification results from NIR measurements, which exhibited a sensitivity of 85% and specificity of 84%. It was shown that nonlinear ultrasound is not only sensitive to molecular properties of a system, but is also influenced by higher level structures.

Subsequently, infection status of women with no or multiple infections was investigated by ultrasound frequency analysis. Changes in frequency profile of measured serum samples were attributed to differences in serum protein concentration and oxidative stress modifications of proteins. Although a number of variables could not be controlled in this study, classification estimates for the no and multiple infection groups showed a sensitivity and specificity of 100% and 86% respectively.

Nonlinear ultrasound was also used to assess the infection status of mice in a controlled design experiment. Variations in measured frequency profiles of mouse serum samples correlated well with infection and were attributed to changes in the viscoelastic properties of the samples. Classification estimates of infected and uninfected mice resulted in a sensitivity of 84% and specificity of 73%. Additionally, unlike previous measurements, the mouse data was collected with a novel heterodyned ultrasound instrument that reduced the cost of instrumentation.

Overall, the ultrasound methodologies presented in this thesis showed advantages over conventional approaches for monitoring chemical changes in biological fluids. For portable diagnostic platforms, ultrasound offers analysis with minimal or no sample preparation and can provide a rapid approach for health assessment. Future directions and refinements for ultrasonic frequency analysis techniques are outlined in the concluding chapter.

Résumé

Le développement des appareils diagnostiques portables pour la mesure rapide des échantillons biologiques complexes est devenu une cible de grand intérêt en chimie analytique. Les protocoles idéals fournissent des résultats précis avec une manipulation minimale de l'échantillon et utilisent des instruments rentables. Cependant, la forte concentration des espèces chimiques qui absorbent et diffusent dans des échantillons biologiques tels que le sang, le sérum, et le lait peut rendre des mesures optiques classiques très difficile. La dilution, la séparation, et l'addition des réactifs sont généralement nécessaires pour réduire la complexité de la matrice de l'échantillon avant l'analyse. Cette thèse présente une méthode de mesure pour la surveillance des marqueurs de santé dans des fluides biologiques par l'analyse des fréquences des ondes d'ultrasons.

Initialement, l'ultrason a été exploré pour la surveillance de pH dans des échantillons homogènes contenant de l'albumine. Au cours de la propagation, les ondes ultrasonores sont soumises à des processus de distorsion qui sont caractéristiques des produits chimiques présents dans l’échantillon. La distorsion non linéaire des ondes peut être mesurée dans le profil de fréquence ultrasonore. La distorsion non linéaire et l'atténuation des ondes ultrasonores se sont révélés être sensibles aux variations de la conformation moléculaire de

l'albumine avec le pH. L'analyse multivariée des profils de fréquence ultrasonore a identifié les fréquences associées aux changements de conformation de l'albumine. Bien qu'un grand nombre d'autres constituants soient présents dans le sérum, le pH des échantillons de sérum humain provenant de plusieurs donneurs peut être estimé avec une erreur de 0,077 unité de pH à partir d'une mesure unique d'ultrasons.

L'application de distorsion non linéaire a été étendu à l'estimation du nombre de cellules dans un système hétérogène, le lait. Le nombre total de cellules a été associées à la présence ou l'absence de mammite subclinique et a été regroupé comme faible ou élevé pour la classification. Malgré la présence de globules gras, des échantillons provenant de femmes à nombre faible ≤150K (cellules/mL) et élevé (cellules ≥600K/mL) de cellules somatiques ont été classés avec une sensibilité et une spécificité de 90% et 79% respectivement. Les estimations par ultrasons étaient comparables aux résultats de classification de mesures NIR, qui présentaient une sensibilité de 85% et une spécificité de 84%. Il a été montré que l'ultrason non linéaire n'est pas seulement sensible aux propriétés moléculaires d'un système, mais est également influencé par des structures de niveau supérieur.

Par la suite, l'état d'infection des femmes sans ou avec plusieurs infections a été étudiée par analyse de fréquence ultrasonore. Les changements

dans le profil de fréquence des échantillons de sérum mesurés ont été attribués à des différences dans la concentration des protéines sériques et des modifications oxydatives des protéines. Bien qu'un certain nombre de variables ne pouvait être contrôlée dans cette étude, les estimations de classification pour groupes sans et avec plusieurs infections ont montré respectivement une sensibilité et une spécificité de 100% et 86%.

L’ultrason non linéaire a également été utilisé pour évaluer l'état d'infection de souris dans une expérience contrôlée. Les variations des profils de fréquence mesurées des échantillons de sérum de souris ont montré une bonne corrélation avec infection et ont été attribués à des changements dans les propriétés viscoélastiques des échantillons. Estimations de classification de souris infectées et non infectées ont donné lieu à une sensibilité de 84% et une spécificité de 73%. De plus, les échantillons de sérum de souris ont été mesurés avec un instrument à ultrason hétérodyne à coût réduit.

Dans l'ensemble, les méthodes ultrasonores présentées dans cette thèse ont montré des avantages par rapport aux méthodes classiques pour surveiller les changements chimiques dans les fluides biologiques. Pour les appareils diagnostiques portables, l’ultrason offre une analyse avec un minimum de préparation d’échantillon et peut fournir une approche rapide pour l'évaluation de

la santé. Les projets futurs et les améliorations des techniques d'analyse de fréquence ultrasonore sont décrits dans le chapitre de conclusion.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

BSA: bovine serum albumin

CV: cervico-vaginal infection

FT: Fourier transform

GA: genetic algorithm

HWT: Haar wavelet transform

IgG: immunoglobulin G

MLR: multilinear regression

NIR: near-infrared

POC: point-of-care

ROS: reactive oxygen species

SCC: somatic cell count

SECV: standard error of cross-validation

SEE: standard error of the estimate

SNV: standard normal variate

UTI: urinary tract infection

A: amplitude

B: bulk modulus

B/A: nonlinearity parameter

c: velocity

d: distance

f: frequency

p: pressure

R: reflection coefficient

s: entropy

t: time

T: transmission coefficient

Z: acoustic impedance

α: attenuation coefficient

φ: phase offset

ρ: density

Original Contributions to Knowledge

1. Developed an ultrasonic frequency analysis approach for determining pH in albumin-containing media based on pH-induced conformational changes of albumin.

2. Demonstrated the usefulness of ultrasound and near infrared spectroscopy for rapid estimation of somatic cell counts in unmodified human milk, a highly scattering medium.

3. Showed that an analysis of nonlinear ultrasound frequency distortions can be used to construct multivariate classification models to distinguish between infected and uninfected serum samples.

4. Developed a novel heterodyned ultrasound detection system based on a low-cost, high-resolution audio digitizer and demonstrated its use for rapid identification of infection.

Contribution of Authors

Articles included in this dissertation are listed below along with the contributions of each author. Prof. David H. Burns acted as both dissertation supervisor and primary reviewer to Andrien A. Rackov.

Chapter 2

- Andrien A. Rackov and David H. Burns, “Ultrasonic Frequency Analysis for Estimating pH in Albumin-Rich Biofluids”, *to be submitted*.

Mr. Rackov designed the procedure, set up the instrumentation, and collected the experimental data. Data analysis was performed by Mr. Rackov and Prof. Burns provided guidance for the interpretation of results. The manuscript was written and prepared by Mr. Rackov and edited by Prof. Burns.

Chapter 3

- Andrien A. Rackov, Hilary M. Wren, Kristine G. Koski, Marilyn E. Scott, and David H. Burns, “Estimation of Somatic Cell Counts in Breast Milk by Ultrasound and Near Infrared Spectroscopy”, *to be submitted*.

Mr. Rackov designed the data collection procedures for both ultrasound and NIR measurements, designed a new ultrasound sample cell, and set up all instrumentation. All ultrasound data were collected by Mr. Rackov and NIR data were acquired by Ms. Wren, under guidance from Mr. Rackov. The somatic cell counting procedure was equally designed and carried out by Mr. Rackov and Ms. Wren. Ms. Wren was responsible for organization of field work and collection of milk samples from rural Guatemalan mothers. Profs. Koski and Scott assisted with the design of the study. All data were analyzed by Mr. Rackov with guidance from Prof. Burns. The manuscript was written and prepared by Mr. Rackov and edited by Prof. Burns.

Chapter 4

- Andrien A. Rackov, Doris Gonzalez-Fernandez, Kristine G. Koski, Marilyn E. Scott, and David H. Burns, “Monitoring Infections in Indigenous Panamanian Mothers by Ultrasound Frequency Analysis of Serum”, *to be submitted*.

Mr. Rackov designed the procedure for ultrasonic measurements of serum samples and collected all spectral data. Mr. Rackov also designed the small-volume sample cell used in the experimental set-up. Ms. Gonzalez-Fernandez

was responsible for organization of all field work pertaining to the collection of serum samples from indigenous mothers in Panama and obtaining infection and nutritional deficiency information. Profs. Koski and Scott provided guidance on the design of the field work. Analysis of ultrasound data was performed by Mr. Rackov with advice from Prof. Burns. The manuscript was written and prepared by Mr. Rackov and edited with Prof. Burns.

Chapter 5

- Andrien A. Rackov, Maurice R. Odiere, Kristine G. Koski, Marilyn E. Scott, and David H. Burns, “A Novel Heterodyned Ultrasound Approach for Assessing *Heligmosomoides bakeri* Infection Levels in Mice”, *to be submitted*.

Mr. Rackov designed the frequency scanning procedure for measurements of serum samples and collected all spectral data. The heterodyned ultrasound instrument was designed, assembled, and tested by Mr. Rackov. Data analysis was performed by Mr. Rackov with recommendations from Prof. Burns. Dr. Odiere designed the 2-by-3 factorial pregnancy/infection experiment and collected the mouse serum samples. Profs. Koski and Scott assisted in the design of the mouse experiment and contributed to the interpretation of the

results. The manuscript was written and prepared by Mr. Rackov and edited by Prof. Burns.

Acknowledgements

I have been very fortunate to enjoy this unique journey with a wonderful support team. First, I would like to thank my supervisor, Professor David Burns, for his guidance, encouragement, and infectious enthusiasm. I had the opportunity to explore a variety of exciting projects and to develop a unique toolbox of skills and interests. His quirky, science-fiction inspired ideas were refreshing and led me to having some serious fun!

My committee members, Professors Eric Salin and Cameron Skinner, were instrumental in providing me with constructive criticisms and educational material, for which they have my thanks! I would also like to sincerely thank Prof. Salin for the opportunities he gave me with his research group, which motivated me to embark on this doctoral adventure.

This work would not have been possible without my gracious collaborators at the Institute of Parasitology and School of Dietetics and Human Nutrition at McGill University: Professors Marilyn Scott and Kristine Koski, Doris Gonzalez-Fernandez, Maurice Odiere, and Hilary Wren. Thank you! I would like to thank Prof. Bruce Lennox for kindly sharing his facilities during my time of flux, ensuring the completion of this dissertation. Often operating behind the scenes, I would also like to thank Chantal Marotte and all the wonderful staff of the Department of Chemistry for orchestrating the details of graduate school.

I will not dismiss the importance of my fellow Burns lab members, who with their good humour and insights made the lab a wonderful place to be. My sincerest thanks to: David Troïani, Fabiano Pandozzi, Jonathan Dion, Michiko Kato, Steven Cassin, Javier Sanchez-Galan, Andrew Dafoe, and Pieter Roos. Jon and Andrew get a bonus thank-you for going above and beyond the call of duty. Thank you for all the sushi, the encouragement, and all your invaluable input that saw me through to submission day.

I am also grateful to the undergraduate students that have passed through the doors of the Burns group and kept life in the lab spicy and interesting. Thank you to Mayan Bar-On, James Duffy, Joshua Potel, Alexei Kazarine, and Yann Desjardins-Langlais.

I would like to extend my gratitude to the Fonds québécois de la recherche sur la nature et les technologies (FQRNT) and McGill University for the years of financial support.

Both inside and outside of the lab, Steven Cassin and Lisa Starr were helpful beyond imagining and kept me on track. To Steve, thank you for your encouragement and being an all-round wonderful friend. To Lisa, you have my sincere gratitude for assistance in the lab and many fruitful discussions on topics of which I knew very little. To my friends Garrett Morgan and Andrew Mescheder,

I thank you for keeping me pleasantly distracted from life in the lab when I needed it.

Finally, I would like to thank my parents Jolanda and Borivoj, siblings Elias and Amelia, grandparents Paul and Hildegard, uncle Paul and aunt Rosita, as well as the rest of my extended family for their unyielding love and support throughout the years.

Chapter 1

Introduction and Background

1.1 Project Overview

Analytical measurements are ubiquitous throughout our society. From routine quality control in industrial settings to medical diagnostics in remote locations, determining the concentration of chemicals within samples is critical in today's world. Advances of methods for chemical quantification have focused on reducing testing times and equipment cost. This has manifested itself over recent years with a surge in point-of-care (POC) technologies that seek to bring the power of the laboratory into the hands of the non-expert user.

The goal of the work presented in this dissertation is the development of simple disease diagnostics using ultrasound spectroscopy. These fall under the category of POC technologies, which seek to provide rapid, on-site results to non-expert users, without compromising affordability, precision, and accuracy. The general approach for all measurements was based on monitoring changes of ultrasound frequency spectra collected for biofluid samples. These changes resulted from attenuation and nonlinear phenomena exhibited by the samples. The ultrasound frequency data was processed using multivariate calibration and classification techniques. This more rigorous data analysis, based on multilinear

regression, was offset by short data collection times (i.e. seconds). All analyses were direct, with minimal sample preparation steps, and yielded straightforward “mix-and-measure” methodologies for health assessment.

The first three manuscripts presented in this thesis introduce methodologies for determining pH, white blood cell counts, and infection states based on ultrasound measurements of biofluids. The final manuscript explores a novel ultrasound instrument for assessing infection states.

1.2 Biofluid Markers and Health

Analysis of biofluids, such as blood, urine, breast milk, cerebrospinal fluid and amniotic fluid, can be critical in the detection of disease and inflammation.[1.1] These analyses are typically done in laboratories and measurements can take hours or days if the laboratories are overwhelmed with samples.[1.2] Further practical concerns in busy laboratories also include improper sample handling and storage, which can affect accuracy of the measurement.[1.3] In contrast, POC technologies offer earlier access to test results and can lead to significant improvements in patient care.[1.4] POC devices can ease laboratory burdens and help medical professionals quickly identify and implement any necessary treatments.

Common markers associated with health are good candidates for non-invasive measurement techniques. These indicators include pH and acid-base balance, serum protein concentrations, leukocyte counts, and oxidative stress markers. Each of these components is a useful indicator in the clinical setting and will be reviewed in the subsequent sections.

1.2.1 pH: The Acid-Base Balance

Humans are complex organisms with many different acid-base buffer systems involved in functional metabolism. Examples are the bicarbonate buffering system, the phosphate buffering system, and the protein buffering system. The acid-base balance in the human body is tightly regulated by these systems, with typical blood pH values found between 7.35 and 7.45.[1.5] Other biofluids and tissues, such as cerebrospinal fluid and amniotic fluid, share a similar pH range.[1.6,1.7] Outside of this range, proteins are denatured, enzymes lose their functionality, and depending on the severity of the imbalance, death may occur. To maintain this pH balance, the various buffering systems work together in different capacities.[1.8] The bicarbonate buffering system is the major buffer system and is responsible for removing about 50% of acid in the body. It allows carbon dioxide and water to be shifted through carbonic acid to produce hydrogen ions and bicarbonate, and can quickly respond to any potential change in pH:

$$H_2O + CO_2 \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \qquad (1.1)$$

Higher pH ⟷ **Lower pH**

The bicarbonate buffer system can regulate pH in the short term through respiration.[1.9] H^+ can be eliminated as CO_2, which is blown off in the expiratory gas from the lungs. Changing the rate of ventilation alters the concentration of

CO_2 in the blood, and shifts the bicarbonate buffer system in the direction necessary to mitigate changes to pH. For instance, hyperventilation decreases the partial pressure of CO_2 in the body and raises the pH of arterial blood, whereas hypoventilation lowers pH. Consequently, changes in respiratory rate can be the source of acidosis/alkalosis. Though slower to compensate, the kidneys can also control pH during acidosis or alkalosis.[1.10] Bicarbonate buffers transport H^+ to the kidneys and transfer the acid to various phosphate buffers in the renal filtrate. These phosphates can then be excreted in the urine. In addition, phosphates also are important intracellular buffers, along with proteins. Together, these buffering systems account for the other 50% of acid removal from the body that is not due to bicarbonate buffering.

Due to high concentrations of protein in the body, the protein buffering system has the greatest buffering capacity and accounts for 60-70% of the total capacity.[1.11] Hemoglobin and albumin are the major circulating protein buffers and can spread the effects of H^+ changes over the entire body. In this way, the overall effect on pH is less in any one tissue.

For any of the buffering systems, the pH is determined by a series of equilibria. When multiple buffers are present in a system, they will be in equilibrium with one another and are tied together by their common reagent, namely H^+. Any change to one buffer will affect the others. The first

determination of blood pH was done by Karl Hasselbalch[1.12], who converted the acid-base equation of Lawrence Henderson to its logarithmic form, now known as the Henderson-Hasselbalch equation.[1.13] This relationship, when applied to relate the pH of blood to constituents of the bicarbonate buffering system (as per equation 1.1) can be written as:

$$pH = pK_{a\,H_2CO_3} + \log_{10}\left(\frac{[HCO_3^-]}{[H_2CO_3]}\right) \qquad (1.2)$$

where $pK_{a\,H_2CO_3}$ is the cologarithm of the acid dissociation constant of carbonic acid. Instead of using H_2CO_3, equation 1.2 may be rewritten in term of the partial pressure of carbon dioxide, pCO_2, to give:

$$pH = pK_{a\,H_2CO_3} + \log_{10}\left(\frac{[HCO_3^-]}{[K_{H\,CO_2} \times pCO_2]}\right) \qquad (1.3)$$

where $K_{H\,CO_2}$ is the Henry's law constant for the solubility of carbon dioxide in blood. Equation 1.3 provides a very useful framework for relating pH to bicarbonate and carbon dioxide concentrations. However, acid-base disorders can also arise from severe imbalances in nonbicarbonate buffering constituents, such as plasma proteins and inorganic phosphates. For example, low albumin levels are associated with an increase in blood pH (alkalosis), whereas high phosphate levels can lead to metabolic acidosis.[1.14] This highlights that pH is dependent on contributions from all the major buffering systems.

Recognizing and managing acid-base disorders is routine practice in critical care settings. Identification of respiratory and metabolic acidosis/alkalosis

may be done by monitoring blood pH.[1.5,1.15] However, current techniques are still mostly invasive: samples are drawn and tested using external pH probes, or blood gas analysers.[1.16] Alternately, pH probes can be embedded for continuous pH monitoring of critically ill patients; this may be done during surgery to monitor potential respiratory acid-base imbalances.[1.17,1.18] In contrast, approaches for determining pH in blood and muscle, such as measurement of exhaled gas concentrations and near infrared (NIR) spectroscopy allow for noninvasive measurements in biological systems.[1.9,1.19,1.20] NIR spectroscopy takes advantage of the 700-900 nm wavelength window in which skin, tissue, and bone are mostly transparent. NIR light passes through skin and subcutaneous fat, whereupon it interacts with strong absorbers such as hemoglobin. The use of multivariate calibration methods (e.g. multilinear regression, principle component analysis, partial least squares regression) has made it possible to extract chemical information from these characteristically complex spectra.

Blood, however, is not the only biofluid of interest to clinicians; the pH of cerebrospinal fluid and amniotic fluid is also an indicator of functional metabolism. Monitoring the pH of cerebrospinal fluid is of clinical importance in the use of ventilation to regulate brain swelling, in tracking adaption to high altitudes, and in the evaluation of brain death.[1.21] However, these measurements

are less routinely performed, since a needle must be inserted into the spine to collect fluid.

Amniotic fluid is also an important fluid. The pH of amniotic fluid offers a simple diagnostic marker for fetal development, as it has been found to decrease from the first to last trimester.[1.22] This decrease in pH can be attributed to an increase in acid bioproducts by the metabolism of the rapidly growing fetus and may provide a means of monitoring fetal metabolism.[1.23] However, collection of amniotic fluid is done by amniocentesis, a medical procedure with some risks to both mother and child.[1.24] A needle is inserted into the abdomen of the mother, passing through the wall of the uterus and into the amniotic sac, where the fluid is collected. Possible complications include injury of the fetus by the needle and risk of infection.[1.25,1.26] Due to the invasive nature of the procedure and the risks posed to mother and child, amniocentesis is not routinely performed; a non-invasive approach would make routine amniotic fluid pH screening feasible.

Constituents of biofluids, particularly the protein components implicated in pH balance, may offer a form of indirect marker for pH. Albumin, for example, is known to undergo a change in conformation with pH.[1.27,1.28] This shift in equilibrium from one form to another is accompanied by changes in viscosity. A detection methodology that could track these viscosity modifications

noninvasively could provide a means to indirectly determining pH in less accessible areas of the body, such as the spine and amniotic sac.

1.2.2 Serum Protein Levels

Determination of protein imbalances in human serum (the fraction of whole blood remaining after cells and clotting factors have been removed) is an important tool for the diagnosis of diseases.[1.29] Serum is composed of thousands of proteins, with concentrations varying over many orders of magnitude (from less than ng/L to g/L).[1.30] Additionally, both relative and absolute protein levels fluctuate in response to specific disease conditions.

Measuring serum proteins individually, especially those present in minute quantities, can be challenging and time consuming. Some common measurement approaches include antibody-binding assays, electrophoresis, and mass spectrometry.[1.31–1.33] Specific recognition of low-concentration proteins requires multiple sample preparation steps and sensitive detection methods to distinguish one protein from thousands of others. In contrast to measuring individual protein levels, inferences regarding health may be made from a total serum protein measurement. Despite the vast number of proteins in serum, 99% of the mass is due to only 20 proteins.[1.34] From this subset, the two most significant proteins are albumin and immunoglobulin G (IgG), comprising 70%

and 14% of total serum protein mass, respectively. Furthermore, many disease states are associated with changes in albumin and/or IgG levels, highlighting the clinical usefulness of a total serum protein test, which is a simple yet powerful diagnostic.

One of the earliest direct methods of protein measurement was by refractometry.[1.35] In 1941, Otto Warburg and Walter Christian reported the use of ultraviolet absorption of tyrosine and tryptophan residues as a means of protein determination.[1.36] Although the UV method requires no reagents, interference from other chromophores, such as nucleic acids, is a concern. Protein determinations have also been done by NIR spectroscopy[1.37], though calibration of the complex spectra requires multivariate techniques. Today, the more popular methods of total protein measurement are colorimetric and include the Coomassie blue method of Bradford, as well as the biuret reaction and its more sensitive variants, due to their greater sensitivity compared to direct methods (i.e. UV, NIR).[1.38–1.41] In these assays, a dye or reagent is added to the protein sample and undergoes a color change upon binding the protein. Limitations of these methods include a low linear range, between 0 and 2 mg/mL, typically requiring dilution of samples before analysis. Additionally, the performance of these colorimetric assays also depends on pH, presence of detergents, and amino acid

composition of the measured protein, often giving subjective concentration values.[1.42]

In addition to total serum protein levels, albumin and immunoglobulin concentrations are also useful in clinical settings. Albumin is the most abundant of the serum proteins with normal levels between 35-50 g/L.[1.34] The many roles of this protein include transport of hormones, fatty acids and other species, regulation of colloid osmotic pressure, and pH buffering. Fluctuations in serum albumin concentrations can be caused by a variety of conditions. Though abnormally high albumin concentrations have been reported, this is typically a relative change observed with dehydration; overproduction of albumin is not known to occur.[1.43] In contrast, lower-than-normal concentrations are associated with a variety of problems including chronic malnutrition, liver cirrhosis, or kidney damage.[1.44] Lower albumin concentrations are also common during pregnancy.[1.45] Dye-binding methods, such as bromcresol green and bromcresol purple, are amongst the most popular approaches for measuring serum albumin levels, as they are simple, fairly specific, and inexpensive.[1.46,1.47] They work very much like the Bradford and biuret reaction methods for total serum protein levels, and share some of the same shortcomings. Dilution of samples is necessary and the presence of interferents, such as uremic toxins, can skew the results of the tests.[1.46] Moreover, the lack of an existing standard reference material for these

measurements gives subjective concentration values.[1.47] Given the clinical significance of albumin, portable tests that can be used in remote settings are much desired[1.48], but are still mostly lacking.

The second most abundant protein in serum is antibodies. Immunoglobulin G is (IgG) the main circulating antibody found in blood and extracellular fluid. Its principle role is to bind viruses, bacteria, and fungi, ultimately protecting the body from infection. The normal reference range of IgG is between 5.7 and 17.7 g/L.[1.34] Elevated levels are associated with chronic infections as well as liver, autoimmune, and parasitic diseases.[1.49,1.50] Reduced levels may be due to immunodeficiency disorders or pregnancy.[1.51,1.52] The most common methods of IgG determination are immunoassays, including enzyme-linked immunosorbent assay, radial immunodiffusion, and turbidimetric immunoassay.[1.53–1.55] In these assays, antibodies specific for IgG are used as the recognition element. Binding results in a color change as measured by a spectrometer, or in agglutination measured visually or by light scattering. Although these tests are very specific, they can take up to 24 hours to perform and require skilled technical personnel to perform. Variants of these methods have also been reported for measuring IgG and albumin in whole blood[1.56,1.57], providing valuable stepping stones towards POC devices for serum protein measurements.

The disadvantages of current methods can complicate clinical analysis in rural settings where access to specialized consumables is limited. For antibody-based assays, cost, storage, and shelf-life of antibodies make these approaches challenging to implement in the field. Moreover, protein assays require dilution steps and are not suited to measuring albumin and IgG levels directly. A direct, reagent-less measurement technique is necessary address these concerns.

1.2.3 Leukocyte Counts

Leukocytes, or white blood cells, are cellular constituents of the immune system that defend the body against infectious pathogens and foreign substances.[1.58] In response to an infection, the number of leukocytes in blood, as well as other biological fluids, increases and these cells attack targets identified by antibodies. An example of a disease characterized by increased leukocytes in milk is mastitis, which has been reported to affect about 20 to 30% of lactating mothers.[1.59]

Mastitis is an inflammation of the breast, typically caused by a blocked duct or bacterial infection.[1.60] The disease is characterized clinically by abnormalities in the mammary gland such as tenderness, swelling, redness, and heat. The leukocyte response in mastitis also contributes to an increase in total cell content of the milk.[1.61] Mastitis is a major cause of reduction in milk

production and approximately 25% of mothers cite it as their reason for early discontinuation of breast feeding.[1.62] Additionally, by altering the cellular composition of milk and local defences within the breast itself, mastitis is a powerful risk factor promoting vertical transmission of infections.[1.62] This also applies to subclinical mastitis, in which infection presents no clinical symptoms.

Subclinical mastitis in humans has been associated with lactation failure, suboptimal growth in infancy, and increased risk of mother-to-child transmission of HIV through breast milk.[1.63,1.64] Early diagnosis is important for successful treatment and extending the period of breastfeeding.[1.65] Flow cytometry has been proposed as a diagnostic tool to allow for fast treatment of mastitis[1.61], though access to this type of testing remains limited in the developing world.

In a laboratory setting, flow cytometry offers the most accurate (±1%) and precise (±5%) white blood cell counts, with the added capability of cell differentiation.[1.66,1.67] In this method, cells are individually funnelled through a narrow opening and counted as they pass a laser. The laser impinges on particles and results in forward- and side-scattered light, which can be used to calculate particle sizes. To distinguish between cells of a similar size, fluorescing probes that bind to the surface of specific cells can be added and measured in addition to the scattering.[1.68] Reference beads of a known concentration are usually added to obtain absolute cell count values. Though well suited to high

throughput screening, the high cost and cumbersome nature of the instrument has limited the use of this technique to laboratories.

A number of portable methods for measuring cell counts have been developed in the dairy industry, where mastitis is the single most costly disease for dairy farmers worldwide.[1.69] The least technical of these methods is a simple, inexpensive, cow-side test known as the California Mastitis Test.[1.70] This test involves mixing the milk with an anionic reagent, which will form a gel if high levels of cells are present. However, the outcome is scored visually and can vary greatly with the individual performing the test. Therefore, as a qualitative test, it cannot be used for counting somatic cells, which is a key factor for determining milk price and quality. For example, in the European Union, milk with somatic cell counts over 4×10^5 cells/mL is deemed unfit for human consumption, with slightly higher thresholds used in the United States.[1.71]

More sophisticated cow-side tests include portable microscope-based cell counters.[1.72] These automated counters rely on disposable cartridges loaded with stains to distinguish cells from non-cellular particles and use image processing techniques to perform cell counts. Label-free cell counting on microfluidic platforms has also been reported.[1.73] This technique uses acoustic waves to separate somatic cells from interfering fat globules, followed by automated microscopy counting. Though this technique does away with the need for any

sample preparation, it can only measure relatively high cell concentrations (>10^6 cells/mL) and cannot provide information for cell counts associated with subclinical mastitis (~3 × 10^5 cells/mL).

Cell counting by NIR spectroscopy also holds much promise, since it requires no disposables and may be incorporated into milking robots for continuous on-line monitoring applications.[1.74,1.75] No sample preparation is needed, though careful calibration is necessary to obtain meaningful estimates. Despite large compositional differences between human and cow milk[1.76], it should be possible to adapt the NIR approach for cell counting in human milk samples.

1.2.4 Oxidative Stress

In recent years, the implication of oxidative stress in a number of diseases has been recognized by both clinicians and researchers. Oxidative stress refers to an increase in pro-oxidant species relative to antioxidant species, with potentially damaging effects. Oxidative stress has been linked to neurodegenerative diseases, cardiovascular diseases, autoimmune diseases, pulmonary disorders, various cancers and tumors, diabetes, liver disease, and aging, among others.[1.77] Oxidative stress is also known to increase during other physically stressful situations, such as acute exercise and pregnancy[1.78,1.79],

where significant changes in levels of lipid hydroperoxides and malondialdehyde, species resulting from lipid peroxidation, have been measured. Oxidative stress can result from abnormally high concentrations of reactive oxygen species (ROS) or deficiencies in antioxidant defences.

ROS, which include free radicals and their non-radical intermediates, are products of normal cellular metabolism. An example of this is the generation of superoxide radicals during oxidative phosphorylation, the process in which adenosine triphosphate (i.e. the major energy currency of cells) is produced. The superoxide radical may be converted to hydrogen peroxide, which in turn can react with iron (II) via the Fenton reaction to give the highly reactive hydroxyl radical.[1.80] Additionally, exposure to UV light, or ionizing radiation, and toxicity by various metals and xenobiotics are linked to production of ROS.[1.81] On the whole, ROS are known for providing beneficial effects at low to moderate concentrations (e.g. part of normal cell signaling), and harmful effects at high concentrations (e.g. damage to lipids, proteins, or DNA).[1.82]

A variety of approaches have been developed to measure either ROS or oxidative stress modifications. However, the reactive nature and short lifetimes of ROS can make them difficult to measure directly, so most approaches rely on trapping ROS for quantification. Fluorescence probes have been developed for imaging ROS, the culprits of oxidative modifications, *in vivo*.[1.83,1.84] These probes

scavenge or react with ROS, and change the fluorescing properties of the sensor. Though fluorescence offers a very sensitive, real-time platform for monitoring ROS, investigations are limited to the laboratory.

Alternately, oxidative stress can be monitored by examining modifications caused by ROS. In proteins, for example, oxidative stress modifications manifest themselves in a multitude of ways: formation of protein hydroperoxides, hydroxylation of aromatic groups and aliphatic amino acid side chains, nitration of aromatic amino acid residues, oxidation of thiol groups and methionine residues, carbonylation of amino acid residues, cleavage of polypeptide chains, and formation of cross-linking bonds.[1.85,1.86] Particularly, the oxidized and reduced forms of glutathione offer a dynamic indicator of oxidative stress of an organism, as this amino acid is an important antioxidant, and have been determined by HPLC[1.87], capillary electrophoresis[1.88], and biochemical methods.[1.89] Alternately, an increase in whole blood viscosity has been reported as an indication of hyperglycemia-induced oxidative stress, as well as exercise-induced oxidative stress.[1.78,1.90] Since testing for specific oxidative stress modifications of proteins is time-consuming, simple diagnostics based on changes in biofluid viscosity may be useful for quick screening of diseases.

1.3 Principles of Ultrasound

This thesis deals with the ultrasonic properties of materials. To understand these ultrasonic properties, it is useful to review the historical and fundamental aspects of acoustics and ultrasonics, as well as the sensors and techniques that allow for these measurements.

1.3.1 A Compressed History of Acoustics

From the low rumblings of earthquakes (infrasound) to the high, inaudible cries of the bat (ultrasound), the world is filled with vibrational disturbances (Figure 1-1). The earliest suspicion that sound is a wave was based on observation of water waves. These oscillatory disturbances propagate away from their source and transport no discernible amount of matter over large distances. The notion that water and sound behave analogously was proposed by the Greek philosopher Chrysippus (c. 240 BC), by the Roman architect and engineer Vetruvius (c. 25 BC), and by the Roman philosopher Boethius (AD 480-524).[1.91] Aristotle (384-322 BC) understood that in order for sound to propagate from one place to another, the air had to move in some way, and that motion is generated by a source, "thrusting forward in like manner the adjoining air, to that the sound travels unaltered in quality as far as the disturbance of the air manages to reach." [1.92]

Some of the earliest known acoustics research was done by Greek philosopher and mathematician, Pythagoras (c. 550 BC), with his investigation of vibrating strings. He noted that a vibrating string, when divided into simple ratios of its length, produced consonant musical intervals, otherwise known as harmonics.[1.91] This concept of harmonics will be particularly important when considering nonlinear phenomena of ultrasound in later sections.

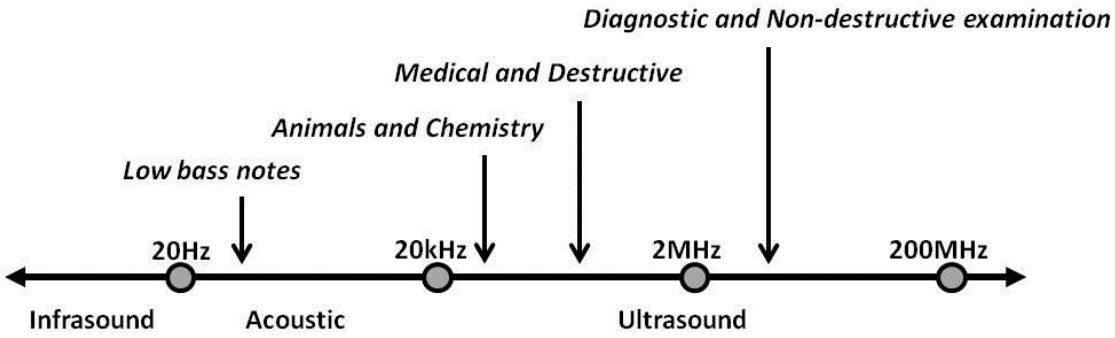

Figure 1-1: The sound spectrum.

These early observations were further pursued towards the end of the Renaissance era, as Galileo Galilei (1564-1642) reviewed the relationship of the pitch of a string to its vibrating length, and related the number of vibrations per unit time to the pitch.[1.93] At the same time, French natural philosopher Marin Mersennes (1588-1648) independently discovered the laws for the frequency of

oscillation of a stretched string, which are described in his *Harmonie Universelle* (1637).[1.94]

In 1640, Robert Boyle confirmed Aristotle's wave description with an experiment in which he observed a decrease in sound intensity as he pumped out the air in a vacuum chamber, demonstrating that sound needs a medium through which to propagate.[1.91] This fundamental condition distinguishes acoustics from optics as sound is a mechanical, rather than electromagnetic, wave motion.

The mathematical underpinnings of sound propagation were tackled by Isaac Newton (1642-1727), as he was the first to attempt to calculate the speed of sound in air. This is outlined in his *Principia* (1687)[1.95], which included a mechanical interpretation of sound as being pressure pulses transmitted through neighboring fluid particles.

A great number of theoretical and experimental contributions on acoustics came from Lord Rayleigh (1842-1919). His notable contributions included the development of an arrangement, known as a Rayleigh disk, which is used to measure sound intensity. Rayleigh also predicted the existence of waves that propagate on the plane surface of an elastic solid.[1.96] These Rayleigh waves can be produced in materials by a localized impact, or by piezoelectric transduction to be used in surface acoustic wave (SAW) filters and sensors.

The discovery of the piezoelectric effect in crystals by Jacques and Pierre Curie (1880) led to significant advancements in acoustic research, including access to more powerful sources of ultrasound.[1.97] They found that applying an electric potential to certain natural crystals resulted in a change of thickness and, furthermore, by applying an alternating electric potential to piezoelectric materials, they could be made to oscillate.

In 1917, Paul Langevin (a student of Pierre Curie), applied the piezoelectric effect to the generation and detection of ultrasound waves for locating submarines.[1.98] This stimulated a great interest in piezoelectric devices, and paved the way for significant developments in acoustic research. These developments, particularly for analytical and biomedical applications, will be considered in greater detail in subsequent sections.

1.3.2 Ultrasound Wave Properties

Sound is a mechanical wave that travels through a given medium (gas, liquid, or solid) by oscillations of the medium itself. In acoustics, waves are typically described by variations in pressure in the medium induced by the propagating wave. A harmonic sound wave can be described in simple terms by the wave equation

$$y(t) = Asin(2\pi ft + \varphi) \qquad (1.4)$$

where A is the amplitude, f is the frequency, and φ is the phase offset.[1.99] The wave function $y(t)$ can be interpreted in terms of the displacement of particles from equilibrium, or as the change in pressure, density, or particle velocity at a given position and time. The wavelength of a wave (λ), as shown in Figure 1-2, can be expressed in terms of the velocity (c) and frequency (f)

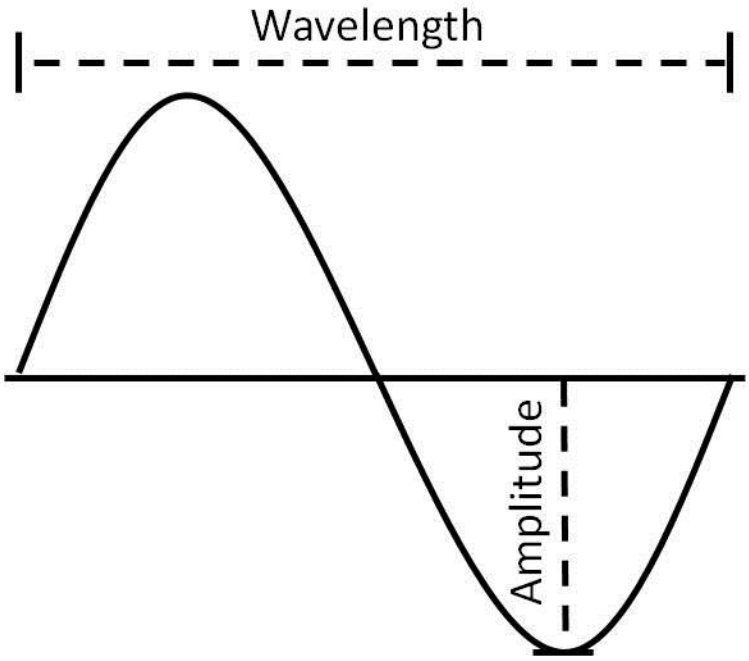

Figure 1-2: Two characteristic features of a sound (pressure) wave. The wavelength denotes the distance for one complete cycle of compression and rarefaction. The amplitude describes the intensity of either the positive or negative pressure maximum, relative to the average local pressure.

$$\lambda = \frac{c}{f} \qquad (1.5)$$

and spans the length of one complete cycle of rarefaction and compression.[1.100] The velocity of the sound wave depends on the density and elasticity of the medium through which it travels. In fluids such as air or water, velocity of a wave can be shown to be given by the Newton-Laplace equation:

$$c = \sqrt{\frac{B}{\rho}} \qquad (1.6)$$

where B is the bulk modulus and ρ is the equilibrium density of the medium.[1.101] As the bulk modulus, also termed coefficient of stiffness, increases so does the speed of sound. Consequently, the physical state of a medium will greatly influence the wave velocity. The speed of sound in air, for example, is much smaller than in water or aluminum.[1.99,1.102] This trend is presented in Table 1-1, which lists various media and their corresponding sound velocity values.

A harmonic wave is characterized by a single frequency and wavelength. Other waves, such as square waves or sawtooth waves, are superpositions of harmonic waves, where two or more waves combine linearly:

$$y(t)_{tot} = y(t)_1 + y(t)_2 \qquad (1.7)$$

This principle of superposition provides a powerful tool for wave analysis.[1.103] Any complex wave signal can be represented mathematically as the sum of simple waves (sine and/or cosine waves) through the process of Fourier decomposition.

The significance of this will be highlighted in subsequent sections dealing with broad-band pulse signal analysis.

Table 1-1: Sound velocity in selected media[1.99,1.102]

Medium	Velocity (m/s)	Density ($kg \bullet m^{-3}$)	Bulk Modulus ($kg \bullet m^{-1} \bullet s^{-2}$)
Air (25°C)	346.1	1.1839	1.42×10^5
Water (25°C, distilled)	1498	997.048	2.2×10^9
Human Tissue (soft)	1540	1060	2.5×10^9
Glass (Pyrex)	5640	2230	3.5×10^{10}
Aluminum (rolled)	6420	2700	7.6×10^{10}

1.3.3 Modes of Propagation

In solids, sound waves can propagate in four principle modes: transverse, longitudinal, surface, and plate waves.[1.99] Transverse waves require a solid material for effective propagation, but can also be observed in highly viscous liquids. Transverse waves are characterized by particles moving perpendicular to the direction of energy displacement. The particles oscillate up and down about their individual equilibrium positions as the wave passes, as shown in Figure 1-3a. An example of this is a beach wave moving towards the shore, with the water molecules bobbing up and down as the wave passes.

In contrast to transverse waves, longitudinal waves have particle displacements parallel to the wave direction, as seen in Figure 1-3b. Longitudinal waves are the most common wave form found in liquids and the only wave mode observed in gases. These waves are characterized by compressed and rarefied regions. The particles do not move indefinitely with the wave; rather, they oscillate back and forth about their individual equilibrium positions. For example, sound waves moving through air feature this type of wave propagation.

In solids, surface waves such as Rayleigh waves, are possible. These waves combine transverse and longitudinal motion and cause surface particles to move in ellipses. When Rayleigh waves are guided in layers they are known as Lamb waves.[1.104] These surface acoustic waves (SAWs) can couple with any

media in contact with the surface. Since this coupling strongly affects the amplitude and velocity of the wave, SAWs are used in a wide range of applications, including physical, chemical and bio-sensing.[1.105]

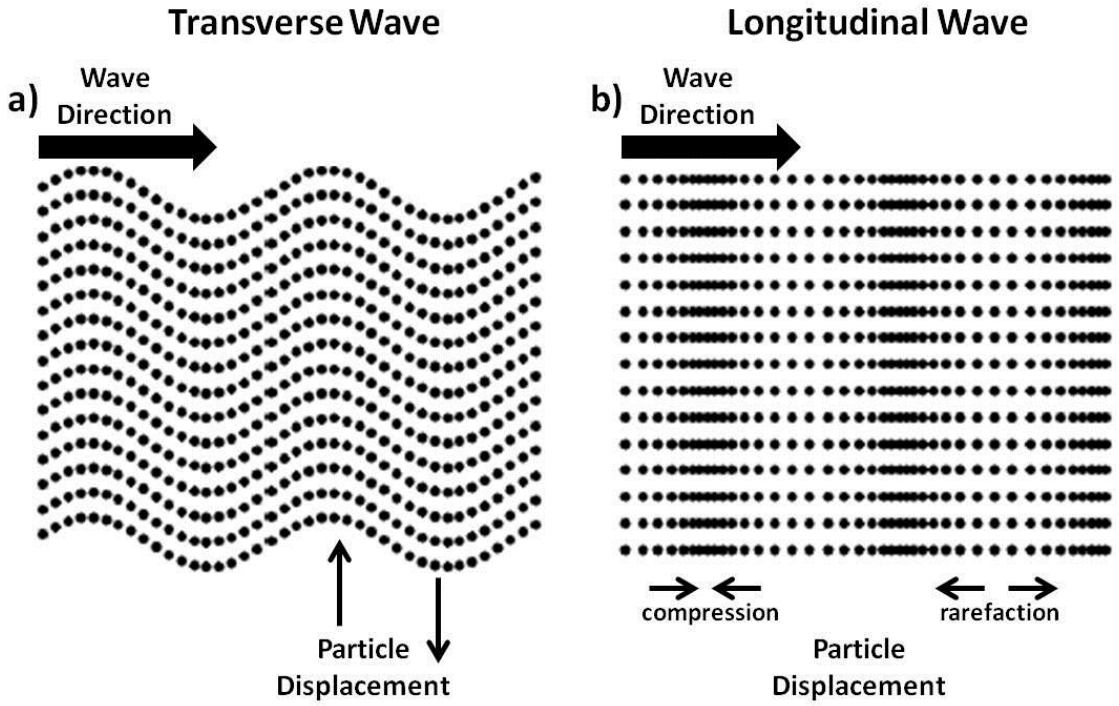

Figure 1-3: a) Transverse waves are characterized by a particle motion perpendicular to the wave direction. b) Longitudinal waves feature particle compression and rarefaction parallel to the wave direction.

