2016-03-01

Discovery of Low-Molecular Weight Novel Serum Biomarkers for Diagnosing Preeclampsia and Alzheimer's Disease

Swati Anand
Brigham Young University

Discovery of Low-Molecular Weight Novel Serum Biomarkers for Diagnosing Preeclampsia and Alzheimer's Disease

Swati Anand

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Steven W. Graves, Chair
David V. Dearden
Daniel E. Austin
John C. Price
John D. Lamb

Department of Chemistry and Biochemistry

Brigham Young University

March 2016

ABSTRACT

Discovery of Low-Molecular Weight Novel Serum Biomarkers for Diagnosing Preeclampsia and Alzheimer's Disease

Swati Anand
Department of Chemistry and Biochemistry, BYU
Doctor of Philosophy

Preeclampsia (PE), a life threatening pregnancy-related disorder, is characterized mainly by new onset of hypertension and proteinuria after 20 weeks of gestation. Currently, PE cannot be predicted prior to onset of symptoms and there is no cure for the disease. There is a clear value in having biomarkers able, early in a pregnancy, to identify women at risk for PE so that proper treatment therapies could be developed. Although a number of serum candidate markers have been proposed to be altered in PE patients, their use is limited due to poor sensitivity and specificity. Therefore, there is ongoing need for better set of novel biomarkers predicting PE.

Consequently, for my first project, we used a serum proteomic approach involving reversed phase capillary-liquid chromatography-electrospray ionization-quadrupole-time of flight mass spectrometry (cLC-ESI-QTOF). Our approach focuses on the less abundant (nM or lower), lower molecular weight peptides and lipids predicting PE. We got previously collected sera from pregnant women at 12–14 weeks gestation. There were 24 controls, having term uncomplicated pregnancies and 24 cases, which developed PE later in the same pregnancy. Many statistically significant serum PE biomarker candidates were found comparing cases and controls. In addition, multimarker combinations having high detection sensitivity and specificity (AUC >0.9) were developed using logistic regression analysis.

For my second project, serum lipidomic analysis of sera from pregnant women was undertaken to determine if useful PE lipid biomarkers exist. A discovery study involving a shotgun lipidomic approach was performed using sera collected at 12-14 weeks of pregnancy from 27 controls with uncomplicated pregnancies and 29 cases that later developed PE. Lipids were extracted using organic solvent and analyzed by direct infusion into a time-of-flight mass spectrometer. Statistically significant lipid markers were found and reevaluated in a second confirmatory study having 43 controls and 37 PE cases. The initial study detected 45 potential PE markers. Of these, 23 markers continued to be statistically significant in the second confirmatory set. Several multi-marker panels with AUC >0.85 and high predictive values were developed from these markers.

My third project also involved the above mentioned approach for detection of novel lipid biomarkers for Alzheimer's disease. Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common cause of age-related dementia. Currently, there are no methods to detect Alzheimer's at an early stage when treatment therapies could be applied. Therefore, there is need for detection of panel of biomarkers for detecting patients at risk to AD at an early stage. In the initial discovery set, sera from 29 different stage AD cases and 32 controls were analyzed using

direct infusion mass spectrometry (ESI-TOF). This study yielded 89 potential lipid biomarkers which were evaluated in another confirmation study. Of these, 35 markers continued to be statistically significant in the second confirmatory set. Using the confirmed markers, several multi-marker panels with AUC > 0.87 were developed for any stage AD cases vs controls. Multi-marker panels with AUCs > 0.90 were developed for each specific CDR vs controls, including the earliest stage of AD. These lipidomic biomarkers are likely to distinguish AD cases regardless of the stage from controls.

In conclusion, we successfully detected, validated and identified low molecular weight novel biomarkers for PE and lipid biomarkers for AD.

Keywords: Preeclampsia, Alzheimer's disease, Proteomics, Lipidomics, Mass spectrometry, diagnosis, biomarkers

ACKNOWLEDGEMENTS

I am grateful to my supervisor, Dr. Steven W. Graves, for giving me this opportunity to work on the exciting research projects involving disease biomarker discovery. His continuous support and guidance made it possible for me to successfully complete these projects.

I would like to thank our collaborators Dr. Sean Esplin, Dr. John Kauwe and Dr. Craig Thulin for their significant contributions in our projects. I would love to express my gratitude towards Bruce Jackson, the manager of the mass spectrometry center, for helping us maintain the instruments throughout our projects. He helped us understand all the important concepts and clarified our doubts regarding the instrument.

I am thankful to Dr. Evan Johnson, Dr. Dennis Tolley, Benjamin Peaden and Justin Barnes for carrying out statistical analysis of our data.

I am grateful to my seniors Dr. Tanielle Alvarez and Dr. Dipti Shah for training us on several lab techniques and teaching me important things throughout my Ph.D work. I would also like to thank Sydney Young and Diana Garcia who helped me with the analysis of my data and my colleagues Komal Kedia, Ying Ding and Huan Kang for their support.

I am thankful to my committee members Dr. David V. Dearden, Dr. Daniel E. Austin, Dr. J. D. Lamb and Dr. John C. Price for their encouragement, critical reviews and suggestions.

Lastly, I would love to thank my parents for their sacrifices, continuous support and motivating me throughout my projects.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

DE	2-dimensional gel electrophoresis
µg	microgram
µL	microliter
µm	micrometer
ACN	acetonitrile
AUC	Area under the curve
Blast	basic local alignment search tool
BMI	Body mass index
CID	collision-induced dissociation
cLC	capillary liquid chromatography
cLC-ESI-QTOFMS	capillary liquid chromatography-electrospray ionization-quadrupole time of flight mass spectrometry
cLC-MS	capillary liquid chromatography-mass spectrometry
CNS	Central nervous system
CRP	C-reactive protein
CV	Coefficient of variation
HELLP	Hemolysis, elevated liver enzymes, low platelet count
HPLC	High performance liquid chromatography
id	inner diameter
MS	Mass spectrometry
MS/MS	tandem MS
MW	Molecular weight
m/z	mass to charge ratio
nanoLC	nano liquid chromatography
NCBI	National Center for Biotechnology Information
P1GF	Placental growth factor
PA	Phosphatidic acid
PAPP-A	Pregnancy-associated placental protein A
PBS	Phosphate buffered saline
PC	Glycerophosphocholine/ phosphatidylcholine
PE	preeclampsia
PE	Phosphatidylethanolamine
rf	Radio frequency
ROC	Receiver operating characteristic
RPM	Rotations per minute
sEng	soluble endoglin
sFlt-1	Soluble VEGF receptor 1/ soluble fms-like tyrosine kinase 1
SGA	Small for gestational age
SIM	Selected ion monitoring
TIC	Total ion chromatogram/current
V	Volts
VEGF	Vascular endothelial growth factor
ESI	Electrospray ionization

NFTs	Neurofibrillary tangles
p-tau	phosphorylated tau
PUFA	Polyunsaturated fatty acids
q	Charge of the ion
AD	Alzheimer's disease
CDR	Clinical dementia rating
CSF	Cerebrospinal fluid
Aβ	Amyloid beta
Aβ 42	42-amino acid amyloid beta peptide

Chapter 1 -Introduction

1.1 Electrospray Ionization-Quadrupole-Time-Of-Flight Mass Spectrometer (ESI-QqTOF)

During the latter half of the twentieth century mass spectrometers (MS) have played a significant role in measuring and identifying biological species used for clinical diagnosis [1-2]. This was because of their high resolution, sensitivity, mass accuracy, dynamic range and tandem MS abilities. We used a quadrupole-time of flight (QqTOF) mass spectrometer having an electrospray ionization source (ESI) for our studies described below.

1.1.1 QqTOF

Although first commercially introduced in 1996, QqTOF-MS have evolved over the years in terms of improved mass accuracy and resolving power and tandem MS abilities. Q is a mass resolving quadrupole; q refers to radio frequency (r.f.) only quadrupole or collision cell and TOF-MS refers to a time-of-flight mass spectrometer. The complete setup is shown in Figure 1.1 [3]. The instrument also has an additional r.f. only quadrupole q_0 to provide collisional damping. Therefore, it consists of three quadrupoles q_0, q_1 and q_2 where q_0 and q_2 are operated in r.f only mode to provide confinement of the ions. Both q_0 and q_2 are operated at several millitorr gas pressue to allow collision of specimen ions with neutral gas molecules like nitrogen or argon effectively resulting in collisional damping and ion focusing. It reduces both the energy and directional spread of ions leading to better transmission.

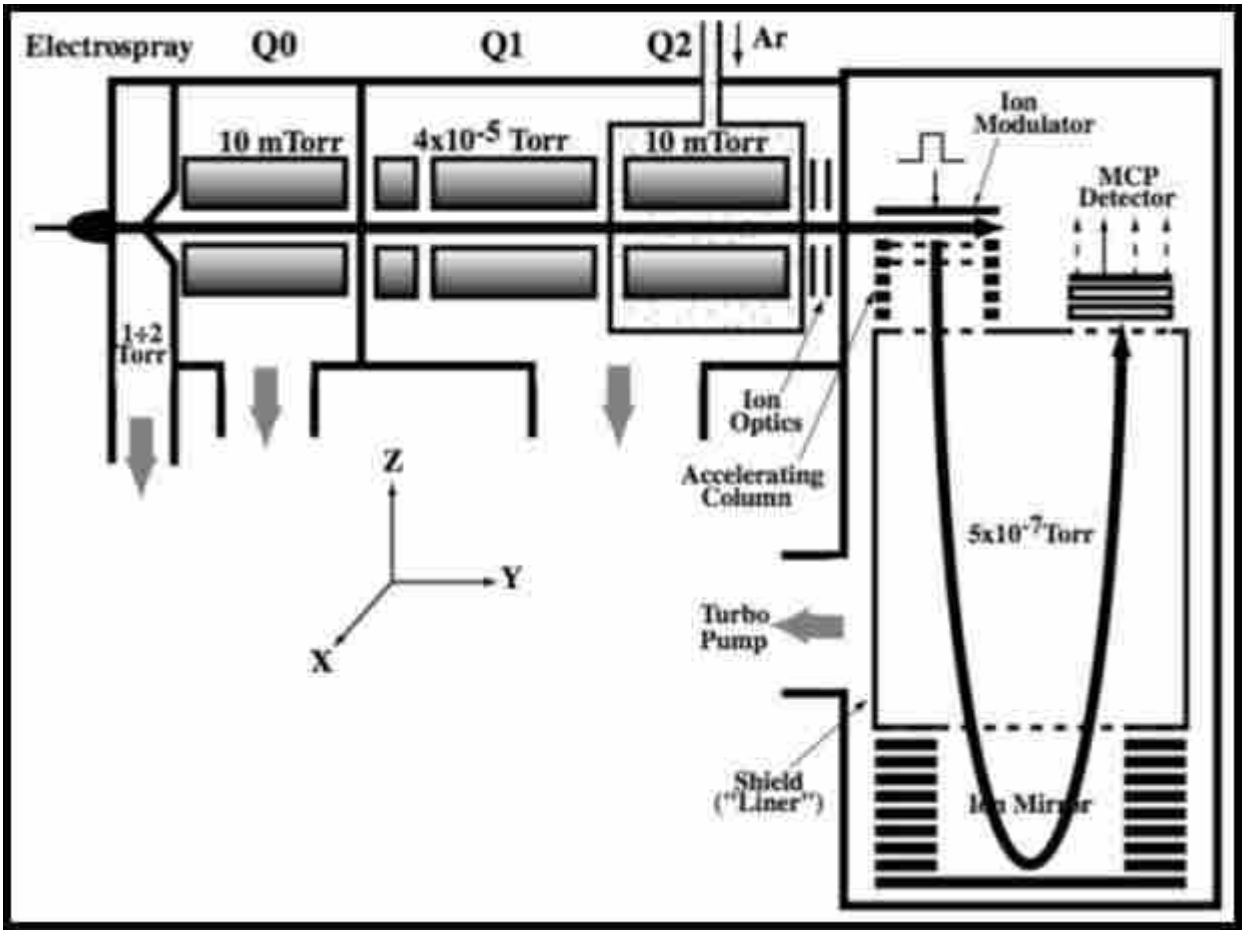


Figure 1.1 Quadrupole time-of-flight mass spectrometer.
(Taken from Chernushevich, Igor V., Alexander V. Loboda, and Bruce A. Thomsoken from. T "An introduction to quadrupole–time-of-flight mass spectrometry." *Journal of Mass Spectrometry* 36.8 (2001): 849-865.)

For MS1 experiments, Q1 is mainly operated in r.f. mode acting as an ion transmission source, while TOF is used for recording the spectra because of its high resolution, mass accuracy and ability to simultaneously record ions without scanning, resulting in faster analysis.

In MS/MS, Q_1 acts as a mass filter selecting a parent ion of interest. Therefore, Q_1 is operated at very low pressures ~10^{-5} torr to successfully isolate the ion of interest. The selected ion is then accelerated at a specified energy before it enters the collision cell where it undergoes collisions with the neutral gas molecules and is fragmented into product ions.

When the quadrupole is used as a mass analyzer, both r.f. and d.c. voltages are applied to the rods. Mathieu parameters q_M and a_M are used to characterize the amplitudes of both the voltages

$$q_M = \frac{4eV}{\left(\frac{m}{z}\right)\acute{\omega}^2 r^2} \qquad a_M = \frac{8eU}{\left(\frac{m}{z}\right)\acute{\omega}^2 r^2}$$

where, e represents the charge of an electron, V and $\acute{\omega}$ are the amplitude and angular frequency of the r.f. voltage. U is the amplitude of the d.c. voltage and r is the radius of the quadrupoles.