1.3.4 Mechanisms of Ultrasound Attenuation

As an ultrasound wave travels through a sample, the amount of energy dissipated is related to the physical and chemical properties of the medium. The principle mechanisms of ultrasound attenuation are absorption, reflection, refraction, and scattering of the waves[1.101], as depicted in Figure 1-4. The absorption component may be further divided to include contributions from thermal conductance effects, viscous effects, chemical effects, and nonlinear distortions.[1.106] In liquids, the viscous force (internal friction) between particles moving with different velocities during compression and rarefaction contributes significantly to acoustic wave absorption. The amplitude of the wave decreases exponentially and is described by

$$A = A_0 e^{-\alpha d} \qquad (1.8)$$

where A_0 is the initial amplitude, α is the attenuation coefficient, and d is the distance the wave has travelled.[1.99] In general terms, the attenuation coefficient may be written as

$$\alpha = a f^b \qquad (1.9)$$

where f is the frequency of the sound wave, and a and b are medium-specific constants.[1.107] The medium-specific parameters are obtained by fitting experimental data, with b between 0 and 2. Acoustic attenuation in water and

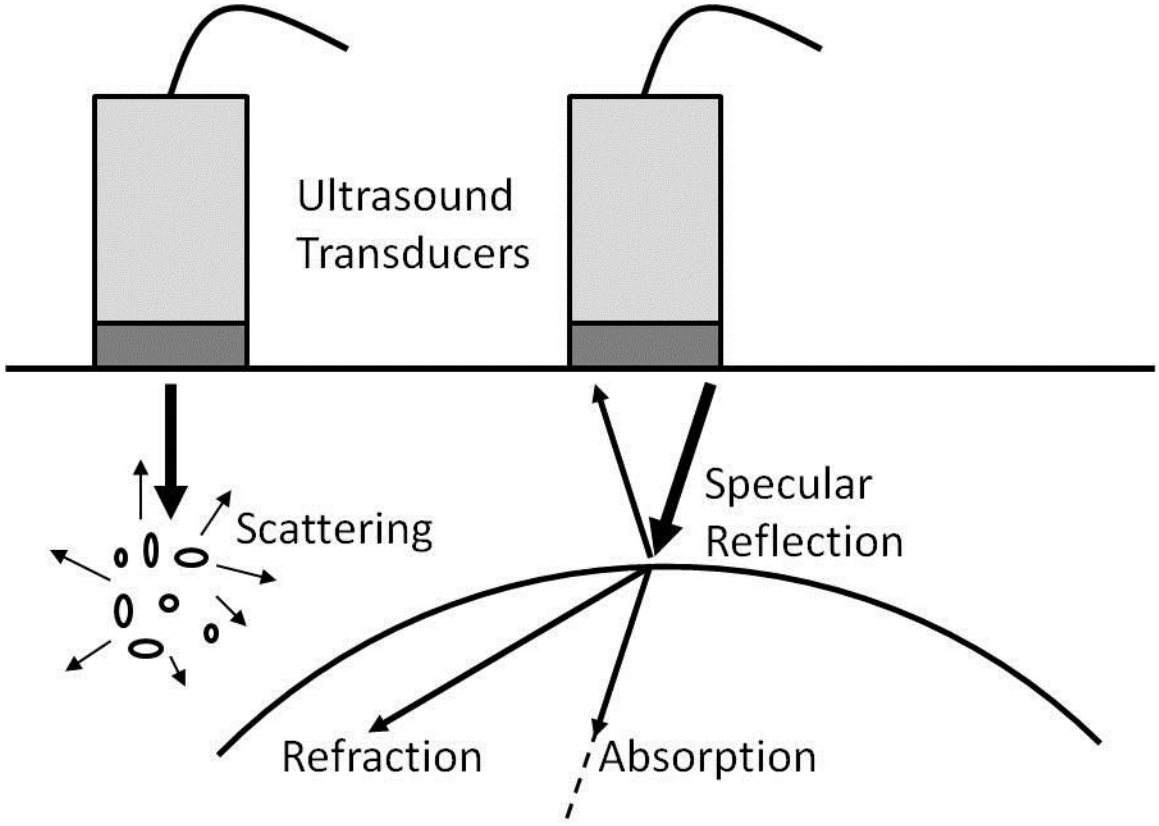

Figure 1-4: Mechanisms of attenuation of ultrasound waves.

many metals have a quadratic frequency dependence, where b = 2. In contrast, the parameter b of soil and rock is around 1, and between 1 and 2 for most soft tissues.[1.108] The frequency dependence of the attenuation coefficient is important for determining which pathlengths can be used for measurements. For example, a 3 MHz wave has a maximum penetration depth of 20 cm, while a 20 MHz wave has a maximum penetration depth of 3 cm, where the penetration depth refers to the distance the wave travels before being completely attenuated. Therefore, in order to measure higher frequencies, it is often necessary to reduce the pathlength of the measurement.

For homogeneous samples, attenuation is predominantly due to absorption phenomena. However, in heterogeneous samples, contributions from reflection and refraction must also be considered. These phenomena, which occur at the interface of two media, can be observed in systems such as colloidal suspensions or tissue interfaces. Note that reflection can be specular (unscattered) or diffuse (scattered).[1.109] Specular reflection takes place when the inhomogeneity is much larger than the wavelength of ultrasound and is characterized by a simple geometric relationship (angle of incidence, θ_i = angle of reflection, θ_r). When the inhomogeneity is the same size or smaller than the ultrasonic wavelength, scattering takes place and the wave deviates from its original path.

Both reflection and refraction are dependent on the characteristic acoustic impedance, Z_0, which is an inherent property of a medium:

$$Z_0 = \rho_0 c_0 \qquad (1.10)$$

where ρ_0 is the density, and c_0 is the speed of sound of the unperturbed medium.[1.101] The part of the energy that is reflected at the interface of two media can be described by the reflection coefficient

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \qquad (1.11)$$

where $Z_1$and Z_2 are the impedances of the two media.[1.100] The transmitted portion of energy from the first to second medium is given by $T = 1 - R$, since all the energy must be divided between either reflection or transmission at the interface.[1.99] To maximize signal transmission and reduce the effects of impedance mismatching, a third layer with acoustic impedance

$$Z_{match} = \sqrt{Z_1 Z_2} \qquad (1.12)$$

can be sandwiched between the two media.[1.110] An example of this is the use of a water-based gel as a coupling agent between probe and tissue in medical ultrasonography. Without the impedance-matched gel, small air pockets between the probe and tissue can significantly decrease the transmitted energy.

When the angle of incidence of an ultrasound wave at the interface of two media is oblique, both reflected and refracted waves are produced (Figure 1-5). The direction of propagation of the transmitted wave is determined by Snell's law:

$$\frac{\sin\theta_i}{c_i} = \frac{\sin\theta_t}{c_t} \qquad (1.13)$$

where θ_i and θ_t are the angles of the incident and transmitted waves respectively.[1.99] Though there is no refraction at normal incidence, it must be considered for *in vivo* applications, where internal structures and multiple phase boundaries may result in significant refraction.[1.111] For focused ultrasound applications, adjustments may be necessary to compensate for refraction phenomena and ensure that the ultrasound wave reaches its target.

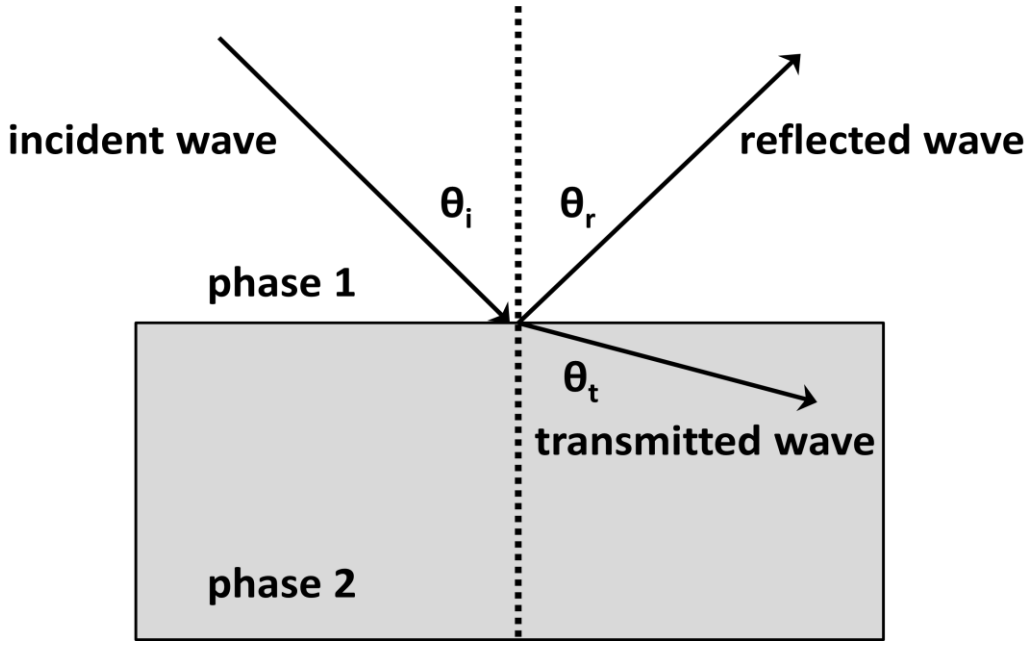

Figure 1-5: Reflection and refraction of sound waves at phase boundaries.

1.4 Ultrasound Instrumentation

The key components of ultrasound instruments include a source and a detector. The source typically consists of a signal generator and a transduction element that can convert the electrical signal into a pressure wave.[1.109] The magnitudes and frequencies of the ultrasonic pressure waves generated are then measured by an acoustic spectrometer, typically consisting of a digital oscilloscope. Single frequency waves can be sent through a sample sequentially and the magnitudes recorded. Alternately, an impulse that generates a distribution of frequencies can be sent through a sample, giving a complex signal over time. In these instances, it is useful to convert the signal from the time domain into frequency components by applying a Fourier transform to the data.

Since the discovery of the piezoelectric effect by the Curie brothers, most ultrasound systems use piezoelectric transducers to both generate and detect ultrasonic waves. Piezoelectric materials develop a potential difference across the surface of the faces to which mechanical stress is applied. Conversely, these materials undergo physical deformation (expansion or contraction) from an applied electrical field. For example, one commonly used material for the active element in ultrasound transducers is lead zirconate titanate (PZT).[1.112] This crystalline material has a unit cell with a fixed dipole moment, which is a key property for piezoelectrics. On a macroscopic level, the unit cells may be

randomly oriented, giving a net polarization of zero. Under these conditions, the piezoelectric effect exhibited will be negligible. It is therefore useful to introduce an artificial anisotropy to the material, such that most dipoles will be more or less aligned in the same direction. These dipoles can be aligned in a process known as poling, in which the material is heated above its Curie temperature—the temperature at which dipoles become randomly oriented.[1.113] A strong electric field is then applied to line up the dipoles, with the material lengthening in the direction of the field. Once the material is cooled and the external field removed, most of the dipoles remain roughly aligned in the same direction, as long as the material remains below its Curie temperature. Applying an external electric field to the poled material will cause it to undergo a change in volume as the unit cells try to orient themselves with the new field.

Though piezoelectric ceramics are widely used as ultrasound transduction elements, flexible piezoelectric polymers, such as polyvinylidene fluoride (PVDF), show promise for creating unique transducer configurations that would be difficult to achieve with brittle ceramics. However, their use is limited to lower temperature regimes, as their piezoelectric activity irreversibly decreases above 80°C.[1.114]

Optical methods can also be used to produce ultrasonic waves in solids. A pulsed laser is fired at an optically absorbing layer, generating ultrasonic waves

through thermal expansion and contraction of the material. The detection of these waves may also be done optically, through the use of an interferometer.[1.115] One possible configuration is a Fabry-Perot interferometer, in which a transparent polymer layer several microns thick is sandwiched between two partially reflecting gold mirrors. Ultrasound waves passing through the polymer induce changes in its thickness and modulate the distance between the two mirrors. A laser beam incident on this sensor configuration creates constructive and destructive interference patterns as the distance between mirrors changes. The measured change in mirror displacement can then be used to extract amplitude and phase information. This approach provides a useful means for non-contact ultrasound measurements, especially for probing high temperature systems above the Curie point of piezoelectric transducers.[1.116]

Typical ultrasound instrument configurations use either one or two transducers to generate and detect ultrasonic waves. The single transducer configuration relies on signal reflecting off the back of the sample or sample cell (Figure 1-6a). Therefore, a large impedance mismatch is desirable to maximize the reflected sound energy. An advantage of this configuration is that it is easy to change the position of the transducer to map out a given area of the system under scrutiny. However, the effective pathlength is doubled; this determines

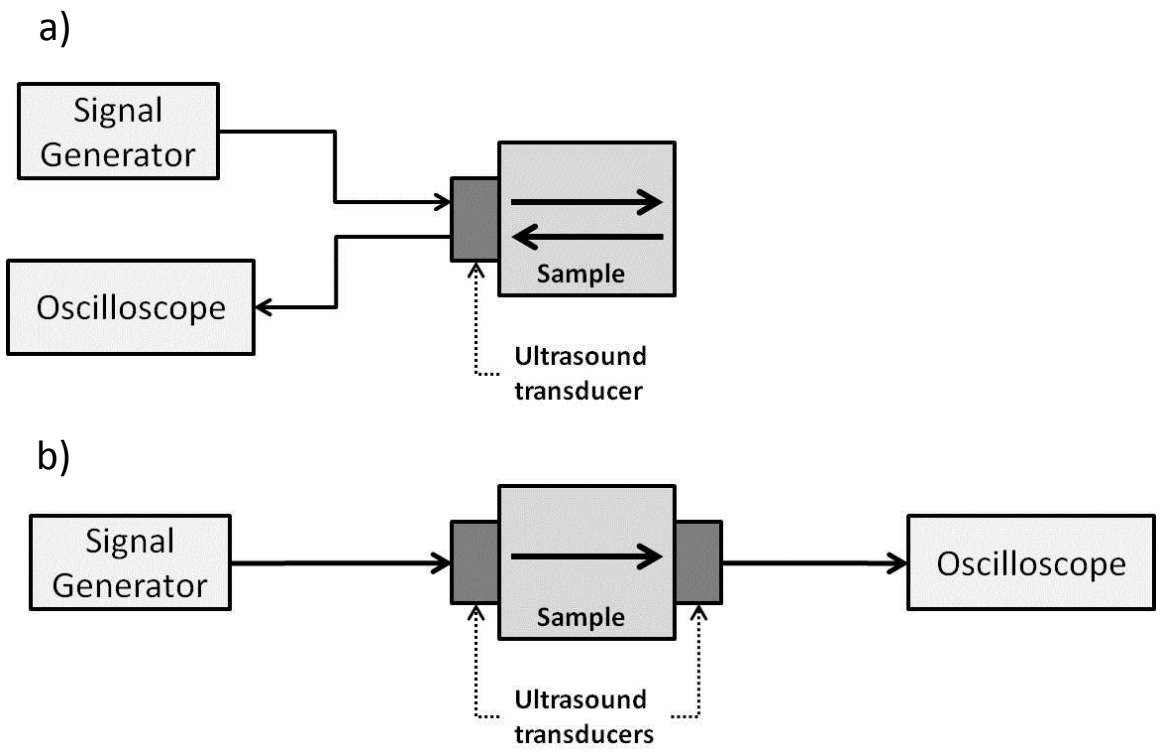

Figure 1-6: Ultrasound instrument configurations. a) A single transducer is used to transmit and receive signal. b) Dedicated transmitting and receiving transducers are used. Both configurations use a signal generator to provide the electric signal impulses to the transmitted and a signal analyzer as the detector of received impulses.

the penetration depth and may need consideration in highly attenuating samples where the measured signal may be very small.

In contrast, two transducers may be used in a through-transmission configuration, with a dedicated transmitter and receiver (Figure 1-6b). This configuration is better suited to measurements that do not require the transducers to be repositioned, such as analysis of liquid samples.

Whether a single or double transducer configuration is more appropriate depends on the type of ultrasonic signal used to probe the system of interest. Typical signals include continuous waveforms, tone-burst pulses, and broad-band pulses (Figure 1-7).[1.117] Continuous wave signals can consist of a single-frequency sinusoidal wave and are used to generate standing waves that are subject to constructive and destructive interference, as the pathlength or frequency of the wave is varied.[1.118,1.119] This interferometric approach yields phase and amplitude values that allow for the determination of ultrasonic velocity (phase) and attenuation (amplitude).

In addition to continuous wave signals, pulsed signals such as tone-burst pulses may also be used to determine ultrasonic velocity and attenuation profile of a sample by measuring the time-of-flight and amplitude of the pulse that has traversed the sample. A tone-burst pulse consists of a few cycles of a single frequency, such that the sample is probed by only one frequency at a time. This

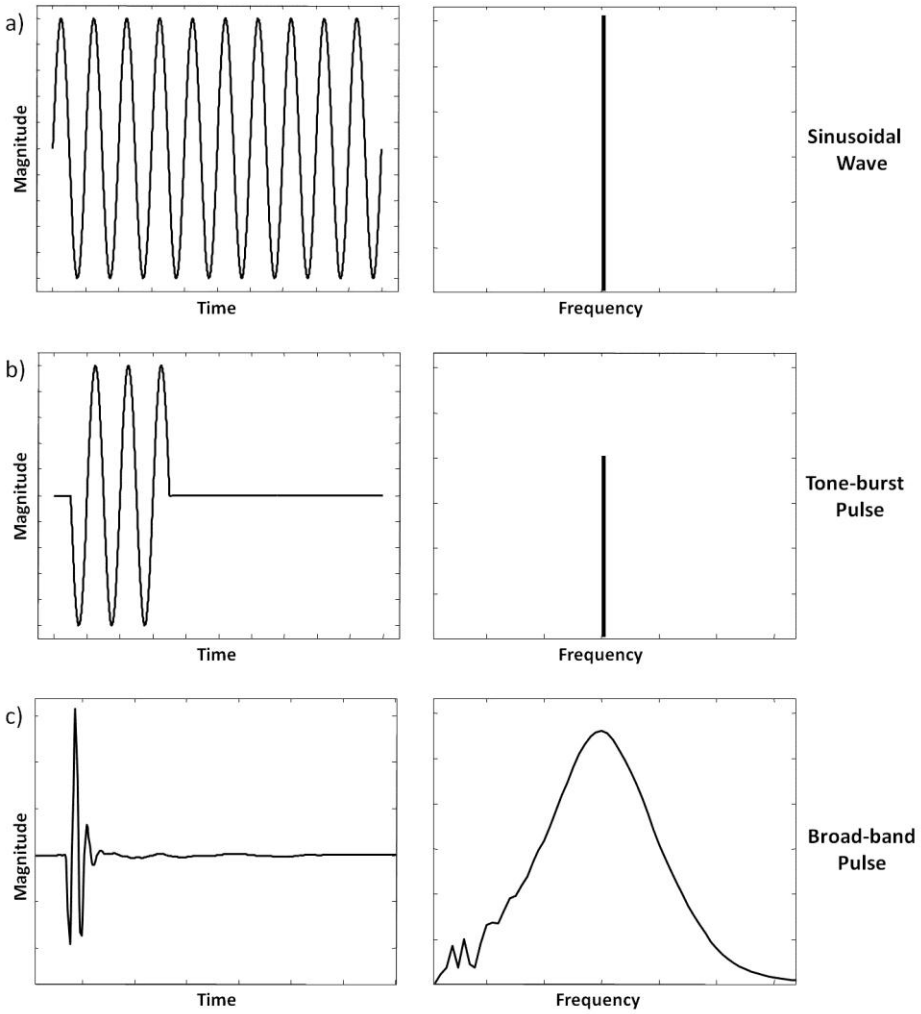

Figure 1-7: Three common wave signals used in ultrasound measurements. a) The continuous sinusoidal wave (left) is characterized by a single frequency component (right). b) The tone-burst pulse (left) is very similar to the continuous wave signal in that it also has only one frequency component (left), but with the distinction that only a few cycles of the wave are sent through as a pulse. c) The broad-band pulse (left) contains a distribution of frequencies (right) that are characteristic of the frequency bandwidth of the transducer.

technique is typically used with a variable pathlength cell, which allows for measurements to be made over a very large frequency range while maintaining a good signal-to-noise ratio. However, the limitation of single frequency signals is that sequential scanning through a range of frequencies can be time-consuming.

In contrast to the single frequency tone-burst pulse, a large negative impulse can be used to generate broad-band signal that contains a wide range of frequencies. When the impulse hits the transducers, similar to a drum stick striking a drum membrane, the piezoelectric device oscillates and exponentially decays to zero, giving a distribution of frequencies. After this frequency-rich pulse has travelled through the sample, the time domain signal is saved and subject to a Fourier transform to yield frequency information. This allows for the collection of data over a range of frequencies with a single measurement.[1.120,1.121] The frequencies that can be observed will ultimately depend on the central frequency and bandwidth characteristics of the transducers used.

1.5 Ultrasound in Analytical Chemistry

Ultrasound was first used for chemical detection by Robert Wood and Alfred Loomis in 1927, who investigated differences in ultrasonic velocities in liquids, solutions, and mixtures. [1.122] Since that time, two principle ultrasound strategies have emerged. Some ultrasound devices relate changes in resonance

frequency of an oscillating sensor to mass of analyte bound to the sensor. However, these ultrasonic resonator sensors must be in direct contact with the medium under investigation in order for mass-exchange to occur. Alternatively, speed of sound and attenuation measurements have been used for elucidating chemical properties in a variety of systems. This approach does not require direct contact with the sample (i.e. measurements can be made through a coupling medium) and offers the ability to probe undiluted samples. To understand the context of ultrasound in analytical chemistry, it is useful to review these applications in more detail.

1.5.1 Chemical Detection with Ultrasonic Resonator Devices

One of the early applications of piezoelectric devices as sensing elements came in the form of quartz crystal microbalances (QCM). The resonance frequency of a vibrating crystal was known to shift with a change in mass, as shown by Lord Rayleigh.[1.123] In 1959, Günter Sauerbrey described a linear relationship between the decrease in frequency of an oscillating quartz crystal, Δf, and the mass loading of deposited metal[1.124]:

$$\Delta f = -\frac{f_0 \cdot \Delta m_s}{d_q \cdot \rho_q \cdot A_q} \qquad (1.14)$$

where f_0 is the natural resonance frequency of the quartz crystal sensor, Δm_s the change in mass of the sensor, d_q the thickness of the quartz crystal, ρ_q the

density of the quartz crystal, and A_q the area of the quartz sensor. These quartz resonators are bulk wave sensors, where the ultrasound waves propagate through the entire crystal. Consequently, the resonance frequency of the crystal is dependent on its thickness, as presented in equation 1.14. Typical QCM frequencies lie between 5 and 30 MHz; manufacturing higher frequency crystals with improved sensitivity remains a challenge, as this necessitates thinner and more fragile substrates.[1.125]

The first QCM sensors were used to measure the mass binding of gas-phase species to the crystal surface and were used for quantifying moisture, volatile organics, and environmental pollutants.[1.126,1.127] The surfaces of these sensors could be modified with recognition elements, such as antibodies or sensing films, allowing for analyte-specific binding.[1.128,1.129] This extended to measurements in solution, with changes in frequency related to changes in viscosity and density in liquid media.[1.130] QCMs have also been investigated as a tool for studying thin polymer films and complex biomolecular systems.[1.125]

Variants of piezoelectric effect sensing devices include SAW devices. Unlike QCMs, the thickness of the substrate does not influence the detection performance of SAW sensors, as the acoustic waves only propagate in a guiding layer along the surface of the substrate. These sensors are used in the frequency range of 25-500 MHz and respond to changes in mass, density, viscosity, and

acoustic coupling phenomena, though sensor fouling remains problematic for both SAW and QCM sensor. [1.105,1.131]

1.5.2 Chemical Detection with Speed of Sound and Attenuation Measurements

Ultrasonic velocity and attenuation measurements offer the potential to probe the molecular properties of matter. While electromagnetic radiation interacts directly with a medium through coupling with electrons and nuclei, mechanical ultrasound waves interact indirectly through vibrational perturbations of the weak bonds between molecules.

One consideration of ultrasound is that waves can be subject to nonlinear phenomena at sufficiently high acoustic pressures. Linear propagation of ultrasound waves assumes the density of a medium changes proportionally with acoustic pressure. However, the linear assumption holds for media in which the pressure amplitude of a propagating wave is sufficiently small to cause weak perturbations in density. Therefore, this section considers ultrasound measurement techniques based on linear wave behaviour.

The speed of sound in liquids depends on both density and compressibility, with changes in viscoelastic properties related to sample composition.[1.119] Consequently, measurements of ultrasonic velocity can be used to determine sample composition in liquid binary mixtures[1.132,1.133] as well as gas

binary mixtures.[1.134,1.135] However, because of the inter-relationship of properties, one velocity measurement is often insufficient for quantification. One example of this is ethanol-water mixtures, where non-monotonic behaviour results in the same velocity for both higher and lower fractions of ethanol. In systems such as this, additional parameters are required. To solve this, researchers have used temperature, concentration, pH, ionic strength, and solvent composition changes to add extra dimensionality to the data that allows for quanification.[1.121] However, this increases the data collection time and the complexity of data analysis.

Ultrasonic spectroscopy has been particularly useful in providing information about proteins in solution. For example, ultrasonic velocity measurements, when combined with density data, allow for the calculation of the hydration of globular proteins.[1.136] Protein conformations have also been studied by looking at ultrasonic absorption at varying pH levels.[1.137,1.138] Ultrasound absorption is sensitive to conformation transitions of myoglobin, bovine serum albumin, and other globular proteins that occur in response to pH changes. One of the main mechanisms of acoustic absorption in protein solutions is attributed to interaction between solvent (i.e. water) and solute (i.e. protein).[1.139] Ultrasound waves perturb the equilibrium distribution of solvent molecules that are weakly bonded to the solute. The perturbation occurs at the expense of acoustic wave energy, resulting in absorption. Ultrasonic absorption by globular protein

solutions have also been found to be linearly dependent on protein concentration.[1.140] Accordingly, ultrasound attenuation profiles can give information about both protein conformation and concentration.

Though more difficult than homogeneous systems, attenuation measurements are also possible in heterogeneous systems, such as emulsions and suspensions.[1.141,1.142] Unlike homogenous systems, in which attenuation is attributed largely to absorption, scattering effects also need to be considered. Determination of particle sizes by ultrasound has also been demonstrated, and allows for the study of concentrated samples, with weight fractions up to 50%.[1.109,1.143,1.144] This property of ultrasound allows for measurements of systems without the need for dilution and makes it particularly appealing for the study of intact biological systems, in which the concentration of cellular, protein, and fat constituents can be very high.[1.145] The minimal sample preparation required with this technique opens up the possibility of using ultrasound for non-invasive measurements of biofluids.

1.6 Medical Acoustics

The non-invasive appeal of ultrasound has been used in a variety of medical applications. One of the early applications of ultrasound in medicine was the therapeutic heating of deep tissue, first demonstrated in 1932 by Hebert

Freundlich *et al*. on bone and muscle tissue.[1.146] At acoustic intensities higher than those used in ultrasound imaging, ultrasound energy absorbed by tissue is converted to heat and has spawned a variety of therapeutic applications, including targeted drug delivery,[1.147] trans-dermal drug delivery,[1.147,1.148] destruction of kidney stones,[1.107] and treatment of solid tumors.[1.149] Ultrasound is also used to measure bone strength and assess fracture risk for osteoporosis, and in the monitoring of fracture healing.[1.104] Use of ultrasound by physical therapists for reducing pain and promoting soft tissue healing is widespread, though this treatment is not supported by effectiveness studies.[1.150,1.151]

Medical ultrasound imaging does not typically measure the speed of sound, but rather interprets the reflected signal in terms of its amplitude or time-of-flight to map spatial variations.[1.152] Ultrasound imaging, commonly associated with prenatal care, has several advantages over other imaging methods, such as computed tomography or magnetic resonance imaging. Ultrasound equipment is an ideal imaging tool due to its portability, relatively low cost, real-time imaging capabilities, and established safety at low intensities (no ionizing radiation).[1.153] Modern systems are even available for home-use and can be operated by non-experts. However, although ultrasound is deemed a safe tool, there is still concern that its use in fetal imaging may put the fetus at some risk, as cells are more susceptible to external influence during periods of rapid division.[1.154]

All anatomic ultrasound imaging systems use pulses, and are commonly known as pulse-echo backscatter systems. They use frequencies in the range of 1-30 MHz and work by sending pulses of ultrasound into tissue and listening for backscattered echoes from various depths, typically with one transducer.[1.153] The depth of a reflector is gauged by the time-of-flight between the transmitted pulse and the received echoes. Ultrasound reflections result from differences in acoustic impedance at tissue interfaces. A reflecting moiety that is far away from the transducer will have a longer time-of-flight than a reflector that is close to the transducer. The axial (depth) resolution is frequency dependent, whereas the lateral resolution is mostly dependent on the spot size of the ultrasound beam.[1.100] Though higher frequencies provide better axial resolution, they cannot penetrate as far into the body, such that lower frequencies are used for deep tissue imaging.[1.104]

Early applications of ultrasound used motion-mode imaging (keeping the transducer in a fixed location) to monitor depth change over time to track movement of the heart and other vascular structures.[1.155] Today, modern ultrasound systems use arrays of transducers to form two-dimensional images in real-time, an alternative to moving a single ultrasound transducer to point in various directions to generate an image.[1.156] Real-time ultrasound has allowed physicians to effectively visualize areas of interest and perform image-guided

biopsies and amniocentesis.[1.100,1.157] This has greatly reduced the risks associated with invasive procedures for collection of tissue or fluid.

In addition to imaging, ultrasound can be used for hemodynamic investigations. These studies of fluid flow take advantage of the Doppler Effect[1.158], where the frequency of a wave is changed by movement of the sound source relative to the detector. Consequently, continuous wave ultrasound can be used to determine speed, direction, and flow characteristics of blood in arteries and veins.[1.159] A fixed-frequency continuous wave impinging on a blood vessel will backscatter, resulting in a frequency-shifted echo from blood moving through the vessel. After measurement by the receiving transducer, the new frequency is mixed with a pair of phase-shifted reference waves at the same frequency as the transmitted wave. The resulting difference frequency, along with the phase information, allow the Doppler system to distinguish between echoes with increased frequency due to blood flowing towards the receiver and those with decreased frequency due to blood flowing away.[1.153] This is particularly useful in identifying blood clots, blocked arteries, and heart valve defects, among others.[1.156]

On the whole, medical ultrasound can be used to image both static and dynamic systems. Interestingly, the pressure amplitudes used for medical imaging are sufficiently large for ultrasound waves to display nonlinear

behaviour. The subsequent section will explore ultrasound within this large signal amplitude regime and show how nonlinear effects can be applied to chemical determinations.

1.7 Nonlinear Ultrasound

In addition to quantification by attenuation measurements, nonlinear distortion of signal can be exploited for quantification as higher power ultrasound travels through a sample. The amount of signal distortion depends on the medium of propagation as well as frequency and pressure amplitude of the wave. This section explores nonlinear distortion of ultrasound and how it can be applied to chemical determinations.

In many systems, the propagation of ultrasound waves cannot be described using linear assumptions alone, particularly when the amplitude of the ultrasound wave is large. Nonlinear phenomena, including distortion in waveform and the generation of frequency harmonics, must also be considered.[1.160] These effects are most prominent in condensed media with comparatively low acoustic attenuation, such as water, amniotic fluid, tissue, and various other biofluids.[1.161] Distortion in waveform arise from a difference in the velocity of propagation between the compression and rarefaction phases of acoustic waves and result from non-negligible differences in density between the two phases. Local particle

velocities are greater in the compression phase than in the rarefaction phase, and lead to a cumulative distortion of a sine wave into a sawtooth wave, as shown in Figure 1-8. Longer propagation times and larger wave amplitudes increase the extent of the observed distortion. This waveform distortion, the degree of which is also dependent on the medium, leads to the generation of new frequency components that are harmonically related.[1.162] Through the principle of superposition, this frequency distortion is also observed in multi-frequency signals, such as broad-band pulses. Every frequency component of the pulse can be subject to distortion and can lead to broadening of the original frequency distribution. This is shown in Figure1- 9, where a broad-band pulse, modeled as an exponentially decaying sine wave, exhibits an increase in frequency content with distortion.

In practice, the harmonics observed will be limited by the bandwidth of the detecting transducer.[1.163] For example, using a transducer with a central frequency of 3 MHz and bandwidth of 1 MHz will allow for detection of frequencies between 2 and 4 MHz. This means that this transducer could detect the first harmonic of a 2 MHz signal, namely 4 MHz. Consequently, to observe higher order harmonics, transducers with larger bandwidths are needed.

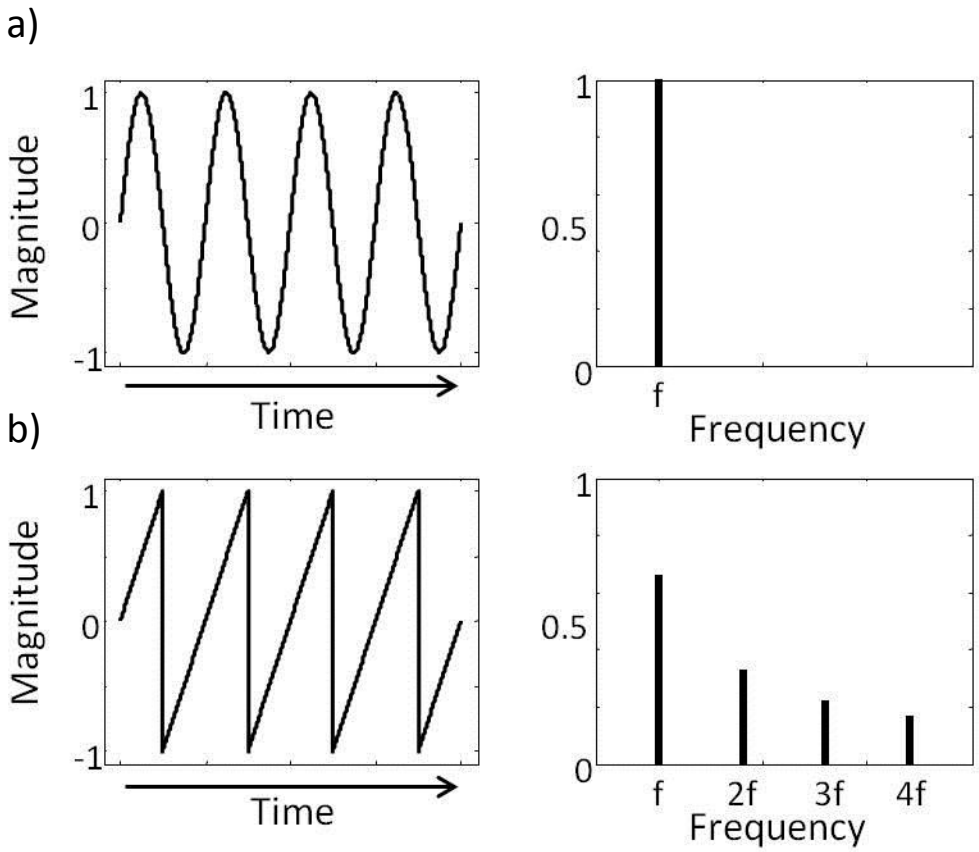

Figure 1-8: Difference in frequency components between (a) sinusoidal and (b) sawtooth waves. Greater particle velocities in the crest of the sinusoidal wave compared to the trough lead to a progressive distortion of the original, single-frequency waveform. This distortion is accompanied by the generation of new frequencies that are harmonically related to the initial frequency.

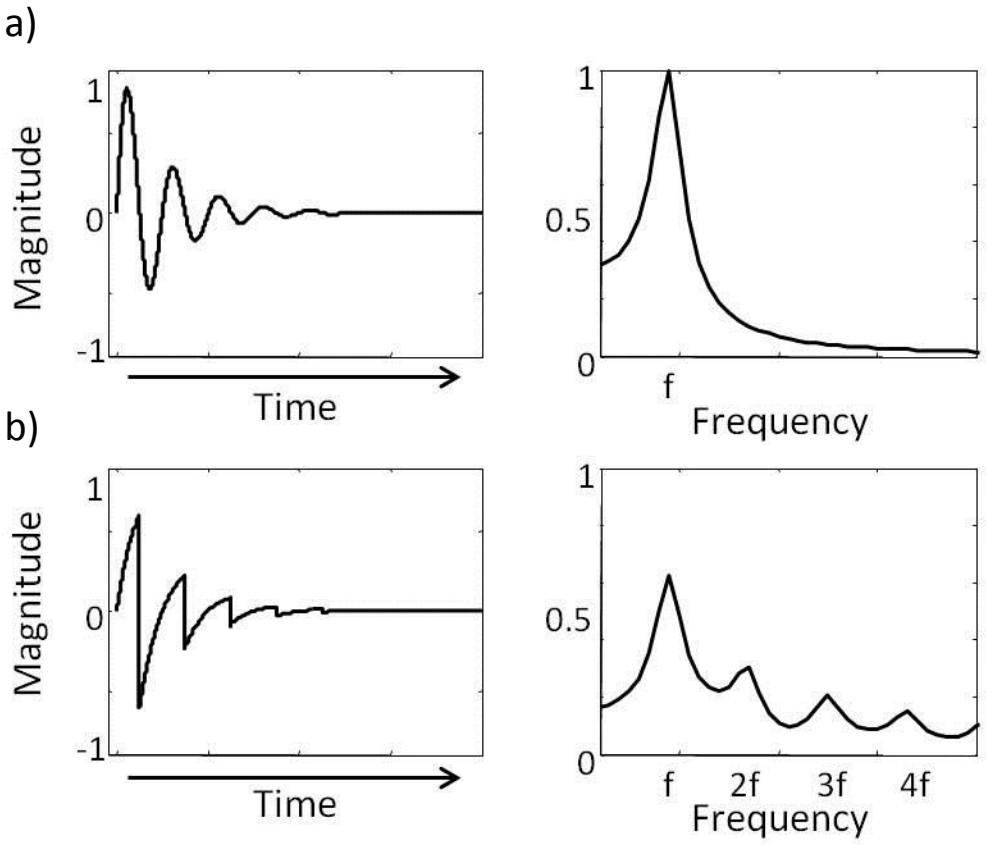

Figure1- 9: Difference in frequency components between (a) undistorted and (b) distorted pulses. A multi-frequency pulse, modeled here as an exponentially decaying sine wave, undergoes a broadening in distribution as harmonic components appear with pulse distortion. This pulsed approach allows for the observation of a range of frequencies with a single measurement.

At very high ultrasound intensities, in the domain of extreme nonlinearities of wave propagation, the highly energetic collapse of sawtooth waves can produce abrupt decreases in pressure. Small cavities, or bubbles, can form in a medium if the pressure change is sufficiently large, in a process known as acoustic cavitation.[1.164] The ultrasound-driven growth and collapse of microbubbles is associated with a variety of special effects. Dirt can be removed from surfaces by cavitation bubbles.[1.165] Light emission may occur upon collapse of the bubbles in a process known as sonoluminescence.[1.166] Acoustic cavitation can also generate large shear forces, which are useful for particle size reduction and emulsification processes.[1.167] In sonochemistry, the large internal temperature of a collapsing bubble (~5000 K) can initiate chemical reactions. Examples of this are bond breakage resulting in highly reactive radical species and pyrolysis products from volatile solvent-solute molecules.[1.168] Given the potentially destructive nature of ultrasound, using appropriate intensities is crucial when investigating biological samples.

The degree of signal distortion observed in a medium can be described by the dimensionless nonlinearity parameter

$$\frac{B}{A} = \frac{\rho}{c_0^2}\left(\frac{\partial^2 p}{\partial \rho^2}\right)_{s,\rho=\rho_0} \qquad (1.15)$$

where ρ is density, c is the sound speed, and p is the pressure in an isentropic system, s. [1.169] Note that changes in entropy are sufficiently small so that the

system can be considered as isentropic.[1.160] This expression has its origins in the equation of state relating variations in pressure in a medium to variations in density:[1.170]

$$p = p(\rho, s) \qquad (1.16)$$

This can be expressed as a Taylor series expansion

$$p = p_0 + \frac{A}{1!}\left(\frac{\rho - \rho_0}{\rho_0}\right) + \frac{B}{2!}\left(\frac{\rho - \rho_0}{\rho_0}\right)^2 + \frac{C}{3!}\left(\frac{\rho - \rho_0}{\rho_0}\right)^3 + \cdots$$

$$\text{with } \begin{cases} A = \rho_0 \left[\left(\frac{\partial p}{\partial \rho}\right)_s\right]_{\rho=\rho_0} \\ B = \rho_0^2 \left[\left(\frac{\partial^2 p}{\partial \rho^2}\right)_s\right]_{\rho=\rho_0} \\ C = \rho_0^3 \left[\left(\frac{\partial^3 p}{\partial \rho^3}\right)_s\right]_{\rho=\rho_0} \end{cases} \qquad (1.17)$$

and coefficients *A*, *B*, and *C* dependent on temperature. The nonlinearity parameter *B/A* is then a ratio of quadratic to linear terms of the Taylor series expansion. Cubic terms, such as *C/A*, and higher are commonly excluded from consideration as they are very small, though they may still be used to describe distortion processes in media.[1.171]

The acoustic nonlinear parameter *B/A* is dependent on the chemical composition of the medium and provides a useful avenue for analytical applications. Table 1-2 lists select *B/A* values for substances of differing chemical composition. Pure liquids, like distilled water, have a *B/A* that differ from

Table 1-2: Nonlinearity parameter *B/A* of select substances[1.99,1.172,1.173]

Substance	Temperature (°C)	*B/A*
Water (distilled)	20	5.0
Sea Water (3.5%)	20	5.25
Methanol	20	9.6
Methanol/water mixture (50%)	20	8.3
Ethanol	20	10.5
Bovine serum albumin (38.8%)	30	6.68
Human breast fat	30	9.9

solutions and mixtures of that liquid. This difference of *B/A* is accompanied by a difference in frequency content of measured ultrasound waves. The nonlinear contribution to the signal, dependent on chemical composition, offers an additional dimension of information along with absorption and scattering phenomena, all of which can be collected in a single ultrasound measurement.

In pure liquids, the acoustic nonlinearity can be related to their molecular properties, namely internal pressure, free energy of binding, the effective van der Waals' constants, the translational diffusion coefficient, and the rotational correlation time.[1.161] Alcohol/water mixtures show distinct variation in *B/A* with changes in concentration.[1.174,1.175] The complex variation in *B/A* is attributed to solute-solvent interactions and the formation of clathrate-like structures between alcohol multimers and surrounding water.[1.174,1.176] The *B/A* parameter of amino acid and protein solutions also depends on concentration and structure.[1.177] Moreover, acoustic nonlinearity measurements also allow for estimates of the ratio of bound to free water molecules in protein solutions.[1.178] In biological systems, the dependence of the *B/A* parameter on structure exists at the molecular, cellular, and tissue level.[1.179] Specifically, for *in vivo* or *ex vivo* measurements, intercellular adhesion, cellular structure, and primary and secondary macromolecular structures need be considered as contributors to the

B/A parameter. In addition to frequency-dependent attenuation, nonlinear ultrasound also exhibits sample-dependent distortions and may allow for chemical determinations in biological systems from a single measurement.

Recently, Dion and Burns reported on the use of nonlinear distortions for determination of volume fractions in 3-component mixtures.[1.180] This frequency analysis of media-induced ultrasound distortion has also been applied to quantification of ethanol and sugar content in commercial beverages.[1.181] In these works, samples were probed with broad-band ultrasound pulses that were recorded for subsequent data analysis. The pulses were subject to a fast Fourier transform to yield frequency data for each sample. Using multilinear regression, in which several frequency components were chosen to develop a model, volume fractions and ethanol and sugar content could be estimated. Differences in frequency profiles between samples were attributed to changes in solute-solvent interactions. It should be possible to extend this frequency analysis approach, which takes advantage of multivariate calibration techniques to estimate chemical properties from a single measurement, to biological media. A brief overview of multivariate calibration techniques can be found in Appendix A.

1.8 Conclusion and Research Objectives

Numerous challenges still exist in the measurement of biological fluids. Most procedures are invasive, require sample preparation, and are time-consuming. Moreover, the need for expensive reagents and expert users, as well as a lack of portability, has limited measurements to laboratory settings.

Nonlinear ultrasound may be able to address some of these current limitations. As a mechanical wave, it can propagate through opaque samples (non-invasive potential) and is established as safe and non-destructive for measurements at diagnostic power levels. Linear and nonlinear phenomena allow for the elucidation of physical and chemical properties of complex systems without the need for sample preparation. With portable, low cost instrumentation available, ultrasound provides a platform with continuous, real-time monitoring capabilities.

This research complements previous acoustic distortion work done by Dr. Jonathan Dion of the Burns research group, which investigated the use of nonlinear frequency analysis for compositional determinations in liquid mixtures and commercial beverages. The current dissertation uses similar techniques to explore nonlinear frequency analysis and multivariate calibration techniques for analytical measurements in biological systems. In particular, modifications in solute-solvent interactions are known to affect the nonlinear propagation of

ultrasound in a sample. Therefore, it should be possible to use frequency analysis to explore changes in protein content and conformation, cell content, and viscosity changes due to oxidative modifications. This could be applied to detection of diseases and inflammation with ultrasound.

The first manuscript, presented in Chapter 2, explores the use of the nonlinear distortion of ultrasound to monitor pH in homogeneous, albumin-containing media. This chapter demonstrates that changes in the molecular conformation of albumin with pH are reflected in the ultrasound frequency domain, allowing for indirect determination of pH.

Subsequently, Chapter 3 extends the application of nonlinear distortion for quantification to a heterogeneous system, milk. The presence of fat globules introduces a scattering contribution to the signal, and measurement of leukocytes demonstrates that nonlinear ultrasound is not only sensitive to molecular properties of a system, but is also influenced by higher level structures.

Chapter 4 considers serum samples from mothers with various infections. Changes in frequency profile with infection are attributed to differences in protein concentration and oxidative modifications of proteins. Despite the large biological variations observed in these samples, the straightforward ultrasound measurement provides useful insights into infection trends.

The concept of measuring infection is further explored in Chapter 5, which uses a controlled set of mouse serum samples. This chapter also presents a novel heterodyned ultrasound instrument to demonstrate the feasibility of ultrasound spectroscopy as a point-of-care technology.

Finally, Chapter 6 highlights the key findings presented in this thesis and outlines future directions for ultrasonic frequency analysis techniques.

1.9 References

1.1. Gowda, G. A. N. *et al.* Metabolomics-based methods for early disease diagnostics. *Expert Rev. Mol. Diagn.* **8,** 617–33 (2008).

1.2. Bissonnette, L. & Bergeron, M. G. Diagnosing infections--current and anticipated technologies for point-of-care diagnostics and home-based testing. *Clin. Microbiol. Infect.* **16,** 1044–53 (2010).

1.3. Zivkovic, A. M. *et al.* Effects of sample handling and storage on quantitative lipid analysis in human serum. *Metabolomics* **5,** 507–516 (2009).

1.4. Yager, P., Domingo, G. J. & Gerdes, J. Point-of-care diagnostics for global health. *Annu. Rev. Biomed. Eng.* **10,** 107–44 (2008).

1.5. Capasso, G. & Unwin, R. Electrolytes and acid–base: common fluid and electrolyte disorders. *Medicine (Baltimore).* **39,** 317–324 (2011).

1.6. Fadel, H. E., Northrop, G., Misenhimer, H. R. & Harp, R. J. Acid-base determinations in amniotic fluid and blood of normal late pregnancy. *Obstet. Gynecol.* **53,** 99–104 (1979).