In general, transmission of ions having very high m/z is compromised as the effective potential is inversely proportional to the m/z.

After leaving the collision cell, the ions are reaccelerated at the given energy, followed by focusing the ion optics into the ion modulator region of the linear TOF. When leaving the acceleration region, all the ions should possess the same kinetic energies while entering the field free drift tube. Because the kinetic energies for all the ions are same, their velocities are inversely proportional to their masses resulting in the lighter ions reaching the detector faster than the heavier ions. The flight time (t_F) is given by

$$t_F = \frac{L}{v} = L\sqrt{\frac{m}{2zeV}}$$

Where L is the distance from the source to the detector, ν is velocity, m is mass, z is charge and V is the acceleration voltage.

A reflector/reflectron time-of-flight analyzer consists of a series of electrodes at the end of the field-free region at high potential. The voltage of the reflectron is higher than the acceleration voltage. This results in the effective penetration of all the ions until they reach zero kinetic energy (KE). The ions having high KE penetrate deeper and spend more time as compared with the ones having lower KE. The ions are then expelled by the reflectron and their kinetic energies remain unchanged. The ions in increasing order of their kinetic energies finally reach the detector. The reflector TOF analyzer allows for correction in time-of-flight as well as spatial spread and angular spread resulting in improved resolving power as compared to a linear TOF analyzer.

The orthogonal acceleration TOF (oa-TOF) analyzer results in the transmission of ions orthogonal to their initial direction due to application of the pulsed electric field across the modulator region. The ions acquire kinetic energy and enter the field-free region where TOF mass separation occurs. The ratio of the velocities (or energies) in the acceleration region and orthogonal region decides the order of the ions reaching the ion mirror, from where the ions are directly deflected to the detector. The ion mirror compensates the energy and spatial spread of the ions.

The oaTOF offers several advantages including high mass resolving power, sensitivity, spectral acquisition rate and mass accuracy.

1.1.2 Electrospray ionization (ESI)

ESI is a soft ionization technique first reported in 1984 [4]. It allows the formation of multiply charged species which allows detection of very large analytes. It takes place at atmospheric pressure and can be coupled to liquid chromatography resulting in the separation of complex mixtures before entering the MS. Analyte solutions are transferred through a metal capillary needle which is maintained at several kV (2-6 kV) with respect to a cylindrical electrode resulting in spray of charged droplets. A coaxial flow of sheath gas (dry N_2) around the capillary results in better nebulization and helps to direct the spray towards the mass spectrometer. The resulting charged spray of droplets from the capillary tip is distorted into a Taylor cone which undergoes rapid evaporation of the solvent. This results in the continuous buildup of charge density until the surface tension of the droplets is balanced by coulombic repulsion. At this point, called the Rayleigh limit, large droplets result in the production of highly charged smaller droplets. These droplets can continue the cycle until all the solvent is evaporated, leaving multiply charged analytes. The resulting analyte ions have low internal energies; therefore, the structure of the analyte generally remains intact (with no fragmentation) provided appropriate instrumental conditions are used. The mechanism of ESI is shown in Figure 1.2.

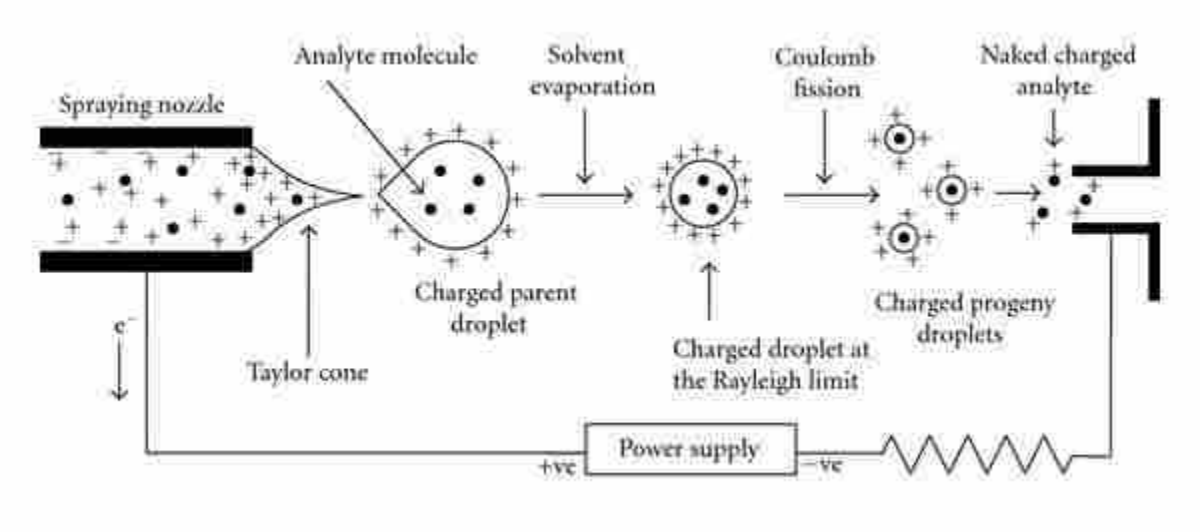


Figure 1.2 Schematic representation of ESI process
(taken from Banerjee, S.; Mazumdar, S., Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International journal of analytical chemistry* 2012.)

1.1.3 Turbomolecular pumps (TMP)

The first turbopump was developed and patented by Dr. Willi Becker at Pfeiffer Vacuum in 1957. It can produce a vacuum ranging from 10^{-3} to 10^{-8} Torr. These pumps usually consist of multiple stacks of rotors with blades and stators in between, which are the fixed discs having blades as shown in Figure 1.3. The rotors and stators are mounted in series**.** The motor makes the rotor spin around the axis.

As the gas molecules enter through the inlet, they get hit by the blades of the rotor. The mechanical energy of the blades is transferred to the gas. Due to the acquired momentum, the gas molecules enter the holes of the stator leading them into the next set of rotors and stators. The transfer of the gas through the series of rotors and stators directs it towards the exhaust where it's collected by the backing pump.

The flow of a gas is viscous flow at atmospheric pressure. This means that the individual gas molecules interact more significantly with each other than they do with the walls of the container. The momentum imparted to gas molecules by the turbopump is less significant compared to the energy acquired after collision between the gas molecules. Therefore, a turbopump cannot function at atmospheric pressure. In order to make it work, a rotary pump is employed to reduce gases before the turbopumps starts to function. These reduce the pressure to ~10^{-3} Torr. At this pressure, the gas molecules do not interact significantly with each other. Thus, the turbopump can effectively impart momentum to each molecule resulting in vacuum of ~10^{-8} Torr. These pumps operate at speeds of up to 20,000 rpm.

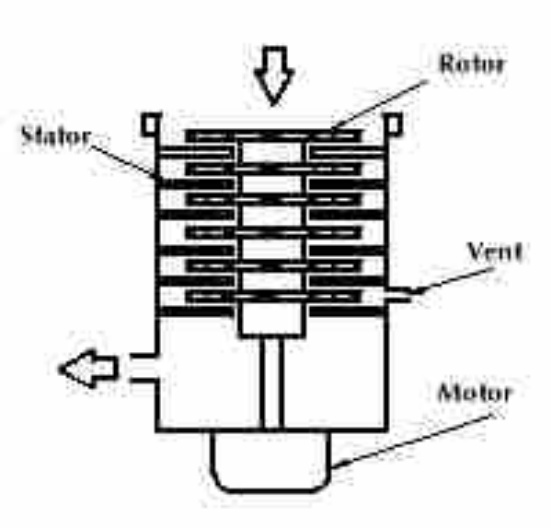


Figure 1.3 Turbomolecular pump diagram
(taken from http://www.repairfaq.org/sam/vacuum/tmpnotes.htm)

1.1.4 Rotary pumps

These pumps attain pressures up to 10^{-4} Torr. They work by taking a volume of gas at the atmospheric pressure, compressing the gas resulting in an increase in pressure and finally expelling the gas to the atmosphere. The intake region of the vacuum chamber is at the same pressure as that of the system, usually atmospheric pressure. Therefore, as the rotor moves, the inlet traps a large volume of gas at system pressure. The gas is then transferred to the next region where the movement of the rotor results in a decrease in volume and an increase in pressure. When the gas pressure exceeds that of the atmosphere, the outlet valve opens and the gas is expelled. They are mainly used as backing pumps for high vacuum pumps like turbopumps or diffusion pumps.

1.1.5 Microchannel plate detectors (MCP)

A MCP consists of an array of tiny detector tubes aligned parallel to one another. These tubes are miniature electron multiplier channels and are usually made of lead glass. Parallel electrical contact between each channel is provided by deposition of a metallic coating on both sides of the array which act as input and output electrodes. The ion beam from the mass analyzer striking the channels results in the production of an electron beam. Thus, a cascade within the MCP results in electron multiplication of $10^4 – 10^7$ with ultrahigh temporal resolution. Spatial resolution is limited by the channel dimensions and spacing.

A phosphorescent screen placed behind the array of the MCP results in the production of light flashes. These flashes of light are directed to an optical array transducer via fiber optics. The transducer then converts the optical signaling to electrical signaling for processing.

1.2 Proteomics

1.2.1 Overview

Proteomics is the study of the expression, structure, function and interactions of proteins. It involves measuring, identifying and quantifying proteins expressed in a biological cell, tissue or other specimen.

It is quite evident in biomedical research that biological/physiological processes are determined by a large set of proteins rather than a single protein. These days chromatography in conjunction with mass spectrometry plays a significant role in the field of proteomics.

In general, the first step is to isolate the proteins from the rest of the biological matrix including lipids, DNA and other non-protein material. Conventionally, for proteomics, two-dimensional gel electrophoresis (2-DGE) in combination with mass spectrometry has been used most extensively. 2-DGE separates the proteins based on their iso-electric point and molecular mass [5]. Although this technique and its variants are extremely useful, they impose some drawbacks. One of them is the limited number of proteins that can be observed on the 2D gels. The expected number of proteins in the human proteome is much higher than the protein spots observed on the 2D gels. Also, the dynamic range offered by the 2D gels is limited by the ability to mainly analyze high M.W. proteins (>30 kDa). Lack of reproducibility is another disadvantage in addition to time consuming procedures.

Currently, the above mentioned conventional techniques are being replaced by capillary liquid chromatography combined with tandem mass spectrometry (LC-MS-MS). It offers speed, sensitivity and can interrogate thousands of different proteins.

For a highly complex sample like serum, a single dimensional separation using reversed phase chromatography might not be sufficient. As a solution, the mixture can be separated using multi-dimensional chromatography including several combinations of size exclusion, reversed phase and ion exchange chromatography. For example, peptides can first be separated based on their charge using an ion exchange column followed by separation based on their hydrophobicities using a reversed-phase HPLC column. This technique is known as multidimensional protein-identification technology (MudPIT) [6]. Such approaches offer very high separation capacities and high throughput analysis allowing for the analysis of complex mixtures.

1.2.2 Potential applications

1. Protein profiling

MS technologies have contributed significantly in the generation of comprehensive proteome maps for a number of organisms. Example, proteomic maps of microorganisms such as yeast or the bacterium Deinococcus radiodurans are almost near completion [6-7]. A P. falciparum proteomic project involved analysis of both the mosquito and human stages of the parasite and 2400 identified proteins were reported [8]

2. Studying protein-protein interactions

For these kinds of studies, a protein itself can be coupled to a stationary phase and be used as an affinity reagent to isolate its own binding partners. Many protein interactions are known to occur under specific cellular environment conditions. Therefore, MS based studies are likely to identify only a part of the interactions that actually occur.

3. Post translational modifications (PTM's)

Post translational modifications (PTM's) are the processing events that change the properties of a protein by proteolytic cleavage or by covalent addition of a modifying group to one or more amino acids [9]. Numerous protein activities are altered by PTM's. Therefore, MS based techniques are applied to studying PTM's. Electron transfer dissociation (ETD) or electron capture dissociation (ECD) are the tandem mass spectrometry techniques most commonly used for characterization of PTM's. These techniques work by mainly fragmenting the peptide backbone and leaving the modifications intact. The combination of ECD with high resolution FTICR has been applied to the analysis of modified intact proteins [10-11].

4. Biomarker discovery

The presence or absence of a particular protein in a specific disease may provide a potential biomarker for that disease. The potential for using shotgun or a global proteomic approached for biomarker discovery was proposed by a number of groups. A group at Bristol-Myers Squibb analyzed urine samples between healthy controls and cases suffering from inflamed pilonidal abcess [12]. Their study resulted in the detection and identification of many more protein biomarkers using LC-MS/MS based approaches as compared to 2D GE methods. Other advantages of LC/MS over 2DGE are the high speed of the experiments and requirement of minute sample amounts.

1.3 Peptidomics

1.3.1 Significance

There are a number of reasons why enrichment of low molecular weight proteins or peptides from biological samples is done for biomarker discovery. Peptides are known to play a role in signal transduction pathways for a number of diseases including convulsive disorders, liver diseases,

metabolic disorders, etc. [13-15]. Also, large proteins are difficult to handle by mass spectrometry and might not all be soluble under the same conditions. Most importantly, if the ultimate goal is to identify the parent protein from the sequence, the mass spectrometer is more efficient at predicting sequence up to 20 residues rather than whole proteins.

1.3.2 Methods for enrichment of peptides

The concentration of total protein in the serum is 60-80 mg/mL with a dynamic range of at least 10 orders of magnitude. The high abundance, high molecular weight proteins including albumins, immunoglobulins and transferrin account for >90% of serum proteins. The higher abundance of these proteins results in the ionization suppression of virtually all low abundance, low M.W. peptides (and small proteins as well) in mass spectrometers making their detection difficult or impossible [16]. Therefore, it's important to deplete the high abundance and generally higher M.W. proteins to enhance the signal from low M.W significant peptides.

A number of techniques for enrichment of low M.W. proteins have been used in the past. Dye ligands and protein A/G columns were tested for their performance to effectively deplete the high abundance proteins [17]. Molecular weight cut off filters have also been used in the past to remove high M.W proteins using ultracentrifugation. Albumin is known to act as a carrier for transporting a number of low M.W. proteins which can serve as potential biomarkers for diseases [18]. A major disadvantage of the above mentioned methods is their inability to release the low M.W. peptides and proteins from carrier proteins such as albumin resulting in their removal as well.

We used organic solvent (acetonitrile) to precipitate and remove high M.W., abundant proteins from the serum. This method resulted in disruption of the interactions between the small peptides

and their carrier proteins resulting in large number of additional peaks in the mass spectra. This provides an opportunity for their analysis to discover potential biomarkers[19].

1.3.3 Mass spectrometry of peptides

The protein depleted specimens are usually passed through a micro capillary column which is coupled online to a mass spectrometer. The peptides are separated on the column using reversed phase chromatography. It involves the use of a gradient elution method involving increasing organic solvent concentration to elute the peptides in the increasing order of their hydrophobicities. The peptides are distinguished by their masses and the signal intensity is directly proportional to the abundance.

1.3.4 Collision induced dissociation (CID)

A technique called tandem MS is used to obtain the primary structure (sequence) of peptides. Collision-induced dissociation (CID) is a soft fragmentation technique. It involves the isolation of the particular species of interest which is collided with inert gases such as nitrogen, argon etc. resulting in the breaking of the peptide at the peptide bond, i.e. between two amino acids. The species that is fragmented is called the precursor ion and the ions produced after the fragmentation are called the product ions. In CID, after the collisions, part of the kinetic energy is converted into internal energy of the ions which finally undergo fragmentation.

The peptide fragmentation pattern produced during CID can be explained by the mobile proton model [20]. According to this model, the fragmentation pathways are charge directed. It proposes that the proton is required at the cleavage site for the fragmentation to occur. Therefore, internal

energy imparted to the peptide ion results in the migration of the proton from the more basic sites (N terminal or basic amino acids) to the peptide backbone leading to cleavage of the peptide bonds.

If a peptide contains no basic amino acids, the only site for protonation is the N terminus which results in singly charged ions. This proton can migrate across the peptide backbone directing the fragmentation reactions at the peptide bonds.

In contrast the presence of a basic amino acid residue in a peptide leads to production of doubly charged species. One proton is attached to the basic amino acid and the other to the N terminus. In this case, the proton attached to the basic amino acid is fixed, while the proton at the N terminus is the mobile proton which can migrate through the peptide backbone as shown in Figure 1.4.

1.3.5 Peptide sequencing

The most informative fragmentation pattern is obtained by cleavage of the peptide bond between the two amino acid residues. If the charge is retained on the N terminal, the resulting ion is called a b ion and a y ion is produced when the charge is retained on the carboxy terminal. For even more in-depth structural information, the product ions can be further fragmented, a method called MS^n. The fragment ions produced, depending on the cleavage site, are shown in Figure 1.5[21].

The tandem MS spectra of peptides is searched using one of the search algorithms explained below:

A) Using databases

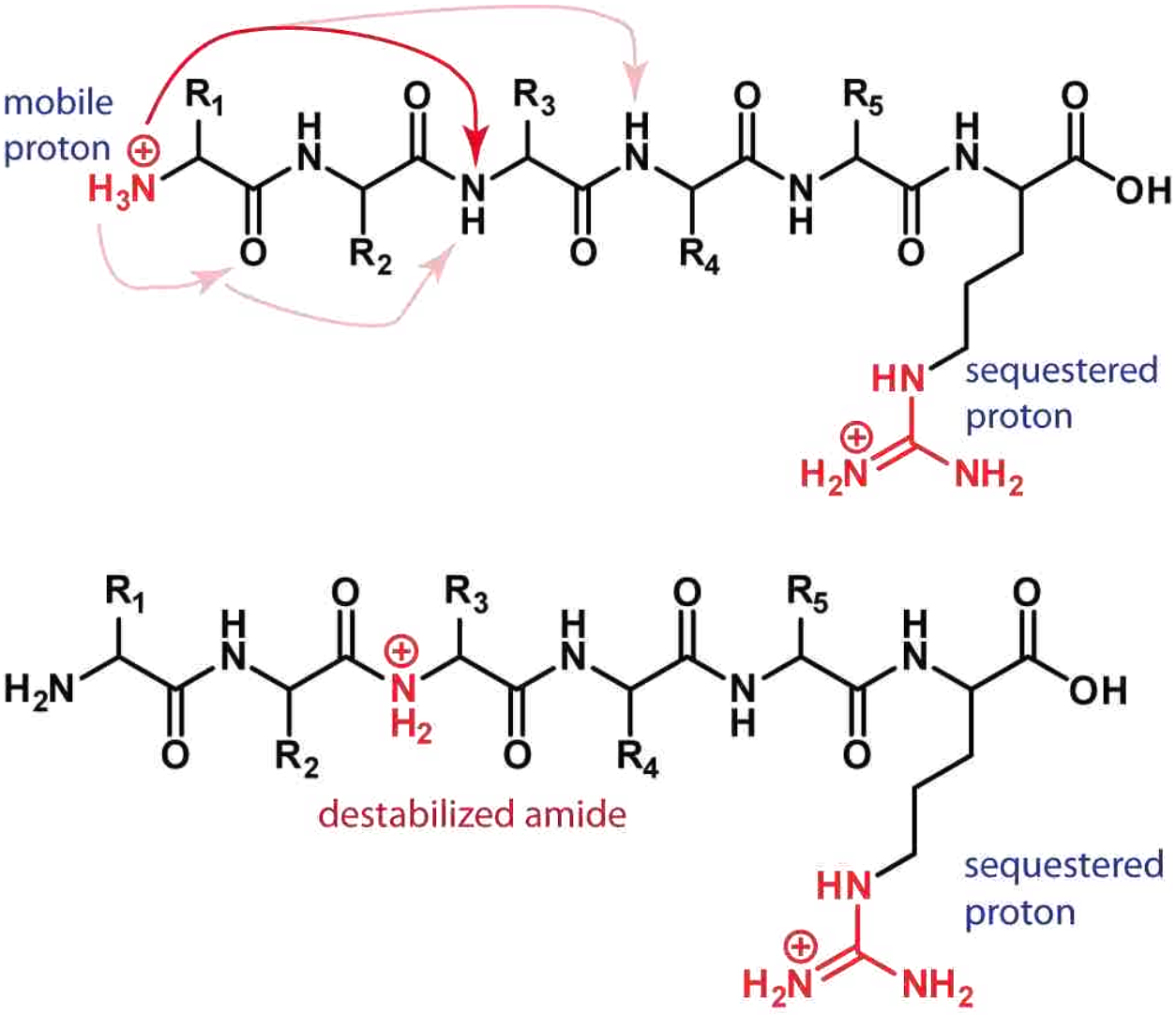


Figure 1.4 Mobile proton hypothesis for doubly charged peptides
(taken from Boyd, Robert, and Árpád Somogyi. "The mobile proton hypothesis in fragmentation of protonated peptides: a perspective." *Journal of the American Society for Mass Spectrometry* 21.8 (2010): 1275-1278.)

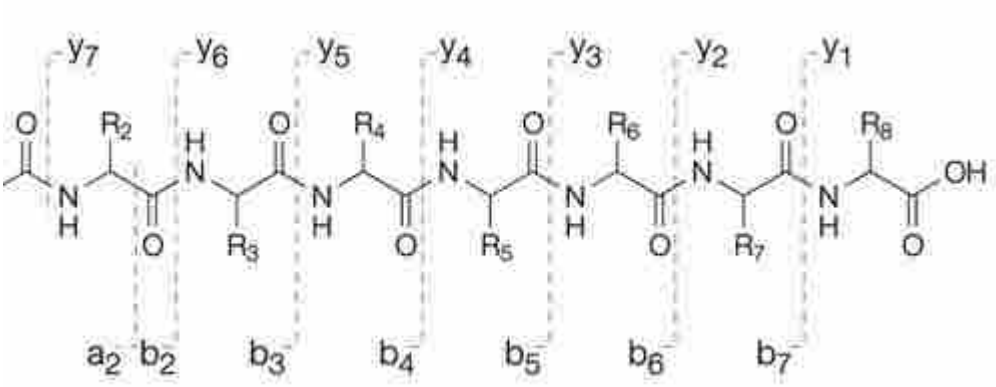


Figure 1.5 Chemical structure of a peptide with b and y fragments labelled (taken from Steen, H.; Mann, M., The ABC's (and XYZ's) of peptide sequencing. *Nature reviews Molecular cell biology* 2004, *5* (9), 699-711).

1) Programs that use a partial amino acid sequence to detect the parent proteins- BLAST, FASTA

These methods match the sequence of the query with the available database and give a probability score based on the matching and non-matching amino acid in different regions. These methods are also used to detect homologous protein sequences.

2) Programs that use peptide molecular weights for the search query- Peptide search, MS-Fit and MOWSE

These are also referred to as mass mapping experiments. The proteins are digested with a specific protease and the M.W. of peptides produced is measured accurately. These are then matched with databases to find out which proteins if digested with same protease would produce one of several of the same M.W. peptides.

3) Programs that use uninterpreted product ion spectra for the search query- SEQUEST, MASCOT

These methods compare the experimental MS/MS spectrum with theoretical spectrum that has been derived from every sequence in the database or from theoretically predicted fragments. A score is given based on the extent of overlap. A best score is given for maximum overlap and scores next to the score for the best matching peptide sequence are also given.

A) *De novo* peptide sequencing

When the computer algorithms are unsuccessful in interpreting a correct peptide sequence, another approach called de novo sequencing can be undertaken. It involves the manual interpretation of the peptide sequence based on the masses of the amino acid residues and m/z's for the b and y ions.

Ultimate confirmation of a peptide sequence requires the combination of de novo sequencing and databases.

1.3.6 Quantitative proteomics

Different proteins have different ionization efficiencies depending upon their structures. Due to this, quantification of proteins is difficult because one cannot infer the amount of a particular protein present in a sample based on the abundance of the ionized species.

The most common way to quantify a protein is to add a known amount of a stable isotope synthetic analogue that is identical otherwise to the compound to be measured but having a heavy isotope containing amino acid substituted for typically one of the naturally occurring amino acids. Because they have exactly the same ionization efficiencies, the concentration of the protein of interest can be determined by adding or using a known amount of the synthetic analogue and referencing to it. Although this method is useful, it is not possible to add a stable isotope for a number of different proteins present in the sample. Therefore, some other approaches might prove to be more useful.

It's mainly done by taking the aliquots of the same sample and isotopically labelling all the proteins from one sample while not labelling the proteins from the other identical sample. Thus the relative amounts of all the proteins can be measured by comparing the intensities of the labelled and the unlabeled peaks for each of the proteins.

There are two methods commonly used for isotopically labelling the proteins. One involves growing the cells in a medium which would incorporate and isotopically label all the proteins in a

cell. Chemical derivatization of all the proteins in a sample using different isotope tags for each sample is another method that is commonly used method.

A number of compounds have been reported for tagging the proteins in a complex mixture. One method is called isotope coded affinity tag (ICAT) and involves a chemical reduction step followed by alkylation to incorporate an isotope tag on cysteine residues in a protein [22-23]. Another tagging system modifies the lysine residues with O-methylisourea to form homoarginine [24].

1.4 Serum: an important specimen for biomarker studies

Useful biomarkers require or involve certain phases including discovery, confirmation and clinical validation which requires the availability of a large number of organized patient samples. Blood specimens including serum and plasma are one of the most easily accessible and commonly used samples for biomarker discovery, while biopsy to obtain tissues is usually less well tolerated. Serum samples are known to be preferred over plasma specimens [25]. Due to the ease of collection, serum samples are often collected and stored in biobanks allowing for research studies. The most important aspect of discovering and validating a biomarker is the quality of specimens used. The factors influencing the quality of the specimen include the conditions employed during collection, minimizing time at each step, the care and precautions used in processing and how specimens are stored. Therefore, blood should be collected without the use of a tourniquet, from the arm opposite any IV lines, allowed to clot at room temperature for no less than 30 minutes and no more than 60 minutes, the clot should be immediately spun out at 4^{o}C and the serum should be immediately aliquoted to freezer-type tightly sealing tubes, immediately snap frozen and stored without thawing

in a freezer until the time of analysis. It is recommended that for short term storage (1-2 weeks), the samples should be stored at -20^{0}C, whereas for long term storage, -80^{0}C should be used.

It has been proposed that the proteome or peptidome of serum may indicate biological events. For example, proteolytic events of the ex vivo coagulation may confirm cancer specific differences [26].

1.5 Lipidomics

1.5.1 Significance

Lipidomics is the comprehensive analysis of lipids in biological systems. Lipids are essential cellular components that play important physiological roles: 1) they form membrane bilayers providing spatial integrity and separating intracellular components from each other and the extracellular environment. 2) They provide a hydrophobic environment facilitating interactions between lipids and membrane associated proteins. 3) They act as a source of messengers for cell signaling produced by action of intracellular enzymes. 4) They act as reservoirs for energy storage.