1.7. Mitchell, R. A. *et al.* Stability of cerebrospinal fluid pH in chronic acid-base disturbances in blood. *J. Appl. Physiol.* **20,** 443–452 (1965).

1.8. Posencheg, M. A. & Evans, J. R. in *Avery's Diseases of the Newborn* 367–389 (Elsevier Inc., 2012). doi:10.1016/B978-1-4377-0134-0.10031-9

1.9. Matsui, T., Hattori, H., Takase, B. & Ishihara, M. Non-invasive estimation of arterial blood pH using exhaled CO/CO2 analyser, microwave radar and infrared thermography for patients after massive haemorrhage. *J. Med. Eng. Technol.* **30,** 97–101 (2006).

1.10. Koeppen, B. M. The kidney and acid-base regulation. *Adv. Physiol. Educ.* **33,** 275–81 (2009).

1.11. Guyton, A. C. & Hall, J. E. in *Textbook of Medical Physiology* 383–401 (Elsevier Inc., 2006).

1.12. Hasselbalch, K. A. Die Berechnung der Wasserstoffzahl des Blutes aus der freien und gebundenen Kohlensäure desselben, und die Sauerstoffbindung des Blutes als Funktion der Wasserstoffzahl. *Biochem. Z.* **78,** 112–144 (1916).

1.13. Henderson, L. J. Concerning the relationship between the strength of acids and their capacity to preserve neutrality. *Am. J. Physiol.* **21,** 173–179 (1908).

1.14. De Morais, H. A., Bach, J. F. & DiBartola, S. P. Metabolic acid-base disorders in the critical care unit. *Vet. Clin. North Am. Small Anim. Pract.* **38,** 559–74 (2008).

1.15. Jones, S. A. & Walter, J. H. Diagnosis and treatment of severe metabolic acidosis. *Paediatr. Child Health (Oxford).* **17,** 260–265 (2007).

1.16. Ganter, M. T. *et al.* How often should we perform arterial blood gas analysis during thoracoscopic surgery? *J. Clin. Anesth.* **19,** 569–75 (2007).

1.17. Venkatesh, B. Continuous Intra-arterial Blood Gas Monitoring. *Crit. Care Resusc.* **1,** 150 (1999).

1.18. Jin, W. *et al.* Continuous intra-arterial blood pH monitoring by a fiber-optic fluorosensor. *IEEE Trans. Biomed. Eng.* **58,** 1232–8 (2011).

1.19. Yang, Y., Soyemi, O. O., Landry, M. R. & Soller, B. R. Noninvasive in Vivo Measurement of Venous Blood pH During Exercise Using Near-Infrared Reflectance Spectroscopy. *Appl. Spectrosc.* **61,** 223–229 (2007).

1.20. Soller, B. R., Yang, Y., Lee, S. M. C., Wilson, C. & Hagan, R. D. Noninvasive determination of exercise-induced hydrodgen ion threshold through direct optical measurement. *J. Appl. Physiol.* **104,** 837–44 (2008).

1.21. Andrews, R. J. M. D., Bringas, J. R. B. S. & Alonzo, G. B. S. Cerebrospinal Fluid pH and PCO2 Rapidly Follow Arterial Blood pH and PCO2 with Changes in Ventilation. *Neurosurgery* **34,** 466–470 (1994).

1.22. Jauniaux, E. *et al.* Investigation of the acid-base balance of coelomic and amniotic fluids in early human pregnancy. *Am. J. Obstet. Gynecol.* **170,** 1365–1369 (1994).

1.23. Seeds, A. & Hellegers, A. Acid-base determinations in human amniotic fluid throughout pregnancy. *Am. J. Obstet. Gynecol.* **101,** 257–60 (1968).

1.24. Underwood, M. A., Gilbert, W. M. & Sherman, M. P. Amniotic fluid: not just fetal urine anymore. *J. Perinatol.* **25,** 341–8 (2005).

1.25. Sundberg, K. *et al.* Randomised study of risk of fetal loss related to early amniocentesis versus chorionic villus sampling. *Lancet* **350,** 697–703 (1997).

1.26. Jeffery, H. E. & Lahra, M. M. in *Fetal and Neonatal Pathology* (eds. Keeling, J. W. & Khong, T. Y.) 379–423 (Springer London, 2007). doi:10.1007/978-1-84628-743-5

1.27. Sadler, P. J. & Tucker, a. pH-induced structural transitions of bovine serum albumin. Histidine pKa values and unfolding of the N-terminus during the N to F transition. *Eur. J. Biochem.* **212,** 811–7 (1993).

1.28. Tanford, C. & Buzzell, J. G. The Viscosity of Aqueous Solutions of Bovine Serum Albumin between pH 4.3 and 10.5. *J. Phys. Chem.* **60,** 225–231 (1956).

1.29. Pulido, R. & Hooft van Huijsduijnen, R. Protein tyrosine phosphatases: dual-specificity phosphatases in health and disease. *FEBS J.* **275,** 848–66 (2008).

1.30. 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,** 1345–64 (2003).

1.31. Badgwell, D. & Bast Jr., R. C. Early Detection of Ovarian Cancer. *Dis. Markers* **23,** 397–410 (2007).

1.32. McPherson, R. A. & Pincus, M. R. *Henry's Clinical Diagnosis and Management by Laboratory Methods*. (Saunders Elsevier, 2007).

1.33. Baggerly, K. A., Morris, J. S. & Coombes, K. R. Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. *Bioinformatics* **20,** 777–85 (2004).

1.34. Anderson, N. L. & Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Mol. Cell. Proteomics* **1,** 845–867 (2002).

1.35. Neuhausen, B. S. & Rioch, D. M. The Refractometric Determination of Serum. *J. Biol. Chem.* **55,** 353–356 (1923).

1.36. Warburg, O. & Christian, W. Isolierung und Kristallisation des Gärungsferments Enolase. *Naturwissenschaften* **29,** 589–590 (1941).

1.37. Van Toorenenbergen, a W., Blijenberg, B. G. & Leijnse, B. Measurement of total serum protein by near-infrared reflectance spectroscopy. *J. Clin. Chem. Clin. Biochem.* **26,** 209–11 (1988).

1.38. Bradford, M. M. A Rapid and Sensitive Method for the Quantitation Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72,** 248–254 (1976).

1.39. Gornall, A. G., Bardawill, C. J. & David, M. M. Determination of serum proteins by means of the biuret reaction. *J. Biol. Chem.* **177,** 751–66 (1949).

1.40. Lowry, O. H., Rosebrough, N. J., Farr, L. & Randall, R. J. Protein Measurement with the Folin Phenol Reagent. *J. Biol. Chem.* **193,** 265–275 (1951).

1.41. Smith, P. K. *et al.* Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150,** 76–85 (1985).

1.42. Noble, J. E. & Bailey, M. J. A. Quantitation of protein. *Methods Enzymol.* **463,** 73–95 (2009).

1.43. Prinsen, B. H. C. M. T. & de Sain-van der Velden, M. G. M. Albumin turnover: experimental approach and its application in health and renal diseases. *Clin. Chim. Acta.* **347,** 1–14 (2004).

1.44. Fanali, G. *et al.* Human serum albumin: from bench to bedside. *Mol. Aspects Med.* **33,** 209–90 (2012).

1.45. Maher, J. E. *et al.* Albumin levels in pregnancy: a hypothesis--decreased levels of albumin are related to increased levels of alpha-fetoprotein. *Early Hum. Dev.* **34,** 209–15 (1993).

1.46. Doumas, B. T. & Peters, T. Origins of Dye-Binding Methods for Measuring Serum Albumin. *Clin. Chem.* **55,** 583–584 (2009).

1.47. Clase, C. M., St Pierre, M. W. & Churchill, D. N. Conversion between bromcresol green- and bromcresol purple-measured albumin in renal disease. *Nephrol. Dial. Transplant* **16,** 1925–9 (2001).

1.48. Coskun, A. F., Nagi, R., Sadeghi, K., Phillips, S. & Ozcan, A. Albumin testing in urine using a smart-phone. *Lab Chip* **13,** 4231–8 (2013).

1.49. Gonzalez-Quintela, a *et al.* Serum levels of immunoglobulins (IgG, IgA, IgM) in a general adult population and their relationship with alcohol consumption, smoking and common metabolic abnormalities. *Clin. Exp. Immunol.* **151,** 42–50 (2008).

1.50. Haswell-Elkins, M. R. *et al.* Immune responsiveness and parasite-specific antibody levels in human hepatobiliary disease associated with Opisthorchis viverrini infection. *Clin. Exp. Immunol.* **84,** 213–8 (1991).

1.51. Aucouturier, P. *et al.* Serum IgG subclass levels in patients with primary immunodeficiency syndromes or abnormal susceptibility to infections. *Clin. Immunol. Immunopathol.* **51,** 22–37 (1989).

1.52. Benster, B. & Wood, E. J. Immunoglobulin levels in normal pregnancy and pregnancy complicated by hypertension. *J. Obstet. Gynaecol. Br. Commonw.* **77,** 518–22 (1970).

1.53. Haroun, M. & M El-Sayed, M. Measurement of IgG levels can serve as a biomarker in newly diagnosed diabetic children. *J. Clin. Biochem. Nutr.* **40,** 56–61 (2007).

1.54. Davis, D. G. *et al.* Measurement of serum IgG in foals by radial immunodiffusion and automated turbidimetric immunoassay. *J. Vet. Intern. Med.* **19,** 93–6 (2005).

1.55. Aoyagi, S., Miyasaka, T., Yoshimi, Y. & Sakai, K. A new reagentless immunosensor for measuring IgG concentration in human plasma based on fluorescence-enhancement immunoassay. *J. Artif. Organs* **5,** 0060–0063 (2002).

1.56. Biagini, R. E. *et al.* Rapid , Sensitive , and Specific Lateral-Flow Immunochromatographic Device To Measure Anti-Anthrax Protective Antigen Immunoglobulin G in Serum and Whole Blood. *Clin. Vaccine Immunol.* **13,** 541–546 (2006).

1.57. Choi, S. *et al.* A rapid, simple measurement of human albumin in whole blood using a fluorescence immunoassay (I). *Clin. Chim. Acta* **339,** 147–156 (2004).

1.58. Guyton, A. & Hall, J. in *Textbook of Medical Physiology* 429–438 (Elsevier Inc., 2006).

1.59. Semba, R. D. *et al.* Mastitis and Immunological Factors in Breast Milk of Lactating Women in Malawi. *Clin. Diagn. Lab. Immunol.* **6,** 671–674 (1999).

1.60. Thomsen, A. C., Espersen, T. & Maigaard, S. Course and treatment of milk stasis, noninfectious inflammation of the breast, and infectious mastitis in nursing women. *Am. J. Obstet. Gynecol.* **149,** 492–495 (1984).

1.61. Hassiotou, F. *et al.* Maternal and infant infections stimulate a rapid leukocyte response in breastmilk. *Clin. Transl. Immunol.* **2,** e3 (2013).

1.62. Michie, C., Lockie, F. & Lynn, W. The challenge of mastitis. *Arch. Dis. Child.* **88,** 818–21 (2003).

1.63. Aryeetey, R. N. O., Marquis, G. S., Timms, L., Lartey, A. & Brakohiapa, L. Subclinical mastitis is common among Ghanaian women lactating 3 to 4 months postpartum. *J. Hum. Lact.* **24,** 263–7 (2008).

1.64. Semba, R. D. *et al.* Mastitis and immunological factors in breast milk of lactating women in Malawi. *Clin. Diagn. Lab. Immunol.* **6,** 671–4 (1999).

1.65. Fetherston, C. Risk Factors for Lactation Mastitis. *J. Hum. Lact.* **14,** 101–109 (1998).

1.66. Shenkin, M., Babu, R. & Maiese, R. Accurate assessment of cell count and viability with a flow cytometer. *Cytometry* **72B,** 427–32 (2007).

1.67. Redelman, D., Butler, S., Robison, J. & Garner, D. Identification of inflammatory cells in bovine milk by flow cytometry. *Cytometry* **9,** 463–8 (1988).

1.68. Pillai, S. R., Kunze, E., Sordillo, L. M. & Jayarao, B. M. Application of differential inflammatory cell count as a tool to monitor udder health. *J. Dairy Sci.* **84,** 1413–20 (2001).

1.69. Reneau, J. K. & Packard, V. L. Monitoring mastitis, milk quality and economic losses in dairy fields. *Dairy, food Environ. Sanit.* **11,** 4–11 (1991).

1.70. Whyte, D., Walmsley, M., Liew, A., Claycomb, R. & Mein, G. Chemical and rheological aspects of gel formation in the California Mastitis Test. *J. Dairy Res.* **72,** 115–121 (2005).

1.71. Schukken, Y. H., Wilson, D. J., Welcome, F., Garrison-Tikofsky, L. & Gonzalez, R. N. Monitoring udder health and milk quality using somatic cell counts. *Vet. Res.* **34,** 579–96 (2003).

1.72. Moon, J. S. *et al.* Application of a new portable microscopic somatic cell counter with disposable plastic chip for milk analysis. *J. Dairy Sci.* **90,** 2253–9 (2007).

1.73. Grenvall, C., Folkenberg, J. R., Augustsson, P. & Laurell, T. Label-free somatic cell cytometry in raw milk using acoustophoresis. *Cytom. Part A* **81,** 1076–83 (2012).

1.74. Tsenkova, R., Atanassova, S., Kawano, S. & Toyoda, K. Somatic cell count determination in cow's milk by near-infrared spectroscopy: a new diagnostic tool. *J. Anim. Sci.* **79,** 2550–7 (2001).

1.75. Kawasaki, M. *et al.* Near-infrared spectroscopic sensing system for on-line milk quality assessment in a milking robot. *Comput. Electron. Agric.* **63,** 22–27 (2008).

1.76. Hassiotou, F., Geddes, D. T. & Hartmann, P. E. Cells in human milk: state of the science. *J. Hum. Lact.* **29,** 171–82 (2013).

1.77. Valko, M. *et al.* Free radicals and antioxidants in normal physiological functions and human disease. *Int. J. Biochem. Cell Biol.* **39,** 44–84 (2007).

1.78. Ajmani, R. S. *et al.* Oxidative stress and hemorheological changes induced by acute treadmill exercise. *Clin. Hemorheol. Microcirc.* **28,** 29–40 (2003).

1.79. Saikumar, P., Jaya, B. & Devi, M. R. R. Oxidative Stress in Pregnancy. *J. Dent. Med. Sci.* **3,** 12–13 (2013).

1.80. Kocha, T., Yamaguchi, M., Ohtaki, H., Fukuda, T. & Aoyagi, T. Hydrogen peroxide-mediated degradation of protein: different oxidation modes of copper- and iron-dependent hydroxyl radicals on the degradation of albumin. *Biochim. Biophys. Acta* **1337,** 319–26 (1997).

1.81. Valko, M., Morris, H. & Cronin, M. T. D. Metals, toxicity and oxidative stress. *Curr. Med. Chem.* **12,** 1161–208 (2005).

1.82. Burton, G. J. & Jauniaux, E. Oxidative stress. *Best Pract. Res. Clin. Obstet. Gynaecol.* **25,** 287–99 (2011).

1.83. Oushiki, D. *et al.* Development and application of a near-infrared fluorescence probe for oxidative stress based on differential reactivity of linked cyanine dyes. *J. Am. Chem. Soc.* **132,** 2795–801 (2010).

1.84. Krumova, K., Greene, L. E. & Cosa, G. Fluorogenic α-tocopherol analogue for monitoring the antioxidant status within the inner mitochondrial membrane of live cells. *J. Am. Chem. Soc.* **135,** 17135–43 (2013).

1.85. Gianazza, E., Crawford, J. & Miller, I. Detecting oxidative post-translational modifications in proteins. *Amino Acids* **33,** 51–6 (2007).

1.86. Karve, T. M. & Cheema, A. K. Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. *J. Amino Acids* **2011,** 207691 (2011).

1.87. Yap, L.-P. *et al.* Determination of GSH, GSSG, and GSNO using HPLC with electrochemical detection. *Methods Enzymol.* **473,** 137–47 (2010).

1.88. Camera, E. & Picardo, M. Analytical methods to investigate glutathione and related compounds in biological and pathological processes. *J. Chromatogr. B* **781,** 181–206 (2002).

1.89. Senft, a P., Dalton, T. P. & Shertzer, H. G. Determining glutathione and glutathione disulfide using the fluorescence probe o-phthalaldehyde. *Anal. Biochem.* **280,** 80–6 (2000).

1.90. Richards, R. & Nwose, E. Blood viscosity at different stages of diabetes pathogenesis. *Br. J. Biomed. Sci.* **67,** 67–70 (2010).

1.91. Rossing, T. D. in *Springer Handbook of Acoustics* (ed. Rossing, T. D.) 9–23 (Springer New York, 2007).

1.92. Pierce, A. D. in *Acoustics: An Introduction to Its Physical Principles and Applications* (Acoustical Society of America, 1994).

1.93. Taylor, S. Galileo's Work in Acoustics. *Nature* **9,** 169–170 (1874).

1.94. Mersenne, M. *Harmonie universelle, contenant la théorie et la pratique de la musique*. (S. Cramoisy, 1637). at <http://gallica.bnf.fr/ark:/12148/bpt6k5471093v>

1.95. Isaac Newton. *Philosophiae naturalis principia mathematica*. (jussu Societatis regiae, 1687). at <http://gallica.bnf.fr/ark:/12148/bpt6k3363w.r=principia+newton.langEN>

1.96. Strutt, J. W. *The Theory of Sound*. (Macmillan, 1878). at <http://gallica.bnf.fr/ark:/12148/bpt6k95131k.r=the+theory+of+sound.langEN>

1.97. Curie, J. & Curie, P. Développement, par pression, de l'électricité polaire dans les cristaux hémièdres à faces inclinées. *Comptes Rendus des Séances l'Académie des Sci.* **91,** 294–295 (1880).

1.98. Manbachi, A. & Cobbold, R. S. C. Development and application of piezoelectric materials for ultrasound generation and detection. *Ultrasound* **19,** 187–196 (2011).

1.99. Breazeale, M. A. & McPherson, M. in *Springer Handbook of Acoustics* (ed. Rossing, T. D.) 207–238 (Springer New York, 2007).

1.100. Lieu, D. Ultrasound physics and instrumentation for pathologists. *Arch. Pathol. Lab. Med.* **134,** 1541–56 (2010).

1.101. O'Brien, W. D. Ultrasound-biophysics mechanisms. *Prog. Biophys. Mol. Biol.* **93,** 212–55 (2007).

1.102. G.D. Ludwig. The velocity of sound through tissues and the acoustic impedance of tissues. *J. Acoust. Soc. Am.* **22,** 862–866 (1950).

1.103. Wentzell, P. D. & Brown, C. D. Signal Processing in Analytical Chemistry. *Encyclopedia of Analytical Chemistry* 1–36 (2006). doi:10.1002/9780470027318.a5207

1.104. Laugier, P. *Bone Quantitative Ultrasound*. (Springer Netherlands, 2011). doi:10.1007/978-94-007-0017-8

1.105. Powell, D. A., Kalantar-zadeh, K., Wlodarski, W. & Ippolito, S. J. Layered Surface Acoustic Wave Chemical and Bio-Sensors. *Encyclopedia of Sensors* **X,** 1–18 (2006).

1.106. Assia, O. & Hakim, D. in *Wave Propagation in Materials for Modern Applications* (ed. Petrin, A.) (InTech, 2010). at <http://www.intechopen.com/books/wave-propagation-in-materials-for-modernapplications/>

1.107. Ter Haar, G. & Coussios, C. High intensity focused ultrasound: Physical principles and devices. *Int. J. Hyperth.* **23,** 89–104 (2007).

1.108. Szabo, T. & Wu, J. A model for longitudinal and shear wave propagation in viscoelastic media. *J. Acoust. Soc. Am.* **107,** 2437–46 (2000).

1.109. Dukhin, A. S. & Goetz, P. J. *Ultrasound for Characterizing Colloids*. (Elsevier, 2002).

1.110. Alvarez-Arenas, T. E. G. Acoustic impedance matching of piezoelectric transducers to the air. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **51,** 624–633 (2004).

1.111. Fan, X. & Hynynen, K. The effect of wave reflection and refraction at soft tissue interfaces during ultrasound hyperthermia treatments. *J. Acoust. Soc. Am.* **91,** 1727–1736 (1992).

1.112. Rouquette, J. *et al.* Pressure tuning of the morphotropic phase boundary in piezoelectric lead zirconate titanate. *Phys. Rev. B* **70,** 014108 (2004).

1.113. Tressler, J. F., Alkoy, S. & Newnham, R. E. Piezoelectric Sensors and Sensor Materials. *J. Electroceramics* **2,** 257–272 (1998).

1.114. Ueberschlag, P. PVDF piezoelectric polymer. *Sens. Rev.* **21,** 118–125 (2001).

1.115. Treeby, B. E., Zhang, E. Z., Thomas, A. S. & Cox, B. T. Measurement of the ultrasound attenuation and dispersion in whole human blood and its components from 0-70 MHz. *Ultrasound Med. Biol.* **37,** 289–300 (2011).

1.116. Davies, S. J., Edwards, C., Taylor, G. S. & Palmer, S. B. Laser-generated ultrasound : its properties , mechanisms and multifarious applications. *J. Phys. D. Appl. Phys.* **26,** 329–348 (1993).

1.117. McClements, D. J. in *Encyclopedia of Analytical Chemistry* (ed. Meyers, R. A.) (John Wiley & Sons, Ltd, 2006). doi:10.1002/9780470027318

1.118. Kaatze, U., Wehrmann, B. & Pottel, R. Acoustical absorption spectroscopy of liquids between 0.15 and 3000 MHz. I. High resolution ultrasonic resonator method. *J. Phys. E.* **20,** 1025–1030 (1987).

1.119. Povey, M. J. W. *Ultrasonic Techniques for Fluids Characterization*. 214 (Academic Press, 1997).

1.120. Kaatze, U., Lautscham, K. & Brai, M. Acoustical absorption spectroscopy of liquids between 0.15 and 3000 MHz: II. Ultrasonic pulse transmission methods. *J. Phys. E.* **21,** 98–103 (1988).

1.121. Eggers, F. & Kaatze, U. Broad-band ultrasonic measurement techniques for liquids. *Meas. Sci. Technol.* **7,** 1–19 (1996).

1.122. Wood, R. W. & Loomis, A. L. The physical and biological effects of high-frequency sound-waves of great intensity. *Philos. Mag.* **4,** 417–436 (1927).

1.123. Rayleigh, Lord. On the free vibrations of an infinite plate of homogeneous isotropic elastic matter. *Proc. London Math. Soc.* **20,** 225–237 (1888).

1.124. Sauerbrey, G. Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung. *Zeitschrift für Phys.* **155,** 206–222 (1959).

1.125. Marx, K. a. Quartz crystal microbalance: a useful tool for studying thin polymer films and complex biomolecular systems at the solution-surface interface. *Biomacromolecules* **4,** 1099–120 (2003).

1.126. King, W. H. Piezoelectric Sorption Detector. *Anal. Chem.* **36,** 1735–1739 (1964).

1.127. Guilbault, G. G. Determination of formaldehyde with an enzyme-coated piezoelectric crystal detector. *Anal. Chem.* **55,** 1682–1684 (1983).

1.128. Guilbault, G. G., Hock, B. & Schmid, R. A piezoelectric immunobiosensor for atrazine in drinking water. *Biosens. Bioelectron.* **7,** 411–9 (1992).

1.129. Jin, X., Huang, Y., Mason, A. & Zeng, X. Multichannel monolithic quartz crystal microbalance gas sensor array. *Anal. Chem.* **81,** 595–603 (2009).

1.130. Parlak, Z., Biet, C. & Zauscher, S. Decoupling mass adsorption from fluid viscosity and density in quartz crystal microbalance measurements using normalized conductance modeling. *Meas. Sci. Technol.* **24,** 085301 (2013).

1.131. Gronewold, T. M. A. Surface acoustic wave sensors in the bioanalytical field: recent trends and challenges. *Anal. Chim. Acta* **603,** 119–28 (2007).

1.132. Atkinson, G. & Baumgartner, E. K. Ultrasonic velocity in nonelectrolyte-water mixtures. *J. Phys. Chem.* **75,** 2336–2340 (1971).

1.133. Van Sint Jan, M., Guarini, M., Guesalaga, A., Ricardo Pérez-Correa, J. & Vargas, Y. Ultrasound based measurements of sugar and ethanol concentrations in hydroalcoholic solutions. *Food Control* **19,** 31–35 (2008).

1.134. Vyas, J. C., Katti, V. R., Gupta, S. K. & Yakhmi, J. V. A non-invasive ultrasonic gas sensor for binary gas mixtures. *Sensors Actuators B Chem.* **115,** 28–32 (2006).

1.135. Toda, H. & Kobayakawa, T. High-speed gas concentration measurement using ultrasound. *Sensors Actuators A Phys.* **144,** 1–6 (2008).

1.136. Chalikian, T. V, Totrov, M., Abagyan, R. & Breslauer, K. J. The hydration of globular proteins as derived from volume and compressibility measurements: cross correlating thermodynamic and structural data. *J. Mol. Biol.* **260,** 588–603 (1996).

1.137. Choi, P. K., Bae, J. R. & Takagi, K. Ultrasonic spectroscopy in bovine serum albumin solutions. *J. Acoust. Soc. Am.* **87,** 874–81 (1990).

1.138. Cho, K. C., Leung, W. P., Mok, H. Y. & Choy, C. L. Ultrasonic absorption in myoglobin and other globular proteins. *Biochim. Biophys. Acta* **830,** 36–44 (1985).

1.139. Kessler, L. W. & Dunn, F. Ultrasonic investigation of the conformal changes of bovine serum albumin in aqueous solution. *J. Phys. Chem.* **73,** 4256–63 (1969).

1.140. Pavlovskaya, G., McClements, D. J. & Povey, M. J. W. Ultrasonic investigation of aqueous solutions of a globular protein. *Food Hydrocoll.* **6,** 253–262 (1992).

1.141. Miles, C. A., Shore, D. & Langley, K. R. Attenuation of ultrasound in milks and creams. *Ultrasonics* **28,** 394–400 (1990).

1.142. Holmes, A. K., Challis, R. E. & Wedlock, D. J. A Wide-Bandwidth Ultrasonic Study of Suspensions: The Variation of Velocity and Attenuation with Particle Size. *J. Colloid Interface Sci.* **168,** 339–348 (1994).

1.143. Bonacucina, G., Misici-Falzi, M., Cespi, M. & Palmieri, G. F. Characterization of Micellar Systems by the Use of Acoustic Spectroscopy. *J. Pharm. Sci.* **97,** 2217–2227 (2008).

1.144. Dukhin, a S., Goetz, P. J. & Travers, B. Use of ultrasound for characterizing dairy products. *J. Dairy Sci.* **88,** 1320–34 (2005).

1.145. Dukhin, A. S., Goetz, P. J. & van de Ven, T. G. M. Ultrasonic characterization of proteins and blood cells. *Colloids Surf. B. Biointerfaces* **53,** 121–6 (2006).

1.146. Freundlich, H., Söllner, K. & Rogowski, F. Einige Biologische Wirkungen von Ultraschallwellen. *Klin. Wochenschr.* **11,** 1512–1513 (1932).

1.147. Mitragotri, S. Healing sound: the use of ultrasound in drug delivery and other therapeutic applications. *Nat. Rev. Drug Discov.* **4,** 255–60 (2005).

1.148. Lewis, G. K. & Olbricht, W. A phantom feasibility study of acoustic enhanced drug perfusion in neurological tissue. in *2007 IEEE/NIH Life Science Systems and Applications Workshop* 67–70 (IEEE, 2007). doi:10.1109/LSSA.2007.4400886

1.149. Kennedy, J. E. High-intensity focused ultrasound in the treatment of solid tumours. *Nat. Rev. Cancer* **5,** 321–7 (2005).

1.150. Robertson, V. J. & Baker, K. G. A Review of Therapeutic Ultrasound: Effectiveness Studies. *Phys. Ther.* **81,** 1339–1350 (2001).

1.151. Baker, K. G., Robertson, V. J. & Duck, F. A. A Review of Therapeutic Ultrasound: Biophysical Effects. *Phys. Ther.* **81,** 1351–8 (2001).

1.152. Kuttruff, H. *Ultrasonics: Fundamentals and Applications*. (Elsevier Science, 1991).

1.153. Beach, K. W. & Barbrina Dunmire. in *Springer Handbook of Acoustics* (ed. Rossing, T. D.) 839–898 (Springer New York, 2007).

1.154. Starritt, H. C. & Duck, F. A. in *Clinical Ultrasound* 51–60 (Elsevier Ltd, 2011). doi:10.1016/B978-0-7020-3131-1.00004-3

1.155. Edler, I. & Lindström, K. The history of echocardiography. *Ultrasound Med. Biol.* **30,** 1565–644 (2004).

1.156. Lawrence, J. P. Physics and instrumentation of ultrasound. *Crit. Care Med.* **35,** S314–22 (2007).

1.157. Seeds, J. W. Diagnostic mid trimester amniocentesis: how safe? *Am. J. Obstet. Gynecol.* **191,** 607–15 (2004).

1.158. McDicken, W. N. & Anderson, T. in *Clinical Ultrasound* 3–15 (Elsevier Ltd, 2001). doi:10.1016/B978-0-7020-3131-1.00001-8

1.159. Hatle, L., Angelsen, B. a & Tromsdal, a. Non-invasive assessment of aortic stenosis by Doppler ultrasound. *Br. Heart J.* **43,** 284–92 (1980).

1.160. Duck, F. A. Nonlinear Acoustics in Diagnostic Ultrasound. *Ultrasound Med. Biol.* **28,** 1–18 (2002).

1.161. Sehgal, C. M. Non-linear ultrasonics to determine molecular properties of pure liquids. *Ultrasonics* **33,** 155–161 (1995).

1.162. Muir, T. G. & Carstensen, E. L. Prediction of nonlinear acoustic effects at biomedical frequencies and intensities. *Ultrasound Med. Biol.* **6,** 345–357 (1980).

1.163. Tranquart, F., Grenier, N., Eder, V. & Pourcelot, L. Clinical use of ultrasound tissue harmonic imaging. *Ultrasound Med. Biol.* **25,** 889–894 (1999).

1.164. Ashokkumar, M. The characterization of acoustic cavitation bubbles - an overview. *Ultrason. Sonochem.* **18,** 864–72 (2011).

1.165. Maisonhaute, E., Prado, C., White, P. C. & Compton, R. G. Surface acoustic cavitation understood via nanosecond electrochemistry. Part III: Shear stress in ultrasonic cleaning. *Ultrason. Sonochem.* **9,** 297–303 (2002).

1.166. Kurcz, A., Capolupo, A. & Beige, A. Sonoluminescence and quantum optical heating. *New J. Phys.* **11,** 053001 (2009).

1.167. Weiss, J., Gulseren, I. & Kjartansson, G. in *Nonthermal Processing Technologies for Food* (eds. Zhang, H. Q. et al.) 109–134 (Wiley-Blackwell, 2010). doi:10.1002/9780470958360

1.168. Ashokkumar, M. & Mason, T. J. Sonochemistry. *Kirk-Othmer Encyclopedia of Chemical Technology* 353–372 (2000). doi:10.1002/0471238961.1915141519211912.a01

1.169. Lauterborn, W., Kurz, T. & Akhatov, I. in *Springer Handbook of Acoustics* (ed. Rossing, T. D.) 257–293 (Springer New York, 2007). doi:10.1007/978-0-387-30425-0

1.170. Khelladi, H., Plantier, F., Daridon, J. L. & Djelouah, H. Measurement under high pressure of the nonlinearity parameter B/A in glycerol at various temperatures. *Ultrasonics* **49,** 668–75 (2009).

1.171. Jayalatha, K. A., Amoros, J. & Ravi, S. Estimation of second-, third- and fourth-order nonlinearity parameters for specific fluorocarbon drug-delivering nanosystems. *Phys. Chem. Liq.* **49,** 309–317 (2011).

1.172. Bjørnø, L. Forty years of nonlinear ultrasound. *Ultrasonics* **40,** 11–17 (2002).

1.173. Law, W. K., Frizzell, L. A. & Street, W. G. Determination of the Nonlinear Parameter B/A of Biological Media. *Ultrasound Med. Biol.* **11,** 307–318 (1985).

1.174. Sehgal, C. M. Ultrasonic nonlinear parameters and sound speed of alcohol–water mixtures. *J. Acoust. Soc. Am.* **79,** 566 (1986).

1.175. Wallace, K. D., Lloyd, C. W., Holland, M. R. & Miller, J. G. Finite amplitude measurements of the nonlinear parameter B/A for liquid mixtures spanning

a range relevant to tissue harmonic mode. *Ultrasound Med. Biol.* **33,** 620–9 (2007).

1.176. Brai, M. & Kaatze, U. Ultrasonic and hypersonic relaxations of monohydric alcohol/water mixtures. *J. Phys. Chem.* **96,** 8946–8955 (1992).

1.177. Sarvazyan, A. P., Chalikian, T. V & Dunn, F. Acoustic nonlinearity parameter B / A of aqueous solutions of some amino acids and proteins. *J. Acoust. Soc. Am.* **88,** 1555–1561 (1990).

1.178. Yang, Y. & Dunn, F. Acoustic non-linearity method for estimating the ratio of bound to free water of biological media. *Ultrasonics* **31,** 35–8 (1993).

1.179. Zhang, J., Kuhlenschmidt, M. S. & Dunn, F. Influences of structural factors of biological media on the acoustic nonlinearity parameter B/A. *J. Acoust. Soc. Am.* **89,** 80–91 (1991).

1.180. Dion, J. R. & Burns, D. H. Determination of volume fractions in multicomponent mixtures using ultrasound frequency analysis. *Appl. Spectrosc.* **65,** 648–56 (2011).

1.181. Dion, J. R. & Burns, D. H. Simultaneous determination of alcohol and carbohydrate content in commercial beverages by ultrasound frequency analysis. *Talanta* **86,** 384–92 (2011).

Chapter 2

Ultrasonic Frequency Analysis for Estimating pH in Albumin-Rich Biofluids

2.1 Foreword

This chapter explores the use of ultrasound frequency analysis as a straightforward, rapid, low-cost tool for monitoring pH in critical care settings. The pH of biofluids, such as blood, is an indicator of respiratory and metabolic acid-base disorders. Whole blood can be separated into its cellular and plasma constituents. Red blood cells are known to undergo changes in size and shape with pH. Likewise, the major protein constituent in blood plasma—albumin—is also known to undergo pH-induced conformational changes. Albumin is the major protein constituent in other whole biofluids, such as amniotic fluid and cerebrospinal fluid, where pH is also an indicator of functional metabolism. We therefore wanted to investigate the role of albumin as a pH indicator.

Nonlinear ultrasound has previously been shown to be sensitive to changes in protein conformation. In protein solutions, for example, conformational changes can lead to modifications in protein hydration. The shift in solute-solvent interactions affects the frequency content of an ultrasound wave propagating through the sample. By analysing the complex frequency spectra

collected for each sample, we can correlate conformational changes of albumin in solution to pH. This indirect method of estimating pH relies on multilinear regression, in which several frequencies are used to develop a calibration model. Additionally, the absolute albumin content can vary between samples and also contributes to the observed frequency profiles. For biological applications, it is necessary to correct for these baseline differences in protein content. Data preprocessing steps can be used to mitigate the differences and allow for pH estimation over a range of protein concentrations. This work starts with an exploration of bovine serum albumin protein solutions and extends to include unfiltered serum samples from several individuals.

2.2 Manuscript

Ultrasonic Frequency Analysis for Estimating pH in Albumin-Rich Biofluids

Authors:

Andrien A. Rackov[1] and David H. Burns[1,2]

[1]Department of Chemistry, McGill University, Montreal, QC, Canada

[2]Department of Chemistry, University of New Brunswick, Fredericton, NB, Canada

2.3 Abstract

Ultrasound is known as a non-invasive imaging modality capable of propagating through highly scattering media such as tissue, blood, and other biological fluids, yet currently provides little chemical information. We have developed a straightforward and rapid methodology for estimating pH in albumin-rich biofluids based on analysis of ultrasonic frequencies. Albumin is the most abundant protein in serum and undergoes conformational changes with pH. It was shown that when ultrasound propagated through albumin solutions, the attenuation of collected ultrasound signals increased with pH. By measuring the ultrasound frequency spectra at several albumin concentrations and pH values, the pH of the solutions could be determined by multilinear regression. Differences in absolute protein content contributed to signal differences in the

frequency profiles and were minimized through normalization of each spectrum by the sum of all its frequency intensities. This strategy was applied to human serum samples from multiple donors, for which a multilinear regression model was developed with a coefficient of determination (R^2) of 0.93 and a standard error of estimate (SEE) of 0.077 pH units. The use of albumin as a pH indicator opens the doors for estimations in other albumin-rich media, such as amniotic fluid and cerebrospinal fluid.

Keywords

Ultrasound frequency analysis, pH, albumin, serum, biofluids, multilinear regression

Abbreviations

BSA: bovine serum albumin

HSA: human serum albumin

SECV: standard error of cross-validation

SEE: standard error of the estimate

2.4 Introduction

2.4.1 Background

In a clinical setting, the pH of biofluids such as blood, amniotic fluid, or cerebrospinal fluid is an indicator of functional metabolism.[2.1–2.4] Current techniques are mostly invasive: samples are drawn and tested using external pH probes, or blood gas analysers.[2.5] Alternately, pH probes can be embedded for continuous pH monitoring applications.[2.6,2.7] This has prompted the exploration of non-invasive approaches for determining pH, including optical absorption spectroscopy and measurement of exhaled gas concentrations.[2.8,2.9] However, optical methods have a limited depth of measurement and estimation of pH from exhaled gas concentrations requires collection of multiple additional parameters. Moreover, these measurements have been restricted to arterial blood pH and do not address the need for measurements of other difficult-to-access fluids, such as amniotic and cerebrospinal fluids.

Ultrasound has proven itself as a non-invasive medical imaging tool [2.10], but it has also been used to provide physical and chemical information about systems.[2.11] Unlike light, ultrasound is a mechanical wave that can propagate through opaque solids and turbid media. Ultrasound spectroscopy can be used in highly concentrated systems, such as undiluted biological samples, that would otherwise prove challenging for existing optical techniques.[2.12–2.16]

In particular, nonlinear ultrasound has shown much promise as a quick, cost-effective technique for chemical determinations. In addition to attenuation of waves, nonlinear phenomena, including distortion in waveform and the generation of frequency harmonics, add to the complexity of the ultrasound frequency spectra.[2.17] Analysis of nonlinear ultrasound frequency components by multivariate calibration techniques, such as multilinear regression, has been successfully applied to the simultaneous determination of multicomponent mixtures of solvents, and also for alcohol and sugar content in commercial beverages.[2.18,2.19] Unlike previous ultrasound attenuation measurements techniques, where it is often necessary to collect frequency spectra while varying additional parameters such as temperature and pH, the nonlinear contribution to the signal has been shown to be sufficient to provide a unique spectrum from a single measurement.

Nonlinear effects are most prominent in liquids with comparatively low acoustic attenuation, such as water and amniotic fluid, and other biological fluids.[2.20] Distortions in waveform arise from a difference in the velocity of propagation between the compression and rarefaction phases of acoustic waves due to non-negligible differences in density between the two phases. Local particle velocities are greater in the compression phase than in the rarefaction phase and lead to a cumulative distortion of a sine wave into a sawtooth wave.

This medium-dependent waveform distortion leads to the generation of new frequency components that are harmonically related. Though ultrasound absorption techniques have been used to study protein solutions[2.21–2.23], exploiting nonlinear ultrasound and these new frequency components may be particularly useful for chemical determinations.

We present a multivariate ultrasonic frequency analysis approach for estimating pH in BSA solutions, and further extend this to pH measurements of unfiltered human serum. We show that pH can be estimated for samples with varying total protein content through use of multilinear regression models, where several frequencies are selected for calibration.

2.4.2 Theory

Albumin, the major protein constituent in serum, amniotic fluid, and cerebrospinal fluid, is responsible for a wide array of functions within the body, including transport and binding of many endogeneous and exogeneous compounds, as well as mediating blood volume and plasma oncotic pressure.[2.24] Due to the great physiological importance the protein, albumin has been subject to in-depth study and is well characterized in the literature. It is known to undergo pH-induced conformational changes [2.25–2.27], which have been studied using ultrasonic absorption and speed of sound measurements[2.28,2.29] , but to our

knowledge, no attempt has been made to use ultrasonic frequency information of biofluids to estimate pH.

Albumin exists in 5 conformations, depending on the pH of its environment: Expanded (E; pH<2.7), Fast (F; pH 2.7-4.3), Normal (N; pH 4.3-8), Basic (B; pH 8-10), and Aged (A; pH>10). Transitions between the F, N and B forms (between pH 2.7 and 10) are known to be reversible. [2.30] Under slightly alkaline conditions, between pH 7.0 and 9.0 (this covers the physiological range), human serum albumin (HSA) and bovine serum albumin (BSA) undergo a conformational change, known as the N-B transition. The B isomerization is characterized by an unfolding of the protein with loss of rigidity, and has a particular effect on the N-terminus region and ligand binding .[2.26] These changes in conformation have been reported to influence the nonlinear behavior of ultrasound waves[2.31] and can be described by a nonlinear parameter, derived below.

Variations in the pressure of a medium can be related to variations in density though the equation of state

$$p = p(\rho, s) \qquad (2.1)$$

where p is pressure, ρ is the density, and s is the entropy of the system.[2.32] This can be expressed as a Taylor series expansion

$$p = p_0 + \frac{A}{1!}\left(\frac{\rho-\rho_0}{\rho_0}\right) + \frac{B}{2!}\left(\frac{\rho-\rho_0}{\rho_0}\right)^2 + \frac{C}{3!}\left(\frac{\rho-\rho_0}{\rho_0}\right)^3 + \cdots \qquad (2.2)$$

$$\text{with} \begin{cases} A = \rho_0 \left[\left(\frac{\partial p}{\partial \rho}\right)_s\right]_{\rho=\rho_0} \\ B = \rho_0^2 \left[\left(\frac{\partial^2 p}{\partial \rho^2}\right)_s\right]_{\rho=\rho_0} \\ C = \rho_0^3 \left[\left(\frac{\partial^3 p}{\partial \rho^3}\right)_s\right]_{\rho=\rho_0} \end{cases}$$

and coefficients *A*, *B*, and *C* dependent on temperature. The strength of the nonlinearity of a medium can be described by a dimensionless nonlinearity parameter, given by the ratio of quadratic to linear terms of the Taylor series expansion:

$$\frac{B}{A} = \frac{\rho}{c_0^2}\left(\frac{\partial^2 p}{\partial \rho^2}\right)_{s,\rho=\rho_0} \qquad (2.3)$$

where ρ is density, c is the sound speed, and p is the pressure in an isentropic system, s.[2.33] The acoustic nonlinear parameter *B/A* is dependent on the chemical composition of solutions and upon the molecular structure of the solute: an increase in the amount of charged groups (e.g. through changes in pH) of the solute molecules increases the nonlinearity parameter.[2.31,2.34] As the nonlinearity parameter for a given sample changes, so does the frequency content of the propagating ultrasound pulse.[2.18] Therefore, we expect to see a change in frequency profile as the charged groups on albumin change with increasing pH.

2.5 Materials and Methods

2.5.1 Ultrasound Apparatus

Ultrasound spectra were collected with the ultrasound system shown in Figure 2-1. Ultrasonic pulses were generated by a 500PR pulser-receiver from Panametrics (Waltham, MA) and collected with a Handyscope HS3 USB oscilloscope (TiePie engineering, Sneek, FR) sampling at 50-MHz with 12-bit resolution. The pulser-receiver generated a <20 ns, 250 V negative impulse with a 0.002% duty cycle. A pulse repetition rate of 1 kHz allowed any echoes in the sample to attenuate to baseline noise levels before the next pulse. Pulses were transmitted and received with two wideband transducers centered at 5 MHz from Technisonic (Fairfield, CT). The two transducers were threaded and fastened into parallel sides of a 3.5 mL anodized aluminum sample cell with a 2.4 cm pathlength, such that the transducers were in direct contact with the sample. This aluminum sample cell was temperature controlled with a thermoelectric cooling device from TE Technology (Traverse City, MI) set to 22.00 ± 0.05°C. For measurements exceeding a pH of 8.5, a sample cell made of polyoxymethylene (Delrin) with the same dimensions as the aluminum cell was used to house the samples, as aluminum is more prone to corrosion outside the 4.5-8.5 pH range.[2.35] The temperature was kept constant at 22.0 ± 0.2°C.

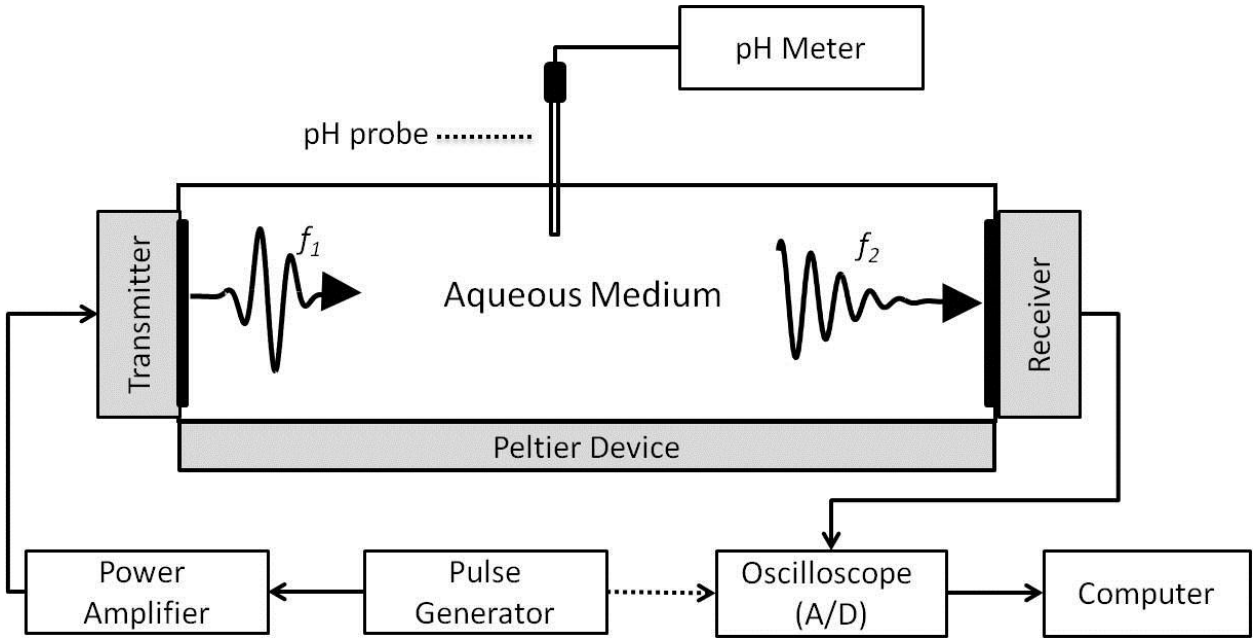

Figure 2-1: A schematic diagram of the experimental setup. Transmitting and receiving ultrasound transducers were embedded in a temperature controlled aluminum cell and were in direct contact with the sample. As the initial pulse (f_1) produced by the transmitter propagated through the sample, it was subject to attenuation and nonlinear phenomena, resulting in a modified waveform (f_2), as recorded by the receiver. The pH probe was positioned out of the direct path of the transducer elements, so as not to interfere with the ultrasound measurements.