Global lipidomics using mass spectrometers was first described by Han and Gross in 2003. They introduced the intrasource separation of lipid species depending upon their ionization efficiencies directly from biological samples using ESI [27]. The approach was termed “Shotgun lipidomics”.

Currently, lipidomics is mainly focused on the following areas: structural characterization of known and unknown lipids from diverse classes, development of methods for the quantitation of lipids at amounts as low as attomole to femtomole quantities and studying lipidomic alterations in diseased states leading to biomarker discovery, aiding in disease diagnosis and progression and also providing insights into the mechanism of diseases.

Studying alterations in the cellular lipidome and quantifying the altered lipids has helped provide new insights into the mechanism of some diseases [28-29]. Specific and important roles for lipids from diverse classes in cellular signaling have also been reported. For example, plasmenylethanolamine (PlsEtn), a subclass of phosphatidylethanolamine (a glycerophospholipid) acts as a reservoir of arachidonic acid, facilitating membrane fusion, preventing its oxidation [30]. Specific triacyglycerols (TAG's) in excess can trigger atherosclerosis and myocardial dysfunction[31].

1.5.2 Classification of lipids

A) Non-Polar lipids

These mainly include cholesterol, cholesterol esters and TAG's. The dominating hydrophobic region makes them highly hydrophobic. The basic structures are shown in Figure 1.6 [32].

B) Polar lipids

1. Phospholipids- These are the most abundant class of polar lipids. They have at least one phosphate group present at the sn-3 position of the glycerol backbone. The subclass of phospholipids vary depending on the head group, namely phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylglycerol (PtdGro), phosphatidylinositols (PI), phosphatidylserines (PS) and diphosphatidylglycerols (cardiolipins). Their structures are shown in Figure 1.7 [32]. The most abundant phospholipids are PC and PE accounting for 75% of total phospholipid mass in eukaryotic cell membranes. The complexity of phospholipids is not just limited to the number of different polar head groups but also depends on the carbon chain length of the fatty acyl components and their number of double bonds.

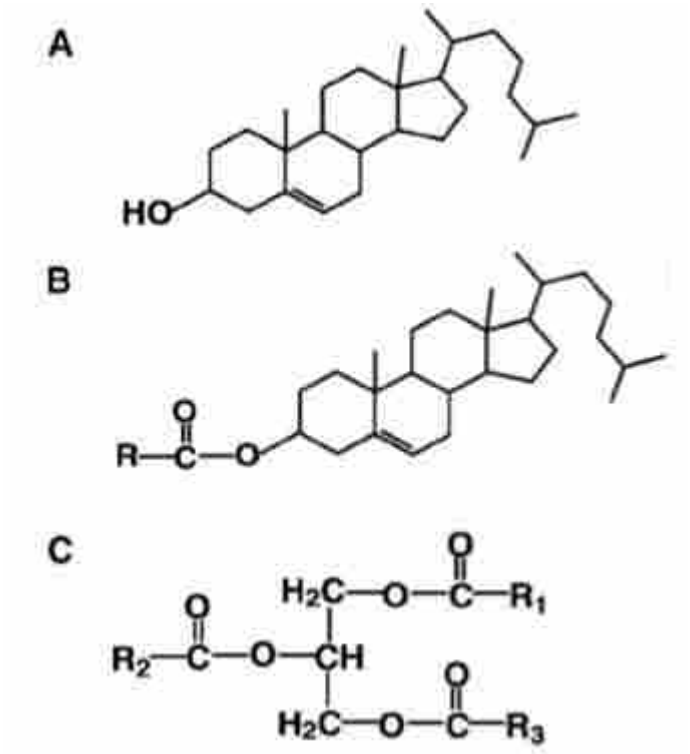


Figure 1.6 general structures for cholesterol (A), cholesterol ester (B) and triacylglycerol (TAG) (C). R, R1, R2 and R3 represent aliphatic chains
(taken from Han, X.; Gross, R. W., Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass spectrometry reviews* 2005, *24* (3), 367-412).

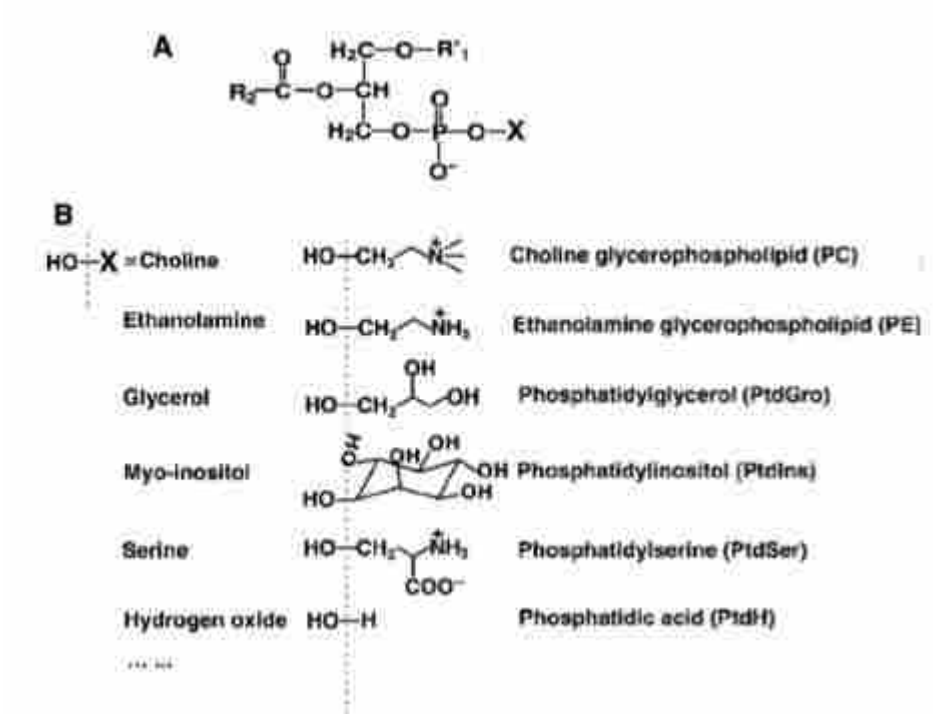


Figure 1.7 General structures for different classes of phospholipids
(taken from Han, X.; Gross, R. W., Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. ***Mass spectrometry reviews*** **2005,** ***24*** **(3), 367-412).**

2. Sphingolipids-These contain a sphingosine backbone or an analog and represent approximately 5-10% of total lipids present in most brain cells and 30% of total lipids in human white matter [33]. Depending on the polar head group attached to the ceramide, sphingolipids can be classified into sphingomyelin (SM), cerebrosides, sulfatides, glucosylceramides, lactosylceramides and other glycosphingolipids having multiple sugar rings as shown in Figure 1.8 [32].

3. Glycolipids-This class includes primarily glycosphingolipids and glycoglycerolipids. The hydrophobic core of the glycosphingolipids is a ceramide backbone, while that of glycoglycerolipids is a diacylglycerol (DAG) backbone.

C) Metabolites

The lipid metabolites are produced from the parent lipid species by enzymatic action. These include acylcarnitines, nonesterified fatty acids (NEFA), fatty acid esters, lysolipids, eicosanoids etc. A lot of lipid metabolites are known to act as second messengers biologically.

1.5.3 Classification of lipids based on their electrical dispositions

The first category is of the anionic lipids which carry one or more negative charges at physiological pH. This group includes PI, PS, PtdGro, cardiolipins, acylcoA, sulfatides, and would carry a negative charge only if exposed to alkaline pH.

Consequently, their ionization efficiencies are lower than those of the anionic lipids. The lipids in this category are PE, lysoPE, ceramides etc. The third category of lipids is the electrically neutral lipids at alkaline or physiological pH. The lipid classes include PC, lysoPC, DAG, TAG, etc.

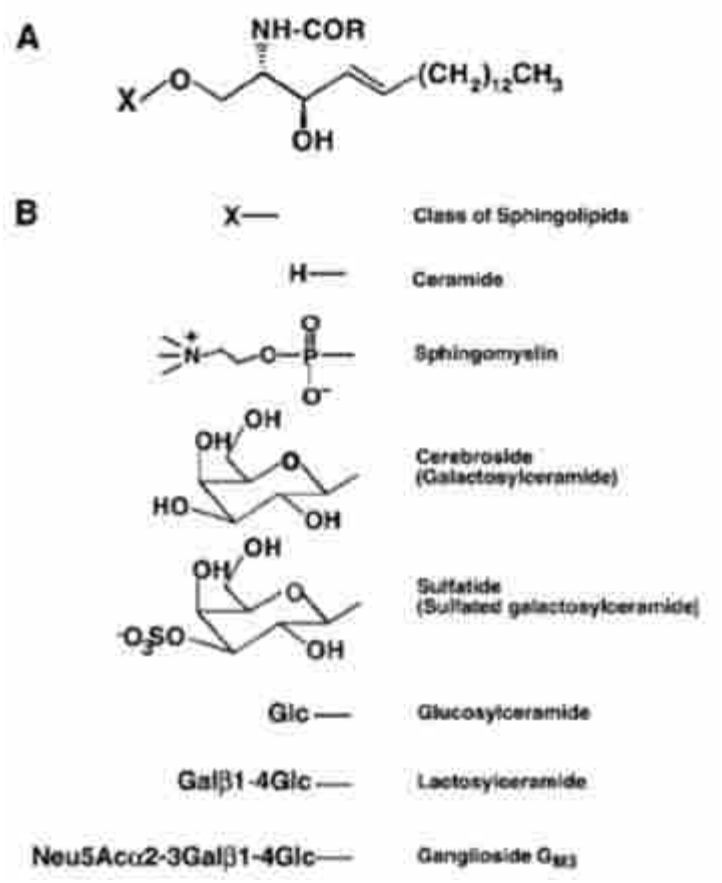


Figure 1.8 General structure and classes of sphingolipids

(taken from Han, X.; Gross, R. W., Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass spectrometry reviews* 2005, *24* (3), 367-412).

1.5.4 Separation of lipids in the positive or negative ion mode through ESI

The ionization efficiency of a lipid depends on its tendency to lose or gain (oxidized or reduced) charge in the electric field. The lipid species carrying an inherent charge ionize more readily than the ones that do not carry an inherent charge. Ionization can also be achieved in the ESI source through the formation of adduct ions using cations or anions. Under the presence of an electric field, a dipole moment sufficient enough could be induced to make the lipid species interact with either an adduct cation or anion to form lipid-adduct ions.

The ionization efficiencies depend upon the degree of the dipole potential induced in the molecular species. For example, a PC contains a large dipole moment in the zwitterionic polar head group. Therefore, it rapidly ionizes to form $M+H^+$ or $M+Na^+$ in the positive ion mode. In the negative ion mode, a PC can form $M+Cl^-$ or $M+OAc^-$. Because head group of a PC contributes in a major way to the dipole moment, the ionization efficiencies of different PC species varying in carbon chain length are nearly identical. On the other hand, TAG's are nonpolar lipids forming ammoniated, sodiated or lithiated adducts. A polar head group is absent and the dipole moment is influenced by the carbon chain length. Therefore, the ionization efficiencies of different TAG species are different depending on the number of carbon atoms and double bonds.

For a lipid extract obtained from a crude biological sample, all the anionic lipid species would be observed in the negative ion mode. After addition of a base to make the extract slightly basic, all the weakly anionic lipids would also be observed in the negative ion mode. All the neutral polar lipid species would show up in the positive ion mode. Setup for intrasource separation of lipids through ESI is shown in Figure 1.9 [32].

1.5.5 Tandem mass spectrometric based lipidomics

The four MS/MS modes used in lipidomics are:

1. Product ion analysis mode

In this mode, the precursor ion of interest is first selected by the first analyzer which acts as a filter to transmit only the ion of interest. The ion of interest is then fragmented in the collision cell to form a number of product ions which are analyzed by the second mass analyzer. The structure of the precursor ion may then be elucidated from the fragment ions.

2. Precursor ion scanning mode

In this mode, the second mass analyzer focusses on the fragment ion of interest and scans different m/z from the first mass analyzer. All the m/z's which produce that particular fragment ion are selected. It is useful to detect different species belonging to the same lipid class.

3. Neutral loss scanning mode

In the neutral loss scanning mode, both the first and second analyzers are scanned simultaneously to find a particular offset mass between the two. For some lipid species, the product ion produced by the fragmentation of the precursor ion corresponds to a neutral loss of a particular mass and likely a particular structural component. All the lipid species belonging to a class producing a particular neutral loss fragment can be identified using this technique.

4. Selected reaction monitoring mode (SRM)

In this mode, the transitions between the precursor and the product ions are previously known. Both the first and second analyzers are focused on selected ions with m/z=M and m/z=p, where p represents the product ion of parent or precursor ion M.