2.5.2 Reagents

Deionized water from a Millipore Milli-Q OM-154 water purification system (Billerica, MA) was used for all experiments. All chemicals were obtained from Sigma-Aldrich (St. Louis, MO), unless otherwise noted.

2.5.3 BSA and Human Serum Samples

A 50 g/L bovine serum albumin stock solution was prepared from lyophilized powder (≥98%) dissolved in buffer (10 mM PBS, 0.14 M NaCl, pH 7.4). This stock solution was used to make all subsequent BSA standard solutions. Unfiltered human serum samples from different donors were obtained from Bioreclamation LLC (Liverpool, NY). Total serum protein levels were determined spectroscopically by means of the Biuret reaction, as described elsewhere.[2.36]

2.5.4 pH and Ultrasound Measurements

An Orion 2 Star pH meter from Thermo Scientific (Ottawa, ON) coupled with a MI-410 Micro-Combination pH Electrode from Microelectrodes (Bedford, NH) was used for all pH measurements. The pH electrode was calibrated with three certified buffer solutions (pH 4.00, 7.00, 10.00) from Fischer Scientific (Nepean, ON) prior to any measurements. The pH value and ultrasound profile were recorded when the pH meter gave a stable reading. For each sample, a total of 500 ultrasound spectra were collected. The pH probe was soaked in a

Tergazyme solution (Alconox, White Plains, NY) for a few minutes between runs to remove protein from the electrode.

2.5.5 Methods of Data Analysis

The raw data for each sample consisted of 500 time-domain spectra, which were aligned to the global maximum of the time series and subsequently averaged to improve the signal-to-noise ratio. Every spectrum spanned a time window of 100 μs, equivalent to 5000 recorded points. This time window was sufficiently large enough to observe the main pulse signal, as well as two reflections. The reflections did not overlap with the main pulse and were excluded from further data analysis to reduce noise contributions. The reflections were excluded by selecting a time interval that contained only the main pulse. The start of pulse was taken as the point where the signal first deviated from baseline and the end point where the signal returned to baseline, in this case the interval from 800 to 2200. This data interval was then subject to a fast Fourier transform, which expressed the pulse in terms of its frequency components. The new frequency-domain spectra were then normalized by the area of each spectrum. This area-normalization consisted of taking the sum of all the points in a spectrum and dividing each point by the sum.

Following these preprocessing steps, the data were divided into calibration and test sets, with roughly 30% of measurements assigned to test sets.

Stagewise multilinear regression was used to determine which linear combination of frequencies best describe the data from the calibration set, in the form:

$$Y = b_0 + b_1X_1 + b_2X_2 + ... + b_nX_n \qquad (2.4)$$

where Y is the dependent variable (here the pH of the sample), $\{X\}$ are the independent variables (the magnitude at a given ultrasound frequency), and $\{b\}$ are the calibration coefficients. These coefficients may be determined when both Y and $\{X\}$ are known. First, the pH was regressed linearly against the ultrasound magnitude at each frequency. The frequency with the highest correlation coefficient was saved and the variance contribution of this frequency subtracted from the spectra. This process was repeated iteratively on the reduced data set to select the frequencies with the highest correlation to pH. An F-test was used to choose the most parsimonious model (α = 0.05). Independent test sets were used to validate calibration models. The stagewise multilinear regression routine is presented in Arakaki and Burns. [2.37] All data processing was done in Matlab (The MathWorks Inc., Natick, MA, R2010a).

2.5.6 Calculation of Albumin Conformation Fractions

In order to understand the distribution of albumin conformations present at any given pH, the fractional compositions of albumin were calculated from dissociation constants of the midpoint pH values of four reported conformational

transitions: $K_{EF} = 10^{-2.7}$, $K_{FN} = 10^{-4.3}$, $K_{NB} = 10^{-8}$, and $K_{BA} = 10^{-10}$.[2.38] The formal concentration, F, is found from the mass balance for albumin:

$$F = [H^+]^4 + K_{EF}[H^+]^3 + K_{EF}K_{FN}[H^+]^2 + K_{EF}K_{FN}K_{NB}[H^+] + K_{EF}K_{FN}K_{NB}K_{BA} \quad (2.5)$$

The fractional composition of any conformation, $\propto_{conf}$, is then simply the ratio of that conformation's concentration, $[conf]$, and the formal concentration:

$$\propto_{conf} = \frac{[conf]}{F} \qquad (2.6)$$

Curves of the fractional compositions were generated by calculating the fractional compositions over a range of pH values. These curves (Figure 2-2b) show what fraction of an albumin conformation is present at any given pH and highlight the pH dependence of albumin.

2.6 Results and Discussion

2.6.1 Dependence of BSA Conformation on pH

In order to determine the effect of pH on the ultrasound profile, a model system consisting of BSA solutions was used for initial investigation. To highlight the effect of pH on the signal, ultrasound profiles were collected for a fixed concentration of albumin solutions (40 g/L), over a pH range of 6-10. Figure 2-2 presents the ultrasound spectra at 5 different pH values, along with a modeled conformation distribution. For pH 6, albumin is predominantly in the N (normal) conformation and a small change in signal is observed in going to pH 7, where

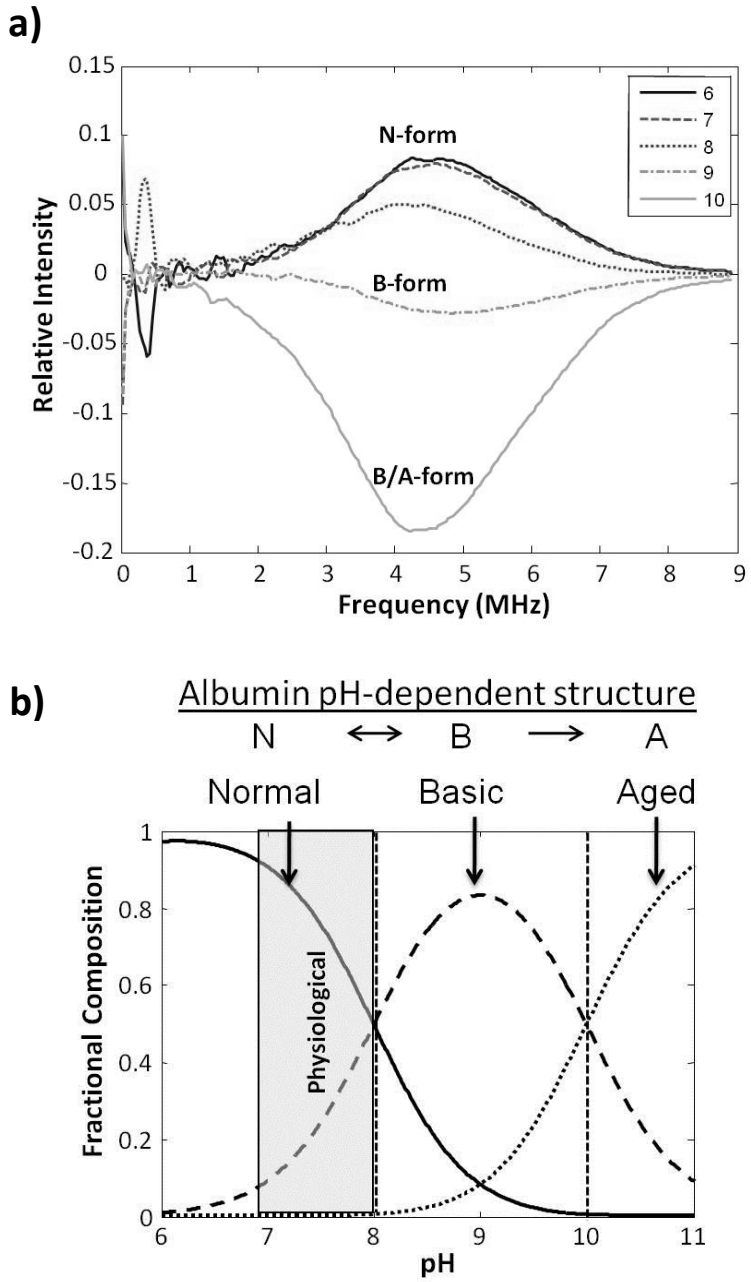

Figure 2-2: a) Mean-centered frequency profiles are shown for a 40g/L BSA solution at 5 different pHs: 6, 7, 8, 9, and 10 respectively. b) The changes in profile can be related to changes in conformation of albumin. The fractional composition curves were generated from the pH values of albumin's conformational transition points.[2.38]

albumin is still mostly in the N conformation. As the fraction of albumin in the B (basic) conformation increases (pH 9), the signal decrease is much more pronounced, and becomes quite large as the fraction of the A (aged; denatured) conformation increases (pH 10).

Interestingly, this signal change is not monotonic for all frequencies over this pH range, especially the lower frequencies. At around 0.4 MHz, there is a distinct increase in signal from pH 6 to 8, but decreases again when going to a higher pH. Interestingly, ultrasound absorption in this 0.4 MHz region has previously been reported to be associated with a proton-transfer reaction at lower pH. [2.29,2.39] This peak may represent a frequency-dependent component of the N-B transition associated with solute-solvent equilibrium. In general, the observed changes in frequency profile with pH are consistent with conformational changes of albumin and the resulting modifications in solute-solvent interactions.

A set of controls without albumin, consisting of saline solutions (0.14M) at pH 6, 8, and 10, were measured to investigate the contribution of changes in hydrogen ion concentration to the signal. No significant differences were observed under these experimental conditions, supporting the role of albumin as a pH indicator.

2.6.2 pH Measurements of BSA

In vivo, pH is strictly regulated to between 7.35 and 7.45, and is dominated by albumin in the N conformation. A trial was designed to investigate the effect of pH on the ultrasound profile in physiological pH range. This was done by varying the pH of 3.00 mL BSA solution with 5 µL aliquots of 0.2 N NaOH between pH 7 and 8 (about 20 additions), at three BSA concentrations, namely 30, 40, and 50 g/L. The pH value and ultrasound profile were recorded when the pH meter gave a stable reading.

Figure 2-3 shows the mean-centered frequency profiles between pH 7 and 8.1 for a 50 g/L BSA solution, along with a calibration model at this concentration. The signal exhibited an overall decrease in intensity as pH increased, consistent with the previous observations from the broad range pH ultrasound profiles. In contrast to the rest of the spectrum, the signal showed an increase in intensity with increasing pH at around 0.4 MHz. Over this narrower pH interval, the non-monotonic behaviour seen over the pH 6 to 10 range was not present.

The figures of merit for pH estimation at individual concentrations, along with model frequencies, are given in Table 2-1. The multilinear regression models for each fixed concentration showed a strong correlation between frequency and pH, with coefficients of determination (R^2) of 0.99, standard errors

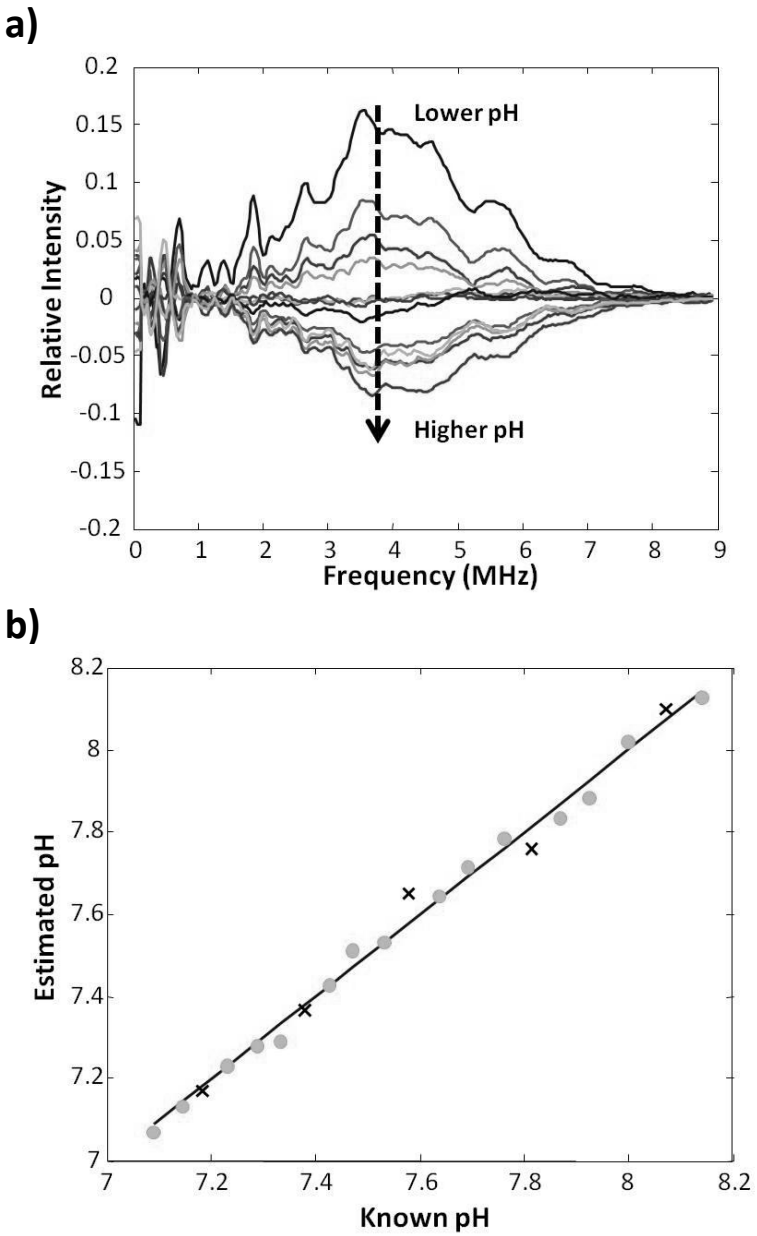

Figure 2-3: a) Mean-centered spectra for a 50 g/L solution of BSA are shown, and exhibit an overall increase in signal attenuation with increasing pH. b) The frequency profiles were used to generate a calibration model (o's) for the 50g/L solution of BSA. A subset of measurements, excluded from the calibration model, was used as an independent test set (x's).

Table 2-1: Figures of merit for the estimation of pH at three different concentrations of BSA.

BSA concentration	R^2	SECV (pH units)	SEE (pH units)	Model frequencies (MHz)		
30 g/L	0.99	0.041	0.064	0.32	0.57	0.61
40 g/L	0.99	0.041	0.065	0.18	0.57	2.68
50 g/L	0.99	0.048	0.065	2.14	0.36	0.29

of cross-validation (SECV) ≤0.048 pH units and standard errors of the estimate (SEE) ≤0.065 pH units. Although these models may be appropriate for estimating pH of albumin solutions of the same concentration, larger SEE values (between 0.140 and 0.350 pH units) were obtained when using a model at one concentration to estimate the pH at another concentration.

In order to accommodate real-world differences in albumin levels of individuals, a two-concentration model (30 and 50 g/L) was developed, with the 40 g/L samples used as an independent test set (see Figure 2-4). This multivariate model (R^2 = 0.94, SECV = 0.101 pH units) was able to estimate the pH of the independent test set with a SEE of 0.102 pH units. The three frequencies chosen for the model were 0.57, 2.68, and 1.07 MHz and included two frequencies that were also selected for some of the previous single concentration BSA models. In contrast to the single concentration models, which selected frequencies optimized for specific concentrations, the mixed concentration model chose frequencies and coefficient weightings appropriate for a BSA concentrations between 30 and 50 g/L. The two-concentration model was developed using data that was normalized by the area of each individual spectrum. The normalization consisted of rescaling individual spectra by dividing the magnitude of each frequency in a spectrum by the sum of all frequency

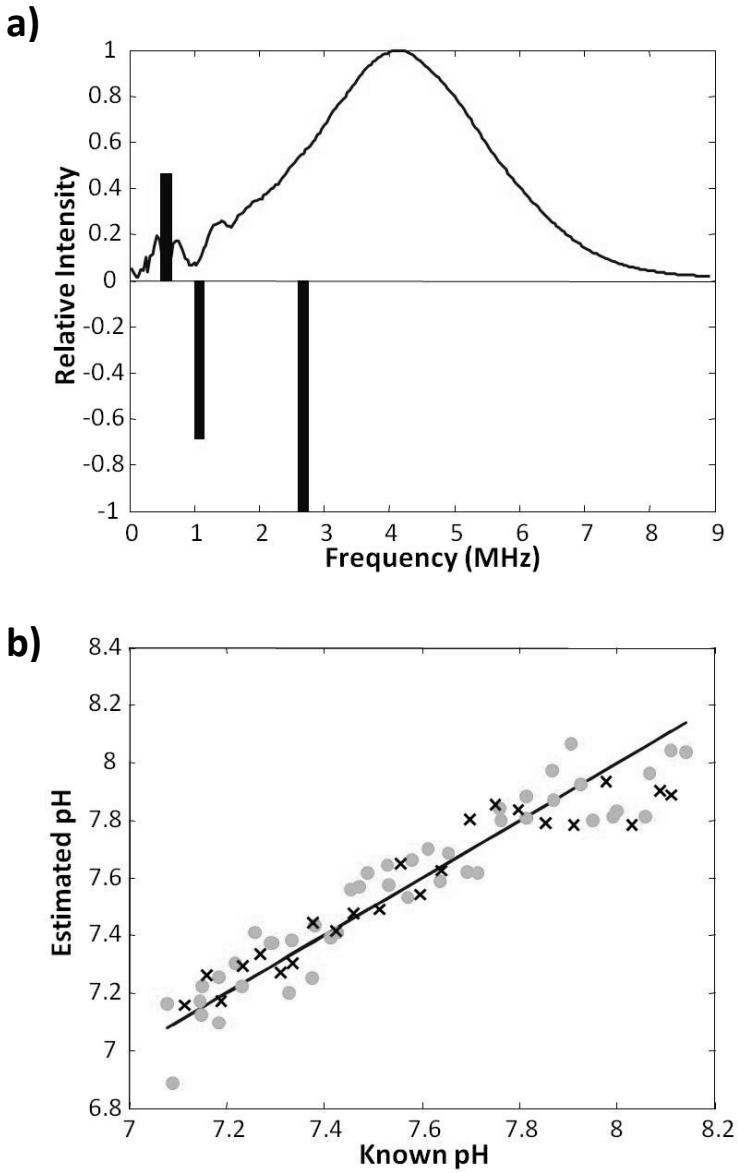

Figure 2-4: a) Frequencies selected for the two-concentration (30 and 50 g/L BSA) calibration model are shown, along with their relative coefficient weightings. b) The calibration model (o's) was used to estimate the pH of samples with a BSA concentration of 40g/L (x's). All spectra used for calibration were area-normalized to minimize baseline differences in concentration.

magnitudes. The effect of this area-normalization is presented in Figure 2-5. This reduced spectral offsets attributed to concentration differences and emphasized changes associated with pH.

Given the standard addition methodology of the experiment, in which roughly 100 µL of NaOH solution was added to 3 mL of albumin solution, a control run was performed to determine the effect of dilution on the signal. This consisted of adding 100 µL of deionized water in place of NaOH to one of the albumin solutions and comparing the total change in frequency magnitude before and after the addition of water. Overall, the total signal change associated with dilution was roughly 10 times smaller than the change observed when NaOH was added and the pH was modified.

2.6.3 pH Measurements of Human Serum

To test the robustness of the ultrasound methodology for pH estimation, additional measurements were made in human serum. Human serum provides a more complex matrix, as it contains a wide array of proteins—of which albumin is still the most plentiful (~70%)–as well as dissolved nutrients, and other metabolites.[2.40] The pH for five serum samples from different donors was varied in the same way as with the BSA solutions, by 5 µL additions of 0.2 N NaOH to 3.00 mL of serum. The pH value and ultrasound profile were recorded when the pH meter gave a stable reading. It should be noted that the starting pH values of

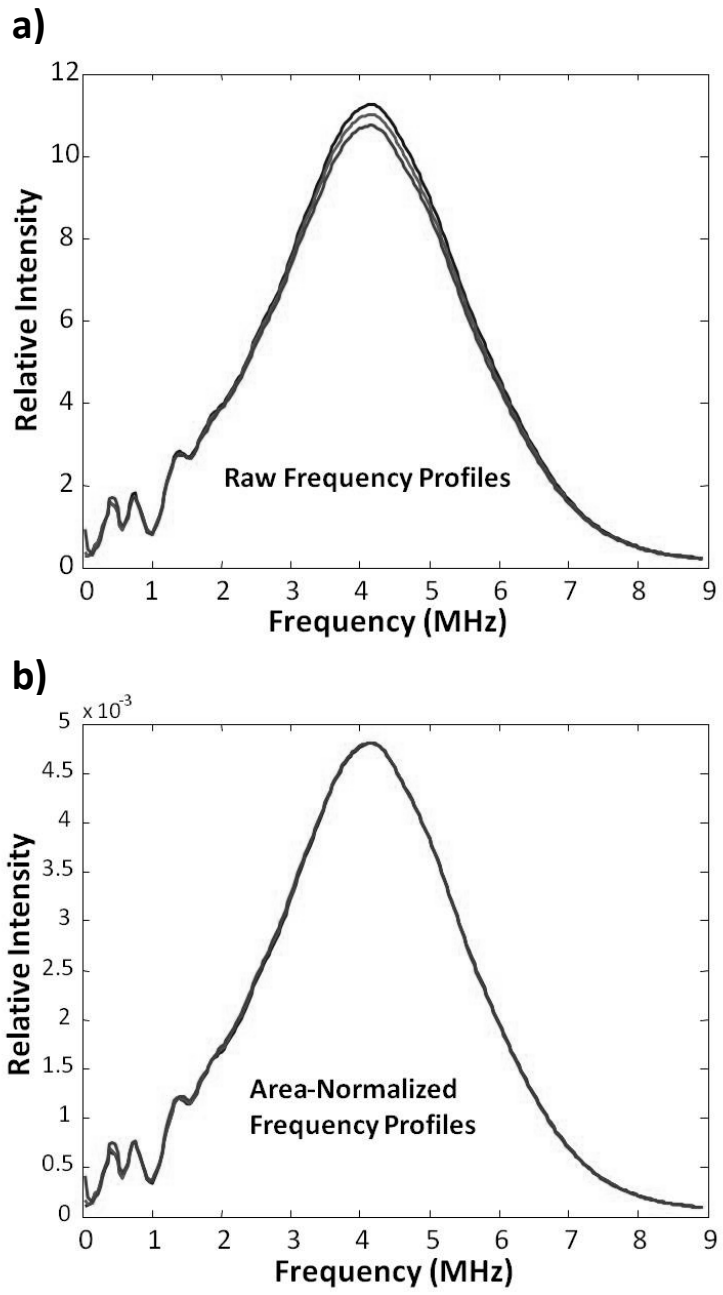

Figure 2-5: a) Raw mean frequency profiles and b) area-normalized frequency profiles of 30, 40, and 50 g/L BSA solutions are shown. Area-normalization was used to rescale the data and minimize spectral differences resulting from concentration changes prior to developing models for the estimation of pH.

the serum samples were higher than the normal physiological pH, at around 7.8 pH units. Exposure of serum to air leads to an escape of CO_2 from the sample and consequently an increase in pH. [2.41] This can be illustrated by the carbonic acid equilibrium present in serum:

$$H_2O + CO_2 \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^- \qquad (2.7)$$

Higher pH ⟷ **Lower pH**

The partial pressure difference between CO_2 in serum (~40 Torr) and dry air (~0.3 Torr)[2.42] drives the loss of CO_2 from the serum sample and shifts the equilibrium in equation 2.7 to the left, resulting in an increase in serum pH.

To examine if the relationship between albumin conformation in solution and pH was also applicable to human serum, the two-concentration BSA model was used to estimate the pH for one of the serum sample runs. Figure 2-6 presents the outcome of the human serum estimates, which are characterized by a distinct offset from the calibration model. Despite the offset, which is attributed to the presence of additional protein constituents in serum, the estimates follow the same trend as the calibration model.

In order to account for the additional components found in serum, a model was developed using serum samples from three donors, with a fourth donor used as an independent test set (Figure 2-7). As with the 2-concentration BSA model,

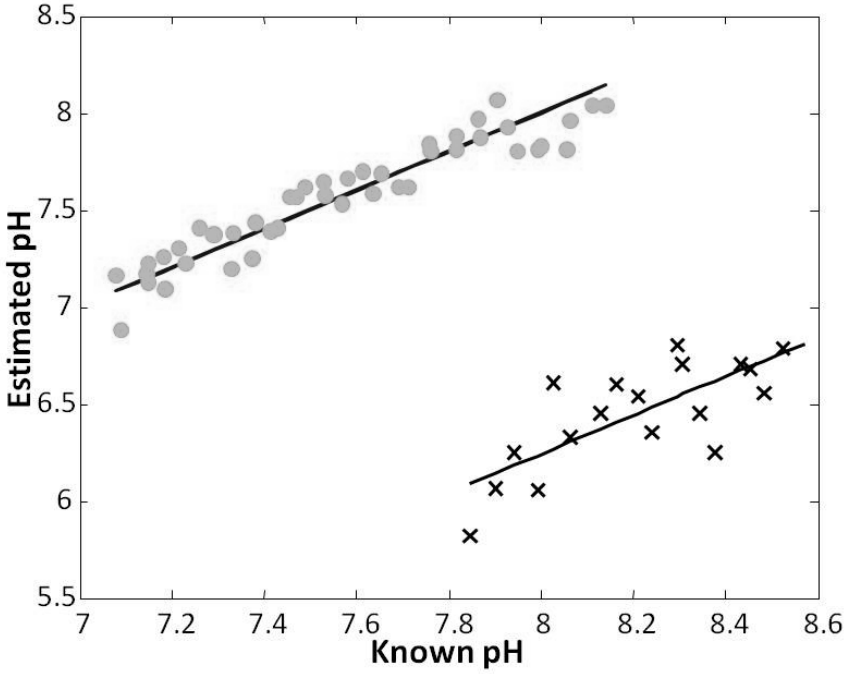

Figure 2-6: The two-concentration BSA model (o's) was used to estimate the pH of human serum samples from one donor (x's). Though the estimated serum sample pH values follow the same trend as the calibration model, there is a distinct offset.

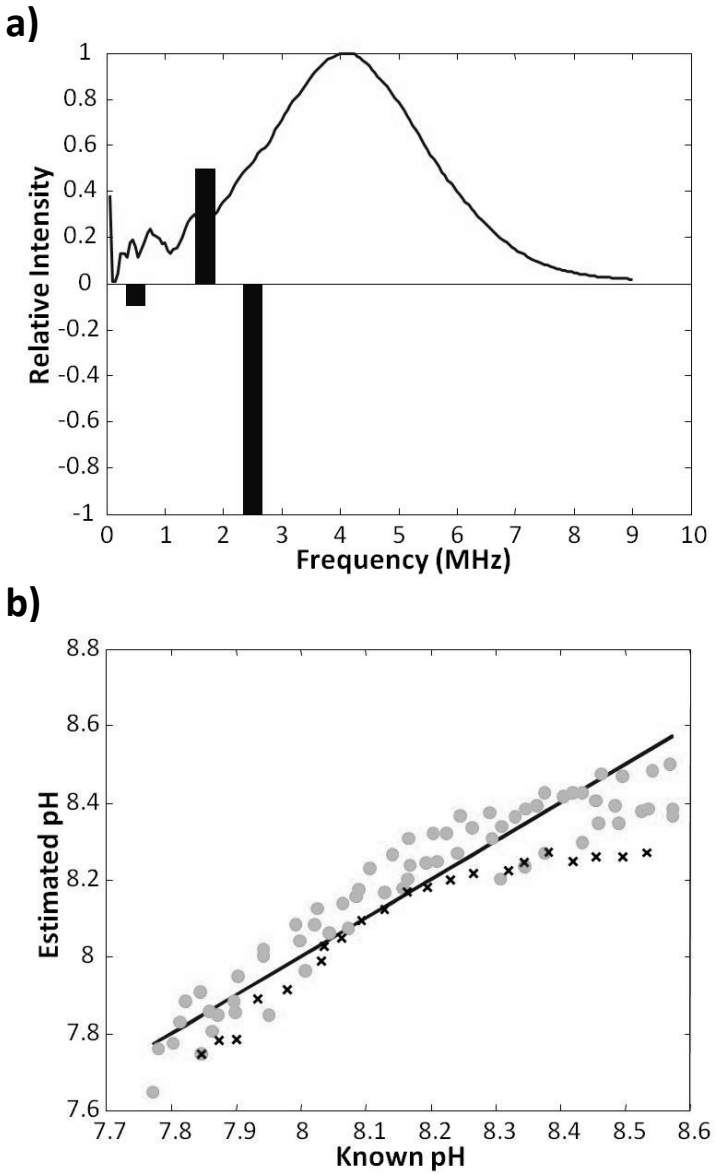

Figure 2-7: a) Frequencies selected for the human serum calibration model, along with their relative coefficient weightings, are presented. b) The calibration model (o's) was developed using data from three different serum samples (different donors). Data from a fourth donor, excluded from the calibration set, was used as an independent test set (x's). All data were area-normalized prior to calibration to minimize signal differences due to concentration.

the data was area-normalized prior to developing a model. Using serum from different donors introduced the variability that is found between total serum protein levels of individuals; in this particular case, the total serum protein concentrations fell between 68 and 78 g/L. The area-normalization step reduced offsets between spectra due to variations in protein concentration so that, despite some of the compositional differences, the multivariate model (R^2 = 0.93, SECV = 0.083 pH units) was still able to estimate the pH values in human serum with a SEE of 0.077 pH units. Compared to the SEE for the serum samples estimated by the BSA model (SEE of 0.182 pH units), this new human serum calibration model reduced the error by more than half.

The three frequencies selected for this model include 2.50, 0.50, and 1.70 MHz. The first two frequencies selected for the human serum model were very similar to those chosen for the two-concentration BSA model. The consistency between the human serum and BSA models highlights the importance of these frequency bands (0.50-0.57 and 2.50-2.68 MHz) for estimating pH. Moreover, these frequency regions may be indicative of changes in hydration sphere of albumin and may be linked to albumin's structure. Interestingly, the human serum model displayed a deviation from linearity at a pH of about 8.4, which also coincides with the appearance of a third albumin conformation, Aged, in addition to the Normal and Basic conformers (see Figure 2b). The deviation may also

have resulted from non-albumin protein contributions to the signal. Albumin, immunoglobulin G, and transferrin make up close to 90% of all serum proteins by mass. Although albumin is the dominant species, immunoglobulin G, transferrin and the remaining plethora of proteins in these serum samples may also be undergoing conformational changes with pH[2.43,2.44] and contributing to the signal. For instance, the immunoglobulin G subclasses IgG_1 and IgG_3 both have isoelectric points around a pH of 8.4,[2.45] which could influence the viscosity of the sample. Nevertheless, the results demonstrate the feasibility of an ultrasound methodology for determining pH in human serum.

2.7 Conclusions

Ultrasound spectra were collected for solutions of BSA and human serum samples at physiological pH and concentration values. Variations in the conformational fractions of albumin with pH, as well as the total concentration of albumin, contributed to changes in the ultrasonic frequency profiles. Applying multivariate calibration techniques allowed for estimation of pH at different baseline concentrations of albumin; an area-normalization was necessary to mitigate concentration effects in multi-concentration models. Similar frequency bands were chosen for the two-concentration BSA model and the human serum model, suggesting that albumin conformation changes might be frequency

dependent. These frequency bands may be particularly sensitive to solute-solvent interactions, namely changes in hydration of albumin.

In addition to albumin solutions, estimates of pH were also possible in human serum, with a SEE < 0.080 pH units and sufficient to detect acid-base imbalances. It should be possible to further refine the models by exploring the contribution of non-albumin components to the signal in order to find frequencies that have minimal overlapping contributions from non-albumin proteins. A reasonable starting point would include the contributions from immunoglobulin G, transferrin, fibrinogen, and α-antitrypsin. Another strategy to improve estimates and reduce the SEE would be to make more measurements over a smaller pH interval. The pH range could be narrowed down further to make this approach suitable for measuring differences in smaller pH ranges, such as 7.35 to 7.45.

On the whole, these results demonstrate the potential of ultrasonic frequency analysis as a tool for measuring pH in albumin-containing systems. Ultrasound frequency analysis provides a platform for direct and quick sample measurements, with real-time monitoring potential. Coupled with the non-invasive capabilities of ultrasound, this methodology is an appealing candidate for point-of-care measurements, meriting further exploration.

2.8 Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Fonds de recherche du Québec—Nature et technologies.

2.9 References

2.1. Jones, S. A. & Walter, J. H. Diagnosis and treatment of severe metabolic acidosis. *Paediatr. Child Health (Oxford).* **17,** 260–265 (2007).

2.2. Capasso, G. & Unwin, R. Electrolytes and acid–base: common fluid and electrolyte disorders. *Medicine (Baltimore).* **39,** 317–324 (2011).

2.3. Fadel, H. E., Northrop, G., Misenhimer, H. R. & Harp, R. J. Acid-base determinations in amniotic fluid and blood of normal late pregnancy. *Obstet. Gynecol.* **53,** 99–104 (1979).

2.4. Andrews, R. J. M. D., Bringas, J. R. B. S. & Alonzo, G. B. S. Cerebrospinal Fluid pH and PCO2 Rapidly Follow Arterial Blood pH and PCO2 with Changes in Ventilation. *Neurosurgery* **34,** 466–470 (1994).

2.5. Ganter, M. T. *et al.* How often should we perform arterial blood gas analysis during thoracoscopic surgery? *J. Clin. Anesth.* **19,** 569–75 (2007).

2.6. Venkatesh, B. Continuous Intra-arterial Blood Gas Monitoring. *Crit. Care Resusc.* **1,** 150 (1999).

2.7. Jin, W. *et al.* Continuous intra-arterial blood pH monitoring by a fiber-optic fluorosensor. *IEEE Trans. Biomed. Eng.* **58,** 1232–8 (2011).

2.8. Yang, Y., Soyemi, O. O., Landry, M. R. & Soller, B. R. Noninvasive in Vivo Measurement of Venous Blood pH During Exercise Using Near-Infrared Reflectance Spectroscopy. *Appl. Spectrosc.* **61,** 223–229 (2007).

2.9. Matsui, T., Hattori, H., Takase, B. & Ishihara, M. Non-invasive estimation of arterial blood pH using exhaled CO/CO2 analyser, microwave radar and infrared thermography for patients after massive haemorrhage. *J. Med. Eng. Technol.* **30,** 97–101 (2006).

2.10. Ophir, J. Elastography: A quantitative method for imaging the elasticity of biological tissues. *Ultrason. Imaging* **13,** 111–134 (1991).

2.11. Dukhin, A. S. & Goetz, P. J. *Ultrasound for Characterizing Colloids*. (Elsevier, 2002).

2.12. Troïani, D., Dion, J. R. & Burns, D. H. Ultrasonic quantification using smart hydrogel sensors. *Talanta* **83,** 1371–5 (2011).

2.13. Dion, J. R. & Burns, D. H. Ultrasonic frequency analysis of antibody-linked hydrogel biosensors for rapid point of care testing. *Talanta* **83,** 1364–70 (2011).

2.14. Dukhin, a S., Goetz, P. J. & Travers, B. Use of ultrasound for characterizing dairy products. *J. Dairy Sci.* **88,** 1320–34 (2005).

2.15. Dukhin, A. S., Goetz, P. J. & van de Ven, T. G. M. Ultrasonic characterization of proteins and blood cells. *Colloids Surf. B. Biointerfaces* **53,** 121–6 (2006).

2.16. Bonacucina, G., Misici-Falzi, M., Cespi, M. & Palmieri, G. F. Characterization of Micellar Systems by the Use of Acoustic Spectroscopy. *J. Pharm. Sci.* **97,** 2217–2227 (2008).

2.17. Duck, F. A. Nonlinear Acoustics in Diagnostic Ultrasound. *Ultrasound Med. Biol.* **28,** 1–18 (2002).

2.18. Dion, J. R. & Burns, D. H. Determination of volume fractions in multicomponent mixtures using ultrasound frequency analysis. *Appl. Spectrosc.* **65,** 648–56 (2011).

2.19. Dion, J. R. & Burns, D. H. Simultaneous determination of alcohol and carbohydrate content in commercial beverages by ultrasound frequency analysis. *Talanta* **86,** 384–92 (2011).

2.20. Muir, T. G. & Carstensen, E. L. Prediction of nonlinear acoustic effects at biomedical frequencies and intensities. *Ultrasound Med. Biol.* **6,** 345–357 (1980).

2.21. Edmonds, P. D. The molecular basis of ultrasonic absorption by proteins. *Bioelectromagnetics* **3,** 157–65 (1982).

2.22. Cho, K. C., Leung, W. P., Mok, H. Y. & Choy, C. L. Ultrasonic absorption in myoglobin and other globular proteins. *Biochim. Biophys. Acta* **830,** 36–44 (1985).

2.23. Pavlovskaya, G., McClements, D. J. & Povey, M. J. W. Ultrasonic investigation of aqueous solutions of a globular protein. *Food Hydrocoll.* **6,** 253–262 (1992).

2.24. Fanali, G. *et al.* Human serum albumin: from bench to bedside. *Mol. Aspects Med.* **33,** 209–90 (2012).

2.25. Tanford, C. & Buzzell, J. G. The Viscosity of Aqueous Solutions of Bovine Serum Albumin between pH 4.3 and 10.5. *J. Phys. Chem.* **60,** 225–231 (1956).

2.26. Dockal, M., Carter, D. C. & Rüker, F. Conformational transitions of the three recombinant domains of human serum albumin depending on pH. *J. Biol. Chem.* **275,** 3042–50 (2000).

2.27. Yadav, S., Shire, S. J. & Kalonia, D. S. Viscosity analysis of high concentration bovine serum albumin aqueous solutions. *Pharm. Res.* **28,** 1973–83 (2011).

2.28. Kessler, L. W. & Dunn, F. Ultrasonic investigation of the conformal changes of bovine serum albumin in aqueous solution. *J. Phys. Chem.* **73,** 4256–63 (1969).

2.29. Choi, P. K., Bae, J. R. & Takagi, K. Ultrasonic spectroscopy in bovine serum albumin solutions. *J. Acoust. Soc. Am.* **87,** 874–81 (1990).

2.30. Peters, T. J. *All about Albumin: Biochemistry, Genetics and Medical Applications.* (Academic Press, 1996).

2.31. Sarvazyan, A. P., Chalikian, T. V & Dunn, F. Acoustic nonlinearity parameter B / A of aqueous solutions of some amino acids and proteins. *J. Acoust. Soc. Am.* **88,** 1555–1561 (1990).

2.32. Khelladi, H., Plantier, F., Daridon, J. L. & Djelouah, H. Measurement under high pressure of the nonlinearity parameter B/A in glycerol at various temperatures. *Ultrasonics* **49,** 668–75 (2009).

2.33. Lauterborn, W., Kurz, T. & Akhatov, I. in *Springer Handbook of Acoustics* (ed. Rossing, T. D.) 257–293 (Springer New York, 2007). doi:10.1007/978-0-387-30425-0

2.34. Zhang, J., Kuhlenschmidt, M. S. & Dunn, F. Influences of structural factors of biological media on the acoustic nonlinearity parameter B/A. *J. Acoust. Soc. Am.* **89,** 80–91 (1991).

2.35. Engineers, U. S. A. C. of. in *EM 1110-2-1614* (USACE, 1995).

2.36. Gornall, A. G., Bardawill, C. J. & David, M. M. Determination of serum proteins by means of the biuret reaction. *J. Biol. Chem.* **177,** 751–66 (1949).

2.37. Arakaki, L. S. L. & Burns, D. H. Multispectral Analysis for Quantitative Measurements of Myoglobin Oxygen Fractional Saturation in the Presence of Hemoglobin Interference. *Appl. Spectrosc.* **46,** 1919–1928 (1992).

2.38. Sadler, P. J. & Tucker, a. pH-induced structural transitions of bovine serum albumin. Histidine pKa values and unfolding of the N-terminus during the N to F transition. *Eur. J. Biochem.* **212,** 811–7 (1993).

2.39. Barnes, C., Evans, J. A. & Lewis, T. J. Ultrasonic absorption of bovine serum albumin solutions in the frequency range 200 kHz-1 MHz. *J. Acoust. Soc. Am.* **78,** 6–11 (1985).

2.40. Anderson, N. L. & Anderson, N. G. The Human Plasma Proteome: History, Character, and Diagnostic Prospects. *Mol. Cell. Proteomics* **1,** 845–867 (2002).

2.41. Loken, H. F., Havel, R. J., Gordan, G. S. & Whittington, S. L. Ultracentrifugal Analysis of Protein-bound and Free Calcium in Human Serum. *J. Biol. Chem.* **235,** 3654–3658 (1960).

2.42. Posencheg, M. A. & Evans, J. R. in *Avery's Diseases of the Newborn* 367–389 (Elsevier Inc., 2012). doi:10.1016/B978-1-4377-0134-0.10031-9

2.43. Calmettes, P., Cser, L. & Rajnavölgyi, E. Temperature and pH dependence of immunoglobulin G conformation. *Arch. Biochem. Biophys.* **291,** 277–83 (1991).

2.44. Baldwin, D. A., De Sousa, D. M. R. & Von Wandruszka, R. M. A. The effect of pH on the kinetics of iron release from human transferrin. *Biochim. Biophys. Acta* **719,** 140–146 (1982).

2.45. Tracy, R. P., Currie, R. M., Kyle, R. A. & Young, D. S. Two-dimensional gel electrophoresis of serum specimens from patients with monoclonal gammopathies. *Clin. Chem.* **28,** 900–907 (1982).

Chapter 3

Estimation of Somatic Cell Counts in Breast Milk by Ultrasound and Near Infrared Spectroscopy

3.1 Foreword

The previous chapter presented the use of nonlinear frequency analysis for probing changes in molecular structures in homogeneous matrices dominated by proteins (i.e. protein solutions and serum samples). Unlike homogeneous serum samples, in which attenuation of ultrasound is mainly due to absorption, this section will explore turbid samples, where both absorption and scattering contribute to the observed attenuation of ultrasound. There are a variety of biological fluids, such as blood and milk, that exhibit scattering of ultrasound.[3.1,3.2] In milk, for example, the presence of fat globules contributes significantly to scattering, which can make quantification in this matrix more difficult. This has been known in the dairy industry for many years, where removal of fat from samples is commonplace prior to analysis.

In addition to fat globules, milk also contains leukocytes (white blood cells), which increase in response to mammary infection. As a result, total cell counts in milk are an indicator of mammary infection and inflammation. Ultrasound has previously been used for counting cells in suspensions of blood cells and in whole blood.[3.3,3.4] Volume fractions of cells in samples have been

estimated from ultrasound attenuation spectra and speed of sound measurements. Estimates of cell concentrations in suspensions have been reported for concentrations as low as 2×10^4 cells/mL.[3.5] This is one to two orders of magnitude lower than cell concentrations in associated with mastitis. Though a number of approaches have been explored for counting cells in dairy milk, a direct measurement technique for human milk is needed. Human milk can provide useful information on maternal health, where elevated cell counts are associated with mastitis.

To get a better understanding of how scattering by fat globules affects ultrasound measurements, a preliminary investigation was performed using commercial dairy samples. Ultrasound spectra were collected for dairy samples with fat contents ranging from 0 to 35% and are presented in Figure 3-0. Over this large range of fat concentrations, attenuation of ultrasound signal increases with fat content. Moreover, higher frequencies are attenuated more strongly than lower frequencies, which is consistent with increased scattering. This is particularly noticeable at fat concentrations of 10% and above, where signal peaks have been shifted to lower frequencies. However, the typical milk fat content in cows and other mammals, including humans, falls in the range of 3-5%.[3.6] The decrease in peak signal between 0 and 3.25% milk fat is fairly linear,

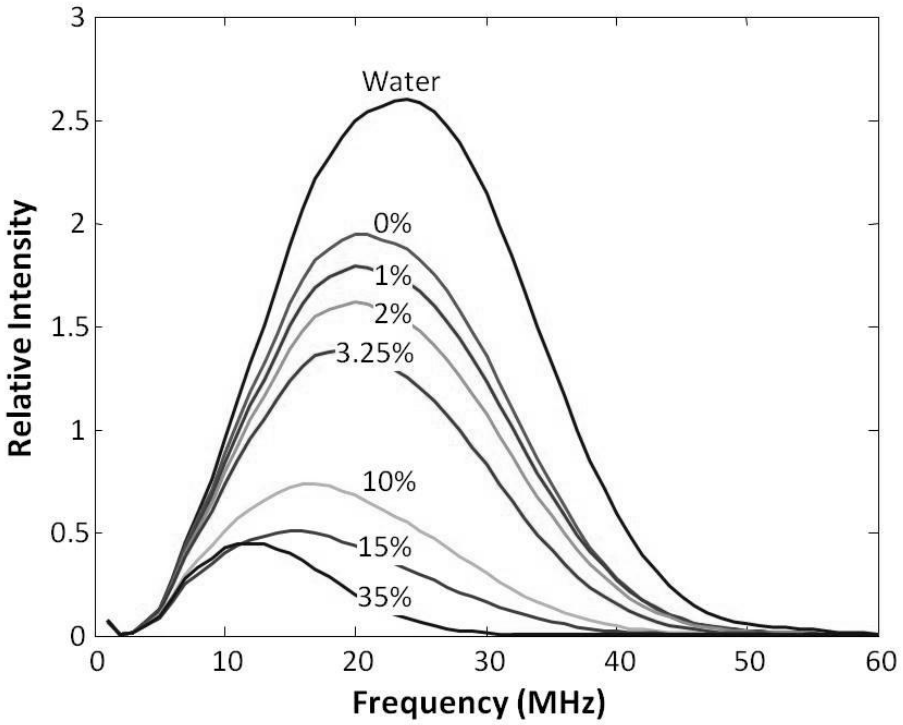

Figure 3-0: Ultrasound frequency spectra for commercial dairy samples. Attenuation of signal increases with fat content and is consistent with scattering of ultrasound by fat globules.

and therefore it should be possible to correct for systematic shifts of spectra from scattering using data scaling techniques.

In this chapter, we present ultrasound frequency analysis as a tool for estimating somatic cell counts in human milk, after correcting for scattering contributions to the signal. A multivariate classification approach, based on multilinear regression, is used to separate samples into low and high cell count groups. Additionally, we compare the ultrasound approach to NIR, which has previously been used in the dairy industry for estimating cell counts, but not in human milk.

3.2 Manuscript

Estimation of Somatic Cell Counts in Breast Milk by Ultrasound and Near Infrared Spectroscopy

Authors:

Andrien A. Rackov[1], Hilary M. Wren[2], Kristine G. Koski[2], Marilyn E. Scott[3], and David H. Burns[1,4]

[1]Department of Chemistry, McGill University, Montreal, QC, Canada

[2]School of Dietetics and Human Nutrition, McGill University (Macdonald Campus), Montreal, QC, Canada

[3]Institute of Parasitology, McGill University (Macdonald Campus), Montreal, QC, Canada

[4]Department of Chemistry, University of New Brunswick, Fredericton, NB, Canada

3.3 Abstract

Leukocyte counts in human milk are a useful indicator of subclinical mastitis, an asymptomatic inflammation of the breast. However, direct measurement of cells is complicated by the presence of fat globules. Near infrared spectroscopy, a technique that has previously been used in the dairy industry for direct measurement of somatic cell counts (SCC) in milk, was compared to ultrasound for estimating SCC in human milk. Ultrasound

spectroscopy has previously been used for measurements in turbid systems, without the need for dilution or reagents. Binary classification models were developed using logistic regression with genetic algorithm searching for selection of wavelets. After correcting NIR frequency spectra for scatter contributions by fat globules and applying a Haar wavelet transform to the data, we show that multivariate classification allows for separation of samples with low (≤150K cells/mL) and high (≥600K cells/mL) SCC. Sensitivity and specificity for cross-validated NIR estimates were 85% and 84%. Similarly, multivariate analysis of ultrasound spectra allowed for cross-validated separation of low and high SCC samples with a sensitivity of 90% and specificity of 79%. Both ultrasound and NIR methods had very low rates of misclassification, with models that used 3 or fewer wavelets for classification. Likewise, these techniques require no sample preparation and provide a rapid method for estimating SCC in human milk, with potential application for identifying subclinical mastitis in lactating mothers.

Keywords

Ultrasound frequency analysis, near infrared spectroscopy, human milk, subclinical mastitis, somatic cell count, multivariate classification

List of Abbreviations

NIR: near infrared

SCC: somatic cell count

SNV: standard normal variate

HWT: Haar wavelet transform

GA: genetic algorithm

3.4 Introduction

Monitoring the composition of human milk can provide useful information regarding the quality of the milk (macronutrient content) and also maternal health (somatic cell count). In the dairy industry, rapid measurement techniques are commonplace. However, current techniques provide semi-quantitative results[3.7] or require the use of reagents.[3.8] In contrast, near infrared (NIR) spectroscopy allows for direct measurement of milk without the need for fat removal, dilution, or the addition of reagents. NIR has been used to obtain information on fat, protein, and lactose content, as well as somatic cell count (SCC) of cow's milk from a single measurement.[3.9–3.11] It was shown that the absorbance at multiple NIR wavelengths correlated with the concentrations of the various milk constituents. Additionally, the analysis of macronutrients in human milk has also been demonstrated by NIR.[3.12] However, an approach for the direct determination of SCC in human milk has not been shown.

It has been proposed that monitoring breast milk leukocyte levels (and SCC) could be used as a diagnostic tool of subclinical mastitis, for which early

diagnosis is important for successful treatment and maintenance of prolonged breastfeeding.[3.13] Subclinical mastitis is an asymptomatic inflammation of the lactating breast and is associated with lactation failure, infant growth faltering during early postpartum period, and increased risk of mother-to-child transmission of HIV.[3.14–3.16] Unlike cow's milk, in which the dominant cells are macrophages, human milk contains predominantly epithelial cells.[3.17] Despite this, SCC levels have been observed to increase tenfold in cases of clinical mastitis and are associated with an increase in leukocyte levels.[3.18] Therefore, monitoring SCC should provide a useful tool for screening for mastitis.

Ultrasound spectroscopy has emerged as a promising technique for quantification in turbid systems and has been previously reported in the literature as a tool for determining fat content in dairy products.[3.2,3.19] As a mechanical wave, ultrasound can propagate through turbid media without the need for dilution or sample preparation. However, the ultrasound signal is still subject to absorption and scattering, which are dependent on the chemical composition of the sample and affect the frequency content of the signal.[3.20] Additionally, nonlinear distortion of ultrasound yields new frequencies and also contributes to the observed frequency profile.[3.21] This progressive distortion of a propagating ultrasound wave arises from a difference in particle velocity between the compression and rarefaction cycles, which results in the generation of

harmonics. Nonlinear distortion of ultrasound in a medium is sensitive to macromolecular structures [3.22] and cellular structures,[3.23] both of which are present in milk. Effectively, the degree of signal distortion is dependent on the chemical composition of the medium such that frequency content will change with sample composition.

Frequency analysis of nonlinear ultrasound spectra has been used in the determination of fractional composition in multicomponent mixtures, as well as estimating alcohol and carbohydrate content in commercial beverages.[3.24,3.25] Specifically, ultrasound pulses were analysed for changes in frequency content and by using several frequencies to develop multilinear calibration models, chemical quantification was possible from a single ultrasound measurement. Furthermore, these selected frequencies correlated well with changes in hydrogen bonding in liquid mixtures. In the case of milk, changes in SCC should contribute to the observed frequency profile.

We demonstrate here that the cellular content of milk can be estimated through analysis of ultrasound frequency data. This is done through multivariate classification, which separates milk samples into high or low cell counts. Additionally, we compare the ultrasound approach to NIR spectroscopy and show that NIR can also be used for estimation in human milk. To our knowledge,

no one has demonstrated determination of SCC in human milk by ultrasound or NIR.

3.5 Methods

3.5.1 Study Design

This cross-sectional study was conducted on an indigenous population living in 7 selected communities in Guatemala in May of 2012. Lactating women with infants aged either between 3 days to 1 month or 4 to 6 months were informed of the study and invited to participate. Eligible mothers who provided informed consent for participation of themselves and their infant were enrolled in the study. Women known to be treated with antibiotics were excluded from participation. The study included 57 mother-infant pairs at infant ages between 3 and 45 days, as well as 60 mother-infant pairs at infant ages between 122 and 183 days.

3.5.2 Ethical Considerations

Ethics approval was obtained from both McGill University in Canada and the Center for Studies of Sensory Impairment, Aging, and Metabolism (CeSSIAM) in Guatemala. All measures were designed to be culturally sensitive within an indigenous population. No risks were determined to be present for

participating mothers and infants due to the minimally-invasive nature of breast milk collection.

3.5.3 Collection of Milk Samples

A single unilateral milk sample was obtained from each participant by manual expression. Participants were asked to provide a full milk expression from the breast not recently used for feeding their infant. Prior to sample collection, the breast was cleaned with sterile cotton gauze soaked with 70% ethanol to reduce microbial contamination. A trained midwife was on site to assist in collection of all milk samples, which were stored in 60 mL plastic vials in a cold chest until subsequent storage at -30°C. Samples were transferred to McGill University on dry ice and divided into three sets of aliquots for flow cytometry (800 µL), ultrasound (800 µL), and NIR (100 µL) measurements. All samples were kept in a -80°C freezer until time of analysis. Milk aliquots analyzed by ultrasound and NIR were unmodified prior to measurements.

3.5.4 Total Cell Count Reference Measurements

A FACSAria flow cytometer (Becton Dickinson, San Jose, CA) equipped with a 488-nm solid state laser was used for cell counting. Green fluorescence was measured with a 500-nm longpass filter. The data were analyzed using FlowJo (Tree Star Inc., Ashland, OR) software.

Frozen milk samples (800 µL) analyzed by flow cytometry were allowed to come to room temperature and were then centrifuged (1000 x *g*) for 3 minutes. The upper cream layer was carefully removed using cotton swabs, and the remaining skim layer was poured off without disturbing the cell pellet. One milliliter of PBS buffer (pH 7.4) was added to the tubes without resuspending the pellets, and subject to another spin cycle, followed by removal of fat and decanting of the aqueous skim layer. The cell pellets were then resuspended in 500 µL of PBS and stained for total cell count as described below.

Cells were stained as described in Dosogne *et al.*[3.26] A 200 nM solution of green fluorescent nucleic acid stain was prepared from 5 mM SYTO 13 (Molecular Probes, Eugene, OR) diluted 1:40 with PBS. A 10 µL aliquot of 200 nM stain was added to each milk sample, and allowed to incubate for 15 minutes prior to cell counting. Thirty microliters of CountBright absolute counting beads (Molecular Probes, Eugene, OR) was also added to each sample. The known number of beads allowed calculation of the sample volume used in each flow cytometry run. This volume could then be used to express cell counts per milliliter of milk.

One million counts were collected for each sample and the fluorescence channel was used to identify and count somatic cells and CountBright beads. A series of controls were used to determine fluorescence thresholds for cells and

reference beads: 30 µL of CountBright beads in 500 µL of PBS; milk sample with 10 µL of diluted SYTO 13 stain and 30 µL of counting beads; and milk sample with 30 µL of counting beads (no stain). The fluorescence intensity of the counting beads was roughly one order of magnitude larger than that of stained cells so that the two distributions did not overlap. The coefficient of variation for triplicate measurements of reference beads was less than 3.5%.

3.5.5 Ultrasound Measurements

Ultrasound spectra were collected with the ultrasound system shown in Figure 3-1. Ultrasonic pulses were generated by a 500PR pulser-receiver from Panametrics (Waltham, MA) and collected with a Handyscope HS5 USB oscilloscope (TiePie engineering, Sneek, FR) sampling at 200 MHz with 12 bit resolution. The pulser-receiver generated a <20 ns, 250 V negative impulse with a 0.002% duty cycle. A pulse repetition rate of 1 kHz allowed any echoes in the sample to attenuate to baseline noise levels before the next pulse. Pulses were transmitted and received with two wideband transducers centered at 30 MHz from Olympus (Waltham, MA), with 4 µs silica delay lines. The two transducers were threaded and fastened into parallel sides of an 0.8 mL anodized aluminum sample cell with an effective 0.6 cm pathlength, such that the transducers were in direct contact with the sample. This aluminum sample cell was temperature-

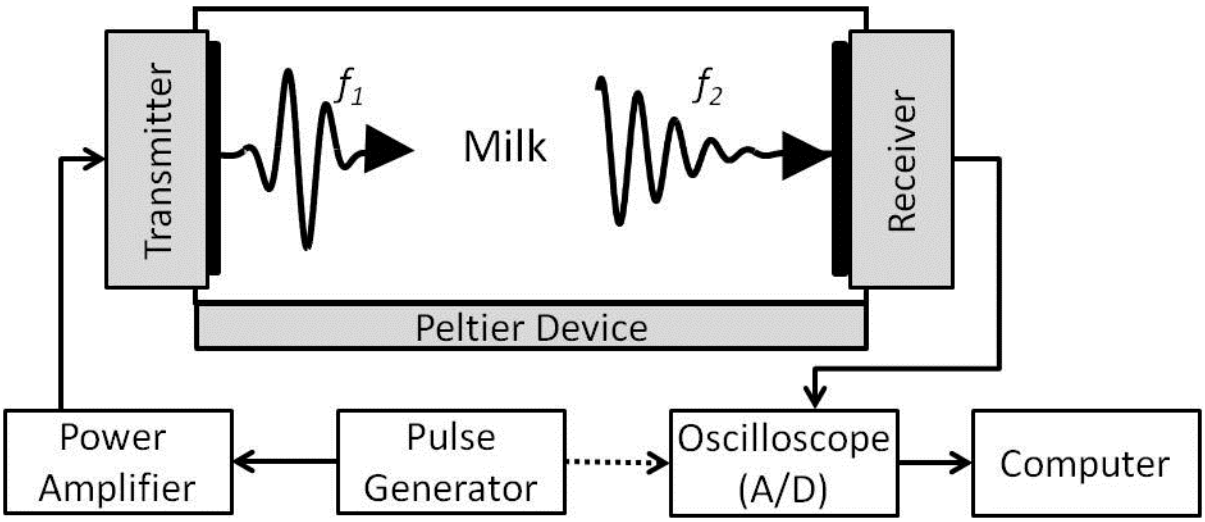

Figure 3-1: Schematic diagram of experimental setup. Transmitting and receiving ultrasound transducers are embedded in a temperature controlled aluminum cell, such that they are in direct contact with the sample. As the initial waveform (f_1) produced by the transmitter propagates through the sample, it is attenuated and distorted, resulting in a modified waveform (f_2), as recorded by the receiver.

controlled with a thermoelectric cooling device from TE Technology (Traverse City, MI) set to 22.00 ± 0.05°C. The sample cell was filled with 700 µL of milk and a total of 500 waveforms were measured. The sample cell was rinsed once with 0.1 M NaOH and four times with deionized water between measurements.

3.5.6 Near-Infrared Measurements

Near-infrared transmittance spectra of human milk samples were acquired with a Cary5000 spectrophotometer (Varian Inc., Palo Alto, CA), fitted with a quartz cuvette with 0.2 mm optical pathlength. The samples were kept at 22.00 ± 0.05°C during the acquisition of spectra. The NIR instrument recorded spectra over the wavelength region of 800-2400 nm in 1 nm steps. A dual-beam configuration allowed for the transmittance spectra to be expressed in terms of absorbance, with air as the reference.

3.5.7 Analysis of Data

The goal of this data analysis was to develop classification models to estimate total cell counts in breast milk based on ultrasound frequency profiles and NIR absorbance spectra, using a minimum number of variables/wavelets. All data processing was done in Matlab (The MathWorks Inc., Natick, MA, R2010a).

The following processing steps were used for all samples, ultrasound and NIR alike, with details given in subsequent sections:

1. Scaling of spectra by Standard Normal Variate (SNV) transformation.

2. Mean-centering of spectra.
3. Haar wavelet transform (HWT) applied to spectra, yielding wavelets of varying size.
4. Variable (wavelet) selection by genetic algorithm (GA) for classification models, with cross-validation of samples.

In addition to these steps, the ultrasound data required a preliminary conversion from the time domain to the frequency domain. The raw ultrasound data for each sample consisted of 250 time-domain spectra, which were aligned to the global maximum of the time series and subsequently averaged to improve the signal-to-noise ratio. Every spectrum spanned a time window of 50 μs, equivalent to 10,000 recorded points. This time window was sufficiently large enough to observe the main pulse signal, as well as seven reflections. The reflections did not overlap with the main pulse and were excluded from further data analysis to reduce noise contributions. The reflections were excluded by selecting a time interval that contained only the main pulse. The start of pulse was taken as the point where the signal first deviated from baseline and the end point where the signal returned to baseline, in this case the interval from 1700 to 2000. This data interval, used for all samples, was then subject to a fast Fourier transform, which expressed the pulse in terms of its frequency components.

3.5.7.1 Standard Normal Variate Transformation

The SNV is a preprocessing treatment often applied to NIR spectral data to reduce scatter effects and is applied to every spectrum individually.[3.27,3.28] First, the mean and standard deviation of all data points in a single spectrum were computed. The mean value was then subtracted from every data point in the spectrum and divided by the standard deviation to generate a corrected profile.

3.5.7.2 Mean-centering

The scatter-corrected profiles were then mean-centered, whereby the mean of all spectra was calculated and subsequently subtracted from each individual spectrum. This treatment emphasized relative differences between spectra.

3.5.7.3 Haar Wavelet Transform

For broad-features spectra, like ultrasound and NIR, wavelet transforms such as the HWT are a useful way to denoise and compress data sets. Moreover, they can simplify interpretation of data by looking at frequency or wavelengths bands, rather than individual frequencies or wavelengths and can be thought of as filters. Wavelet models obtained could easily be implemented in instruments using filters or slits to choose relevant spectral regions and can lead to simplified instruments for specific applications.

The HWT uses two functions, commonly known as the father and mother wavelets, as the basis set to break up and reconstruct a signal. This is similar to a Fourier transform, which employs sine and cosine functions to deconstruct a signal. The father wavelet is akin to a moving average function, while the mother wavelet provides the difference between the original signal and the average given by the father wavelet. For data defined over the range $0 \leq x < 1$, the father(φ) and mother (ψ) wavelets are given by:

$$\varphi(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.1)$$

$$\psi(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1/2 \\ -1 & \text{if } 1/2 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

The HWT allows for the retention of important spectral features in a potentially small number of wavelets and is useful for compressing large data sets. In this context, the wavelets used may be thought of as integrated spectral regions or bands, in which the averaging serves to reduce noise. Interpretation of the spectral data as wavelets allows for identification of broad trends in the spectra, as well as at individual wavelengths and frequencies. The HWT is further described in Nievergelt[3.29] and presented as a NIR preprocessing modality in Gributs and Burns.[3.30]

3.5.7.4 Classification Models

To relate the SCC to ultrasound and NIR spectra, a multivariate classification approach was used. This consisted of assigning samples as having either a low (≤150K cells/mL) or high (≥600K cells/mL) cell count for the purposes of developing a low/high classification model. Dummy variables were assigned to the two classes, such that samples from the low group were represented by 0's and those from the high group by 1's. Since mastitis is only clear at higher levels, samples with cell counts that fell between these two extremes (i.e. grey zone) were excluded from the model.[3.31] The classification model was based on a multivariate least squares function, given by

$$Y = b_0 + b_1X_1 + b_2X_2 + ... + b_nX_n \qquad (3.3)$$

where Y is the dependent variable (here the classification score), $\{X\}$ are the independent variables (the magnitude of a given wavelet), and $\{b\}$ are the calibration coefficients. These coefficients may be determined when both Y and $\{X\}$ are known. The combination of wavelets that most parsimoniously estimated Y were determined by multilinear regression and GA searching. The version of the GA used in this work was based on a program written by Gributs and Burns in Matlab[3.30], which uses principles such as selection, cross-over, and mutation to screen many models and is described in detail elsewhere.[3.30,3.32]

Briefly, the GA is an iterative process that tries to find a good solution for optimization problems. In this case, the objective was to find which variables (wavelets) were best for separating low and high SCC groups. Solutions were generated for the best 1-wavelet model, up to the best 7-wavelet model. The search process started with a random initial population of variables (100) and tried to find the best variables from this pool. Variables were chosen based on a fitness score. For each variable in the population, the coefficients b_1 to b_n of eq. 3.4 were calculated by multilinear regression, where one sample at a time was withheld. Estimates of Y were then obtained for the withheld sample, and the process repeated for all samples. Consequently, the score of each sample was estimated independently of the others in a procedure known as a leave-one-out cross-validation.

This leave-one-out cross-validated result was used to calculate the fitness of each variable, which was defined as the sum of the sensitivity (true-positive rate) and specificity (true-negative rate) of the cross-validated model. A larger fitness score was associated with a better variable. The two fittest variables were passed on to the next generation unchanged, while the rest of the population was subject to cross-over and mutation to introduce diversity (i.e. add in variables that were not included in the initial population). Within this new population, the fitness of each variable was recalculated and the whole process

repeated for 500 generations. This was a sufficient number of generations to converge to a stable solution. Once a good solution was found for a 1-variable model, the cycle was repeated up to a 7-variable model. A Wald test was used to test the statistical significance of each wavelet in the model ($\alpha = 0.05$) and to choose the most parsimonious model (i.e. model with least number of wavelets necessary). Typically, a model that uses fewer variables is considered more robust and should be chosen unless the inclusion of additional variables significantly increases the success of the model.

3.5.7.5 Interpretation of Models

Low and high SCC groups were represented by box-and-whisker plots, which make no assumptions about the underlying distribution of points (i.e. suitable for non-parametric data). The data is split into quartiles, such that the bottom, middle, and upper lines of the box denote the 25th (q_1), 50th (q_2, median), and 75th (q_3) percentiles. The whiskers extending from the box are drawn to the farthest points that fall within $1.5\times(q_1\text{-}q_3)$. Points outside this minimum-maximum range are considered outliers.

The performance of the classification models was described by sensitivity and specificity, which provide a percentage measure of how often low and high cell count samples were classified correctly. Sensitivity, or true positive rate, represents the correct number of high cell counts identified above a threshold.

Likewise, specificity, or true negative rate, represents the correct number of low cell counts identified below a threshold. These performance metrics were calculated as

$$\text{Sensitivity} = \frac{\text{Number of high cell counts identified (above threshold)}}{\text{Actual number of high cell counts (above threshold)}} \quad (3.4)$$

$$\text{Specificity} = \frac{\text{Number of low cell counts identified (below threshold)}}{\text{Actual number of low cell counts (below threshold)}} \quad (3.5)$$

The threshold used to discriminate low cell count samples from those with a high cell count was 0.5.

3.6 Results and Discussion

Ultrasound frequency profiles and NIR absorbance spectra were collected for 117 thawed breast milk samples from Guatemalan mothers, for the purposes of correlating changes in spectra to SCC. Flow cytometry was used as the reference method for determining SCC of samples. Though forward- and side-scattering data collected by flow cytometers can provide size information of counted particles, fat globules present in milk (0.1 – 15 μm)[3.33] have a size distribution that overlaps with leukocytes (7 – 15 μm).[3.34] Moreover, freeze-thaw cycles can compromise cell integrity through deformation and rupture and make size-based particle counting challenging. However, it has been previously shown that if cells are stained with a cell-permeable nucleic dye and counted using the fluorescence channel of the flow cytometer, cell counts are not significantly

affected by multiple-freeze thaw cycles, nor by the presence of fat globules.[3.35] For this reason, cells were stained and counted by fluorescence. Additionally, to reduce the number of fat globules counted, most of the fat was removed from each sample prior to staining and analysis. Forward- and side-scattering plots and fluorescence intensity plots for select milk samples are provided in Appendix B.

The SCC distribution for the milk samples, as determined by fluorescence-based cell counting, is presented in Figure 3-2. More than 80% off samples had SCC values below 600k cells/mL, with the peak number of samples falling in the range of 300k – 400k cells/mL. Samples were assigned to either a low (≤150K cells/mL) or high (≥600K cells/mL) cell count group for the purposes of developing a low/high classification model, with high cell counts associated with presence of mastitis. Samples with cell counts that fell between these two extremes represent a grey zone and were excluded from the model. Though breast conditions are a continuum with no clear boundaries[3.36], the low/high thresholds chosen are similar to those previously reported in the literature.[3.31] Interestingly, using 600K cells/mL as an upper threshold results in 17% of women having high cell counts, which is consistent with previous studies of the prevalence of mastitis in lactating women in Malawi (15.5%)[3.37] and Australia (20%).[3.38]

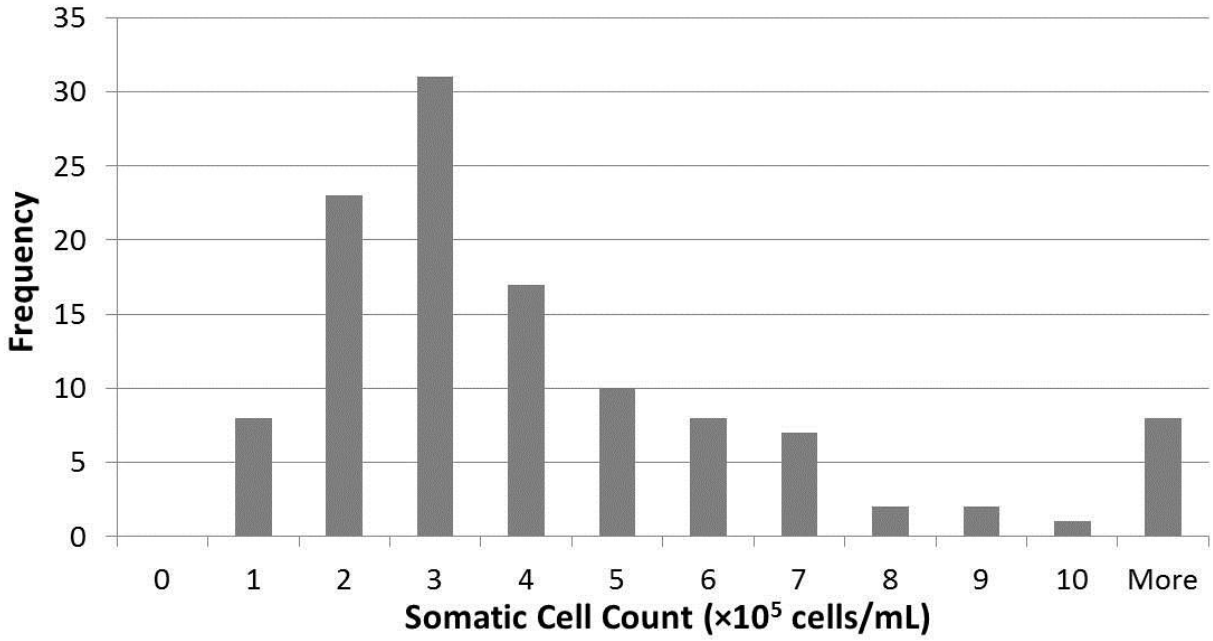

Figure 3-2: Distribution of SCC for 117 breast milk samples, as determined by fluorescence cell counting.

Prior to developing classification models, spectra were treated with a SNV transformation to reduce contributions from fat globules to the signal.[3.27] Much like the reference method of cell counting, which required a cell stain to address the problem of fat globules, spectra collected by ultrasound and NIR required a correction before being subject to further processing. In the case of ultrasound spectra, scattering of waves by fat globules is known to contribute to the observed attenuation of signal.[3.2] Likewise, the presence of fat globules is known to contribute to scatter effects in NIR spectra.[3.39] To make the spectra comparable in terms of their absorbance contributions, a standard normal variate transformation was applied to raw ultrasound and NIR spectra to reduce baseline offsets due to scatter. The results of the SNV transform are shown in Figure 3-3. The broad attenuation distribution of raw ultrasound spectra was greatly reduced after transformation and similarly, the baseline offsets of NIR spectra were also curbed.

The scatter-corrected spectra were then mean-centered to emphasize relative differences between spectra and subject to a HWT to express data as wavelets. As previously described in the Methods section, wavelets are a useful way to compress and denoise spectral data. Moreover, wavelets are simple to interpret for broad-featured spectra, where they can be thought of as integrated regions.

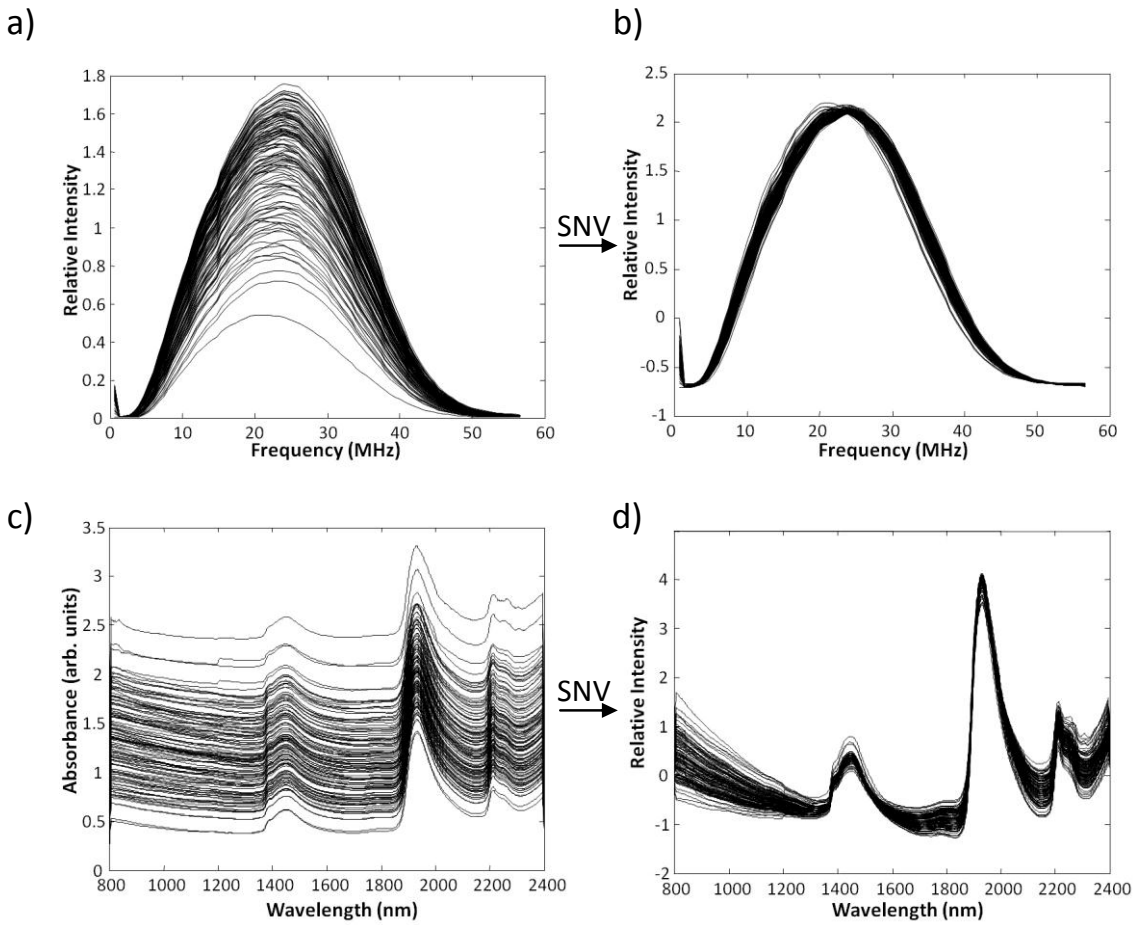

Figure 3-3: Effect of SNV transformation on ultrasound and NIR spectra. Differences in attenuation due to the presence of fat globules in raw ultrasound frequency spectra (a) are mitigated by rescaling the data through a SNV transformation (b). Likewise, this processing treatment reduces baseline offsets in raw NIR spectra (c) to give scatter-corrected spectra (d).

Subsequently, ultrasound wavelet data were used to develop a classification model to separate low SCC samples from those with a high SCC. Classification of samples was done by multilinear regression, which used GA searching to identify which combination of wavelets gave the best separation. A cross-validated model was developed in which each sample was estimated independently of the others using the wavelets identified by the GA. This means that for the 39 samples, an equivalent number of unique models were used to estimate the score of each sample. The results are presented in Figure 3-4. As per a Wilcoxon rank sum test (α=0.05), the medians of the low and high SCC groups were found to be significantly different. The most parsimonious model, as determined by a Wald test (α=0.05), contained three wavelets centered at 28.3, 32.3, and 33.7 MHz. These higher frequency components overlap with harmonics of lower frequencies and suggest that nonlinear distortion of ultrasound may play an important role in estimating SCC. Furthermore, sensitivity and specificity of the estimates were 90% and 79% respectively. For the high SCC group, only 2 out of 20 samples were misclassified, whereas 4 out 19 samples were misclassified for the low SCC group.

In a similar fashion to the ultrasound data, NIR wavelets were also used to develop a classification model to separate low and high SCC samples. Multilinear regression with GA searching was used to develop a cross-validated model,

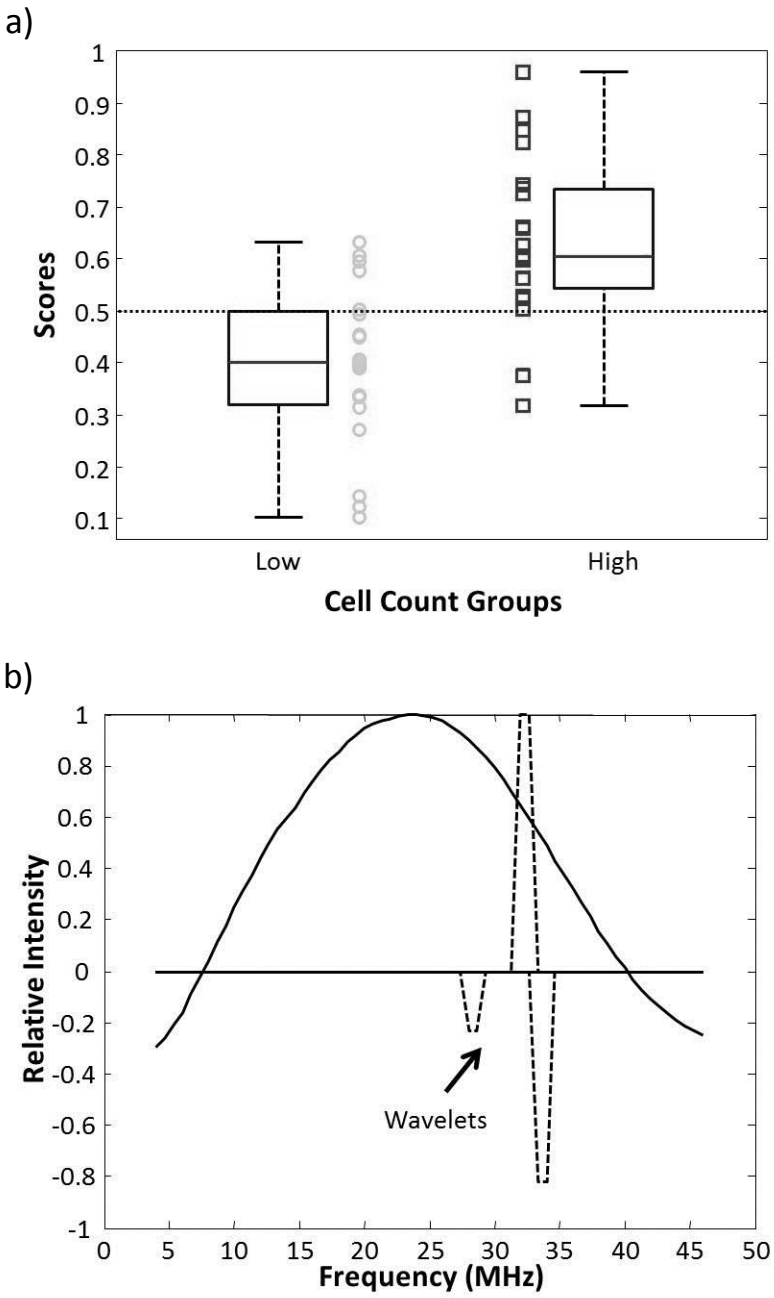

Figure 3-4: Ultrasound classification model. a) Cross-validated estimates of low and high SCC samples. The dotted line (0.5) denotes the separation threshold between groups. b) Wavelets with relative coefficient weightings.

shown in Figure 3-5. Using a Wilcoxon rank sum test (α=0.05), the medians of the low and high SCC groups were found to be significantly different. Unlike the ultrasound model, where 3 wavelets were used for separation of the two groups, only 2 wavelets were used in the most parsimonious NIR model, as determined by a Wald test (α=0.05). The wavelets chosen for the model were centered at 1793.5 and 1801.5 nm, just above the first overtone region for S-H (1740 nm) and C-H (1730 nm) bonds and below the water band region (1930 nm). Absorption bands around 1700-1730 nm have previously been associated with the first overtone C-H stretch of fatty acid chains,[3.40] whereas absorption at 1900 nm by O-H stretch/C-O stretch combination vibrations have been associated with carbohydrate content.[3.41] However, further study is needed to understand the band assignment for the two chosen wavelets. Nevertheless, the model separated samples with a sensitivity of 85% and a specificity of 84%. This means that for the high SCC group, 3 out of 20 samples were misclassified, whereas only 3 out of 19 samples was misclassified for the low SCC group.

The low/high SCC models were then used to estimate the previously withheld samples, with SCC between 150K and 600K cells/mL. The intermediate range contained 78 of the 117 samples analyzed, which were assigned to 4 groups with the following SCC ranges: [150K - 262.5K], [262.5K - 375K], [375K-487.5K], and [487.5K - 600K] cells/mL. Estimates for ultrasound and NIR models

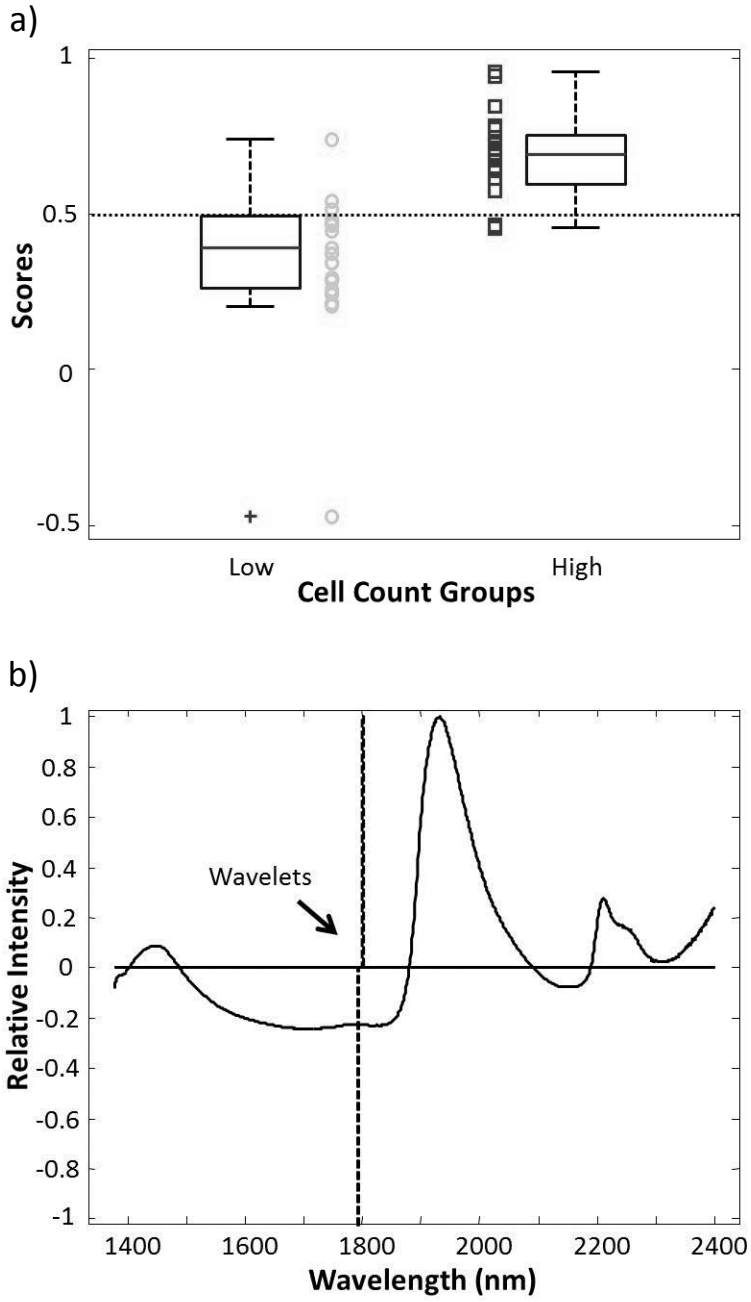

Figure 3-5: NIR classification model. a) Cross-validated estimates of low and high SCC samples. The dotted line (0.5) denotes the threshold used for separation of groups. b) Wavelets with relative coefficient weightings.

are shown in Figure 3-6. To emphasize the relative positions of the medians of the intermediate groups to the low/high groups, a dashed line was used to mark the mean of the intermediate group medians. In the case of the ultrasound model, the medians of the intermediate groups fell between the low and high groups, though they were closer to the low group. The apparent jump in going from intermediate to high SCC for ultrasound estimates may be due to compositional changes in milk with mastitis. For example, the onset of mastitis has been shown to decrease lactose concentrations, which can leak out of the alveolus between epithelial cells.[3.42] Additionally, the increased capillary permeability from inflammation can shift the protein distribution in milk, with reductions in casein concentrations and increases in serum albumin and immunoglobulin concentrations.[3.43] With regards to the NIR model, the medians of the intermediate groups also fell between the low and high groups, though roughly halfway between the two. Whereas the ultrasound estimates seem to jump up suddenly in going from the intermediate to high group, the NIR estimates show a more progressive increase. The positive trend for the NIR estimates suggests that NIR absorbance may be more sensitive to increases in SCC than ultrasound attenuation and distortion processes.

Overall, analysis of ultrasound frequency profiles for determining SCC gave very encouraging results, with very few samples misclassified. Likewise, the

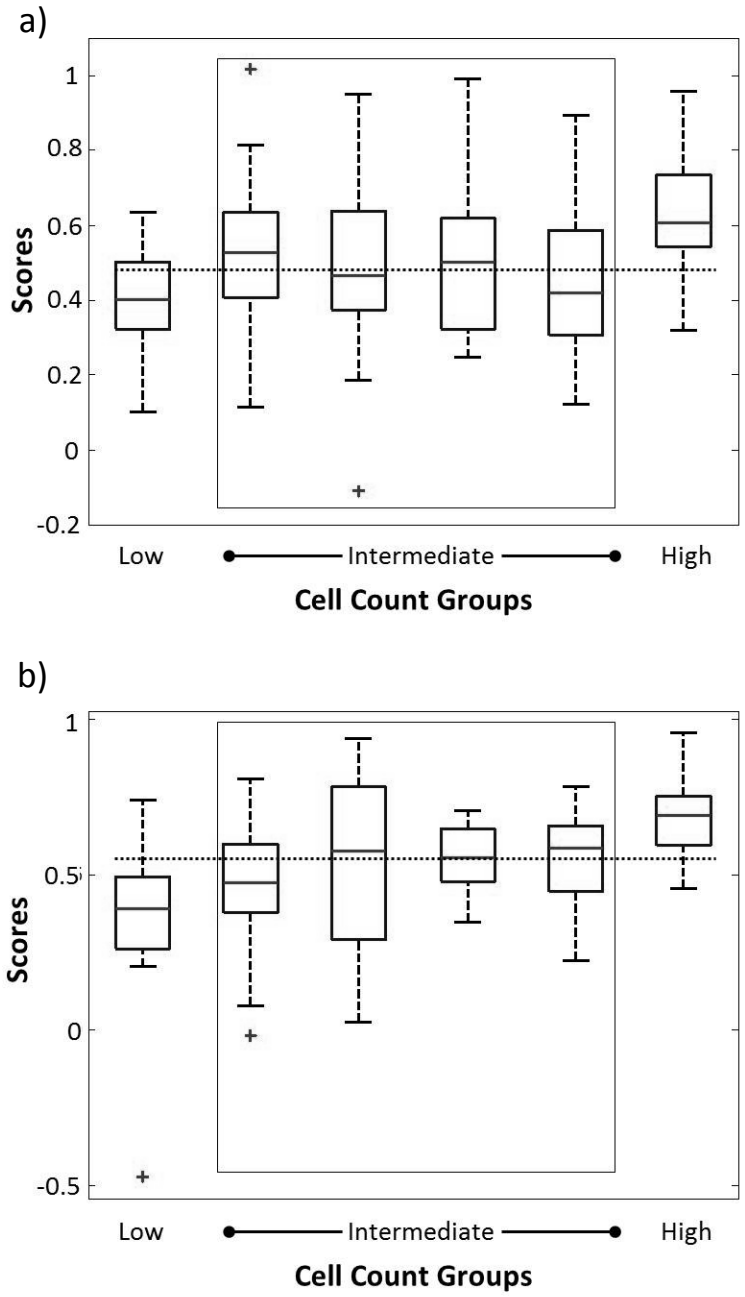

Figure 3-6: Estimates of intermediate SCC samples for (a) ultrasound and (b) NIR classification models. The dashed line represents the mean of the intermediate group medians.

NIR approach had very similar sensitivity and specificity but used one wavelet less than the ultrasound model. Although the medians of the low/high groups were significantly different for ultrasound and NIR estimates, the fairly broad distribution of outcomes, especially for the intermediate SCC samples, can be attributed to the grouping of a continuous variable (i.e. SCC) into fixed classes. In contrast, continuous-variable NIR models based on partial-least-squares regression have been reported in the dairy literature with between 7 to 13 factors used for estimation of SCC.[3.9,3.11] However, the distinct advantage of our classification approach is that models with as few as two variables (wavelets) can be used to categorize SCC as low or high. In practice, models with fewer variables are extremely useful for development of simplified instrumentation. Designing an instrument to collect a few specific wavelength or frequency bands allows for reductions in cost and optimization of performance for the required spectral regions.

3.7 Conclusion

Ultrasound frequency analysis is a promising technique for estimating cell counts in turbid media. While the presence of fat globules can be a problem for measurement of cells in milk, treating frequency spectra for scatter contributions and using multivariate analysis revealed systematic variations in frequency

profiles can be used to address the issue. Changes in frequency profiles were measured and used to develop multilinear classification models to independently estimate low and high SCC samples with a sensitivity and specificity of 90% and 79% respectively. NIR spectroscopy, a technique previously used to estimate SCC in dairy milk, was used to gauge the performance of the ultrasound approach. Similarly, multivariate analysis of NIR spectra showed a separation of low and high SCC samples, with a sensitivity of 85% and specificity of 84%. Overall, the NIR model had fewer sample misclassifications with a simpler model (2 wavelets). However, future models could be refined by including the withheld groups into the model, at the cost of a more complicated k>2 classification model. Additionally, use of fresh instead of frozen milk samples in future studies should also help reduce misclassification of samples. One underlying assumption was that the freeze-thaw cycle affected all samples in the same way. Though the fluorescing cell counting procedure was taken to be robust to deformed and ruptured cells, this may not have carried over to the NIR and ultrasound methodologies, which may have been more sensitive to structural features of cells. This could be side-stepped in future studies by using fresh milk samples.

There is a need for portable diagnostics, especially in the developing world, where early identification of subclinical mastitis in lactating mothers can reduce the likelihood of transmitting diseases like HIV to nursing infants. The key

advantage of the techniques outlined in this work is that following sample collection, spectra can be collected in seconds, allowing for rapid estimation of SCC in milk. Moreover, no sample preparation is necessary prior to measurement, which is important because it would allow for measurements to be made in remote settings, such as rural Guatemala, from where the samples used in this work originate.