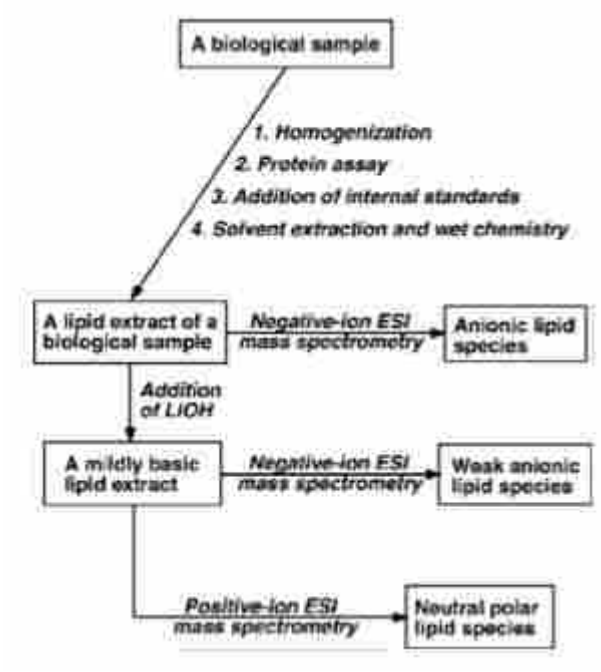


Figure 1.9 Intrasource separation of lipid species through ESI directly from the crude extract. (taken from Han, X.; Gross, R. W., Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass spectrometry reviews* 2005, *24* (3), 367-412).

When either one or both the analyzers are set to record multiple ions, it's called multiple reaction monitoring (MRM). For SRM/MRM mode, the mass spectrometer is coupled with HPLC. These techniques provide high sensitivity and specificity because of the high duty cycle of transitions of interest and are commonly used for the quantitative analysis of a lipid species of interest.

These different modes are shown in Figure 1.10 [34].

1.5.6 Mass spectrometric analyses of the lipid species

1. Analysis of anionic lipid molecular species

Those species possessing a negative charge inherently can be analyzed in the negative ion mode at neutral pH. For example, abundant deprotonated peaks for anionic lipids including PI, PtdGro, PS and cardiolipins were observed in the ESI mass spectrum of a crude lipid extract from mouse myocardium, whereas it displayed low abundance pseudo-molecular ions for otherlipid classes like PC and PE [32]. In the negative ion mode, the molecular ion species of PI yielded a fragment ion at m/z 241. Therefore, all the species corresponding to PI can be identified by precursor ion scanning or product ion analysis. Similarly, lipids from the PS class can be identified by neutral loss scanning of 87 amu. Also, a product ion at m/z 153 corresponding glycerophosphate are produced from deprotonated anionic phospholipid species including PtdGro, cardiolipins etc.

2. Analysis of weakly anionic lipid molecular species

Addition of a weak base in low concentrations will result in deprotonation of weakly anionic lipid species including PE, lysoPE and eicosanoids, therefore displaying abundant deprotonated peaks.

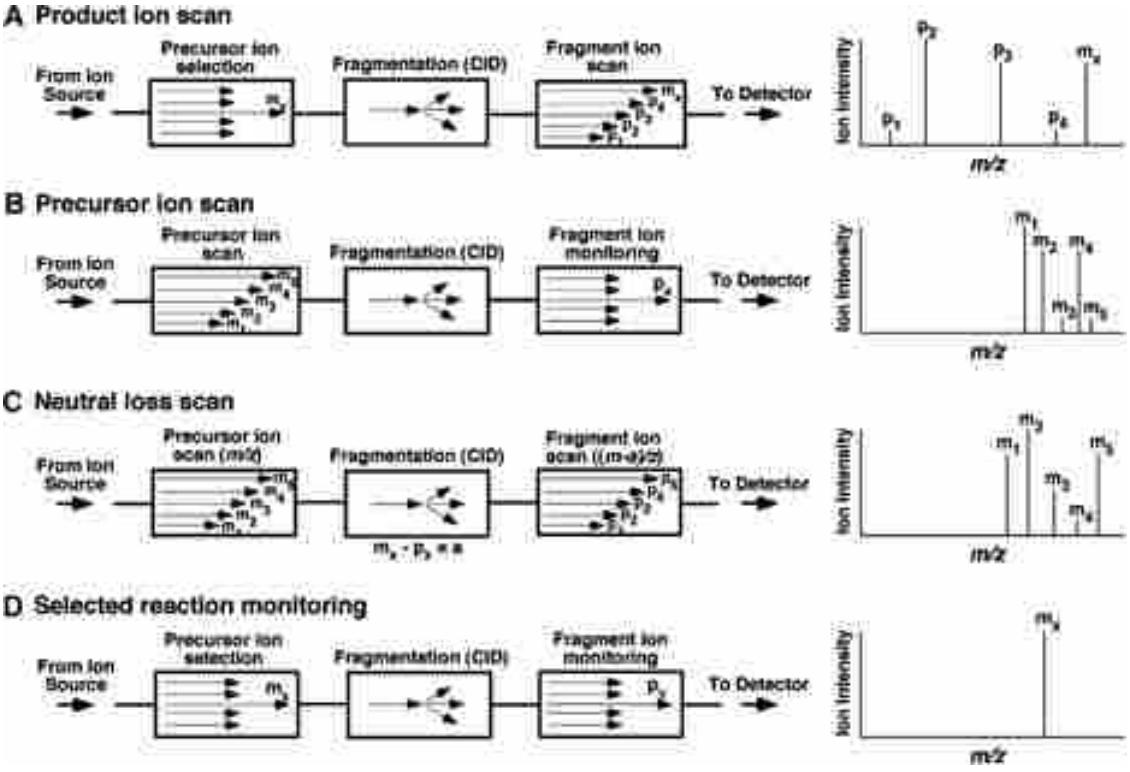


Figure 1.10 Tandem mass spectrometry modes for analyses of lipid species (taken from Multi-dimensional mass spectrometry-based shotgun lipidomics and novel strategies for lipidomic analyses. *Mass spectrometry reviews)*

The combinations of acyl chains present in PE can be identified by looking for the individual fragment ions. For example, m/z 255.2 is for 16:0, m/z 279.2 is for 18:2, m/z 281.2 for 18:1 and m/z 303.3 for 20:4. Although the anionic lipids ionize much more readily as compared to weakly anionic lipids, the ion suppression of PE species is not affected by anionic species. This is mainly because of the low abundance of anionic lipids relative to PE molecular species resulting in minimum interference during the ionization process.

3. Analysis of electrically neutral lipid species

These species are mainly detected in the positive ion mode. The highly polar lipid species including PC, lysoPC and SM are easily ionized with protons, whereas relatively less polar lipids e.g. TAG require ammonium, sodium or lithium ions to form adducts. The fragmentation spectrum of choline containing lipid species like PC, lysoPC and SM display product ions with m/z 59, 104 and 184 corresponding to trimethylamine and choline and phosphocholine moieties. The parent SM species can be distinguished from PC species by the nitrogen rule. The nitrogen rule states that organic compounds containing exclusively H, C, N, O, Si, P, S and the halogens either have odd nominal mass that indicates an odd number of nitrogen atoms or even an even nominal mass that indicates an even number of nitrogen.

The combinations of acyl chains can also be determined by analyzing the neutral loss of fatty acids from the parent in the product ion spectrum.

TAG's contain three fatty acid chains attached to the glycerol backbone. Although non-polar, abundant peaks are observed for TAG adducts (sodium, lithium or ammonium) in MS1 spectrum. Because of the absence of a head group, the ionization efficiencies for TAG's mainly depend on the acyl chain length and the number of double bonds.

1.5.7 Approaches for lipidomic analysis

The first approach used for lipidomics is an LC-MS based approach where separation of lipids takes place on a column prior to ESI-MS. The other major approach involves direct infusion of a lipid extract into the mass spectrometer through ESI. This approach is termed as shotgun lipidomics.

A) LC/MS based approaches:

The combination of ESI-MS with LC separation provides high sensitivity and resolution. Column separation also facilitates the analysis of isobaric compounds with different hydrophobicities. The LC-MS based approaches includes SIE (selective ion extraction), SRM/MRM and data-dependent analysis. In SIE, selected ions are extracted from the chromatogram after the LC run. The ion peak area of all the species of interest can be compared either to a standard curve of the same molecular species or ion peak area of an internal standard under identical experimental conditions. This approach is useful for quantitation of targeted low abundant species whose standard curve could be generated. However, this approach cannot be used for large scale quantification of lipids. The interference caused by coexisting lipid species in the same peak having similar hydrophobicities is a drawback of this approach. In contrast to SIE, SRM/MRM is preferable because the quantification is dependent on the fragment ion intensities specific to the parent ion of interest.

A large number of studies have been carried out in the past to quantify lipid species using LC/MS. For example, quantification of ceramides from human stratum corneum has been successfully performed using normal phase LC/MS [35]. In a study LC/MS/MS facilitated the quantification of > 50 fatty acyl amino acids (35) [36].

B) Shotgun lipidomics

Direct infusion of a lipid extract into ESI-MS was first described by Han and Gross in 1994. This approach is widely termed as shotgun lipidomics. It facilitates high throughput analysis of a lipidome directly from the lipid extract of complex biological samples. An advantage of the direct infusions over LC/MS is this high throughput sample analysis requires less time. The methods most commonly used for shotgun lipidomics are:

1. Tandem MS based shotgun lipidomics

As described earlier, most of the lipid species belonging to a particular class produce a characteristic fragment ion associated with the head group specific to that class. Therefore, all the lipid species belonging to a particular class can be isolated, measured and identified using precursor ion scan or neutral loss scan. This offers the advantages of simplicity, efficiency, high sensitivity, ease of management and faster analysis. The filtering process of MS/MS results in a high S/N ratio.

There are several limitations associated with this approach. Firstly, the aliphatic constituents associated with different lipid species belonging to the same class are not identified. Secondly, the detection with MS/MS approach might not always be specific to the particular class. For example, the fragment ion at m/z 184 can be produced by species belonging to PC or SM; therefore, it can easily produce assignment artifacts. Lastly, accurate quantification of the lipid species might not be simple because of slightly different fragmentation pathways of the species belonging to the same lipid class.

Despite the described disadvantages, many laboratories have adopted this approach for lipid analysis. For example, it's widely used for plant lipidomics, detecting changes in lipids in response

to individual conditions or stresses [37-38]. The identification of PC subclasses and the individual molecular species belonging to PC have also been achieved using NLS and PIS from their lithiated molecular ions [39].

2. High mass accuracy lipidomics

Currently, the Q-TOF mass spectrometer offers an increased duty cycle improving the detection sensitivity. TOF analyzers can detect ions simultaneously, providing high mass accuracy and resolution resulting in exact mass determination of the precursor or fragment ions minimizing false positives.

1.5.8 Application of lipidomics in disease biomarker discovery

Lipids are known to play important physiological functions including acting as a source of energy, maintaining cell membrane integrity and being involved in cell signaling pathways. Therefore, it's important to consider and detect lipids involved in disease processes which may reflect disease states and indicate about the mechanism of the disease.

A number of studies have proven that dyslipidemia or abnormal lipid metabolism can lead to a number of human diseases including diabetes [40], obesity [41], atherosclerosis [42] and brain disorders [43]. Therefore, studying alterations of individual molecular species of lipids in diseased specimens can help in detection of potential lipid biomarkers indicative of that disease.

The most common workflow for detecting lipid biomarkers includes the following steps. Lipid extraction from biological samples including serum, tissue or other body fluids is carried out for a

set of cases and controls. The crude lipid extract is either first separated using LC or GC and then passed through ESI to MS or directly infused into the MS through ESI. Following that, a peak list of individual molecular species of lipids belonging to diverse classes from the mass spectrum is generated. Next, the data is normalized and subjected to various statistical tests to detect peaks which are statistically different between the cases and controls. Finally, the detected lipid biomarkers are correlated with metabolic pathways and their possible role in the mechanism of the disease is interpreted.

A number of lipidomic studies have been helpful in detecting novel lipid species that could serve as potential biomarkers for a number of diseases. For example, in a study, phospholipid profiling in plasma samples of diabetic mellitus-2 (DM-2) cases and controls was done using an LC-MS/MS approach to detect potential lipid markers. This study led to discovery of four lipid markers including two PE and two lyso-PC molecular species that discriminated DM-2 cases from controls [44]. In another study, lipid profiling of liver tissues obtained from wild type and APOE*3-Leiden transgenic mice suffering from diet induced hyperlipidemia and atherosclerosis was done using LC-MS on both cases and controls [45]. The results found higher levels of two lysoPC species in the APOE*3-Leiden transgenic mice. Also, a number of TAG species were found to be higher in these transgenic mice. Ceramides have been found to be increased in neuropathic diseases [33] and a number of TAG's were found to be elevated in cardiovascular diseases [46].

1.6 Preeclampsia

1.6.1 Background

Preeclampsia (PE) is a serious life threatening gestational disorder. It occurs in 5-8% of pregnancies worldwide and is a leading cause of fetal and maternal deaths and morbidity [47]. As established by International Society for the Study of Hypertension in Pregnancy (ISSHP), PE is a state of new onset hypertension after 20 weeks of gestational age together with proteinuria, followed by the postpartum normalization of the symptoms within 3 months. The thresholds defined for hypertension in PE are a systolic blood pressure of >140 mm Hg and a diastolic blood pressure of >90 mm Hg measured twice within an interval of at least 4 h, whereas proteinuria is defined to be the presence of >300 mg/day of urinary protein [48]. Furthermore, the ISSHP has defined PE to be severe if the blood pressure is >160 mm Hg systolic and >110 mm Hg diastolic and proteinuria of >5 g/day [49]. Despite the fact that proteinuria has been conventionally considered a requisite for diagnosis of PE, its role as a marker is debatable [50]. Also, the appearance of additional complications in PE can occur and is termed HELLP (hemolysis, elevated liver enzymes and low platelet counts) syndrome characterized by thrombocytopenia, microangiopathic hemolytic anemia and hepatic dysfunction [51]. In some cases severe PE can progress to eclampsia which involves the development of seizures antepartum, intrapartum or postpartum, probably due to cerebral vasoconstriction or vasospasm, ischaemia, cerebral edema and/or cerebral haemorrhage. Other maternal symptoms considered for clinical diagnosis of PE include renal insufficiency (creatinine >0.09 mmol/L), liver disease, neurological problems (hyperreflexia with severe headaches, visual disturbances), hematological disturbances and fetal growth restriction [49]. Fetal complications include prematurity and low birth weight. Collectively, taken together, a wide range of symptoms makes PE a syndrome rather than a disease.