3.8 Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Fonds de recherche du Québec—Nature et technologies.

3.9 References

3.1. Savéry, D. & Cloutier, G. High-frequency ultrasound backscattering by blood: analytical and semianalytical models of the erythrocyte cross section. *J. Acoust. Soc. Am.* **121,** 3963–71 (2007).

3.2. Miles, C. A., Shore, D. & Langley, K. R. Attenuation of ultrasound in milks and creams. *Ultrasonics* **28,** 394–400 (1990).

3.3. Dukhin, A. S., Goetz, P. J. & van de Ven, T. G. M. Ultrasonic characterization of proteins and blood cells. *Colloids Surf. B. Biointerfaces* **53,** 121–6 (2006).

3.4. Treeby, B. E., Zhang, E. Z., Thomas, A. S. & Cox, B. T. Measurement of the ultrasound attenuation and dispersion in whole human blood and its components from 0-70 MHz. *Ultrasound Med. Biol.* **37,** 289–300 (2011).

3.5. Chen, S.-H., Lin, Y.-H., Li, W.-T., Wang, S.-H. & Huang, C.-C. Estimation of Cell Concentration Using High-Frequency Ultrasonic Backscattering. *J. Med. Biol. Eng.* **32,** 157–162 (2011).

3.6. Zou, X. *et al.* Lipid composition analysis of milk fats from different mammalian species: potential for use as human milk fat substitutes. *J. Agric. Food Chem.* **61,** 7070–80 (2013).

3.7. Whyte, D., Walmsley, M., Liew, A., Claycomb, R. & Mein, G. Chemical and rheological aspects of gel formation in the California Mastitis Test. *J. Dairy Res.* **72,** 115–121 (2005).

3.8. Moon, J. S. *et al.* Application of a new portable microscopic somatic cell counter with disposable plastic chip for milk analysis. *J. Dairy Sci.* **90,** 2253–9 (2007).

3.9. Kawasaki, M. *et al.* Near-infrared spectroscopic sensing system for on-line milk quality assessment in a milking robot. *Comput. Electron. Agric.* **63,** 22–27 (2008).

3.10. Tsenkova, R., Atanassova, S., Kawano, S. & Toyoda, K. Somatic cell count determination in cow's milk by near-infrared spectroscopy: a new diagnostic tool. *J. Anim. Sci.* **79,** 2550–7 (2001).

3.11. Tsenkova, R., Meilina, H., Kuroki, S. & Burns, D. Near infrared spectroscopy using short wavelengths and leave- one-cow-out cross-validation for quantification of somatic cells in milk. *J. Near Infrared Spectrosc.* **17,** 345 (2009).

3.12. Sauer, C. W. & Kim, J. H. Human milk macronutrient analysis using point-of-care near-infrared spectrophotometry. *J. Perinatol.* **31,** 339–43 (2011).

3.13. Hassiotou, F. *et al.* Maternal and infant infections stimulate a rapid leukocyte response in breastmilk. *Clin. Transl. Immunol.* **2,** e3 (2013).

3.14. Aryeetey, R. N. O., Marquis, G. S., Timms, L., Lartey, A. & Brakohiapa, L. Subclinical mastitis is common among Ghanaian women lactating 3 to 4 months postpartum. *J. Hum. Lact.* **24,** 263–7 (2008).

3.15. Willumsen, J. F. *et al.* Breastmilk RNA viral load in HIV-infected South African women: effects of subclinical mastitis and infant feeding. *AIDS* **17,** 407–14 (2003).

3.16. Semba, R. D. *et al.* Human immunodeficiency virus load in breast milk, mastitis, and mother-to-child transmission of human immunodeficiency virus type 1. *J. Infect. Dis.* **180,** 93–8 (1999).

3.17. Hassiotou, F., Geddes, D. T. & Hartmann, P. E. Cells in human milk: state of the science. *J. Hum. Lact.* **29,** 171–82 (2013).

3.18. Hunt, K. M. *et al.* Mastitis is associated with increased free fatty acids, somatic cell count, and interleukin-8 concentrations in human milk. *Breastfeed. Med.* **8,** 105–10 (2013).

3.19. Dukhin, a S., Goetz, P. J. & Travers, B. Use of ultrasound for characterizing dairy products. *J. Dairy Sci.* **88,** 1320–34 (2005).

3.20. Dukhin, A. S. & Goetz, P. J. *Ultrasound for Characterizing Colloids*. (Elsevier, 2002).

3.21. Duck, F. A. Nonlinear Acoustics in Diagnostic Ultrasound. *Ultrasound Med. Biol.* **28,** 1–18 (2002).

3.22. Sarvazyan, A. P., Chalikian, T. V & Dunn, F. Acoustic nonlinearity parameter B / A of aqueous solutions of some amino acids and proteins. *J. Acoust. Soc. Am.* **88,** 1555–1561 (1990).

3.23. Zhang, J., Kuhlenschmidt, M. S. & Dunn, F. Influences of structural factors of biological media on the acoustic nonlinearity parameter B/A. *J. Acoust. Soc. Am.* **89,** 80–91 (1991).

3.24. Dion, J. R. & Burns, D. H. Determination of volume fractions in multicomponent mixtures using ultrasound frequency analysis. *Appl. Spectrosc.* **65,** 648–56 (2011).

3.25. Dion, J. R. & Burns, D. H. Simultaneous determination of alcohol and carbohydrate content in commercial beverages by ultrasound frequency analysis. *Talanta* **86,** 384–92 (2011).

3.26. Dosogne, H., Vangroenweghe, F., Mehrzad, J., Massart-Leën, a M. & Burvenich, C. Differential leukocyte count method for bovine low somatic cell count milk. *J. Dairy Sci.* **86,** 828–34 (2003).

3.27. Barnes, R. J., Dhanoa, M. S. & Lister, S. J. Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Appl. Spectrosc.* **43,** 772–777 (1989).

3.28. Dhanoa, M., Lister, S., Sanderson, R. & Barnes, R. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. *J. Near Infrared Spectrosc.* **2,** 43 (1994).

3.29. Nievergelt, Y. *Wavelets Made Easy*. *The Mathematical Gazette* **85,** 379 (Birkhäuser Boston, 1999).

3.30. Gributs, C. E. W. & Burns, D. H. Parsimonious calibration models for near-infrared spectroscopy using wavelets and scaling functions. *Chemom. Intell. Lab. Syst.* **83,** 44–53 (2006).

3.31. Meilina, H., Kuroki, S., Jinendra, B. M., Ikuta, K. & Tsenkova, R. Double threshold method for mastitis diagnosis based on NIR spectra of raw milk and chemometrics. *Biosyst. Eng.* **104,** 243–249 (2009).

3.32. Yang, X.-S. in *Nature-Inspired Optimization Algorithms* 77–87 (Elsevier, 2014). doi:10.1016/B978-0-12-416743-8.00005-1

3.33. Rüegg, M. & Blanc, B. The fat globule size distribution in human milk. *Biochim. Biophys. Acta - Lipids Lipid Metab.* **666,** 7–14 (1981).

3.34. Wheater, P. R., Burkitt, H. G. & Daniels, V. G. *Functional histology: A text and colour atlas.* (Churchill Livingstone, 1979).

3.35. Barkema, H. W. *et al.* Effect of freezing on somatic cell count of quarter milk samples as determined by a Fossomatic electronic cell counter. *J. Dairy Sci.* **80,** 422–6 (1997).

3.36. Betzold, C. M. An update on the recognition and management of lactational breast inflammation. *J. Midwifery Womens. Health* **52,** 595–605 (2007).

3.37. Semba, R. D. *et al.* Mastitis and immunological factors in breast milk of lactating women in Malawi. *Clin. Diagn. Lab. Immunol.* **6,** 671–4 (1999).

3.38. Kinlay, J. R., O'Connell, D. L. & Kinlay, S. Incidence of mastitis in breastfeeding women during the six months after delivery: a prospective cohort study. *Med. J. Aust.* **169,** 310–2 (1998).

3.39. Sasić, S. & Ozaki, Y. Short-wave near-infrared spectroscopy of biological fluids. 1. Quantitative analysis of fat, protein, and lactose in raw milk by

partial least-squares regression and band assignment. *Anal. Chem.* **73,** 64–71 (2001).

3.40. Sato, T., Kawano, S. & Iwamoto, M. Near infrared spectral patterns of fatty acid analysis from fats and oils. *J. Am. Oil Chem. Soc.* **68,** 827–833 (1991).

3.41. Shenk, J. S., Workman, J. & Westerhaus, M. . in *Handbook of Near Infrared Analysis* (eds. Burns, D. A. & Ciurczak, E. W.) 347–387 (CRC Press, 2008).

3.42. Harmon, R. J. Physiology of mastitis and factors affecting somatic cell counts. *J. Dairy Sci.* **77,** 2103–12 (1994).

3.43. Poutrel, B., Caffin, J. P. & Rainard, P. Physiological and pathological factors influencing bovine serum albumin content of milk. *J. Dairy Sci.* **66,** 535–41 (1983).

Chapter 4

Monitoring Infections in Indigenous Panamanian Mothers by Ultrasound Frequency Analysis of Serum

4.1 Foreword

The previous chapters presented methodologies for determining pH and estimating cell counts in biofluids using ultrasound frequency analysis. In chapter 2, we saw that ultrasound measurements were sensitive to total serum protein concentration. Increases in protein concentration were accompanied by increased attenuation of ultrasound. Additionally, ultrasound frequency profiles were also sensitive to protein conformation.

Changes in serum protein concentration and conformation are both associated with infection and are good candidates for markers of health by ultrasound frequency analysis. For instance, infection is typically accompanied by changes in serum protein concentrations, such as increases in immunoglobulins. Likewise, infection has also been linked with oxidative stress. Increases in reactive free radicals can lead to modifications and damage of circulating proteins. For example, hydroxyl radicals have been shown to fragment albumin and induce conformational changes through oxidation of amino acid

residues.[4.1,4.2] However, it is unclear if oxidative stress modifications lead to measurable changes in the ultrasound spectrum.

To investigate the effect of oxidative stress on proteins, a trial was designed in which ultrasound spectra from controls and oxidized protein samples were compared. A model system, consisting of bovine serum albumin (BSA) solutions of equal concentration (40 g/L), was exposed to highly reactive hydroxyl radicals. Hydroxyl radicals can damage a variety of macromolecules, including lipids, carbohydrates, nucleic acids, and proteins. Radicals were generated with Fenton's reagent,[4.3] in which hydrogen peroxide in the presence of a metal catalyst (e.g. iron or copper) produces hydroxyl radicals:

$$\mathrm{Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + OH^- + HO\cdot}$$

Figure 4-0 shows the relative differences between the BSA control and oxidized BSA ultrasound spectra. Five replicate measurements were collected for each group and were averaged together. Over the frequency interval of 1.5 to 9 MHz, oxidized BSA attenuated the ultrasound signal more strongly than the BSA control, which is consistent with an increase in sample viscosity. Therefore, elevated oxidative stress due to infection may also contribute to observed differences in frequency profiles. In addition to increased serum protein concentration, oxidative stress modifications of proteins may provide a complementary marker for infection.

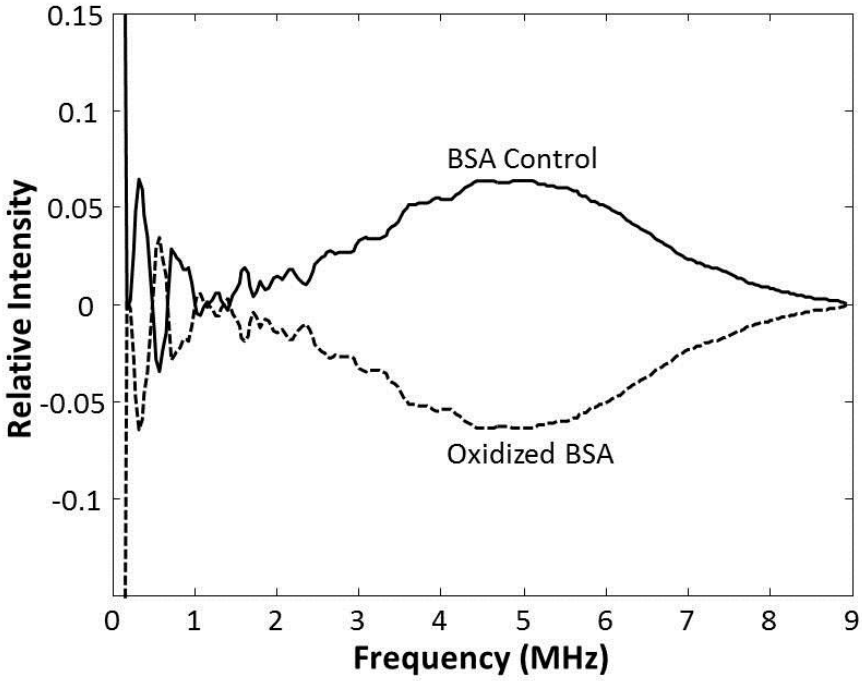

Figure 4-0: Frequency profiles of unmodified BSA (solid line) and oxidized BSA (dashed line). The spectra for each group have been averaged together and the mean of all spectra has been subtracted from them.

In this chapter, we explore differences in human serum frequency profiles as they relate to infection. Specifically, ultrasound is used to track infection-induced changes in serum samples from indigenous Panamanian mothers, which present as modifications to viscoelastic properties of the samples. The following manuscript also uses the multivariate classification approach presented in chapter 3 for assessment of infection.

4.2 Manuscript

Monitoring Infections in Indigenous Panamanian Mothers by Ultrasound Frequency Analysis of Serum

Authors:

Andrien A. Rackov[1], Doris Gonzalez-Fernandez[2], Kristine G. Koski[3], Marilyn E. Scott[2], and David H. Burns[1,4]

[1]Department of Chemistry, McGill University, Montreal, QC, Canada

[2]Institute of Parasitology, McGill University (Macdonald Campus), Montreal, QC, Canada

[3]School of Dietetics and Human Nutrition, McGill University (Macdonald Campus), Montreal, QC, Canada

[4]Department of Chemistry, University of New Brunswick, Fredericton, NB, Canada

4.3 Abstract

Due to limited access to equipment, diagnosing infections in the field can be challenging and is based mostly on clinical exams. Many infections are accompanied by changes in serum protein concentrations (i.e. increased immunoglobulins), which offers a general marker of infection. We present an ultrasound approach that is sensitive to viscoelastic changes in serum to screen for infection and which could be adapted for use in the field. Ultrasound spectra

were collected for 208 serum samples from pregnant indigenous women in Panama with a high prevalence of vaginal (92%), urinary tract (56%), oral (23%) and skin (20%) infections. Analysis of ultrasound frequency spectra by multivariate classification showed a significant separation ($\alpha = 0.05$) between samples from women with no clinically diagnosed infections and those with multiple infections (i.e. vaginal, urinary tract, skin, oral). The cross-validated classification model, developed with only two wavelets, had a sensitivity of 100% and a specificity of 86%. Attenuation and distortion of ultrasound signals were attributed to changes in total serum protein level. This simple and rapid ultrasound frequency analysis approach may be a valuable complementary tool in the field, where clinical exams are often the only means of assessing the health of patients.

Keywords

Ultrasound frequency analysis, human serum, infection, multivariate classification

List of Abbreviations

UTI: urinary tract infection

CV: cervico-vaginal infection

4.4 Introduction

Infection is culpable for a large proportion of maternal, fetal, and neonatal mortality and morbidity worldwide.[4.4] In the developing world, the prevalence of infections is often a function of poverty, crowding, and malnutrition, with a health cost to the mother and risks to the fetus. These risks translate to spontaneous abortion, stillbirth, preterm birth, low birth weight, and infection, among others. Access to adequate maternal healthcare to mitigate the toll of infections remains a challenge in the developing world.[4.5] A variety of infections prevalent in the developing world, such as asymptomatic bacteriuria,[4.6] periodontal disease,[4.7] and abnormal vaginal flora[4.8] are known to contribute to preterm birth and are targets for intervention. However, clinical exams are often the only means of assessing the health of patients in remote settings, where access to diagnostic testing is limited.[4.9]

Biofluids, such as serum, can provide important information about infection. Though detection of low concentrations of specific proteins can be time-consuming and costly, measuring more abundant proteins, such as immunoglobulins, can offer a quick and general marker of infection. After albumin, which accounts for roughly 70% of proteins by mass in serum, immunoglobulin G accounts for about 14%.[4.10] Many bacterial, viral, and other infections are accompanied by increased immunoglobulins in serum.[4.11,4.12]

Serum viscosity is known to increase with total serum protein concentration and can be used as an indicator of infection.[4.13] Additionally, serum viscosity has been shown to increase with oxidative stress, which has also been implicated with infection.[4.14,4.15] Oxidative stress is associated with increases in reactive free radicals, which can react with proteins to either induce conformational change or cause fragmentation.[4.1,4.2] Measuring serum protein concentration and oxidative stress modifications through changes in viscoelastic properties of serum may provide a useful way to screen for infection.

A good candidate for measurements of infection-induced changes in serum is ultrasound. Ultrasound is commonly associated with ultrasonography (e.g. fetal imaging) and portable versions of this technology have emerged, with great potential for use in remote locations and the developing world.[4.16,4.17] As a mechanical wave, ultrasound is not limited by the opacity of a sample and can travel through highly turbid systems without the need for dilution. This has prompted the use of ultrasound technology for investigating the chemical properties of biological fluids such as milk, blood, and serum.[4.18–4.20] Attenuation and nonlinear distortion of ultrasound signals have been shown to be sensitive to the viscoelastic properties of the medium through which they propagate. For instance, attenuation, which comprises absorption and scattering of ultrasound, is dependent on the chemical composition of the sample and affects the

frequency content of the ultrasound wave.[4.21] Additionally, medium-dependent nonlinear distortions, which arise from a difference in velocity between particles in the compression phase and the rarefaction phase, lead to the generation of harmonics and add to the observed frequencies.[4.22,4.23] Analyzing frequency spectra for changes in attenuation and for nonlinear contributions has been used to characterize multicomponent liquid mixtures and commercial beverages.[4.24,4.25] Multivariate analysis of frequency spectra showed that determination of composition was possible from a single ultrasound measurement and frequencies selected for calibration correlated strongly with changes in viscoelastic properties of the samples.

This work considers serum samples from indigenous Panamanian mothers with varying levels of vaginal, urinary tract, skin, and oral infections. We propose that women with multiple infections will have a higher serum protein concentration than women with no infection and that this is reflected in the viscoelastic properties of their sera. We present here a strategy for multivariate classification of women according to infection status based on ultrasound measurements of their serum. Moreover, we provide a framework for a simple, portable, and cost-effective methodology to address the pressing need for portable diagnostics in low resource settings.

4.5 Methods

This work was undertaken as a complementary investigation to address the need for diagnostic testing in remote locations, as highlighted by a large cross-sectional study on anemia, infection, and fetal growth conducted in Panama by McGill University, in collaboration with the University of Panama and the Panamanian Ministry of Health.[4.26,4.27] Biobanked human serum samples were analyzed by ultrasound and were integrated with existing data from this larger study, with a summary presented below.

4.5.1 Study Design

This study was conducted in the Ngäbe-Buglé indigenous population living in Western Panama, of which over 90% lives in extreme poverty. The inclusion criteria for the study were pregnancy as determined by a pregnancy test or by physical exam, as well as homes within a 2 hr walk of 14 health centers accessible by car from the regional hospital in San Felix. Women with twin pregnancies, abnormal pregnancy, or critical illness were not included. Overall, 208 women were included in the study.

4.5.2 Ethical Considerations

Ethical approval was obtained from McGill University in Canada, the Gorgas Memorial Institute Ethics Board in Panama, the Panamanian Ministry of Health, provincial and local health authorities, and indigenous authorities.

Pregnant women were invited to an information session where details of the study and confidentiality were explained. Participants signed a consent form, with the right to withdraw from the study at any time. Participants received no financial compensation, but received complete medical evaluations and were prescribed necessary treatments based on the exams.

4.5.3 Clinical exams

A clinical exam was performed on all participants (n = 208) to diagnose oral, respiratory, skin, and vaginal infections. Oral infections were determined by visual inspection for presence/absence of dental cavities (caries) or inflammation of the gum (gingivitis). Symptoms of upper and lower respiratory infection and complicated upper urinary tract infection were recorded as present or absent. Among skin infections, lesions typical of fungal infection (dermatomycosis), impetigo, and scabies were recorded as present or absent. Presence of vaginitis and cervicitis were determined by genital examination.

4.5.4 Diagnositc tests

During genital examination, a vaginal smear was prepared for Gram staining to detect *Trichomonas vaginalis,* gonococcal infection, vaginal candidiasis, and bacterial vaginitis. Urine samples were analyzed using dipstick URISCAN® strips (YD Diagnosis, Kyunggi- Do, Korea) on a Miditron-M semi-automated reflectance photometer with reagents for semi-quantitative

measurement of urinary infection markers (leukocyte esterase, nitrites, hemoglobin (Hb), and urinary pH). Details of diagnosis are given elsewhere.[4.26]

A 10 mL blood sample was collected by venupuncture and a complete blood cell count (BC-5500 Mindray Auto Hematology Analyzer, Diamond Diagnostics, Massachusetts, USA) was conducted at the San Felix Hospital laboratory. C-reactive Protein (CRP) was assayed using enzyme immunoassays (MP Biomedicals, Orangeburg,NY) at the Gorgas Memorial Institute in Panama City. Cytokines and cortisol were determined on a Luminex 200 analyzer (Luminex Corp., Austin, TX) using Human 10-plex Cytokine/Chemokine Magnetic Bead Panel and Steroid/Thyroid Hormone Magnetic Bead Panel respectively (Millipore, Billerica, MA). Standards, controls, and samples were run in duplicate. See Zambo[4.27] for more details. Ultrasound frequency profiles were collected in Montreal, Canada, on serum samples stored at -80°C.

4.5.5 Ultrasound Measurements

Ultrasonic pulses were generated by a 500PR pulser from Panametrics (Waltham, MA) and recorded with a Handyscope HS5 USB oscilloscope (TiePie engineering, Sneek, FR) sampling at 100-MHz with 14-bit resolution. A schematic of the instrument used to collect spectra is presented in Figure 4-1. The pulser generated a <20 ns, 250 V negative impulse with a 0.002% duty

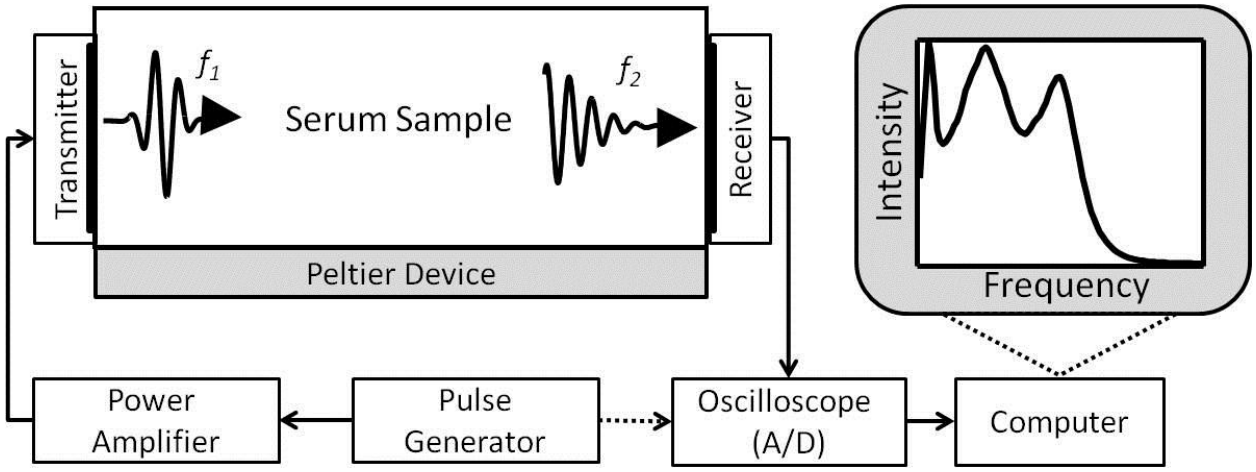

Figure 4-1: Schematic diagram of the ultrasound instrument. Transmitting and receiving ultrasound transducers are on opposite faces of a temperature controlled aluminum cell, coupled to acetate windows with petroleum jelly. As the initial waveform (f_1) produced by the transmitter propagates through the sample, it is attenuated and distorted, resulting in a modified waveform (f_2), as recorded by the receiver.

cycle. A pulse repetition rate of 1 kHz allowed any echoes in the sample to attenuate to baseline noise levels before the next pulse. Pulses were transmitted and received with two wideband transducers centered at 5 MHz from Technisonic (Fairfield, CT). The two transducers were placed on opposing sides of an anodized aluminum sample cell with a 9 mm pathlength and nominal volume of 90 µL. Thin acetate windows (60 µm) on both ends of the sample channel prevented the sample from leaking out of the cell. A thin layer of petroleum jelly was applied to the transducer faces to improve coupling to the acetate windows. Given the dependence of serum viscosity on temperature, the sample cell was regulated with a thermoelectric cooling device from TE Technology (Traverse City, MI) set to 22.00 ± 0.05°C. For all samples, a total of 500 waveforms were measured. The sample cell was cleaned between runs with 0.1 M NaOH solution, followed by several washes with distilled water. A blank measurement, consisting of distilled water, was made between subsequent serum sample measurements to ensure stability and reproducibility of the instrument. The relative standard deviation of water measurements between 0 and 10 MHz was less than 1%, with the lowest value of 0.5% found at the central frequency of the transducers (5 MHz).

4.5.6 Analysis of Data

All data processing was done with Matlab (The MathWorks Inc., Natick, MA, R2010a). Several preprocessing steps were applied to the raw data prior to development of a classification model. The raw data for each serum sample consisted of 500 spectra in the time domain. The spectra were aligned to the global maximum of the time series and subsequently averaged to improve the signal-to-noise ratio. Every spectrum spanned a time window of 100 μs, equivalent to 10,000 recorded points. This time window was large enough to observe the main pulse signal, as well as two reflections. The reflections did not overlap with the main pulse and were excluded from further data analysis to reduce noise contributions. The reflections were excluded by selecting a time interval that contained only the main pulse. The start of pulse was taken as the point where the signal first deviated from baseline and the end point where the signal returned to baseline, in this case the interval from 600 to 1600. A fast Fourier transform was applied to this data interval to obtain frequency profiles for each sample.

The data were then mean-centered to emphasize relative differences between spectra. This was done by subtracting the mean of all spectra from each individual spectrum. Subsequently, the data was subject to a Haar wavelet transform that denoised and compressed the data. The Haar wavelet transform

uses two functions, commonly known as the father and mother wavelets, as the basis set to break up and reconstruct a signal. This is similar to a Fourier transform, which employs sine and cosine functions to deconstruct a signal. In this context, the wavelets used may be thought of as integrated spectral regions or bands of allowed widths (2, 4, 8… 2^n), in which averaging serves to reduce noise. The Haar transform allows for the retention of important spectral features in a potentially small number of wavelets and is useful for compressing large data sets. Interpretation of the spectral data as wavelets allows for identification of broad trends in the spectra, as well as at individual wavelengths and frequencies. The Haar transform is further described in Nievergelt[4.28] and presented as a preprocessing technique in Gributs and Burns.[4.29]

The goal of this data analysis was to develop classification models to identify women with multiple infections from those with none based on ultrasound frequency profiles, using a minimum number of wavelets. Samples were assigned as having either no clinical infections or multiple infections (concurrent vaginal, urinary tract, oral, and skin infections) for the purposes of developing a no/multiple infection classification model. Dummy variables were assigned to the two classes, such that samples from the no infection group were represented by 0's and those from the multiple infection group by 1's. Samples from mothers with between one and three infections were not used in development of the

model. The classification model was based on a least squares regression, given by:

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n \qquad (4.1)$$

where Y is the dependent variable (here the classification score), $\{X\}$ are the independent variables (the magnitude at a given frequency), and $\{b\}$ are the calibration coefficients. These coefficients may be determined when both Y and $\{X\}$ are known.

To optimize selection of wavelets for the classification model, genetic algorithm searching was used to identify which wavelets to use for the best predictive model. In contrast to screening all possible wavelet combinations, which is computationally intensive, the genetic algorithm reduces computation times and avoids human selection bias. Details of the genetic algorithm can be found elsewhere.[4.29]

The best model maximized the sum of the sensitivity (true-positive rate) and specificity (true-negative rate) for a leave-one-out cross-validation, in which the score of each sample was calculated independent of all the other samples. Sensitivity and specificity, using a threshold of 0.5, were calculated as

$$\text{Sensitivity} = \frac{\text{Number of infected women identified (above threshold)}}{\text{Actual number of infected women (above threshold)}} \qquad (4.2)$$

$$\text{Specificity} = \frac{\text{Number of uninfected women identified (below threshold)}}{\text{Actual number of uninfected women(below threshold)}} \qquad (4.3)$$

A Wald test was used to test the statistical significance of each wavelet in the model ($\alpha = 0.05$) and to choose the most parsimonious model. The most parsimonious model included only those variables that significantly improved the model and was considered more robust than models that used additional variables.

4.6 Results and Discussion

In order to establish a relationship between infection and frequency spectra, ultrasound spectra were collected for 208 serum samples from Panamanian women with varying degrees of clinically diagnosed infections. The prevalence of vaginal (92%), urinary tract (56%), oral (23%) and skin (20%) infections in this study population was very high, with only 3.5% of women showing no clinical signs of infection. The distribution of infections is presented in Figure 4-2. Coinfection (two or more infections) was also very common and contributed to the complexity of the data set. To address the diverse distribution of infection in this population, a binary classification strategy was adopted to investigate the extremes of the data set. Effectively, serum samples from women with no infection (n=7) and those with multiple infections (n=6) were used to develop a model to separate healthy from infected women. The multiple infection group was defined as women diagnosed with concurrent cervico-vaginal (CV),

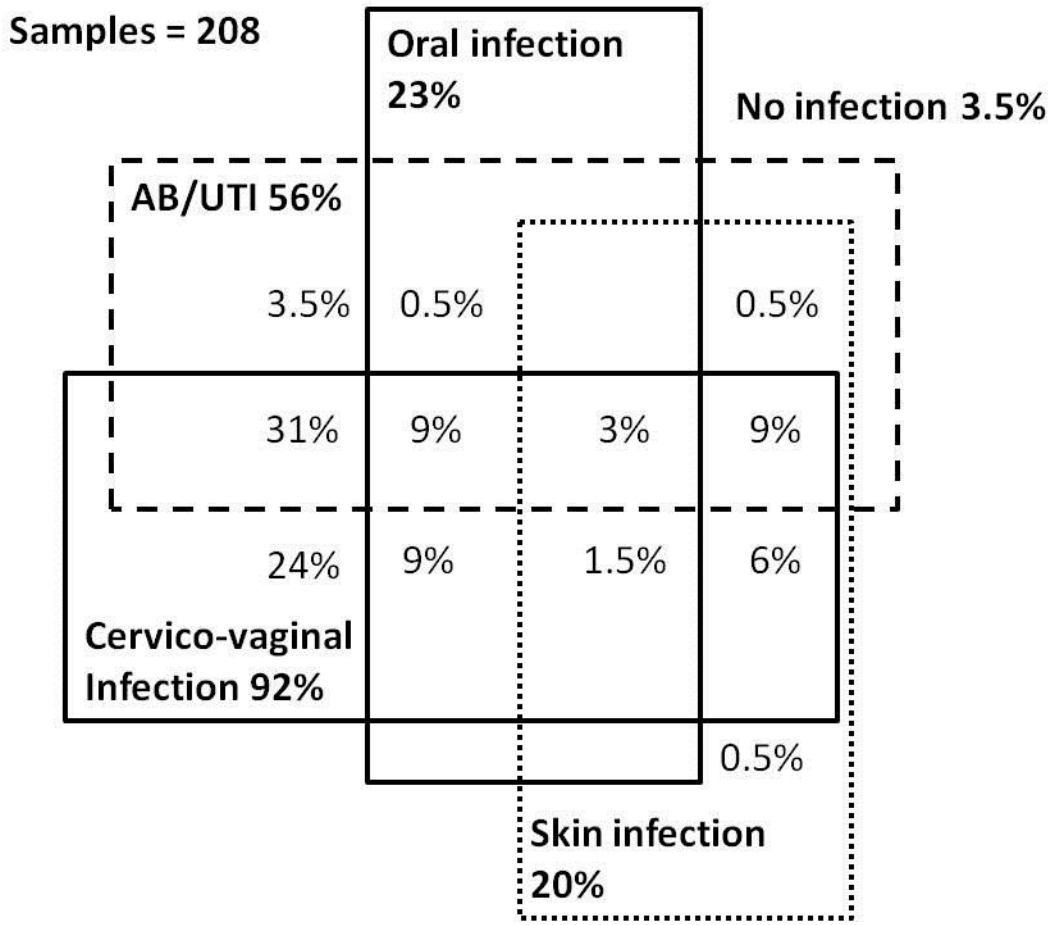

Figure 4-2: Infections (n = 208); skin infection (scabies, dermatomycosis or impetigo), oral infection (caries or gingivitis), AB/UTI (asymptomatic bacteriuria/urinary tract infection), cervico-vaginal infection (bacterial vaginitis, vaginal trichomoniasis, vaginal candidiasis, gonococcal infection). Note that the least prevalent infection (respiratory) has been omitted.

urinary tract (UTI), oral, and skin infections. The binary classification approach presented below provides a straightforward way for looking at this complex data set.

First, each raw ultrasound spectrum was subtracted from the mean of all spectra to emphasize relative differences between the no and multiple infection groups. Figure 4-3 shows spectra before and after subtraction. Over the observed 0 to 10 MHz frequency range, the intensity of the multiple infection group was always lower, except for two regions below 1 MHz. Overall, the lower signal intensity of the multiple infection group is consistent with an increase in serum viscosity, associated with an increase in immunoglobulins.

Prior to developing the classification model, mean-centered spectra were subject to a Haar wavelet transform to express data as wavelets. Wavelets are useful for compressing and filtering data and can capture broad spectral trends in few variables. Wavelet data were used to develop a classification model using logistic regression to separate the no infection group from the multiple infection group. A cross-validated model was developed in which each sample was estimated independently of the others using the wavelets identified by the genetic algorithm, as shown in Figure 4-4. Effectively, for the 13 samples used for classification, an equivalent number of unique models were used to estimate the score of each sample. Given the small sample size, a cross-validation was a

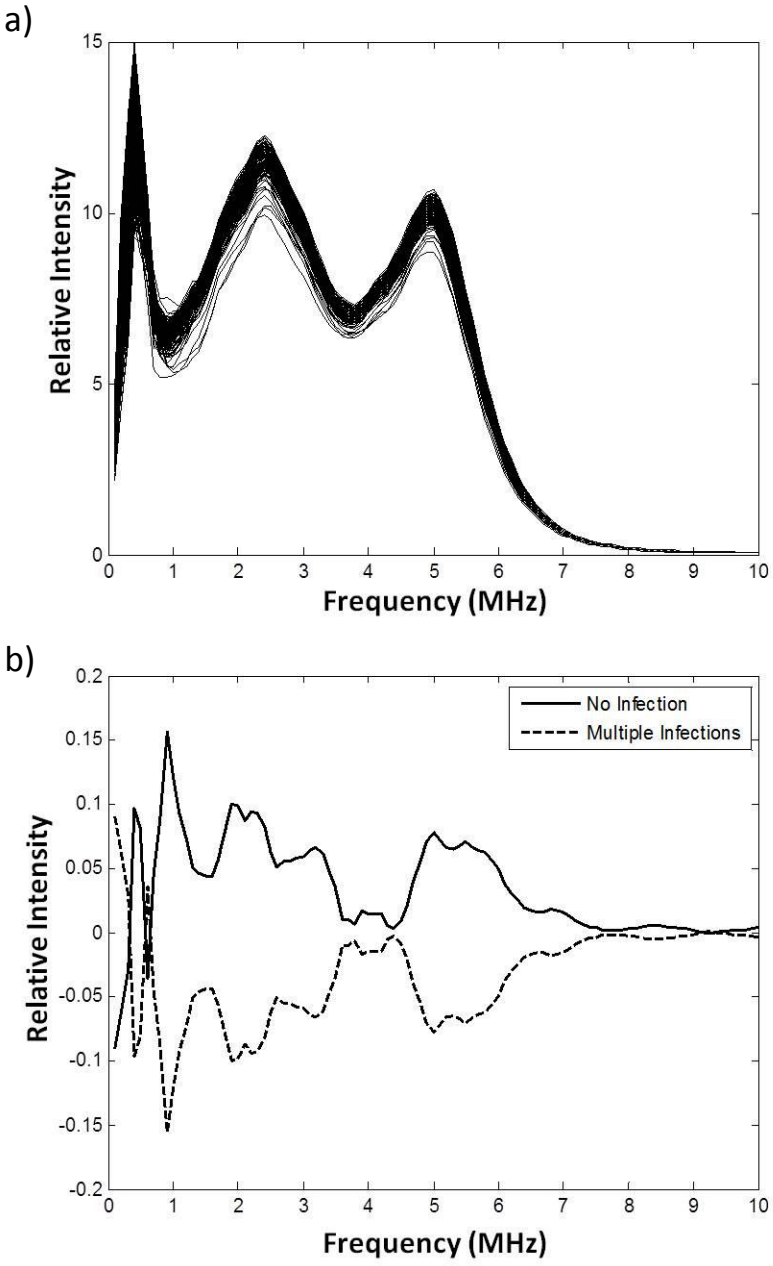

Figure 4-3: a) Smoothed frequency profiles of all serum samples. b) Frequency spectra of serum samples with no (solid) and multiple (dashed) infections. The spectra for each group have been averaged together and the mean of all spectra has been subtracted from them.

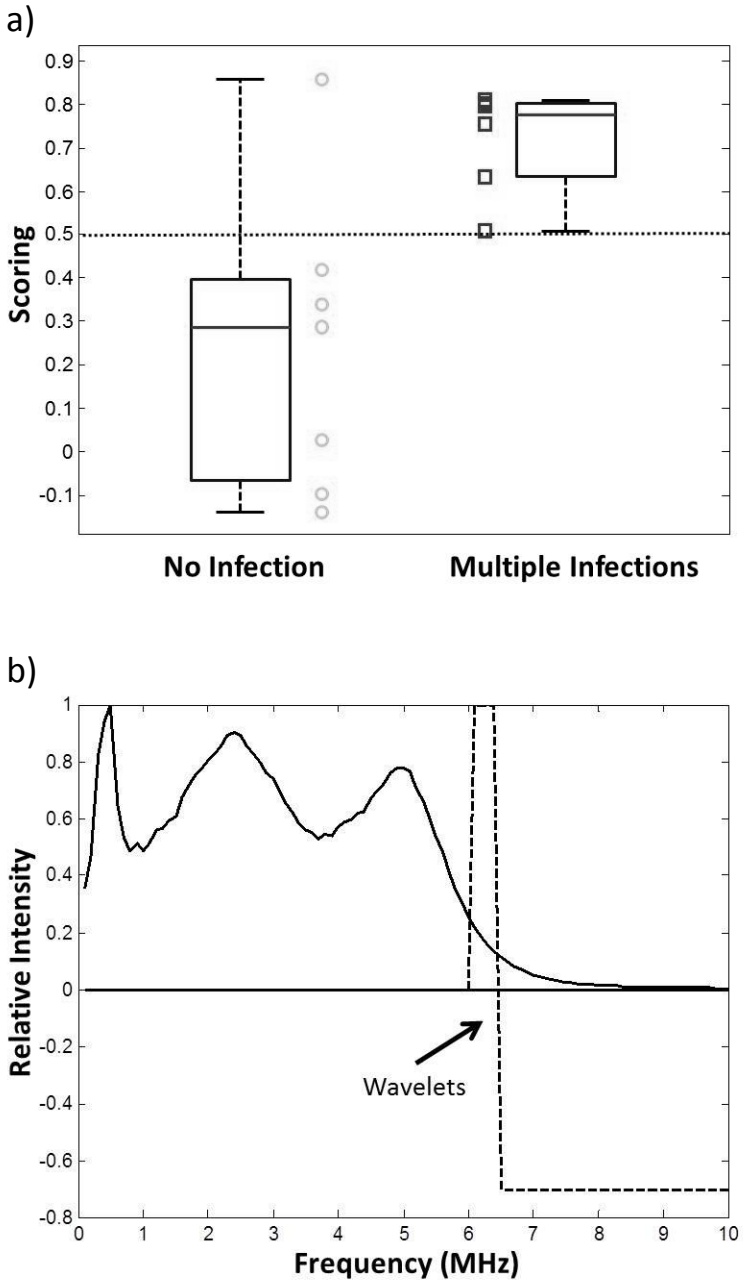

Figure 4-4: a) Classification model for the no infection (n=7) and multiple infection (n=6) groups, with a sensitivity and specificity of 100% and 86% respectively. b) The two wavelets chosen for the model are shown, centered at 6.25 and 8.25 MHz.

useful way to evaluate the fit of the model using a hypothetical test set since an independent test set was not available. For the multiple infection group, no samples were misclassified (sensitivity of 100%), whereas only one sample was misclassified for the no infection group (specificity of 86%). To reiterate the separation of the two groups, a Wilcoxon rank sum test (α=0.05) was performed, where the medians of the no and multiple infection groups were found to be significantly different. The most parsimonious model, as determined by a Wald test (α=0.05), consisted of only two wavelets centered at 6.25 and 8.25 MHz. The selected higher frequency regions overlap with harmonics of lower frequencies and suggest that nonlinear distortion of ultrasound may play an important role in tracking changes in viscoelastic properties of samples.

The no/multiple infection model was then used to estimate the outcomes of samples from women with between 1 and 3 infections. These intermediate infection samples were split into two groups that encompassed the two most prevalent infections, namely CV and UTI. The first group consisted of samples from women with CV and/or skin and/or oral infections (i.e. CV, CV & skin, CV & oral, CV & skin & oral). The second group was made up of samples from women with CV and UTI, with or without skin or oral infections (i.e. CV & UTI, CV & UTI & oral, CV & UTI & skin). The outcomes are shown in Figure 4-5. The medians of the two intermediate infection groups fell between the values of the no/multiple

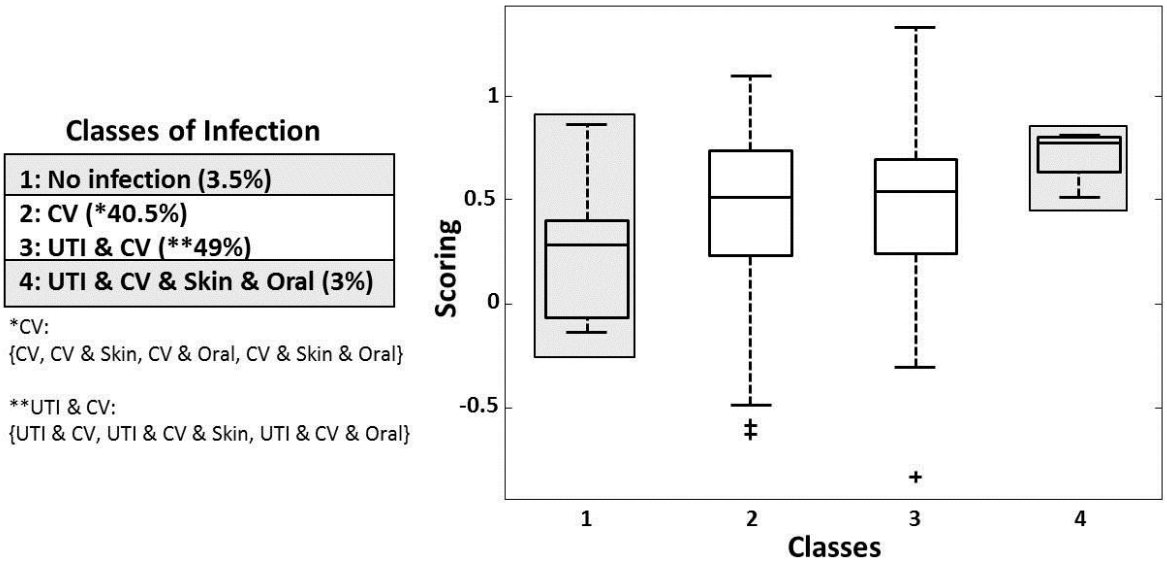

Figure 4-5: Estimates of intermediate infection groups (classes 2 and 3) using the no/ multiple infection model (classes 1 and 4) are shown. Intermediate infection groups were defined as samples with between 1 and 3 clinically diagnosed infections.

infection model. Interestingly, the outcomes suggest an increased immune response with increasing number of infections. CV infections, such as vaginal trichomoniasis and gonoccocal infections, stimulate both a local and systemic immune response with increased immunoglobulins in serum.[4.30,4.31] Likewise, UTI are also accompanied by increased serum antibody levels.[4.32,4.33] Skin infections including impetigo,[4.34] scabies,[4.35] and dermatomycosis[4.36] have all been associated with increases in serum proteins. In contrast, oral infections like gingivitis have not been associated with significant changes in immunoglobulins,[4.37] but more subtle changes in serum have been observed, such as increased endotoxin levels[4.38] and oxidative stress modifications.[4.39] However, coinfection often results in an immune response beyond the simple additive effects of the concurrent infections. Immune response to one infection may be suppressed[4.40] or aggravated[4.41] by other infections. The fairly broad distribution of the estimates may in part reflect the complex interplay of coinfection.