Despite years of research, the etiology of the disease is not known. Indeed, PE has been known since the times of ancient Egypt around 3000 years ago, where it was first reported as a pregnancy related disorder [52]**.** However, it has been accepted that the presence of the placenta is responsible and/or necessary for the disease which is supported by the fact that the symptoms are relieved following its removal [53]**.** It is defined as the "disease of theories" as exhibited by its having no decided cause [54].

1.6.2 Pathogenesis of PE

The first step that is proposed to initiate the PE process is the inadequate remodeling of the uterine spiral arteries leading to improper perfusion of the fetal-placental unit. This further results in the placenta becoming hypoxic. The alternate periods of hypoxia and re-oxygenation within the placenta triggers oxidative stress and increases placental apoptosis and necrosis [55]. This shedding of placental debris into the maternal circulation leads to systemic endothelial dysfunction and an adverse inflammatory response [51].

1.6.3 Biomarkers for PE

Despite a number of clinical studies, there is no single reliable parameter for early diagnosis of PE. Currently, a number of biomarkers have been proposed for PE. The most commonly studied are 1. Angiogenic factors- vascular endothelial growth factor (VEGF) and placental growth factor (PlGF). 2. Anti-angiogenic factors- soluble fms like tyrokinase (sflt1) and soluble endoglin.

Other proposed candidate markers for PE are Placental Protein 13, P-Selectin and Pregnancy-Associated Plasma Protein A (PAPP-A).

Although a number of evaluations have been performed on these markers, their use is limited due to lack of sensitivity and specificity and inability to diagnose PE at an early stage [56].

Therefore, there is a need for detection of set of novel biomarkers which can differentiate patients at risk for PE from controls at a very early stage. We carried out a study using our serum proteomic approach to detect low M.W. biomarkers for PE as elaborated in Chapter 2.

We also used a lipidomic approach to detect and identify potential lipid biomarkers for PE specifically (details in Chapter 3).

1.7 Alzheimer's disease (AD)

1.7.1 Background

Alzheimer's disease (AD) is a chronic neurodegenerative disorder and is the most common form of age related dementia in modern society [57]. It is characterized by progressive loss of cognitive functions. The early signs mainly include memory loss and mild behavioral changes with gradual progression to dementia. It is estimated that approximately 25 million people worldwide currently have AD, and the number of affected people is estimated to be increased to 80 million by 2040[58].The risk factors mainly linked to the incidence of AD include age, gender (females are more likely to be affected), genetic factors, head injury and Down's syndrome.

Mild cognitive impairment (MCI) is proposed to be an early phase of cognitive decline that precedes dementia. MCI patients may progress to AD, vascular disease and other kinds of

dementia. A study showed that people with MCI were 6.7 times more likely to develop AD than cognitively normal individuals [59].

1.7.2 Clinical Dementia Rating Scale (CDR) for AD

The 5 stages of dementia used to evaluate the progression of symptoms in patients with dementia are

1. CDR 0 or no impairment: This stage represents no memory problems, proper orientation in time and space, normal judgement and the ability to take care of the all of the personal needs.

2. CDR 0.5 or very mild dementia: It represents minor memory inconsistencies but the ability to manage personal needs is not affected.

3. CDR 1 or mild dementia: Symptoms include moderate memory loss, geographic disorientation and difficulty in handling problems. Social judgement is usually maintained.

4. CDR 2 or moderate dementia: It includes severe memory loss, severe difficulty with time and space orientation, social function usually impaired and there is no pretense of independent function outside home.

5. CDR 3 or severe dementia: It is accompanied by severe memory loss, complete judgment impairment, can't solve problems and inability to take care of personal needs.

1.7.3 Pathophysiology of AD

Currently, AD is believed to be driven by two biochemical processes. The first involves the extracellular deposition of beta amyloid (Aβ) which is a 36-43 amino acid peptide cleaved from the parent protein amyloid precursor protein (APP) by action of enzymes β and γ secretase. Further, the Aβ monomer polymerizes to soluble oligomers and then to insoluble fragments such as Aβ 42 which finally precipitates to amyloid fibrils resulting in deposition.

The second biochemical process associated with AD is the formation of neurofibrillary tangles due to abnormal processing and accumulation of tau protein. Tau protein forms aggregates by twisting filaments around each other. These aggregates then interfere with normal cellular processes by displacing organelles and impairing axonal transport.

1.7.4 Current Neuroimaging techniques

Disease pathogenesis may affect the brains of patients well before the actual clinical symptoms of AD occur. Therefore, early diagnosis may help in the prevention of disease progression and also benefit development of effective treatment.

Neuroimaging techniques used for diagnosing AD are described below.

1) Structural neuroimaging

Magnetic resonance imaging (MRI) allows measurement of hippocampal volume, which is a known established AD marker [60]. Multiple studies have confirmed hippocampal neuron loss at post-mortem in AD patients[61-62].

2) Functional neuroimaging

Decrease in regional blood flow and glucose metabolism preceding cerebral atrophy is known to occur in AD patients [63]. Functional imaging with fluorodeoxyglucose positron emission tomography (FDG-PET) shows characteristic glucose hypometabolism in temporoparietal areas of AD brain due to synaptic loss [64]. Multiple studies have distinguished established AD from other healthy controls with 90% sensitivity and specificity using FDG-PET [65].

PET can also detect binding of labelled molecules to Aβ indicating Aβ plaques in AD brains [66]. Although these imaging techniques can improve diagnosis, their use is limited to small scale studies. This is because of expensive instrumentation, their invasive nature and laborious methodologies.

1.7.5 Current proposed biomarkers for AD

A number of candidate markers proposed for AD in CSF, serum, plasma and urine are shown in Table 1.1 [67].

Despite a number of proposed candidate biomarkers for AD, the definite diagnosis can only be made by post mortem histopathological examination of brain tissue having amyloid plaques containing a core of Aβ peptide and neurofibrillary tangles of tau protein.

Therefore, there is a need for detection of panels of novel biomarkers that could aid in the early diagnosis of AD and help in the development of treatment therapies.

1.7.6 Lipid biomarkers for AD

Lipids could be directly involved in the biochemical pathway of AD or could be produced as a result of the disease. A number of studies have shown alteration of lipids in AD mentioned in Chapter 4. Therefore, lipids could prove useful as potential markers for diagnosing AD at an early stage. We carried out a study to characterize the alterations in lipid species from different classes between AD cases and controls using direct infusion ESI mass spectrometry.

The methods and results are elaborated in detail in Chapter 4.

Table 1.1 Select candidate fluid and imaging biomarkers of AD
(taken from Cell-based compositions and methods for treating conditions of the nervous system. Google Patents: 2008)

Biomarker	Biological Sample Type	Correlation to Dementia Risk	Reference
tau protein	cerebrospinal fluid (CSF)	increased	Hampel *et al.* (2004), *Mol Psychiatry*, 9:705-710
phospho-tau protein	CSF	increased	Hampel *et al.* (2004), *Arch Gen Psychiatry*, 61:95-102 Hansson *et al.* (2006), *Lancet Neurol*, 5(3):228-234
β-amyloid$_{1-42}$ peptide	CSF	decreased	Hampel *et al.* (2004), *Mol Psychiatry*, 9:705-710
Ratio of β-amyloid$_{1-42}$ peptide to β-amyloid$_{1-40}$ peptide	plasma CSF	decreased decreased	Graff-Radford *et al.* (2007), *Arch Neurol*, 64(3):354-362; Hansson *et al.* (2007), *Dement Geriatr Cogn Disord*, 23(5):316-20
C1q protein	CSF	decreased	Smyth *et al.* (1994), *Neurobiol Aging*, 15(5):609-614
IL-6 protein	plasma CSF	increased increased	Licastro *et al.* (2000), *J Neuroimmunol*, 103:97-102; Sun *et al.* (2003), *Dement Geriatr Cogn Disord*, 16(3):136-44
APOE protein	CSF	increased	Fukuyama *et al.* (2000), *Eur Neurol*, 43(3):161-169
α-1-antichymotrypsin protein	plasma	increased	Dik *et al.* (2005), *Neurology*, 64(8):1371-1377.
oxysterol	CSF	increased	Papassotiropoulos *et al.* (2002), *J Psychiatr Res*, 36(1):27-32
isoprostane	CSF	increased	Montine *et al.* (2005), *Antioxid Redox Signal*, 7(1-2):269-275
3-nitrotyrosine	CSF	increased	Tohgi *et al.* (1999), *Neurosci Lett*, 269(1):52-54
homocysteine	plasma	increased	Seshadri *et al.* (2002), *N Engl J Med*, 346(7):476-83
cholesterol	plasma	increased	Panza *et al.* (2006), *Neurobiol Aging*, 27(7):933-940

1.8 References

1. Elin, R. J., Instrumentation in clinical chemistry. *Science* **1980,** *210* (4467), 286-289.

2. Chace, D. H., Mass spectrometry in the clinical laboratory. *Chem. Rev.* **2001,** *101* (2), 445-478.

3. Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A., An introduction to quadrupole–time-of-flight mass spectrometry. *J. Mass Spectrom.* **2001,** *36* (8), 849-865.

4. Yamashita, M.; Fenn, J. B., Electrospray ion source. Another variation on the free-jet theme. *The J. Phys.Chem.* **1984,** *88* (20), 4451-4459.

5. Romijn, E. P.; Krijgsveld, J.; Heck, A. J., Recent liquid chromatographic–(tandem) mass spectrometric applications in proteomics. *J. Chromatogr. A* **2003,** *1000* (1), 589-608.

6. Washburn, M. P.; Wolters, D.; Yates, J. R., Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001,** *19* (3), 242-247.

7. Lipton, M. S.; Paša-Tolić, L.; Anderson, G. A.; Anderson, D. J.; Auberry, D. L.; Battista, J. R.; Daly, M. J.; Fredrickson, J.; Hixson, K. K.; Kostandarithes, H., Global analysis of the Deinococcus radiodurans proteome by using accurate mass tags. *Proc. Natl. Acad.Sci.* **2002,** *99* (17), 11049-11054.

8. Florens, L.; Washburn, M. P.; Raine, J. D.; Anthony, R. M.; Grainger, M.; Haynes, J. D.; Moch, J. K.; Muster, N.; Sacci, J. B.; Tabb, D. L., A proteomic view of the Plasmodium falciparum life cycle. *Nature* **2002,** *419* (6906), 520-526.

9. Sokolowska, I.; Wetie, A. G. N.; Woods, A. G.; Darie, C. C., Applications of mass spectrometry in proteomics. *Aust. J. Chem.* **2013,** *66* (7), 721-733.

10. Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W., Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue. *Proc. Natl. Acad.Sci.* **2002,** *99* (4), 1774-1779.

11. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sørensen, E. S.; Zubarev, R. A., Complete characterization of posttranslational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage. *Anal. Chem.* **2003,** *75* (10), 2355-2361.

12. Pang, J. X.; Ginanni, N.; Dongre, A. R.; Hefta, S. A.; Opiteck, G. J., Biomarker discovery in urine by proteomics. *J. Proteome Res.* **2002,** *1* (2), 161-169.

13. Bottini, N.; Saccucci, P.; Piciullo, A.; Iannetti, P.; Lucarini, N.; Lucarelli, P.; Gloria-Bottini, F.; Curatolo, P., Convulsive disorder and the genetics of signal transduction; a study of a low molecular weight protein tyrosine phosphatase in a pediatric sample. *Neurosci. Letters* **2002,** *333* (3), 159-162.

14. McClain, C. J.; Song, Z.; Barve, S. S.; Hill, D. B.; Deaciuc, I., Recent advances in alcoholic liver disease IV. Dysregulated cytokine metabolism in alcoholic liver disease. *American Journal of Physiology-Gastrointestinal and Liver Physiology* **2004,** *287* (3), G497-G502.

15. Fontana, L.; Eagon, J. C.; Trujillo, M. E.; Scherer, P. E.; Klein, S., Visceral fat adipokine secretion is associated with systemic inflammation in obese humans. *Diabetes* **2007,** *56* (4), 1010-1013.

16. Adkins, J. N.; Varnum, S. M.; Auberry, K. J.; Moore, R. J.; Angell, N. H.; Smith, R. D.; Springer, D. L.; Pounds, J. G., Toward a Human Blood Serum Proteome analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Prot.* **2002,** *1* (12), 947-955.

17. Govorukhina, N.; Keizer-Gunnink, A.; Van der Zee, A.; de Jong, S.; De Bruijn, H.; Bischoff, R., Sample preparation of human serum for the analysis of tumor markers: comparison of different approaches for albumin and γ-globulin depletion. *J. Chromatogr. A* **2003,** *1009* (1), 171-178.

18. Petricoin, E. F.; Paweletz, C. P.; Liotta, L. A., Clinical applications of proteomics: proteomic pattern diagnostics. *J. Mammary gland Bol. Neoplasia* **2002,** *7* (4), 433-440.

19. Merrell, K.; Southwick, K.; Graves, S. W.; Esplin, M. S.; Lewis, N. E.; Thulin, C. D., Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry. *J. Biomol. Tech.* **2004,** *15* (4), 238.