To better understand this complex behavior, several biomarkers relevant to infection were compared to establish whether these differed between the no/multiple infection groups. The eight markers considered were total white blood cell count and subtypes (lymphocytes, neutrophils, eosinophils), cortisol (stress hormone), C-reactive protein (marker of inflammation), IL-6, and IL-1β (pro-

inflammatory cytokines). Given the small sample sizes of the groups, a non-parametric Wilcoxon rank sum test (α=0.05, left-tailed) was used to establish whether biomarker levels were higher in the multiple infection group. The results are summarized in Table 4-1. Interestingly, none of the biomarker values were significantly different. This may be a result of the limited samples available for each group. Alternately, in this study population, malnutrition (protein, vitamin, and mineral deficiencies) may play a key role in the observed null result for biomarkers. Protein deficiency, known to suppress immune response and increase susceptibility to infection, is hypothesized to be the leading contributor to immune deficiency globally.[4.42,4.43] Recent data from this study population have revealed multiple micronutrient deficiencies of iron, folic acid, and vitamins A, D and B12[4.26], which are also implicated in altered immune responses.[4.44] Additionally, the maternal immune response during pregnancy is complex and still not well understood and may also contribute to the null result.[4.45]

Despite the possibility of a suppressed immune response, ultrasound frequency analysis still gave a significant separation of the no/multiple infection groups. This suggests that if immunoglobulin increases were suppressed, oxidative stress modifications may have played a key role in modifying the viscoelastic properties of the serum samples. However, further work is needed to better understand the effects of oxidative stress on the viscoelastic properties of

Table 4-1: Biomarker values for no/multiple infection groups

Biomarker	No Infections (n=7)	Multiple Infections (n=6)	P –value
White Blood Cells (10^3 cells/μL)	7.5±1.4	8.4±2.7	0.3141
Lymphocytes (10^3 cells/μL)	1.8±0.3	1.7±0.4	0.8170
Neutrophils (10^3 cells/μL)	4.9±1.1	6.0±2.3	0.2226
Eosinophils (10^3 cells/μL)	0.4±0.2	0.3±0.2	0.7675
Cortisol (ng/mL)	76.3±66.3	99.1±52.9	0.1830
C-Reactive Protein (μg/mL)	2.8±2.5	7.1±7.8	0.2541
IL-6 (pg/mL)	20.5±25.1	2.1±1.9	0.9965
IL-1β (pg/mL)	8.5±8.2	5.9±10.2	0.6346

Biomarker values are reported as mean ± 1 standard deviation. P-values were calculated via a Wilcoxon rank sum test (α=0.05, one-sided, left-tailed test) to test whether biomarker values were higher in the infected group compared to the uninfected group.

serum. Nevertheless, monitoring viscoelastic changes in serum with ultrasound may provide a means of identifying infection in patients with a suppressed immune response.

4.7 Conclusions

Field research often relies on diagnosis of infectious diseases based on clinical exams, given the limited availability of more precise, portable diagnostic tools. We presented an ultrasound methodology suitable for identifying infection in remote settings. Ultrasound frequency analysis of serum samples from indigenous pregnant Panamanian women was able to separate the no/multiple infection groups with a simple two-wavelet model. Whereas biomarkers associated with infection did not show significant differences between the no/multiple infection groups, changes in viscoelastic properties as measured by ultrasound allowed for classification of samples with a very high sensitivity and specificity.

Future work should address some of the potential confounding factors to the infection model including malnutrition, infection severity, and pregnancy. This work highlighted the health challenge that exists in the developing world and a simple and rapid serum measurement approach may be a valuable complimentary tool to clinical measurements, especially in settings in which

access to laboratory equipment is limited. A quick screening procedure could help identify women that would benefit from more rigorous examinations, and with implementation of proper treatment, pregnancy-associated complications could be minimized.

4.8 Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Fonds de recherche du Québec—Nature et technologies.

4.9 References

4.1. Wolff, S. P. & Dean, R. T. Fragmentation of proteins by free radicals and its effect on their susceptibility to enzymic hydrolysis. *Biochem. J.* **234,** 399–403 (1986).

4.2. Meuccis, E. *et al.* Metal-catalyzed Oxidation of Human Serum Albumin : Conformational and Functional Changes. *J. Biol. Chem.* **266,** 4692–4699 (1991).

4.3. Kocha, T., Yamaguchi, M., Ohtaki, H., Fukuda, T. & Aoyagi, T. Hydrogen peroxide-mediated degradation of protein: different oxidation modes of copper- and iron-dependent hydroxyl radicals on the degradation of albumin. *Biochim. Biophys. Acta* **1337,** 319–26 (1997).

4.4. Jeffery, H. E. & Lahra, M. M. in *Fetal and Neonatal Pathology* (eds. Keeling, J. W. & Khong, T. Y.) 379–423 (Springer London, 2007). doi:10.1007/978-1-84628-743-5

4.5. Say, L. & Raine, R. Public health reviews A systematic review of inequalities in the use of maternal health care in developing countries : examining the scale of the problem and the importance of context. *Bull. World Health Organ.* **85,** 812–819 (2007).

4.6. Smaill, F. M. & Vazquez, J. C. Antibiotics for asymptomatic bacteriuria in pregnancy. *Cochrane database Syst. Rev.* CD000490 (2007). doi:10.1002/14651858.CD000490.pub2

4.7. Lopez, N. J., Smith, P. . & Gutierrez, J. Higher Risk of Preterm Birth and Low Birth Weight in Women with Periodontal Disease. *J. Dent. Res.* **81,** 58–63 (2002).

4.8. Kiss, H., Petricevic, L. & Husslein, P. Prospective randomised controlled trial of an infection screening programme to reduce the rate of preterm delivery. *BMJ* **329,** 371 (2004).

4.9. Bissonnette, L. & Bergeron, M. G. Diagnosing infections--current and anticipated technologies for point-of-care diagnostics and home-based testing. *Clin. Microbiol. Infect.* **16,** 1044–53 (2010).

4.10. De, M. *et al.* Sensing of proteins in human serum using conjugates of nanoparticles and green fluorescent protein. *Nat. Chem.* **1,** 461–5 (2009).

4.11. Fearon, D. T. & Locksley, R. M. The instructive role of innate immunity in the acquired immune response. *Science* **272,** 50–3 (1996).

4.12. Medzhitov, R. Recognition of microorganisms and activation of the immune response. *Nature* **449,** 819–26 (2007).

4.13. Késmárky, G., Kenyeres, P., Rábai, M. & Tóth, K. Plasma viscosity : A forgotten variable. *Clin. Hemorheol. Microcirc.* **39,** 243–246 (2008).

4.14. Ciragil, P. *et al.* The effects of oxidative stress in urinary tract infection during pregnancy. *Mediators Inflamm.* **2005,** 309–11 (2005).

4.15. Barth, K. R., Isabella, V. M., Wright, L. F. & Clark, V. L. Resistance to peroxynitrite in Neisseria gonorrhoeae. *Microbiology* **155,** 2532–45 (2009).

4.16. Harris, R. D. & Marks, W. M. Compact Ultrasound for Improving Maternal and Perinatal Care in Low-Resource Settings: Review of the Potential Benefits, Implementation Challenges, and Public Health Issues. *J. Ultrasound Med.* **28,** 1067–1076 (2009).

4.17. Harris, R. D., Cho, J. Y. & Deneen, D. R. Compact Ultrasound Donations to Medical Facilities in Low-Resource Countries: A Survey-Based Assessment of the Current Status and Trends. *J. Ultrasound Med.* **31,** 1255–1259 (2012).

4.18. Dukhin, a S., Goetz, P. J. & Travers, B. Use of ultrasound for characterizing dairy products. *J. Dairy Sci.* **88,** 1320–34 (2005).

4.19. Dukhin, A. S., Goetz, P. J. & van de Ven, T. G. M. Ultrasonic characterization of proteins and blood cells. *Colloids Surf. B. Biointerfaces* **53,** 121–6 (2006).

4.20. Choi, P. K., Bae, J. R. & Takagi, K. Ultrasonic spectroscopy in bovine serum albumin solutions. *J. Acoust. Soc. Am.* **87,** 874–81 (1990).

4.21. Bonacucina, G., Misici-Falzi, M., Cespi, M. & Palmieri, G. F. Characterization of Micellar Systems by the Use of Acoustic Spectroscopy. *J. Pharm. Sci.* **97,** 2217–2227 (2008).

4.22. Law, W. K., Frizzell, L. A. & Street, W. G. Determination of the Nonlinear Parameter B/A of Biological Media. *Ultrasound Med. Biol.* **11,** 307–318 (1985).

4.23. Sarvazyan, A. P., Chalikian, T. V & Dunn, F. Acoustic nonlinearity parameter B / A of aqueous solutions of some amino acids and proteins. *J. Acoust. Soc. Am.* **88,** 1555–1561 (1990).

4.24. Dion, J. R. & Burns, D. H. Determination of volume fractions in multicomponent mixtures using ultrasound frequency analysis. *Appl. Spectrosc.* **65,** 648–56 (2011).

4.25. Dion, J. R. & Burns, D. H. Simultaneous determination of alcohol and carbohydrate content in commercial beverages by ultrasound frequency analysis. *Talanta* **86,** 384–92 (2011).

4.26. Gonzalez-Fernandez, D. Combined impact of infections and nutritional deficiencies on maternal health and fetal parameters in a rural community in Panama. (McGill University, 2012).

4.27. Zambo, Z. Vitamin deficiency, infection and cytokine interactions in pregnant Panamanian women: impact on cortisol-IGF-1-fundal height pathway. (McGill University, 2013).

4.28. Nievergelt, Y. *Wavelets Made Easy*. *The Mathematical Gazette* **85,** 379 (Birkhäuser Boston, 1999).

4.29. Gributs, C. E. W. & Burns, D. H. Parsimonious calibration models for near-infrared spectroscopy using wavelets and scaling functions. *Chemom. Intell. Lab. Syst.* **83,** 44–53 (2006).

4.30. Mutwiri, G. K. & Corbeil, L. B. Genital and Systemic Immune Responses in a Murine Model of Tritrichomonas foetus Infection. *J. Parasitol.* **84,** 321 (1998).

4.31. Brooks, G. F. & Lammel, C. J. Humoral Immune Response to Gonococcal Infections. *Clin. Microbiol. Rev.* **2,** 5–10 (1989).

4.32. Spencer, J. D., Schwaderer, A. L., Becknell, B., Watson, J. & Hains, D. S. The innate immune response during urinary tract infection and pyelonephritis. *Pediatr. Nephrol.* **29,** 1139–49 (2014).

4.33. Sanford, B. A., Thomas, V. L., Forland, M., Carson, S. & Al, S. E. T. Immune Response in Urinary Tract Infection Determined by Radioimmunoassay and Immunofluorescence : Serum Antibody Levels Against Infecting Bacterium and Enterobacteriaceae Common Antigen. *J. Clin. Microbiol.* **8,** 575–579 (1978).

4.34. Darmstadt, G. L. & Lane, A. T. Impetigo: an overview. *Pediatr. Dermatol.* **11,** 293–303 (1994).

4.35. Arlian, L. G., Morgan, M. S., Vyszenskimoher, D. L. & Stemmer, B. L. Sarcoptes scabiei: The Circulating Antibody Response and Induced Immunity to Scabies. *Exp. Parasitol.* **78,** 37–50 (1994).

4.36. Grappel, S. F., Bishop, C. T. & Blank, F. Immunology of Dermatophytes and Dermatophytosis. *Bacteriol. Rev.* **38,** 222–250 (1974).

4.37. Baehni, P. *et al.* Studies of host responses during experimental gingivitis in humans. *J. Periodontal Res.* **14,** 279–288 (1979).

4.38. Wahaidi, V. Y., Kowolik, M. J., Eckert, G. J. & Galli, D. M. Endotoxemia and the host systemic response during experimental gingivitis. *J. Clin. Periodontol.* **38,** 412–7 (2011).

4.39. Pradeep, A. R., Ramchandraprasad, M. V, Bajaj, P., Rao, N. S. & Agarwal, E. Protein carbonyl: An oxidative stress marker in gingival crevicular fluid in healthy, gingivitis, and chronic periodontitis subjects. *Contemp. Clin. Dent.* **4,** 27–31 (2013).

4.40. Salgame, P., Yap, G. S. & Gause, W. C. Effect of helminth-induced immunity on infections with microbial pathogens. *Nat. Immunol.* **14,** 1118–26 (2013).

4.41. Weng, M. *et al.* Alternatively Activated Macrophages in Intestinal Helminth Infection: Effects on Concurrent Bacterial Colitis. *J. Immunol.* **179,** 4721–4731 (2007).

4.42. Field, C. J., Johnson, I. R. & Schley, P. D. Nutrients and their role in host resistance to infection. *J. Leukoc. Biol.* **71,** 16–32 (2002).

4.43. Ruth, M. R. & Field, C. J. The immune modifying effects of amino acids on gut-associated lymphoid tissue. *J. Anim. Sci. Biotechnol.* **4,** 27 (2013).

4.44. Chandra, R. K. Nutrition and the immune system: an introduction. *Am. J. Clin. Nutr.* **66,** 460S–463S (1997).

4.45. Mor, G. & Cardenas, I. The immune system in pregnancy: a unique complexity. *Am. J. Reprod. Immunol.* **63,** 425–33 (2010).

Chapter 5

A Novel Heterodyned Ultrasound Approach for Assessing *Heligmosomoides bakeri* Infection Levels in Mice

5.1 Foreword

The previous chapter introduced the concept of using frequency analysis of serum samples from expecting mothers to monitor infection and highlighted some of the challenges of working with biological samples. Many factors (i.e. pregnancy, infection severity, malnutrition) were not controlled and may have affected the outcome of the previously presented model for infection classification. Despite some of the challenges highlighted in chapter 4, the findings suggested the feasibility of monitoring infection-induced changes in viscoelastic properties of serum with ultrasound. Particularly, elevations in immunoglobulins and oxidative stress modifications of serum proteins were attributed to observed differences between uninfected and infected women.

The objectives in this chapter were twofold. First, we wanted to address the challenges and limitations of the previous study by using a controlled experimental design. To further investigate ultrasound frequency analysis of serum for monitoring infection, mice infected with the gastrointestinal parasite *H. bakeri* are used as a model for nematode infections in humans. The experimental

design controls for pregnancy status, malnutrition, infection severity, and type of infection. Moreover, the number of samples available for each infection group is larger.

Concurrently, this chapter introduces a modified ultrasound system consistent with the criteria of point-of-care devices. Whereas the previous chapters presented a pulsed ultrasound instrument, where a whole frequency spectrum could be collected in a single measurement, an alternate approach is explored where frequencies are collected sequentially. Instead of using high-intensity ultrasound pulses, high-intensity continuous wave signals are used along with a detection scheme based on the principle of heterodyning. This allows ultrasound signals to be shifted to the audible frequency domain and recorded by a sound card, providing a very cost-effective detector compared to high-speed oscilloscopes. Even though high-intensity ultrasound pulses are substituted for high-intensity continuous wave signals, nonlinear ultrasound propagation is still observed. Effectively, nonlinear distortion is dependent on the intensity of propagating waves.

With the new detection scheme, only the fundamental frequency is observed and the harmonics are not recorded. The heterodyning process acts as a filter and isolates the original (fundamental) wave frequency. Although the harmonics not are recorded, the change (decrease) in intensity of the

fundamental is still related to the emergence of harmonics. That is, the fundamental frequency changes as a result of both attenuation and nonlinear distortion. This is in contrast to conventional ultrasound spectrometers, which operate at lower intensities and do not exhibit nonlinear wave phenomena.

To illustrate that ultrasound waves are still subject to nonlinear phenomena for the modified instrument, a high-intensity 3 MHz continuous wave signal was recorded before (reference) and after propagation through a 40 g/L bovine serum albumin solution. Likewise, a lower-intensity 3 MHz continuous wave signal was also recorded after travelling through the sample solution. A high-speed oscilloscope was used for these measurements to confirm the generation of harmonics. The before-and-after waveforms, along with their associated frequency components, are shown in Figure 5-0. A peak shift is discernible for the high-intensity signal that travelled through the BSA solution, relative to the reference. The distortion of the sine wave is accompanied by the generation of new frequency components at integer intervals of the original frequency. The emergence of harmonics decreases the intensity of the fundamental (3 MHz) frequency. In contrast, the lower-intensity sine wave is undistorted, confirmed by a single peak at 3 MHz in the frequency domain.

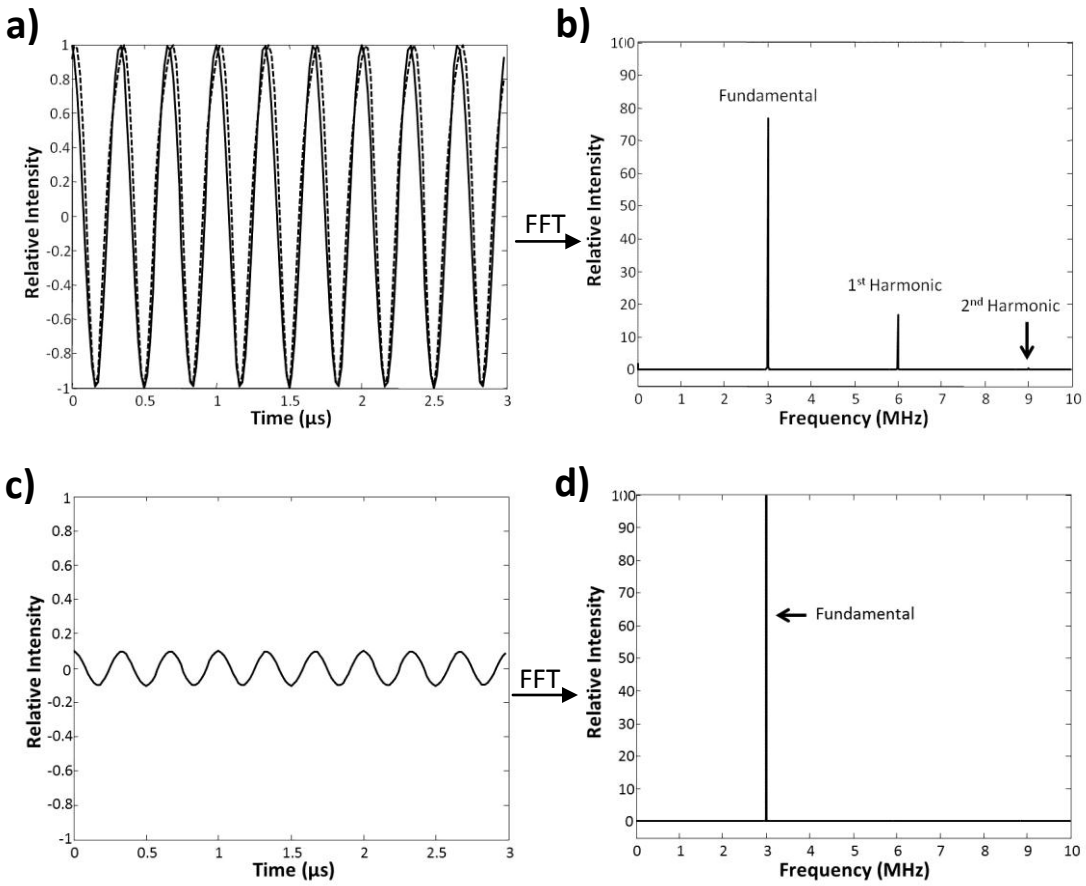

Figure 5-0: a) Overlay of an undistorted (computer-generated) 3 MHz sine wave (solid line) and a distorted (recorded) sine wave (dashed line), generated with high amplification. b) FFT spectrum of distorted sine wave after propagating through a 40 g/L BSA solution. c) Undistorted (recorded) 3 MHz sine wave, generated with low amplification. d) FFT spectrum of undistorted sine wave after propagating through a 40 g/L BSA solution.

Although the new detection system records only one frequency at a time (no harmonics), attenuation of the fundamental is still dependent on the generation of harmonics. Consequently, nonlinear distortion still plays a crucial role as a contributor to the observed signal and the strategies developed in the previous chapters are still applicable to the ultrasound spectra collected with the new instrumentation.

5.2 Manuscript

A Novel Heterodyned Ultrasound Approach for Assessing *Heligmosomoides bakeri* Infection Levels in Mice

Authors:

Andrien A. Rackov[1], Maurice R. Odiere[2], Kristine G. Koski[3], Marylin E. Scott[2], and David H. Burns[1,4]

[1]*Department of Chemistry, McGill University, Montreal, QC, Canada*

[2]*Institute of Parasitology, McGill University (Macdonald Campus), Montreal, QC, Canada*

[3]*School of Dietetics and Human Nutrition, McGill University (Macdonald Campus), Montreal, QC, Canada*

[4]*Department of Chemistry, University of New Brunswick, Fredericton, NB, Canada*

5.3 Abstract

Given the global prevalence of parasite infections, there is a pressing need for Point-of-Care diagnostic tests for high- and low-resource settings. We developed an inexpensive ultrasound methodology to distinguish between levels of nematode infection in serum. Serum samples were collected from pregnant and non-pregnant CD1 mice infected 3 times with 0, 50, or 100 *H. bakeri* larvae. Ultrasound intensities for 51 frequencies between 2 and 7 MHz were converted

to audible frequencies through heterodyning and recorded by a sound card. On average, samples from infected mice attenuated the ultrasound signal more strongly than samples from uninfected mice. A multivariate classification model was developed to separate samples with no and high infection based on differences in ultrasound frequency spectra. The cross-validated model, in which each sample was estimated independent of the others, showed a sensitivity and specificity of 84% and 73% respectively and used only two wavelets centered at 3.1 and 5.8 MHz for classifying samples. We attribute the differences between groups to changes in total serum protein levels. We present a novel heterodyned continuous wave ultrasound instrument for determining infection status based on frequency profiles of mouse serum, independent of pregnancy status. Ultrasound has potential as a non-invasive diagnostic test and may provide insight into the mechanisms of response to infection.

Keywords

Ultrasound frequency analysis, heterodyning, *Heligmosomoides bakeri*, nematode parasite infection

5.4 Introduction

Gastrointestinal nematode parasites are among the neglected tropical diseases afflicting more than 1 billion humans and causing widespread disease in livestock.[5.1,5.2] Although the body can fight off many of these infections in a few weeks, repeated exposure to the parasites due to environmental circumstances often leads to reinfection and typically results in a chronic infection. Given the prevalence of parasitic infections, we consider the intestinal parasite *Heligmosomoides bakeri* (=*Heligmosomoides polygrus* and *Nematospiroides dubius*), commonly used as a mouse model for gastrointestinal nematode infections in humans.[5.3] The infection elicits an immune response typically characterized by splenomegaly, mesenteric lymphadenitis, eosinophilia, mastocytosis, and marked elevations in serum IgE and IgG1.[5.4–5.6] In addition, oxidative stress has been associated with parasite infection and has been proposed as an immune-activated defense mechanism against gastrointestinal nematodes.[5.7–5.9] This shift in the pro-oxidant/anti-oxidant balance in favour of reactive oxygen species has been implicated in oxidative protein modification and degradation.[5.10–5.12] Increases in immunoglobulins and oxidative protein modifications are both known to contribute to changes in serum viscosity[5.13,5.14] and may provide a useful avenue for tracking parasite infection.

There is a pressing need for point-of-care tests for high- and low-resource setting, given the prevalence of parasite infections. Criteria for the ideal point-of-care device include portability, low per-test cost, fast results (minutes), good precision/accuracy (sensitivity/specificity), and non-expert use.[5.15] These platforms can ease the burden on laboratories, avoid sample mix-ups, and guide the decision making process of health care professionals, ultimately leading to better patient care.[5.16] Though some point-of-care tests based on immunochromatographic methods have emerged for malaria, there is still a large dependence on microscopy for most parasite detection.[5.17] Alternative methods of diagnosis are based on immunoassays, molecular approaches, and proteomics using mass spectrometry, though these are still labor-intensive and not practical for field measurements.[5.18,5.19]

Ultrasound spectroscopy has garnered great interest as a tool for non-destructive and non-invasive determination. As a mechanical wave, ultrasound is not limited by the opacity of a sample, and can travel through highly turbid systems without the need for dilution. Moreover, ultrasound waves are sensitive to the chemical properties of the system through which they propagate. For example, increasing protein concentration in solutions has been shown to increase attenuation of ultrasound. [5.20] Additionally, nonlinear distortion of ultrasound waves, which results in the generation of harmonics, has been shown

to depend on amino acid composition of proteins. [5.20,5.21] Changes in molecular structures were found to affect solute-solvent interactions and determined the hydration sphere of proteins. Consequently, changes in frequency content of ultrasound signals can be related back to chemical properties of the sample. Recently, multivariate analysis of attenuation and nonlinear contributions in ultrasound spectra has been used for estimation of volume fraction in multicomponent liquid mixtures[5.22] and determination of alcohol and carbohydrate content in commercial beverages.[5.23] This suggests that once frequencies that correlate strongly with the sample property of interest have been identified, a simplified instrument that only measures the few frequencies of interest could be used. For targeted applications, a simplified instrument could significantly reduce costs.

To address the need for cost-efficient instrumentation, we explore the use of frequency mixing to produce frequencies that can be recorded with an inexpensive computer sound card, as compared to high-bandwidth oscilloscopes. In addition to the reduction in cost, the sound card offers a higher resolution (24 bit) than many digital oscilloscopes (8-16 bit). Frequency mixing, also referred to as heterodyning, involves the mixing of two frequencies to produce both the sum and difference of the two frequencies.[5.24] For example, two ultrasound

frequencies can be chosen so that the resulting difference produces a frequency in the audible range, which can then be recorded with a sound card.

We present a novel heterodyned continuous wave ultrasound instrument for determining infection status based on frequency profiles of mouse serum. Parasite infection stimulates a strong immune system response with marked elevation of immunoglobulins. We anticipate that samples of mice with infection will attenuate the ultrasound signal more strongly than samples of uninfected mice, and should allow for multivariate classification of mice according to infection status.

5.5 Materials and Methods

5.5.1 Samples

All mouse serum samples were obtained from a study by Odiere and colleagues, described elsewhere.[5.25] In summary, 67 mouse serum samples were collected from 67 female mice that were either pregnant or non-pregnant, and had either no, low, or high levels of infection with the murine gastrointestinal nematode, *Heligmosomoides bakeri*. The number of mice within each group is summarized in Table 5-1. Maternal serum samples were frozen at -20 °C until later analysis by ultrasound. All procedures were approved by the McGill Animal Care Committee, in accordance with the guidelines of the Canadian Council on Animal Care.

Table 5-1: Pregnant and non-pregnant mice with 3 doses of trickle infection with 3rd stage larvae (L_3) of *H. bakeri* (Uninfected: 3 x 0 L_3; Low: 3 x 50±3 L_3; High: 3 x 100±3 L_3).

	INFECTION STATUS		
	Uninfected	Low	High
Non-pregnant	11	15	17
Pregnant	7	9	8

5.5.2 Ultrasound Instrument

Ultrasonic frequencies were produced by a Novatech 409A signal generator (Novatech Instruments, Inc. Seattle, WA) that was programmed to sweep over a 2 to 7 MHz range in 0.1 MHz increments. After the signals were amplified with a ZHL-6A power amplifier (MiniCircuits, Brooklyn, NY), they were transmitted and received by a pair of wideband transducers centered at 5 MHz (Technisonic, Fairfield, CT). These transducers were positioned on parallel faces of a 9 mm pathlength sample cell with a 1 mm channel diameter, fabricated in-house out of PEEK plastic. A schematic of the ultrasound instrument is shown in Figure 5-1. To minimize signal loss due to the high acoustic impedance of air, the transducers were coupled to the sample cell's thin acetate windows (60 μm) with a thin layer of petroleum jelly. The nominal sample cell volume was 28 μL. The attenuated ultrasound signal was fed into a ZAD-1 mixer (MiniCircuits), along with a known offset ultrasonic frequency. This mixing of two frequencies may be expressed as the product of two sine waves:

$$\sin(2\pi f_1 t)\sin(2\pi f_2 t) = \frac{1}{2}\cos[2\pi(f_1 - f_2)t] - \frac{1}{2}\cos[2\pi(f_1 + f_2)t] \qquad (5.1)$$

A device with a nonlinear response, such as a transistor or diode, is needed to perform the multiplication and results in the generation of the sum $(f_1 + f_2)$ and difference $(f_1 - f_2)$ frequencies, as well as the original frequency components. If the offset frequency is appropriately chosen, the difference component, $(f_1 - f_2)$,

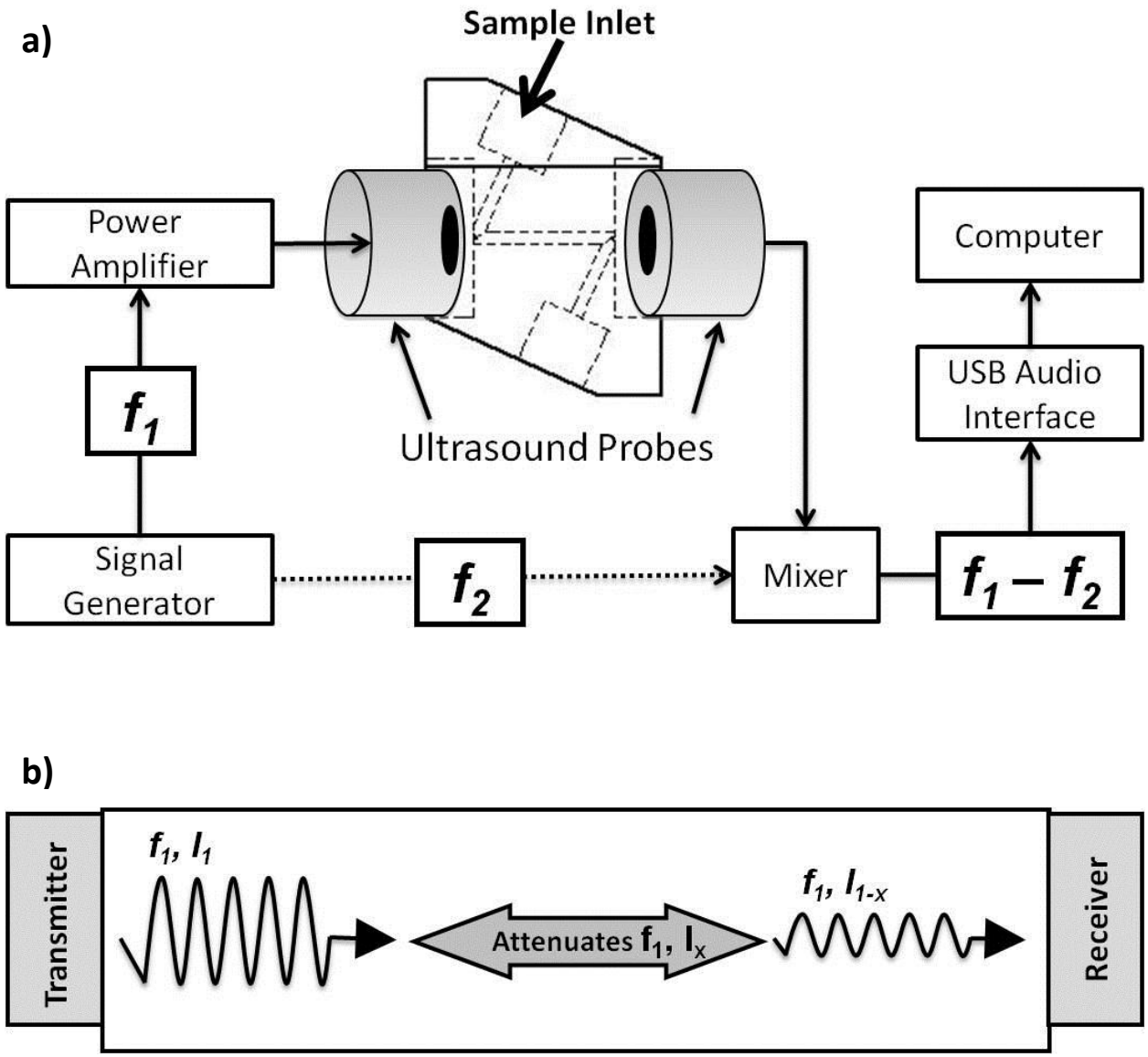

Figure 5-1: a) Schematic diagram of the continuous wave ultrasound apparatus with a heterodyned frequency recorder. b) As the ultrasound wave, f_1, with an initial intensity, I_1, propagates through the sample, it is attenuated by some amount, I_x, related to the physico-chemical properties of the sample.

can be an audible frequency with an amplitude proportional to the original frequency, f_1. The mixing process produced difference frequencies in the audible frequency range (5 to 15 kHz) that were recorded with a UA-1EX USB sound card (Roland, Los Angeles, CA), with a sampling frequency of 96 KHz and 24 bit resolution.

5.5.3 Ultrasound Measurements

Serum samples were thawed on wet ice, and were allowed to reach room temperature (22.0 ± 0.5°C) prior to measurement. The sample cell was cleaned between measurements with 0.1 N NaOH solution, followed by deionized water. An air purge was used to ensure no residual liquid remained in the cell prior to injecting any sample. To minimize sampling bias, frequency spectra of serum samples (~40 µL) were collected in randomized order. A blank measurement, consisting of distilled water, was made between serum sample measurements to ensure stability and reproducibility of the instrument. The relative standard deviations of blank measurements fell between 0.4% and 1.2% over a frequency range of 2 to 7 MHz. After passing through the sample, ultrasound signals were converted to audible frequencies by mixing with an offset frequency. This offset frequency was chosen so that the frequency difference would be between 5 and 15 kHz, increasing in 0.2 kHz increments. For example, a 2 MHz signal, after passing through the sample, was mixed with an offset frequency of 2.005 MHz,

the difference of which is 5 kHz and in the audible frequency range. As such, each unique frequency in the audible domain corresponded to a value in the ultrasound range. The samples were probed one frequency at a time, with an automated script that swept through 51 frequencies. These audio recordings were approximately 90 seconds in length and were considered as the raw data used for sample classification.

5.5.4 Data Analysis

All data processing used custom programs written in Matlab (The MathWorks Inc., Natick, MA, R2010a). Fourier transforms were carried out on the time domain audio recordings to obtain frequency spectra. The frequency data was then subject to a Haar-wavelet transform[5.26] that denoised and simplified the data. The Haar wavelet is a square wave that can be used to decompose the data into step functions, or wavelets, analogous to integrated spectral regions. Wavelets retain important spectral features and were used as the variables in model development, rather than individual frequencies.

The goal of this data analysis was to develop a classification model to distinguish serum samples of mice with a high infection dose from non-infected mice based on frequency profiles, using a minimum number of variables (wavelets). Dummy variables were assigned to the two classes, such that samples from the no infection group were represented by 0's and those from the

high group by 1's. The low infection samples were not used in the development of the model. The classification model was based on a logistic regression method, analogous to a stepwise multilinear regression, given by

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n \qquad (5.2)$$

where Y is the dependent variable (here the classification score), $\{X\}$ are the independent variables (the magnitude at a given frequency), and $\{b\}$ are the calibration coefficients. These coefficients may be determined when both Y and $\{X\}$ are known. For a perfect classification, scores of samples from the no infection group would be as close to 0 as possible, whereas scores for the high infection group would fall as close to 1 as possible.

To optimize selection of wavelets for the classification model, genetic algorithm searching was used to identify which wavelets to use for the best predictive model. In contrast to screening all possible wavelet combinations, which is computationally intensive, the genetic algorithm avoids human selection bias and reduces computation time. Details of the genetic algorithm can be found elsewhere.[5.27]

The best model maximized the sum of the sensitivity (true-positive rate) and specificity (true-negative rate) for a leave-one-out cross-validation, in which the score of each sample was calculated independent of all the other samples. A Wald test was used to test the statistical significance of each wavelet in the

model (α = 0.05) and to choose the most parsimonious model. The most parsimonious model includes only those variables that significantly affect the success of the model and is considered more robust than models that use additional variables.

5.6 Results and Discussion

A continuous wave ultrasound instrument with heterodyne detection was used to collect frequency profiles of sera from pregnant and non-pregnant mice with either no, low, or high infection with *H. bakeri* nematodes. The most significant changes in sera of mice infected with *H. bakeri* are increases in IgE and IgG1 due to the host immune response.[5.6,5.28] The parasites also secrete immunomodulatory proteins, which may be found in the host serum[5.29]; however, these proteins are present at minute levels compared to immunoglobulins. After albumins, immunoglobulins are the most abundant serum proteins; this increase in protein level with infection is reflected in the frequency profiles of mouse sera, as shown in Figure 5-2. The difference between samples with no infection and infection is marked by a decrease in signal, particularly at around 3.2 and 5.6 MHz. There is also a discernible difference between low and high infection status in the same region. An increase in serum viscosity due to raised immunoglobulin levels would account for the observed signal attenuation.

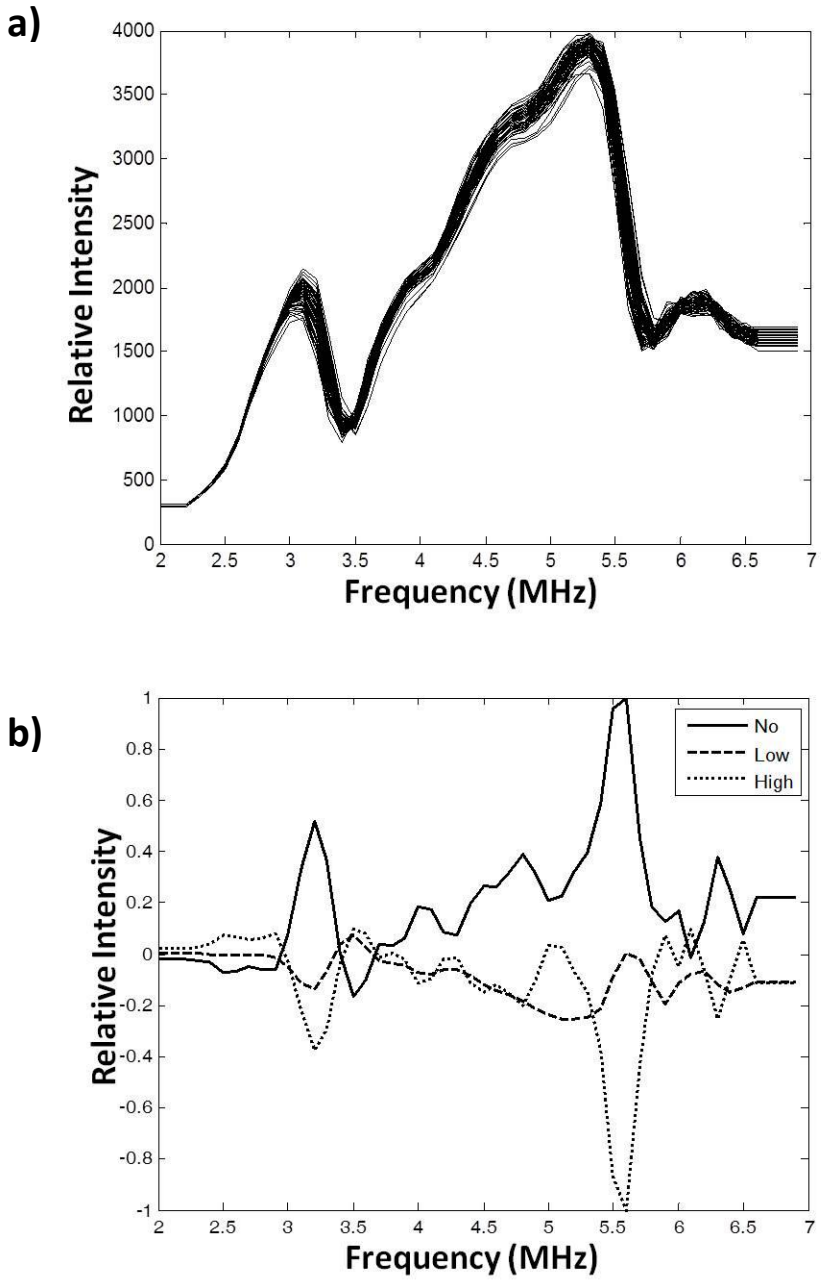

Figure 5-2: a) Smoothed frequency spectra of all mouse samples. b) Frequency spectra of mouse samples with no (solid), low (dashed), and high (dotted) infection. The spectra for each group have been averaged together, and the mean of all spectra has been subtracted from them.

To investigate if samples could be classified as infected or non-infected on the basis of their ultrasound spectra, the two infection extremes were considered, namely the no infection (n = 18) and high infection (n = 25) groups. The two groups included samples from both pregnant and non-pregnant mice to increase the sample size for classification. Samples from the low infection group (n = 24) were reserved as an independent test set. Before developing classification models, frequency spectra were treated with a Haar wavelet transform to express data as wavelets. Wavelets are useful for compressing and denoising data and can capture broad spectral trends in few variables. Wavelet data was used in place of single frequencies to develop a classification model to separate the no infection group from the high infection group.

Classification of samples was done by logistic regression, which used GA searching to identify which combination of wavelets gave the best separation. A cross-validated model was developed in which the classification score of each sample was estimated independently of the others, as shown in Figure 5-3. Effectively, for the 43 samples used for classification, an equivalent number of unique models were used to estimate the score of each sample. The most parsimonious model, as determined by a Wald test (α=0.05), consisted of only two wavelets centered at 3.1 and 5.8 MHz. Interestingly, the chosen wavelets were reminiscent of the frequency regions with the largest magnitude of signal

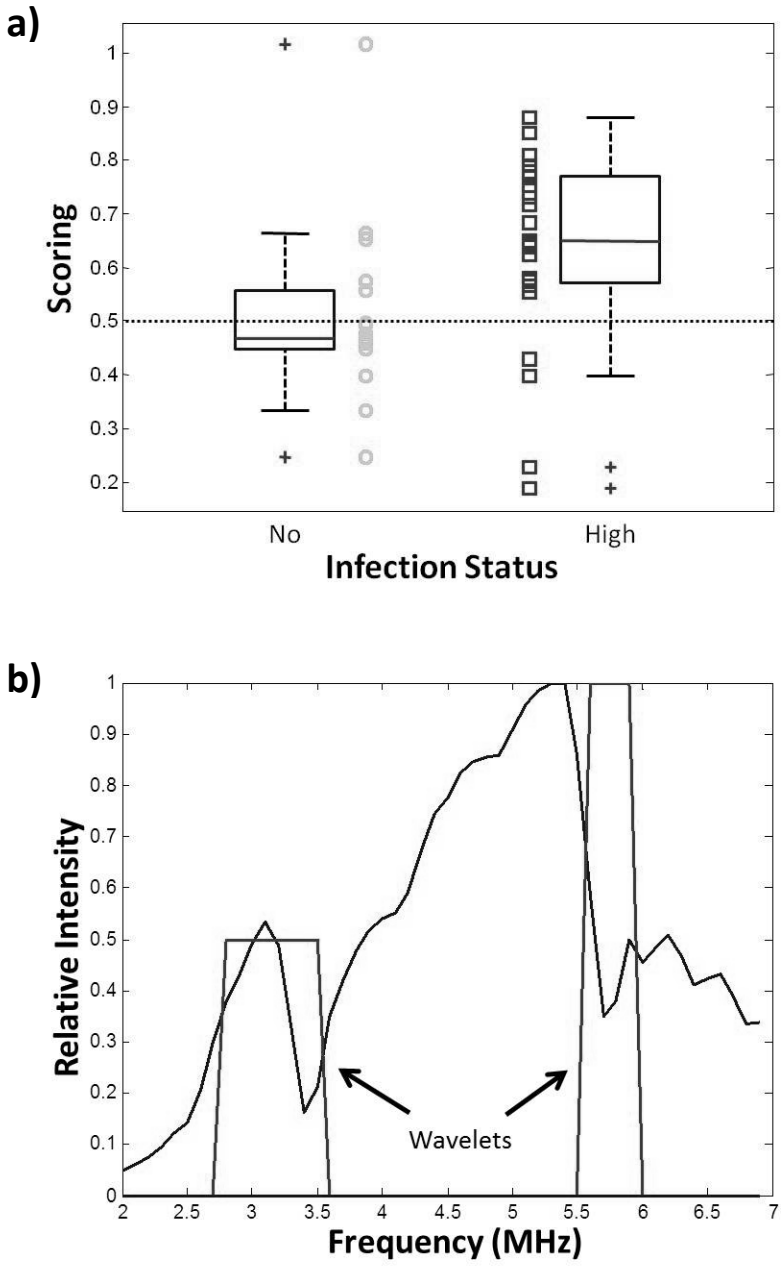

Figure 5-3: a) Classification model for no infection and high infection. Pregnant and non-pregnant samples were included in this model. Samples with low infection were estimated. b) Wavelets used in the classification model, centered at 3.1 and 5.8 MHz.

change with presence of infection in the mouse serum spectra, as shown in Figure 5-2. Using a Wilcoxon rank sum test (α=0.05), the medians of the no and multiple infection groups were found to be significantly different. Furthermore, the sensitivity and specificity of the model were 84% and 73% respectively. Sensitivity represents the correct number of infected mice identified above the model threshold, which was 0.5. Likewise, specificity represents the correct number of uninfected mice identified below the threshold. This means that for the high infection group only 4 out of 25 samples were misclassified, whereas only 5 out of 18 samples were misclassified for the no infection group.

The classification model was subsequently used to estimate the low infection samples, which were not included in the development of the classification model. The outcomes of the estimates are shown in Figure 5-4. The median classification score of the low infection group fell between the no and high infection group medians, although closer to the high group. Even with a low level of infection, a difference is apparent from the uninfected group, with a sensitivity of 71%. Additionally, the intermediate median score of the low infection group supports that the degree of the inflammatory response following infection is be dependent on the amount of parasite to which the host is exposed.[5.30] Effectively, the high infection group, which was given a total parasite dose of 300,

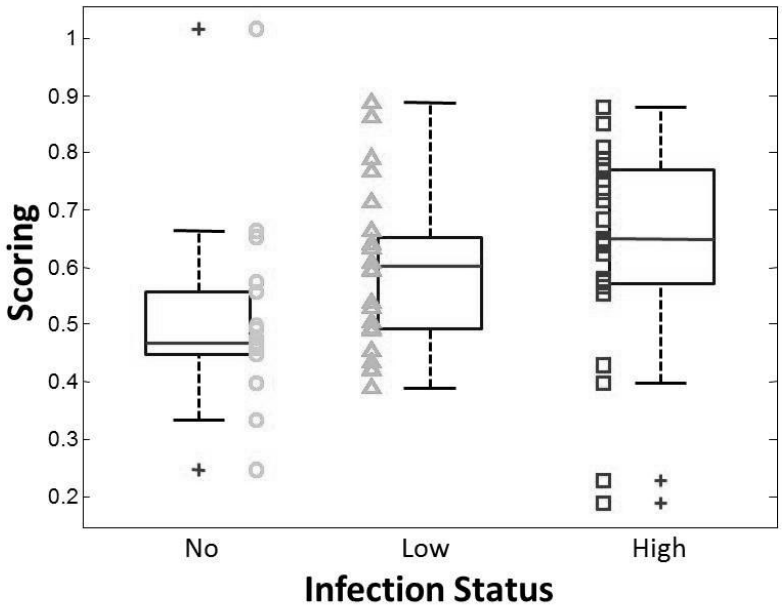

Figure 5-4: Estimate of the low infection group using the no/high infection group classification model.

had a higher median score compared to the low infection group, which was given a total parasite dose of 150.