20. Dongre, A. R.; Jones, J. L.; Somogyi, Á.; Wysocki, V. H., Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. *J. Am. Chem. Soci.* **1996,** *118* (35), 8365-8374.

21. Steen, H.; Mann, M., The ABC's (and XYZ's) of peptide sequencing. *Nat. Rev. Mol. Cell Biol.* **2004,** *5* (9), 699-711.

22. Griffin, T. J.; Han, D. K.; Gygi, S. P.; Rist, B.; Lee, H.; Aebersold, R.; Parker, K. C., Toward a high-throughput approach to quantitative proteomic analysis: expression-dependent protein identification by mass spectrometry. *J. Am. Soci. Mass Spectrom.* **2001,** *12* (12), 1238-1246.

23. Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999,** *17* (10), 994-999.

24. Cagney, G.; Emili, A., De novo peptide sequencing and quantitative profiling of complex protein mixtures using mass-coded abundance tagging. *Nat. Biotechnol.* **2002,** *20* (2), 163-170.

25. Lundblad, R., Considerations for the use of blood plasma and serum for proteomic analysis. *The Internet Journal of Gastroenterology* **2005,** *1* (2).

26. Villanueva, J.; Shaffer, D. R.; Philip, J.; Chaparro, C. A.; Erdjument-Bromage, H.; Olshen, A. B.; Fleisher, M.; Lilja, H.; Brogi, E.; Boyd, J., Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J. Clin. Invest.* **2006,** *116* (1), 271.

27. Han, X.; Gross, R. W., Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry a bridge to lipidomics. *J. Lipid Res.* **2003,** *44* (6), 1071-1079.

28. Han, X.; Holtzman, D. M.; McKeel, D. W., Plasmalogen deficiency in early Alzheimer's disease subjects and in animal models: molecular characterization using electrospray ionization mass spectrometry. *J. Neurochem.* **2001,** *77* (4), 1168-1180.

29. Han, X.; Gubitosi-Klug, R. A.; Collins, B. J.; Gross, R. W., Alterations in individual molecular species of human platelet phospholipids during thrombin stimulation: Electrospray ionization mass spectrometry-facilitated identification of the boundary conditions for the magnitude and selectivity of thrombin-induced platelet phospholipid hydrolysis. *Biochemistry* **1996,** *35* (18), 5822-5832.

30. Murphy, R. C., Free-radical-induced oxidation of arachidonoyl plasmalogen phospholipids: antioxidant mechanism and precursor pathway for bioactive eicosanoids. *Chem. Res. Toxicol.* **2001,** *14* (5), 463-472.

31. Carmena, R.; Duriez, P.; Fruchart, J.-C., Atherogenic lipoprotein particles in atherosclerosis. *Circulation* **2004,** *109* (23 suppl 1), III-2-III-7.

32. Han, X.; Gross, R. W., Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass Spectrom. Rev.* **2005,** *24* (3), 367-412.

33. Han, X.; M Holtzman, D.; W McKeel, D.; Kelley, J.; Morris, J. C., Substantial sulfatide deficiency and ceramide elevation in very early Alzheimer's disease: potential role in disease pathogenesis. *J. Neurochem.* **2002,** *82* (4), 809-818.

34. Han, X.; Yang, K.; Gross, R. W., Multi-dimensional mass spectrometry-based shotgun lipidomics and novel strategies for lipidomic analyses. *Mass Spectrom. Rev.* **2012,** *31* (1), 134-178.

35. Masukawa, Y.; Narita, H.; Sato, H.; Naoe, A.; Kondo, N.; Sugai, Y.; Oba, T.; Homma, R.; Ishikawa, J.; Takagi, Y., Comprehensive quantification of ceramide species in human stratum corneum. *J. Lipid Res.* **2009,** *50* (8), 1708-1719.

36. Tan, B.; Yu, Y. W.; Monn, M. F.; Hughes, H. V.; O'Dell, D. K.; Walker, J. M., Targeted lipidomics approach for endogenous N-acyl amino acids in rat brain tissue. *J. Chromatogr. B* **2009,** *877* (26), 2890-2894.

37. Welti, R.; Li, W.; Li, M.; Sang, Y.; Biesiada, H.; Zhou, H.-E.; Rajashekar, C.; Williams, T. D.; Wang, X., Profiling Membrane Lipids in Plant Stress Responses ROLE OF PHOSPHOLIPASE Dα IN FREEZING-INDUCED LIPID CHANGES IN ARABIDOPSIS. *J. Biol. Chem.* **2002,** *277* (35), 31994-32002.

38. Welti, R.; Wang, X.; Williams, T. D., Electrospray ionization tandem mass spectrometry scan modes for plant chloroplast lipids. *Anal. Biochem.* **2003,** *314* (1), 149-152.

39. Hsu, F. F.; Turk, J.; Thukkani, A. K.; Messner, M. C.; Wildsmith, K. R.; Ford, D. A., Characterization of alkylacyl, alk-1-enylacyl and lyso subclasses of glycerophosphocholine by tandem quadrupole mass spectrometry with electrospray ionization. *J. Mass Spectrom.* **2003,** *38* (7), 752-763.

40. Abrams, J. J.; Ginsberg, H.; Grundy, S. M., Metabolism of cholesterol and plasma triglycerides in nonketotic diabetes mellitus. *Diabetes* **1982,** *31* (10), 903-910.

41. Yetukuri, L.; Katajamaa, M.; Medina-Gomez, G.; Seppänen-Laakso, T.; Vidal-Puig, A.; Orešič, M., Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. *BMC Systems Biology* **2007,** *1* (1), 12.

42. Goldberg, I. J., Lipoprotein lipase and lipolysis: central roles in lipoprotein metabolism and atherogenesis. *J. Lipid Res.* **1996,** *37* (4), 693-707.

43. Adibhatla, R. M.; Hatcher, J.; Dempsey, R., Lipids and lipidomics in brain injury and diseases. *The AAPS journal* **2006,** *8* (2), E314-E321.

44. Wang, C.; Kong, H.; Guan, Y.; Yang, J.; Gu, J.; Yang, S.; Xu, G., Plasma phospholipid metabolic profiling and biomarkers of type 2 diabetes mellitus based on high-performance liquid chromatography/electrospray mass spectrometry and multivariate statistical analysis. *Anal. Chem.* **2005,** *77* (13), 4108-4116.

45. Clish, C. B.; Davidov, E.; Oresic, M.; Plasterer, T. N.; Lavine, G.; Londo, T.; Meys, M.; Snell, P.; Stochaj, W.; Adourian, A., Integrative biological analysis of the APOE* 3-leiden transgenic mouse. *Omics: a journal of integrative biology* **2004,** *8* (1), 3-13.

46. Brindle, J.; Antti, H.; Holmes, E.; Tranter, G.; Nicholson, J.; Bethell, H.; Clarke, S.; Scho, P., eld, E. McKilligin, DE Mosedale and DJ Grainger. *Nat. Med* **2002,** *8*, 1439-1444.

47. Organization, W. H., The World health report: 2005: make every mother and child count: overview. **2005**.

48. Ray, J. G.; Vermeulen, M. J.; Schull, M. J.; Redelmeier, D. A., Cardiovascular health after maternal placental syndromes (CHAMPS): population-based retrospective cohort study. *The Lancet* **2005,** *366* (9499), 1797-1803.

49. Meekins, J.; Pijnenborg, R.; Hanssens, M.; McFadyen, I.; Asshe, A. v., A study of placental bed spiral arteries and trophoblast invasion in normal and severe pre-eclamptic pregnancies. *BJOG,***1994,** *101* (8), 669-674.

50. Alexander, B. T.; Llinas, M. T.; Kruckeberg, W. C.; Granger, J. P., L-arginine attenuates hypertension in pregnant rats with reduced uterine perfusion pressure. *Hypertension* **2004,** *43* (4), 832-836.

51. Roberts, J. M.; Taylor, R. N.; Musci, T. J.; Rodgers, G. M.; Hubel, C. A.; McLaughlin, M. K., Preeclampsia: an endothelial cell disorder. *Am. J. Obstetr. Gynecol.* **1989,** *161* (5), 1200-1204.

52. Stevens, J. M., Gynaecology from ancient Egypt: The papyrus Kahun: A translation of the oldest treatise on gynaecology that has survived from the ancient world. *Med. J. Austr.* **1974,** *2* (25-26), 949-952.

53. Huppertz, B., Placental origins of preeclampsia challenging the current hypothesis. *Hypertension* **2008,** *51* (4), 970-975.

54. Roberts, J.; Cooper, D., Pathogenesis and genetics of pre-eclampsia. *The Lancet* **2001,** *357* (9249), 53-56.

55. Hung, T.-H.; Burton, G. J., Hypoxia and reoxygenation: a possible mechanism for placental oxidative stress in preeclampsia. *Taiwanese J. Obstetr. Gynecol.* **2006,** *45* (3), 189-200.

56. Grill, S.; Rusterholz, C.; Zanetti-Dällenbach, R.; Tercanli, S.; Holzgreve, W.; Hahn, S.; Lapaire, O., Potential markers of preeclampsia—a review. *Reprod. Biol. Endocrinol.* **2009,** *7* (70), 10.1186.

57. Duff, K.; Suleman, F., Transgenic mouse models of Alzheimer's disease: how useful have they been for therapeutic development? *Briefings in Functional Genomics & Proteomics* **2004,** *3* (1), 47-59.

58. Ferri, C. P.; Prince, M.; Brayne, C.; Brodaty, H.; Fratiglioni, L.; Ganguli, M.; Hall, K.; Hasegawa, K.; Hendrie, H.; Huang, Y., Global prevalence of dementia: a Delphi consensus study. *The lancet* **2006,** *366* (9503), 2112-2117.

59. Boyle, P.; Wilson, R.; Aggarwal, N.; Tang, Y.; Bennett, D., Mild cognitive impairment Risk of Alzheimer disease and rate of cognitive decline. *Neurology* **2006,** *67* (3), 441-445.

60. Hampel, H.; Bürger, K.; Teipel, S. J.; Bokde, A. L.; Zetterberg, H.; Blennow, K., Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. *Alzheimer's Dement.* **2008,** *4* (1), 38-48.

61. Nagy, Z.; Jobst, K.; Esiri, M.; Morris, J.; King, E.-F.; MacDonald, B.; Litchfield, S.; Barnetson, L.; Smith, A., Hippocampal pathology reflects memory deficit and brain imaging measurements in Alzheimers disease: clinicopathologic correlations using three sets of pathologic diagnostic criteria. *Dementia and Geriatric Cognitive Disorders* **1996,** *7* (2), 76-81.

62. Bobinski, M.; De Leon, M.; Wegiel, J.; Desanti, S.; Convit, A.; Saint Louis, L.; Rusinek, H.; Wisniewski, H., The histological validation of post mortem magnetic resonance imaging-determined hippocampal volume in Alzheimer's disease. *Neuroscience* **1999,** *95* (3), 721-725.

63. Grady, C. L.; Haxby, J.; Horwitz, B.; Sundaram, M.; Berg, G.; Schapiro, M.; Friedland, R.; Rapoport, S., Longitudinal study of the early neuropsychological and cerebral metabolic changes in dementia of the Alzheimer type. *J. Clin. Exp. Neuropsychol.* **1988,** *10* (5), 576-596.

64. Panegyres, P. K.; Rogers, J. M.; McCarthy, M.; Campbell, A.; Wu, J. S., Fluorodeoxyglucose-positron emission tomography in the differential diagnosis of early-onset dementia: a prospective, community-based study. *BMC neurology* **2009,** *9* (1), 41.

65. Bloudeka, L. M.; Spackmanb, D. E.; Blankenburgc, M.; Sullivand, S. D., Review and Meta-Analysis of Biomarkers and Diagnostic Imaging in Alzheimer's. *J. Alzheimer's Dis.* **2011,** *26*, 1-5.

66. Jack, C. R.; Lowe, V. J.; Senjem, M. L.; Weigand, S. D.; Kemp, B. J.; Shiung, M. M.; Knopman, D. S.; Boeve, B. F.; Klunk, W. E.; Mathis, C. A., 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. *Brain* **2008,** *131* (3), 665-680.

67. Kerr, D. A.; Gorelik, M.; Levy, M., Cell-based compositions and methods for treating conditions of the nervous system. Google Patents: 2008.

Chapter 2 Serum Biomarkers Predictive of Preeclampsia

Disclaimer: This chapter is mainly reproduced from published research article: Anand, S., Bench Alvarez, T. M., Johnson, W. E., Esplin, M. S., Merrell, K., Porter, T. F., & Graves, S. W. (2015). Serum biomarkers predictive of pre-eclampsia, *Biomarkers in medicine*, *9*(6), 563-575.

2.1 Abstract

Materials and methods: Sera obtained at 12-14 weeks of pregnancy from 24 cases who later developed PE and 24 controls with uncomplicated pregnancies were processed and analyzed using a serum proteomic approach.

Results: Many statistically significant serum PE biomarker candidates (>60) were found comparing cases and controls. In addition, logistic regression analysis modeled biomarker data resulted in 14 different multi-marker combinations having high detection sensitivity and specificity (AUC>0.9).

Conclusion: Developed panels of serum biomarkers appeared effective in identifying pregnant women at 12-14 weeks gestation at risk of PE later in their pregnancy.