Furthermore, classification was based only on infection status, with pooled pregnant and non-pregnant samples. Despite differences in pregnancy status, the model still showed good separation between the no/high infection groups and estimates of the low infection group. Pregnancy is known to be accompanied by changes in serum protein concentration[5.31,5.32] and increased oxidative stress.[5.33,5.34] Nevertheless, the results suggest that infection has a noticeably greater effect on serum composition in the selected frequency regions and that estimates of infection can be done independent of pregnancy status.

5.7 Conclusions

We have presented a straightforward, rapid, and cost-effective device for measuring infection severity in a model system of mice with the gastrointestinal nematode *H. bakeri*. Mouse serum samples were measured with a continuous wave ultrasound system, using an inexpensive heterodyned detection configuration. The frequency profiles reflected changes in viscosity associated with fluctuations in total serum protein levels. A classification approach was presented that separated no and high infections samples independent of pregnancy status using a very simple two-wavelet model. Models with few

variables are very useful for development of simplified instrumentation. For example, designing an instrument to generate and collect a few specific frequency bands allows for reductions in cost and optimization of performance for the required spectral regions. Additionally, the use of a sound card for recording ultrasonic signals demonstrates the feasibility for porting the ultrasound system to smartphones and tablets in future iterations. Ultrasound offers the potential for non-invasive measurements which makes this a particularly appealing avenue for parasite diagnostics and merits further exploration.

5.8 Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Fonds de recherche du Québec—Nature et technologies.

5.9 References

5.1. Bethony, J. M. *et al.* Vaccines to combat the neglected tropical diseases. *Immunol. Rev.* **239,** 237–70 (2011).

5.2. Sargison, N. D. Pharmaceutical treatments of gastrointestinal nematode infections of sheep--future of anthelmintic drugs. *Vet. Parasitol.* **189,** 79–84 (2012).

5.3. Behnke, J. & Harris, P. D. Heligmosomoides bakeri: a new name for an old worm? *Trends Parasitol.* **26,** 524–9 (2010).

5.4. Scott, M. E. & Koski, K. G. Zinc Deficiency Impairs Immune Responses against Parasitic Nematode Infections at Intestinal and Systemic Sites. *J. Nutr.* **130,** 1412–1420 (2000).

5.5. Behm, C. a & Ovington, K. S. The role of eosinophils in parasitic helminth infections: insights from genetically modified mice. *Parasitol. Today* **16,** 202–9 (2000).

5.6. Gause, W. C., Joseph Jr, F. U. & Stadecker, M. J. The immune response to parasitic helminths: insights from murine models. *Trends Immunol.* **24,** 269–277 (2003).

5.7. Van de Crommenacker, J., Richardson, D. S., Koltz, A. M., Hutchings, K. & Komdeur, J. Parasitic infection and oxidative status are associated and vary with breeding activity in the Seychelles warbler. *Proc. Biol. Sci.* **279,** 1466–76 (2012).

5.8. Schirmer, R. H. & Eisenbrand, G. Oxidative stress as a defense mechanism against parasitic infections. *Free Radic. Res. Commun.* **3,** 3–12 (1987).

5.9. Brophy, P. M., Ben-Smith, a, Brown, a, Behnke, J. M. & Pritchard, D. I. Differential expression of glutathione S-transferase (GST) by adult Heligmosomoides polygyrus during primary infection in fast and slow responding hosts. *Int. J. Parasitol.* **25,** 641–5 (1995).

5.10. Karve, T. M. & Cheema, A. K. Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. *J. Amino Acids* **2011,** 207691 (2011).

5.11. Kocha, T., Yamaguchi, M., Ohtaki, H., Fukuda, T. & Aoyagi, T. Hydrogen peroxide-mediated degradation of protein: different oxidation modes of copper- and iron-dependent hydroxyl radicals on the degradation of albumin. *Biochim. Biophys. Acta* **1337,** 319–26 (1997).

5.12. Gianazza, E., Crawford, J. & Miller, I. Detecting oxidative post-translational modifications in proteins. *Amino Acids* **33,** 51–6 (2007).

5.13. Lefebvre, J. Viscosity of concentrated protein solutions. *Rheol. Acta* **21,** 620–625 (1982).

5.14. Górnicki, A. The influence of oxidative stress on microviscosity of hemoglobin-containing liposomes. *Gen. Physiol. Biophys.* **22,** 121–7 (2003).

5.15. Yager, P., Domingo, G. J. & Gerdes, J. Point-of-care diagnostics for global health. *Annu. Rev. Biomed. Eng.* **10,** 107–44 (2008).

5.16. Kost, G. J. *et al.* Point-of-care Testing : Principles , Practice , and Medicine. *Encyclopedia of Analytical Chemistry* (2008). doi:10.1002/9780470027318.a0540.pub2

5.17. Clerc, O. & Greub, G. Routine use of point-of-care tests: usefulness and application in clinical microbiology. *Clin. Microbiol. Infect.* **16,** 1054–61 (2010).

5.18. Fox, J. C. *et al.* An overview of serological tests currently available for laboratory diagnosis of parasitic infections. *Vet. Parasitol.* **20,** 13–29 (1986).

5.19. Ndao, M. Diagnosis of parasitic diseases: old and new approaches. *Interdiscip. Perspect. Infect. Dis.* **2009,** 278246 (2009).

5.20. Sarvazyan, A. P., Chalikian, T. V & Dunn, F. Acoustic nonlinearity parameter B / A of aqueous solutions of some amino acids and proteins. *J. Acoust. Soc. Am.* **88,** 1555–1561 (1990).

5.21. Law, W. K., Frizzell, L. A. & Street, W. G. Determination of the Nonlinear Parameter B/A of Biological Media. *Ultrasound Med. Biol.* **11,** 307–318 (1985).

5.22. Dion, J. R. & Burns, D. H. Determination of volume fractions in multicomponent mixtures using ultrasound frequency analysis. *Appl. Spectrosc.* **65,** 648–56 (2011).

5.23. Dion, J. R. & Burns, D. H. Simultaneous determination of alcohol and carbohydrate content in commercial beverages by ultrasound frequency analysis. *Talanta* **86,** 384–92 (2011).

5.24. Poon, T.-C. in *Encyclopedia of Modern Optics* (ed. Guenther, B. D.) 201–206 (Elsevier, 2005).

5.25. Odiere, M. R., Koski, K. G., Weiler, H. A. & Scott, M. E. Concurrent nematode infection and pregnancy induce physiological responses that impair linear growth in the murine foetus. *Parasitology* **137,** 991–1002 (2010).

5.26. Nievergelt, Y. *Wavelets Made Easy*. *The Mathematical Gazette* **85,** 379 (Birkhäuser Boston, 1999).

5.27. Gributs, C. E. W. & Burns, D. H. Parsimonious calibration models for near-infrared spectroscopy using wavelets and scaling functions. *Chemom. Intell. Lab. Syst.* **83,** 44–53 (2006).

5.28. Mishra, P. K., Palma, M., Bleich, D., Loke, P. & Gause, W. C. Systemic impact of intestinal helminth infections. *Mucosal Immunol.* 1–10 (2014). doi:10.1038/mi.2014.23

5.29. Hewitson, J. P. *et al.* Secretion of protective antigens by tissue-stage nematode larvae revealed by proteomic analysis and vaccination-induced sterile immunity. *PLoS Pathog.* **9,** e1003492 (2013).

5.30. Borges, D. C., Araújo, N. M., Cardoso, C. R. & Lazo Chica, J. E. Different parasite inocula determine the modulation of the immune response and outcome of experimental Trypanosoma cruzi infection. *Immunology* **138,** 145–56 (2013).

5.31. Maher, J. E. *et al.* Albumin levels in pregnancy: a hypothesis--decreased levels of albumin are related to increased levels of alpha-fetoprotein. *Early Hum. Dev.* **34,** 209–15 (1993).

5.32. Benster, B. & Wood, E. J. Immunoglobulin levels in normal pregnancy and pregnancy complicated by hypertension. *J. Obstet. Gynaecol. Br. Commonw.* **77,** 518–22 (1970).

5.33. Burton, G. J. & Jauniaux, E. Oxidative stress. *Best Pract. Res. Clin. Obstet. Gynaecol.* **25,** 287–99 (2011).

5.34. Saikumar, P., Jaya, B. & Devi, M. R. R. Oxidative Stress in Pregnancy. *J. Dent. Med. Sci.* **3,** 12–13 (2013).

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis has explored the application of ultrasound for the determination of general markers of health in biological fluids. It was found that ultrasound transmission measurements allowed for rapid and accurate determination of sample pH and total cell counts, as well as presence of infection indicated by changes in total serum protein concentration and possible structural modifications from oxidative stress. Changes in biofluid properties were found to be related to the attenuation and nonlinear distortion of the waveform propagating through the sample. The quick, non-destructive ultrasound approach is suitable for routine measurements, particularly point-of-care diagnostics. A summary of the work and concluding remarks are given in the subsequent sections.

6.1.1 Research Summary

Multivariate analysis of ultrasound frequency data was found to be useful for the development of calibration and classification models. In the case of pH, multilinear regression models were developed to estimate the pH of albumin solutions and human serum samples. Binary classification models were developed for low and high cell counts in breast milk, as well as for no/multiple

infections in human serum and no/high infection in mouse serum. These classification models were then used to estimate intermediate states for the various samples. The simplicity of these classification models in describing complex samples, often with only two factors, underlined the usefulness of a yes/no type approach.

The effect of the medium on the ultrasound frequency profile was investigated by considering albumin solutions of varying concentrations at different pH values. It was shown that when ultrasound propagated through the solutions, the nonlinear distortion and attenuation depended on the albumin conformation. By measuring the ultrasound frequency spectra at several albumin concentrations and pH values, the pH of the solutions could be determined. Differences in absolute protein content contributed to baseline offsets in the frequency profiles and were minimized via area normalization. This strategy was extended to include human serum samples, with a standard error of estimate less than 0.08 pH units over the 7.7 to 8.6 pH range. These results show that protein conformation can be used as an indicator of pH.

Furthermore, we have shown the usefulness of ultrasound spectroscopy for assessing highly scattering samples. We have demonstrated that the nonlinear contribution of cellular structure to the ultrasound frequency profile can be used to estimate total cell counts in human milk. After correcting for

contributions from fat to the ultrasound signal, separation of low/high cell counts was possible with 90% sensitivity and 79% specificity. Additionally, we compared the ultrasound technique to NIR spectroscopy, which has previously been used to estimate cell counts in dairy. The classification model based on NIR, which used one less wavelet than the ultrasound model, had a lower sensitivity (85%) but higher specificity (84%). The direct analysis capabilities of both ultrasound and NIR for determining total cell counts in milk offer a straightforward and rapid method for monitoring maternal health.

The applicability of ultrasound to screening for general presence of infection was shown for both human and mouse serum samples. Observed differences in sera profiles of infected and non-infected groups were attributed to differences in total serum protein levels (increased immunoglobulins) and oxidative stress modifications of proteins. The rural indigenous Panamanian study population highlighted the prevalence of co-infections and malnutrition in the developing world and underlined the need for portable diagnostic tools to complement clinical exams. A general framework for separating women with no infection from those with multiple infections based on ultrasound frequency profiles was presented, with a sensitivity of 100% and specificity of 86%. Though the results were encouraging, they also revealed the challenges of developing an infection model for a real-world study population, in which many factors, such as

infection severity, presence of undiagnosed diseases, and malnutrition were not controlled for. These concerns were addressed in a controlled mouse experiment, which showed that mice with no or high levels of parasite infection could be identified with a sensitivity and specificity of 84% and 73% respectively.

Concurrently, we have addressed the need for more cost-effective instrumentation under a point-of-care paradigm with the development of an ultrasound device featuring a heterodyned detection system. The novel detection scheme, based on a sound card and signal mixer, offered higher resolution measurements compared to more expensive ultrasound oscilloscopes. These instrument refinements offer opportunities for developing more affordable diagnostic platforms.

6.1.2 Concluding Remarks

On the whole, nonlinear ultrasound frequency analysis was demonstrated to allow rapid and simple determination of general health indicators in complex, real-world biological samples. Ultrasound offers a variety of advantages compared to conventional methods of measurement, including the potential for non-invasive measurements. Ultrasound was found to be sensitive to molecular conformation, concentration, and structure of proteins, particularly to those present in high concentrations (g/L), such as albumins and immunoglobulins without the need for dilution. While ultrasound is well-suited to measuring

concentrated samples, proteins present in low concentrations do not significantly contribute to the observed ultrasound frequency profile. For low-concentration biomarker measurements, techniques that make use of dyes and reagents may be more suitable. At higher concentrations, though, ultrasound offers direct measurement of proteins without the need for consumables. While albumin and immunoglobulin concentrations provide useful markers of health, when considered on their own, they may not be sufficient to differentiate one type of infection or disease from another. Assays sensitive to infection-specific biomarkers would be more suitable. However, together with a clinical exam, the quick ultrasound measurement may provide very useful information for identification of infection.

With multiple parameters changing simultaneously, such as protein concentration and conformation, multivariate analysis techniques were necessary to identify frequency regions correlated with sample properties of interest. Further investigation is needed to understand the interplay between these parameters. Whereas the nonlinear ultrasound data was more complicated to interpret, requiring the use of multivariate calibration tools, the additional data processing was offset by minimal sample preparation. However, there is a conceivable limit to how much information can be gleaned using multivariate analysis before sample preparation techniques would be needed.

For biomedical diagnostics, portability coupled with real-time and accurate measurements could improve patient healthcare in both low- and high-resource settings. This methodology is suited to personal health monitoring, where routine measurements of total serum protein, for instance, can provide a general indicator of an individual's health and can signal when there are deviations from the norm, as determined by an initial baseline measurement.

6.2 Future Work

Based on the findings in this dissertation, there are two paths to consider in refining the ultrasound frequency analysis methodology. The first is expanding our understanding of the nonlinear contributions of proteins and cells to the signal in biofluids. This could lead to refined models with better predictive capabilities. The second route entails exploring modifications to the instrument for further simplification and cost-effectiveness. Preliminary results are presented where applicable.

6.2.1 Refining Nonlinear Ultrasound for pH measurements

The first manuscript in this thesis presented an approach for estimating pH based on conformational changes of albumin. Though albumin is the dominant protein constituent in serum, it would be useful to explore contributions from other abundant proteins, such as immunoglobulin G and transferrin, to the signal.

To explore this, a preliminary trial was designed in which two structurally different proteins, BSA and ovalbumin, were mixed together in varying proportions while keeping the total protein concentration fixed at 40 g/L. It was found that signal magnitude decreased with increasing weight fraction of BSA. This suggests that it should be possible to measure the concentration of individual proteins in a mixture using ultrasound frequency analysis. It would be interesting to extend this to 3-component systems, such as an albumin/IgG/transferrin mixture. This could allow for a total serum protein level measurement that would include albumin and immunoglobulin information and could be used to refine models for estimating pH in serum.

It would also be useful to perform pH measurements in blood. The milk measurements presented in this thesis were done on a whole biofluid and considered contributions from cellular components to the ultrasound signal. In contrast, serum is not a whole biofluid and is free of cellular constituents. Looking at changes, such as pH, in whole blood instead of serum would allow for direct sample analysis by eliminating a sample preparation step. Red blood cells, like albumin, are known to undergo a shape change with pH.[6.1] Preliminary measurements made on whole rat blood show a shift in frequency profile with pH, as presented in Figure 6-1. A cross-validated calibration model was developed with a coefficient of determination of 0.99 and a standard error of 0.03

pH units. These preliminary measurements are very encouraging, but were restricted to one blood sample and do not explore differences in total protein levels or total cell counts. Further investigation is needed to understand the concurrent contribution of the protein conformation and the erythrocytes shape to the signal.

6.2.2 Instrument Modifications

Instrumentation for fixed applications typically does not require the flexibility offered by equipment used in a research setting and can be tailored for affordability. With regards to transducers, it may be feasible to use piezoelectric film sensors made of polyvinylidene difluoride (PVDF) in place of ceramic transducers. To investigate the feasibility of this, rectangular elements of PVDF with silver ink screen printed electrodes (typically used as contact microphones and strain gauges) were coupled to the sides of a sample cell filled with water and used to successfully transmit and receive ultrasound signals. The significant difference in cost between PVDF compared to ceramic transducers (e.g. $1 vs. $500) opens up the possibility for creating disposable transducer/sample cell modules. However, variability in transducer response would need to be addressed. Manufacturing transducers with closely matched responses is difficult

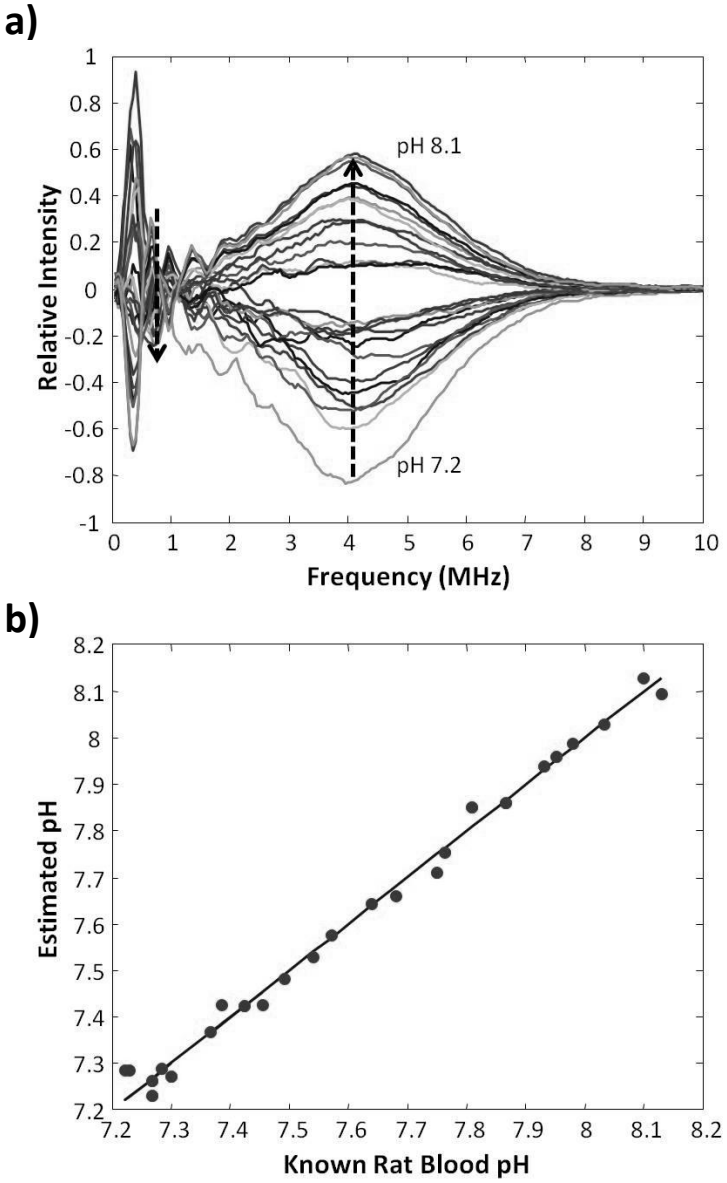

Figure 6-1: a) Mean-centered ultrasound frequency profiles of rat blood show an increase in signal intensity with increasing pH. Interestingly, this is the opposite of the trend observed for albumin solutions and serum samples. b) A multilinear regression model for rat blood pH shows a coefficient of determination of 0.99 and a standard error of cross-validation of 0.03 pH units over the 7.2 to 8.1 pH range.

and is one of the major reasons high performance ceramic transducers are so costly. In order to make comparable measurements between devices using inexpensive PVDF transducers, calibration will be necessary. One way to do this is through external calibration, where one or several reference samples are measured and used to establish the response of the device. Integrating these disposable transducer modules with refined ultrasound detectors will also need to be considered.

The heterodyned detection system presented in this dissertation made use of an external sound card and relied on a computer for data collection and processing. However, portable devices such as smartphones and tablets offer compact alternatives, as they are already equipped with soundcards and have sufficient processing power to handle data acquisition and manipulation. Overall, frequency mixing offers many exciting opportunities for developing the next generation of handheld ultrasound diagnostic devices.

6.3 References

6.1. Gedde, M. M., Davis, D. K. & Huestis, W. H. Cytoplasmic pH and human erythrocyte shape. *Biophys. J.* **72,** 1234–46 (1997).

Appendix A

Chemometrics

Abbreviations:

FT: Fourier transform

GA: genetic algorithm

HWT: Haar wavelet transform

MVC: multivariate calibration

MLR: multiple linear regression

PCR: principle component regression

PLSR: partial least-squares regression

PCA: principle component analysis

PC: principle component

SECV: standard error of cross-validation

A.1 Introduction to Chemometrics

The continued emergence of more powerful computers over the last few decades has allowed for the development of advanced data processing and calibration techniques. These computational data analysis techniques incorporate elements of multivariate statistics, applied mathematics, signal

processing, and computer science and have been applied to analytical chemistry problems under the rubric of chemometrics.[a.1]

The calibration of an instrument in analytical chemistry seeks to establish a quantitative relationship between instrument signals collected for a series of analytical samples and one or more properties of the samples. A univariate calibration would, for example, consider the absorption of light at a certain wavelength and relate this signal to the concentration of one analyte. In more complex samples, such as biofluids, using one signal is often insufficient to describe changes in the property of interest. In these cases multivariate calibration (MVC), which uses multiple signals to determine the property of interest, has much improved precision and selectivity. MVC has unlocked the potential of fast, inexpensive, and non-destructive analytical measurement techniques to estimate sample properties, even in the presence of unknown interfering chemical species. It works even when there is no selective wavelength region for the analyte. The broad and nonselective features of near-infrared spectra, as compared to infrared or Raman spectra, have been used successfully to develop multivariate calibration models for a broad range of applications.[a.2,a.3] Likewise, the complex and broad-featured spectra typical of nonlinear ultrasound spectroscopy have also benefited from the use of MVC.[a.4,a.5] The application of MVC can be extended to relate any instrumental profile, including

chromatograms, kinetic curves, and acoustic spectra to any property of the analyzed sample, such as , viscosity, particle size, taste, presence of infection etc.[a.6]

A.2 Multivariate Calibration Methods

Various MVC techniques exist and include multiple linear regression (MLR), principal component regression (PCR), and partial least squares regression (PLSR). The first of these, MLR, is a linear regression fit that has been extended to several variables. As with other MVC approaches, it is useful when several factors contribute to the overall observed response. MLR determines which linear combination of m independent variables best describes the sample property of interest, and is given by:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m + e \qquad \text{(A.1)}$$

where y is the dependent variable (e.g. concentration), $\{x\}$ are the independent variables (e.g. the magnitude at a certain wavelength), $\{b\}$ are the calibration coefficients, and e is the error. Equation A.1 describes the multilinear dependencies for only one sample (n=1). Extending this to n samples, we can write this in matrix form:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{e} \qquad \text{(A.2)}$$

where $\boldsymbol{y}$ and $\boldsymbol{b}$ are column vectors and $\{x\}$ form the rows of a matrix $\boldsymbol{X}$. The coefficients $\boldsymbol{b}$ may be determined when both $\boldsymbol{y}$ and $\boldsymbol{X}$ are known.

Two approaches can be used for variable selection. The first screens all possible combinations of variables and can be very computationally-intensive if the number of variables is large. For a data set with n variables, there are 2^n possible models to screen using an all possible combinations approach, which is typically not very practical. Instead, an alternate strategy that considers subsets of variables to arrive at a good solution is often used. However, since not all possible combinations of regressor variables are considered, better models may be missed. An example of this is stagewise MLR, where the concentration is regressed linearly against the signal magnitude at each wavelength. [a.7] The wavelength with the highest correlation coefficient is saved and the variance contribution of this wavelength is then subtracted from the spectra. This process is repeated iteratively on the reduced data set to select the subsequent wavelengths with the highest correlation to concentration. The optimal, or most parsimonious, model contains the minimum number of wavelengths such that the addition of another variable does not significantly change the prediction accuracy of the model.[a.8] The optimal number of wavelengths for the calibration model that balances adequate modeling and overfitting may be determined objectively by an F-test.[a.9] Once the calibration model has been developed, data from an

independent set, known as a validation set, is used to evaluate the performance of the model. Data is typically divided into a calibration and validation set prior to developing the model, with the intent to test the model with unbiased data points.

An alternate method for screening subsets of regressor variables is via a genetic algorithm (GA).[a.10] The GA uses principles such as selection, cross-over, and mutation to screen many models and find a good MLR model.[a.11,a.12] In an iterative process, solutions are generated for the best 1-variable model, up to the best n-variable model, which can be specified by the user.

The search process starts with a random initial subset population of variables and tried to find the best variables from this pool. Variables are chosen based on a fitness score. For each variable in the population, the coefficients b_1 to b_n of eq. A.1 are calculated by multilinear regression, where one sample at a time is withheld. Estimates of Y are then obtained for the withheld sample, and the process repeated for all samples. Consequently, y for each sample is estimated independently of the others in a procedure known as a leave-one-out cross-validation.

The standard error of the leave-one-out cross-validated result (i.e. SECV) is then used to calculate the fitness of each variable, which was defined as the 1/SECV. A larger fitness score is associated with a better variable. Subsequently, the two fittest variables are passed on to the next generation unchanged, while

the rest of the population is subject to cross-over and mutation to introduce diversity (i.e. add in variables that were not included in the initial population). Within this new population, the fitness of each variable is recalculated and the whole process repeated for many generations (i.e. 1000), so as to converge to a stable solution. Once a good solution is found for a 1-variable model, the cycle is repeated up to an n-variable model. An F-test can be used to choose the parsimonious model (i.e. model with least number of variables necessary). Typically, a model that uses fewer variables is considered more robust and should be chosen unless the inclusion of additional variables significantly increases the success of the model.

For very large data sets, methods such as PCR and PLSR may be more appropriate, as they can take data sets with thousands of variables and reduce the dimensionality prior to performing a regression. This is done by transforming to a new set of variables which are arranged so that the first few components retain most of the variation of the original data set.[a.13]

PCR is performed in two steps: the first step, a Principal Component Analysis (PCA) decomposes a data matrix X (m samples by n variables) into a new matrix of reduced rank (m samples by f principal components, with $f << n$) by defining new variables known as principal components (PC). These principal components are a linear combination of n variables in the original matrix X. The

first PC is chosen so that it captures the most variance in the data, with subsequent PCs ordered according to the amount of variance they describe. In this way, it may be possible to capture the majority of the total variance in the original data set with only a handful of PCs. Each PC can be defined by a loading vector, which describes the direction of the PC (direction cosines) and by a score vector, which contains the coordinates of the points along the PC line. A data matrix X can be thus be written as the outer product of the scores t and the loadings p':

$$X = t_1 p_1' + t_2 p_2' + \cdots + t_\alpha p_\alpha' \qquad \text{(A.3)}$$

or equivalently as $X = TP'$, where P' is made up of p' as rows and T of t as columns. The scores and loadings can be calculated pair-by-pair by an iterative procedure known as the nonlinear iterative partial least squares algorithm.[a.14] The second step of PCR uses the reduced rank score matrix $T = XP$ and the dependent variable for the MLR, which can be written in a similar form to Eq. A.2:

$$Y = TB + E \qquad \text{(A.4)}$$

Replacing the variables of X by score vectors that are orthogonal addresses the issue of collinearity, which can be problematic when using classic MLR.[a.15] Eliminating lesser PCs can also lead to noise reduction, though this elimination step also carries the risk that useful information in lesser PCs will be discarded. Methods such as PLS address this issue.

PLS is very similar to PCA, though with a key distinction. PCA finds components that describe the most variance in the data matrix X. These components may not be the best for describing the response variable, Y. PLS finds components in X that best describe the variance in Y.[a.16] The independent and dependent variables can be written in terms of scores and loadings:

$$X = TP' + E \qquad \text{(A.5)}$$

$$Y = UQ' + F \qquad \text{(A.6)}$$

where T and U are the X- and Y-scores respectively, P and Q are the loadings, and E and F are the error terms. The mixed relation connecting X and Y is given by

$$Y = TBQ' + F \qquad \text{(A.7)}$$

and allows for the estimation of the score and loading matrices. The first factor in PLS encompasses the maximum nonrandom variance in the data X that also correlates with Y, with subsequent factors describing the remaining variance in descending order.

A.3 Classification Approaches

The dependent variable in chemical analysis may not always be continuous. In the case of a categorical response variable, Y will represent a group or state. This is especially true in human disease diagnosis, where the

goal is to determine whether a certain disease is present or absent.[a.17] Using classification techniques, which are a form of pattern recognition, samples can be classified according to a specific property based on measurements indirectly related to that property. The previously described MVC methods, namely PCA, MLR, and PLS are easily adapted to tackle the problem of sample classification. PCA can provide an overview of how samples group together .This is done by plotting the data in a coordinate system defined by two or three PCs and looking for differences and similarities among the samples. Outliers can also be identified by this approach. For samples for which there is no a priori knowledge about which classes the samples may belong to, exploratory techniques under the umbrella of cluster analysis can determine structural characteristics of a data set by separating the data into groups, clusters, or hierarchies.[a.18]

MLR and PLS can also be used in cases where the response variable Y is categorical rather than continuous.[a.19] The only modification that needs to be done to the previously described methods is to code Y with dummy variables. For example, a two-state system could be represented by “0” for one state and “1” for the other.

A.4 Preprocessing of data

Preprocessing of data prior to applying calibration methods can be critical to obtaining meaningful models. This involves taking raw data collected by an instrument and manipulating it to make it more suitable for calibration. These preprocessing steps can include filtering methods (e.g. boxcar averaging, differentiation), baseline corrections , domain transformations (e.g. Fourier transform), and wavelet transforms.[a.20] Applying filters can reduce noise in data, while baseline corrections can minimize offsets to make spectra more comparable. Fourier transforms (FT) allows for time domain signals to be interpreted in terms of their frequency content. Any signal, s_x, can be decomposed into the sum of simple sine and cosine functions:

$$s(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\,x) + \sum_{n=1}^{\infty} b_n \sin(n\,x) \quad \text{(A.8)}$$

where a_n and b_n are Fourier coefficients and $n = 1,2,3, \ldots$ .The FT is particularly useful for ultrasound signals when nonlinear distortions and generation of harmonic frequencies are of interest.

Wavelet transforms are gaining increased acceptance as a preprocessing tool for analytical measurements and are useful for denoising, smoothing, and compressing data.[a.20,a.21] Much like the Fourier transform, the Haar wavelet transform (HWT) projects the data onto a basis set of two functions, though instead of using sine and cosine waves, a “father” and “mother” wavelet are used

to break up and reconstruct a signal.[a.22] The father wavelet is comparable to a moving average function, while the mother wavelet provides the difference between the original signal and the average given by the father wavelet. Over the range $0 \leq x < 1$, the father(φ) and mother (ψ) wavelet functions are defined as:

$$\varphi(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad \text{(A.9)}$$

and

$$\psi(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1/2 \\ -1 & \text{if } 1/2 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \qquad \text{(A.10)}$$

The HWT allows for the retention of important spectral features in a potentially small number of wavelets and is useful for compressing large data sets. Within this framework, wavelets may be thought of as integrated spectral regions or bands, in which the averaging serves to reduce noise. This type of preprocessing step may be used in place of rank reducing techniques, such as PCA and PLS, to reduce the size of the data sets. The advantage of the HWT is that wavelets can be directly interpreted in terms of the wavelengths and frequencies which they represent and allow for identification of broad trends in the spectra.[a.11] Preprocessing steps can also be combined to potentially improve outcomes, such as the blending of PCR, PLS, or MLR with wavelet compression.[a.11,a.23,a.24]

In chemical analysis, extensive sample preparation can reduce the complexity of a sample to the point where determinations are possible with

univariate techniques. Although interpretation of the collected data is very simple, the labor-intensive and time-consuming nature of sample preparation can be impractical. Multivariate tools provide an exciting means to potentially forego sample preparation steps and to perform chemical determinations in complex media. Whereas the initial development of a MVC may be tedious, once a suitable calibration model has been developed and tested, sample analysis can be done rapidly. Eliminating the need for special sample treatment would allow tests to be performed by non-experts. The advantages of MVC highlight its potential for the development of point-of-care diagnostic platforms.

A.5 References

a.1. Wold, S. Chemometrics; what do we mean with it, and what do we want from it? *Chemom. Intell. Lab. Syst.* **30,** 109–115 (1995).

a.2. Iñón, F. a, Llario, R., Garrigues, S. & de la Guardia, M. Development of a PLS based method for determination of the quality of beers by use of NIR: spectral ranges and sample-introduction considerations. *Anal. Bioanal. Chem.* **382,** 1549–61 (2005).

a.3. Armenta, S., Garrigues, S. & de la Guardia, M. Determination of edible oil parameters by near infrared spectrometry. *Anal. Chim. Acta* **596,** 330–7 (2007).

a.4. Dion, J. R. & Burns, D. H. Determination of volume fractions in multicomponent mixtures using ultrasound frequency analysis. *Appl. Spectrosc.* **65,** 648–56 (2011).

a.5. Dion, J. R. & Burns, D. H. Simultaneous determination of alcohol and carbohydrate content in commercial beverages by ultrasound frequency analysis. *Talanta* **86,** 384–92 (2011).

a.6. Wold, S. & Josefson, M. Multivariate Calibration of Analytical Data. *Encyclopedia of Analytical Chemistry* 1–27 (2006). doi:10.1002/9780470027318.a5205

a.7. Arakaki, L. S. L. & Burns, D. H. Multispectral Analysis for Quantitative Measurements of Myoglobin Oxygen Fractional Saturation in the Presence of Hemoglobin Interference. *Appl. Spectrosc.* **46,** 1919–1928 (1992).

a.8. Hosmer Jr., D. W., Lemeshow, S. & Sturdivant, R. X. *Applied Logistic Regression*. (John Wiley & Sons, Inc., 2013). doi:10.1002/9781118548387

a.9. McClave, J. T. & Sincich, T. in *Statistics* (Pearson Education, 2006).

a.10. Leardi, R., Boggia, R. & Terrile, M. Genetic Algorithms As A Strategy for Feature-Selection. *J. Chemom.* **6,** 267–281 (1992).

a.11. Gributs, C. E. W. & Burns, D. H. Parsimonious calibration models for near-infrared spectroscopy using wavelets and scaling functions. *Chemom. Intell. Lab. Syst.* **83,** 44–53 (2006).

a.12. Yang, X.-S. in *Nature-Inspired Optimization Algorithms* 77–87 (Elsevier, 2014). doi:10.1016/B978-0-12-416743-8.00005-1

a.13. Kumar, N., Bansal, A., Sarma, G. S. & Rawal, R. K. Chemometrics tools used in analytical chemistry: an overview. *Talanta* **123,** 186–99 (2014).

a.14. Geladi, P. & Kowalski, B. R. Partial least-squares regression: a tutorial. *Anal. Chim. Acta* **185,** 1–17 (1986).

a.15. Wold, S., Ruhe, A., Wold, H. & Dunn, III, W. J. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J. Sci. Stat. Comput.* **5,** 735–743 (1984).

a.16. Haaland, D. M. & Thomas, E. V. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Anal. Chem.* **60,** 1193–1202 (1988).

a.17. Madsen, R., Lundstedt, T. & Trygg, J. Chemometrics in metabolomics--a review in human disease diagnosis. *Anal. Chim. Acta* **659,** 23–33 (2010).

a.18. Kaufman, L. & Rousseeuw, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis*. (John Wiley & Sons, Inc., 1990). doi:10.1002/9780470316801

a.19. Barker, M. & Rayens, W. Partial least squares for discrimination. *J. Chemom.* **17,** 166–173 (2003).

a.20. Wentzell, P. D. & Brown, C. D. Signal Processing in Analytical Chemistry. *Encyclopedia of Analytical Chemistry* 1–36 (2006). doi:10.1002/9780470027318.a5207

a.21. Boisvert, M. R., Koski, K. G., Burns, D. H. & Skinner, C. D. Prediction of gestational diabetes mellitus based on an analysis of amniotic fluid by capillary electrophoresis. *Biomark. Med.* **6,** 645–653 (2012).

a.22. Nievergelt, Y. *Wavelets Made Easy*. *The Mathematical Gazette* **85,** (Birkhäuser Boston, 1999).

a.23. Leger, M. N. & Wentzell, P. D. Maximum likelihood principal components regression on wavelet-compressed data. *Appl. Spectrosc.* **58,** 855–62 (2004).

a.24. Trygg, J. & Wold, S. PLS regression on wavelet compressed NIR spectra. *Chemom. Intell. Lab. Syst.* **42,** 209–220 (1998).

Appendix B

Flow Cytometry Plots

Plots of forward and side scatter intensities (FSC and SCC, respectively) and plots of fluorescence intensities (FITC) are provided for select human milk samples from Chapter 3 (“Estimation of Somatic Cell Counts in Breast Milk by Ultrasound and Near Infrared Spectroscopy”). The first sample presented consisted of CountBright reference beads suspended in a solution of PBS (Figure B-2). These reference beads had a very distinct FSC-SCC profile and exhibited very strong fluorescence intensities, along them to be easily identified from fat globules and cells present in milk.

The next sample presented consisted of human milk with reference beads added (Figure B-3). Defatting of the milk prior to analysis reduced the amount of fat globules present in the sample. However, the remaining fat globules and cells had overlapping size distributions and could not be distinguished in the FSC-SSC plot, whereas the reference beads could still be identified. Likewise, the fluorescence channel showed a clear separation of the reference beads from the weak scatter signal produced by the fat and cell constituents.

In order to separate milk cells from fat globules, a cell-permeable nucleic fluorescent stain was added to the same milk sample (Figure B-4). Although no differences were observed in the FSC-SSC plot, a new fluorescence signal was

observed. This signal fell between the weak scattering of fat globules and the strong fluorescent intensities of the reference beads. The observed fluorescing cell intensities were one to two orders of magnitude lower than those of the reference beads and corresponded to the manufacturer's specifications of fluorescent intensities. Given the presence of fat globules, cell counting was done on the basis of the fluorescence channel. This cell-counting strategy was applied to the remainder of the milk samples.

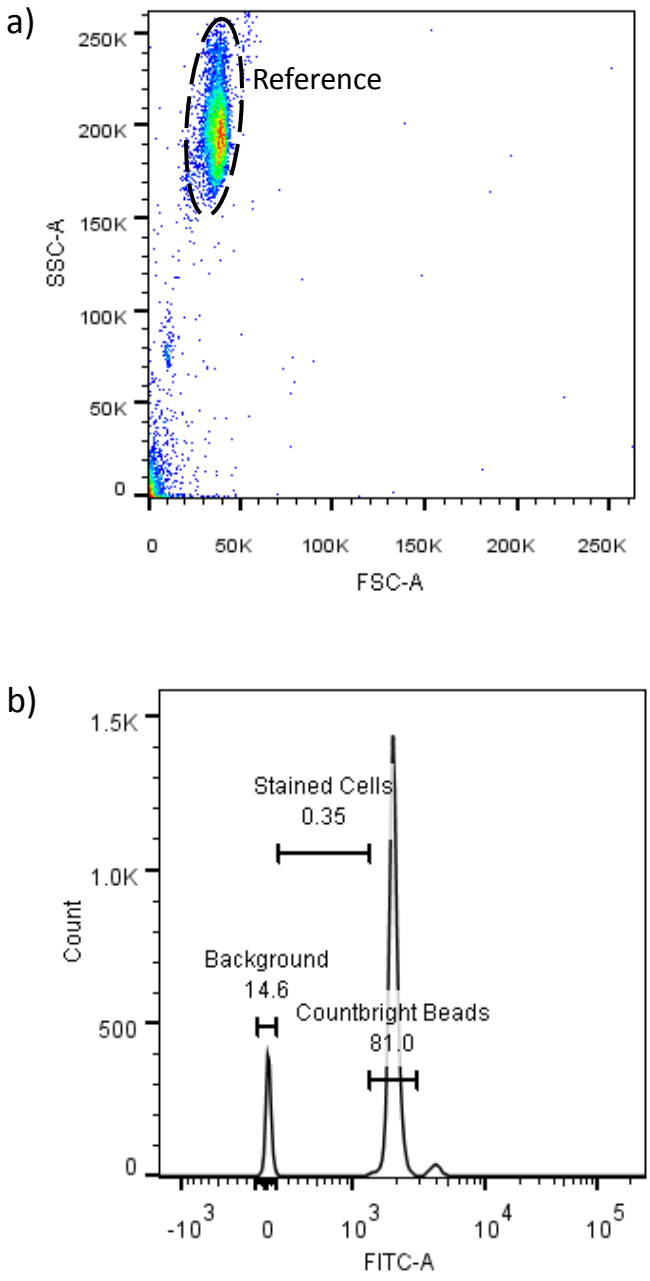

Figure B-2: Reference counting (CountBright) beads in a PBS solution. a) Plot of forward scatter (FSC) and side scatter (SSC) intensities. b) Plot of fluorescence channel (FITC) intensities.

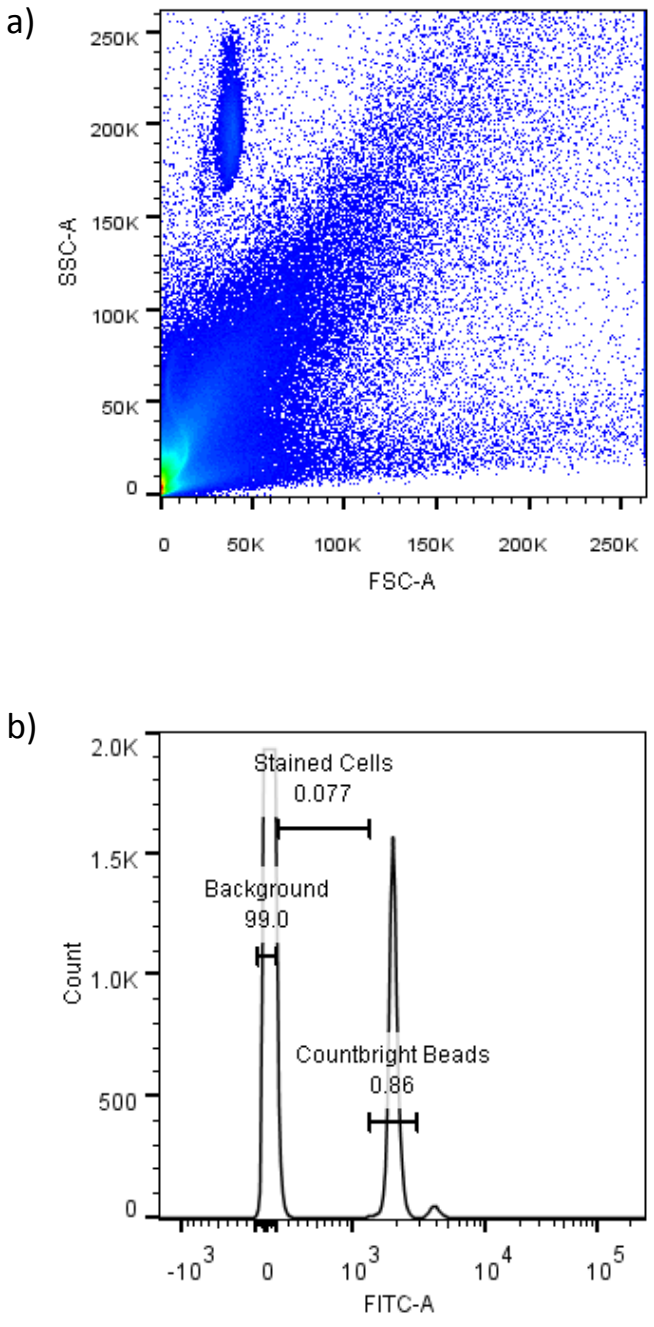

Figure B-3: Unstained human milk sample with reference counting beads. a) Plot of FSC-SSC intensities. b) Plot of fluorescence channel intensities.

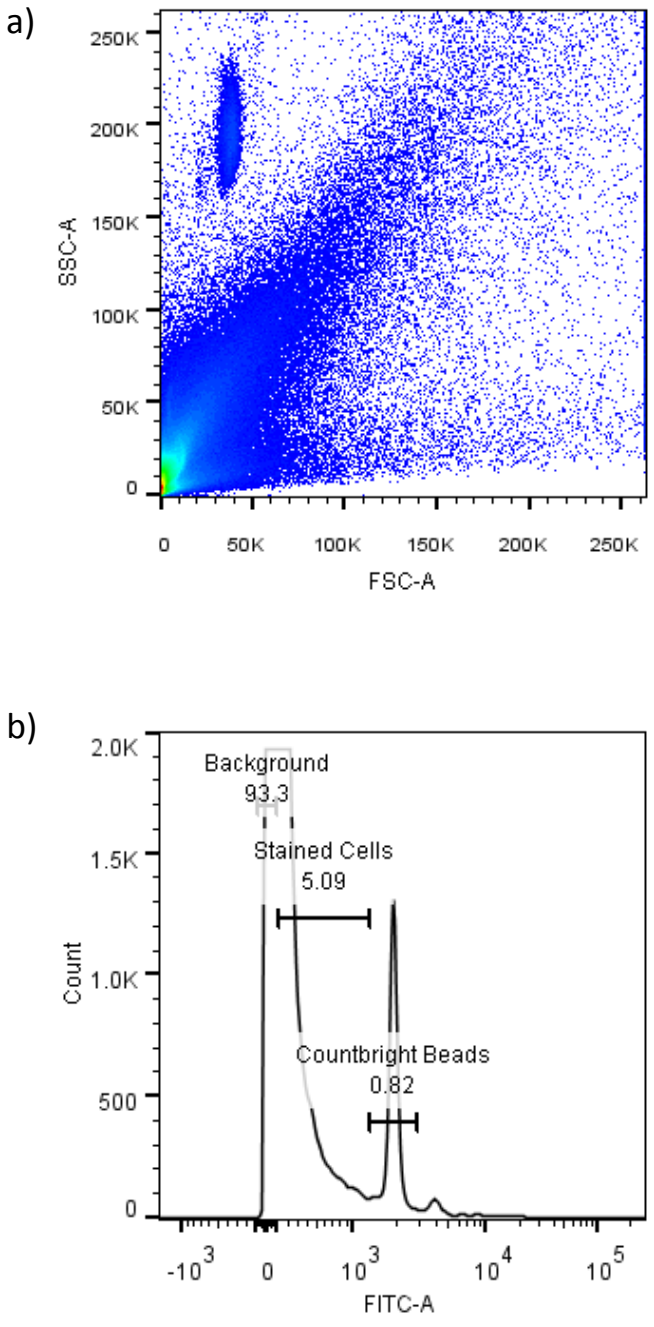

Figure B-4: Stained human milk sample with reference counting beads. a) Plot of FSC-SSC intensities. b) Plot of fluorescence channel intensities.