2.2 Introduction

Preeclampsia (PE) is a potentially life-threatening pregnancy disorder. It occurs in 5-8% of pregnancies worldwide and is a primary cause of fetal and maternal morbidity and death [1]. PE has been defined by new onset hypertension and proteinuria, but can worsen to involve several systems.

Although there has been extensive research, the early etiology of PE is incompletely understood. It is widely accepted that the placenta is required for the disease which is supported by the observation that symptoms reverse rapidly following its removal[2],[3]. PE historically has been called the "disease of theories" based its many proposed causes[4].

Despite no definitive cause, and potentially no single cause, there is evidence for a number of potentially contributory problems in PE that occur prior to clinical presentation of the disease. For example PE is associated with an incomplete invasion and remodeling of the maternal uterine spiral arterioles by extravillous trophoblasts leading, it is thought, to ischemia and hypoperfusion of the placenta [5]. Further, it has been proposed that alternating recurrence of hypoxia-reoxygenation in the uteroplacental region leads generation of reactive oxygen species (ROS) [6]. There also appears to be altered expression of angiogenic and anti-angiogenic factors in the maternal circulation prior to established PE [7]. Such early changes hold out the possibility of biochemical changes that may foreshadow the disease and could be used as biomarkers.

While knowing that a pregnant woman is at risk of PE, presently, treatment or prevention of PE remains limited. Several clinical trials have investigated the use of various agents to prevent or reduce the incidence of PE, including low dose aspirin (as an anti-inflammatory and anti-coagulant); vitamins C and E (as antioxidants and agents countering ROS) and calcium supplementation (as a potential mediator of hypertension) but these did not provide a general reduction of PE incidence overall although there is evidence of each providing a benefit to a small subset of pregnant women [8-10]. Again, the potential heterogeneity of PE may obscure selected benefits. Drug interventions are limited to blood pressure control in symptomatic PE women and

steroids to hasten fetal lung maturation when premature delivery may be required. Other drug therapies have yet to be established and approved. Nevertheless, having a reliable set of predictive biomarkers might allow for more appropriate levels of clinical surveillance for both at-risk and not-at-risk pregnant women.

Biomarkers have been proposed for PE. They have been chosen most often because of their being altered in clinically established, active disease. The most highly studied of these are markers of angiogenic imbalance, with elevated serum levels of the anti-angiogenic factors, soluble fms-like tyrosine kinase-1, sFlt-1, and soluble endoglin, sEng, and decreased levels of pro-angiogenic factors, i.e. placental growth factor, PlGF and vascular endothelial growth factor, VEGF [7, 11-12]. However, as predictive markers, they have shown limited accuracy in unselected populations[13-16]. Furthermore, they are not entirely specific to PE with comparable changes found in other pregnancy-related complications, e.g. intrauterine growth restriction without PE [17]. Additionally, it has been shown that prophylactic use of anti-oxidant vitamins, while failing to reduce the incidence of PE, did surprisingly normalize these pro-angiogenic/anti-angiogenic factors in those women who later had PE, suggesting that environmental factors, such as nutrition, can directly affect them without changing risk for disease. Other less studied, potentially predictive PE biomarkers are placental protein 13 (PP13), P-selectin and pregnancy associated plasma protein A, but they also demonstrate low predictive value. Consequently, there is on-going need for a better set of PE biomarkers.

Proteomics is an analytical approach that couples chromatographic separations with mass spectrometry. These methods can survey hundreds to thousands of different molecules

simultaneously, creating the possibility of a more global, unbiased assessment of potential diagnostic or predictive biochemical changes accompanying or predating clinical disease.

The use of serum as part of proteomic approaches to finding biomarkers is challenging. Not only is the target specimen complex, a more significant problem is the presence of ~30 highly abundant proteins. High abundance species cause ion suppression in MS-based methods obscuring thousands of lower abundance species. Hence, there is always a tradeoff: focus on large proteins and suffer ion suppression losing thousands of potentially informative biomolecules or focus on lower molecular weight species and eliminate potentially informative proteins. Previously, there have been few, if any, attempts at global serum proteomic approaches to identify serum or plasma PE biomarkers, certainly none that have fully eliminated high abundance proteins and none that have interrogated lower molecular weight, low abundance species. Our approach is new and focuses on the less abundance (nM or lower), lower molecular weight (<10,000 daltons) proteins, peptides and lipids predicting PE. In doing so, the approach increases dramatically the number of biomarkers interrogated by MS (>5000 additional novel analytes), increasing the possibility of finding important predictive PE biomarkers. This approach at a minimum complements previous studies. In this study, this approach has been applied to serum from pregnant women collected at 12-14 wks pregnancy and markers have been sought to allow for prediction of later PE. These results are described here.

2.3 Methods

2.3.1 Patient Population

Serum specimens utilized for this PE study were obtained from the Department of Obstetrics and Gynecology, University of Utah School of Medicine and from Intermountian Healthcare and represent sera previously collected from pregnant women at 12-14 weeks gestation as part of a prospective study that considered universal potential complications. There were 24 controls, having term uncomplicated pregnancies and 24 cases, which developed PE later in the same pregnancy. The diagnosis of PE followed guidelines established by the American Congress of Obstetricians and Gynecologists and the International Society for the Study of Hypertension in Pregnancy [1]. Women on medications or with major intercurrent disease were excluded, including those with diabetes, renal disease, or preexisting hypertension. Demographic information on the women whose speciments were included in the study in summarized in Table 2.1. All women were Caucasian subjects reflecting the racial make-up of the general hospital population. The study was approved by representative Institutional Review Boards at both the University of Utah and Brigham Young University. We note that at the outset of this research we had anticipated completing a follow-up, confirmation study, but found that there were no additional specimens available from cases that had not been thawed and refrozen, perhaps repeatedly. We were unaware of other studies that collected specimens prospectively at 12-14 gestation, processed expeditiously and kept frozen until analysis available to us.

2.3.2 Sample Preparation

As part of the original protocol, blood was collected, allowed to clot (~60 min) and the serum was separated immediately from the clot by centrifugation and stored at -80^{0}C until processed further.

Table 2.1 Demographic Data for Study Groups

	Maternal Age (yr)	Gestational Age (wks)	Body mass index (BMI)	Birth weight (gm)	Max. SBP mm Hg	Max. DBP mm Hg
Cases	28 ± 1.2	35 ± 0.8	25.5 ± 1.2	2467 ± 200	159 ± 3.4	103 ± 1.6
Controls	28 ± 0.97	38 ± 0.35	21.2 ± 1.4	3447 ± 108	132 ± 2.8	83 ± 2.3

SBP: Systolic blood pressure, DBP: Diastolic blood pressure

In order to selectively focus on low MW peptides, a protocol was followed to deplete proteins from the sera by adding acetonitrile in a ratio to serum of 2:1 [18]. This removed high abundance, high MW proteins. The apparent protein concentration of protein-depleted samples was determined using a Bio-Rad microtiter plate protein assay following the manufacturer's method (Bio-Rad Laboratories, Hercules, CA). An aliquot containing 4 µg apparent protein was transferred into a fresh microcentrifuge tube. The aliquot was lyophilized (CentriVap Concentrator Labconco Corporation, Kansas, City, MO) to less than 20 µL, and the volume was brought up to 20 µL with HPLC grade water and acidified with 20 µL of 88% formic acid. The specimens were loaded onto a capillary liquid chromatography-tandem mass spectrometer system (cLC-MS) via autosampler injection (Dionex Corporation, Sunnyvale, CA).

2.3.3 cLC-MS Analysis of Protein Depleted Specimens

Specimens were introduced onto an LC system utilizing a 1 mm (16.2 µL) microbore guard column (Upchurch Scientific, Oak Harbor, WA) coupled to a 15 cm x 250 µm i.d. in-house packed capillary column (POROS R1 reversed-phase media, Applied Biosystems, Framingham, MA). The instrument employed an LC Packings Ultimate Capillary HPLC pump system, with a FAMOS® autosampler (Dionex Corporation, Sunnyvale, CA) controlled by Analyst QS® software (Applied Biosystems, Foster City, CA). The chromatographic separation gradient used an aqueous phase (98% HPLC grade H_2O, 2% acetonitrile, 0.1% formic acid) and an organic phase (2% H_2O, 98% acetonitrile, 0.1% formic acid). Details of the elution gradient used and other parameters of the method have been described previously [18].

During the chromatographic separation, eluate from the fractionated samples was introduced into a QSTAR Pulsar I quadrupole orthogonal time-of-flight mass spectrometer through an ESI (electrospray ionization) source. The MS data for each sample was collected for m/z 500 to 2500 from 5 to 55 min of the cLC elution with mass spectra obtained every 1 sec. The elution profile of each cLC fractionated sample was reported as the total ion chromatogram (TIC) and the results analyzed using Analyst QS® 1.1 software (Applied Biosystems, Foster City, CA). Following the void volume (~12-15 min), ~25 min of the remaining 55 min gradient was divided into ten ~2 minute regions or windows based on previously described serum time markers [19]. Within each two-minute window, the spectra for cases and controls were overlaid and color-coded to distinguish the two comparison groups for analysis. Visual inspection was conducted and peaks with apparent quantitative differences between cases and controls were recorded for further analysis. The intensities of these potential markers were recorded for each of the specimens in the initial specimen set using the extracted ion chromatogram (XIC) function and were used for statistical analyses. Additionally, a co-eluting species having comparable intensity for the cases and controls were chosen as the reference peaks. Their intensities were determined and used for the normalization of the potential biomarkers.

2.3.4 Statistical Analyses

The peak heights of potential markers of interest from the initial sample set were subjected to Student's T-test to determine statistical significance of case versus control differences observed in the visual inspection step. A p-value of <0.05 was considered statistically significant, although markers with a p-value of less than 0.1 were also included for modeling biomarker panels with higher specificity and sensitivity. Thus, all the potential markers with a p-value of less than 0.1

(66 markers) were subjected to biomarker panel development using a forward selection, leave one out, logistic regression approach [20-21]. It is recognized that a simple T-test is not a sufficient statistical test. However, the field of proteomics is emerging and there is no agreed upon statistical approach. Typically, multi-marker modeling and/or retesting of markers increase confidence in their usefulness. The forward selection process was done in conjunction with logistic regression analysis and receiver operator characteristic (ROC) curves were generated to allow for sensitivities and specificities to be considered. The ROC would ideally be used with larger numbers but provide a sense of the usefulness of biomarker combinations.

2.3.5 Tandem MS Identification of Significantly Different Candidate Biomarkers

Previously assayed samples having high abundance of a candidate PE biomarker were selected for fragmentation studies. A quantity of 5-20 μg of the sample was treated with 2.5 μL of 88% formic acid and loaded onto cLC-MS-MS system. Prior to fragmentation experiments, the specimen was submitted to two cLC-MS runs to determine the exact elution time of the molecule of interest. If a consistent elution time was obtained, a two minute window comprising one minute before and one minute after the candidate peak height maximum was directed into the fragmentation cell and fragmentation data collected. The same guard column, capillary column and column packing were used as described above. The aqueous phase was 98% HPLC grade H_2O, 2% acetonitrile, 0.1% formic acid and the organic phase was 2% H_2O, 98% acetonitrile, 0.1% formic acid. The initial cLC-MS runs lasted for 40 min starting with 3 min of 95% aqueous phase and 5% organic phase, followed by a linear increase to 60% organic phase over 20 min. Following that, the gradient was then further increased to 95% organic phase over 5 min and held at 95% organic for 6 min. Finally, the gradient was then returned to 95% aqueous phase over 3 min and

held there to re-equilibrate the column. The mass range for collecting the scans was dependent on the parent mass fragmented. Fragmentation was produced by ion collisions with nitrogen or argon. Collision product ions were then analyzed by time-of-flight MS and a composite spectrum for the 2 min of fragmentation spectra were obtained. Fragmentation at different collision energies was also carried out until fragment ion coverage was as complete as possible. As one spectrum was collected per second, 120 MS/MS spectra were collected in 2 minutes and the MCA (multi-channel analyzer) function was turned on which summed all 120 MS/MS spectra resulting in increased signal to noise. The MCA spectra for different energies were overlaid together using the 'overlay' feature, followed by summing of all these spectra using the 'sum overlays" feature to provide a single spectrum with optimal fragmentation coverage.

The summed spectrum was visually inspected and compared to the exported peak list to check for any miss-assigned charge states. After charge state correction, all the peaks in the peak list higher than +2 were transformed into their +1 m/z values using the formula: +1 mass = m/z value * charge – (charge – $1H^+$) to simplify the database search. The corrected mass list was exported as a tab-delimited text file and submitted to Mascot (Mascot 2.3, www.matrixscience.com) database search.

2.3.6 De Novo Sequencing of Protein or Peptide Markers

In instances when Mascot database comparisons of candidate marker MS/MS spectra did not yield any promising results, *de novo* sequencing of these markers was attempted. MS/MS spectra were first examined for b_1 and/or y_1 ions to indicate the start of the peptide sequence at either the N- or C-terminus. These ions occur in the lower m/z range of the spectra. Peak

comparisons were then made in an attempt to identify sequential amino acids. This process was continued until a continuous chain of amino acid sequence was determined. The predicted peptide sequences were also confirmed by independently subjecting them to a BLAST search available through NCBI website which gave possible parent proteins for a peptide.

When a complete sequence of contiguous amino acids representing b- or y-series ions cannot be obtained, sometimes the parent protein can still be determined. In general a sequence of at least 5-6 consecutive amino acids must be identified before the parent protein can be determined by comparison to a protein data base.

2.3.7 Mass Spectrometric Fragmentation and Manual Chemical Characterization of Lipid Markers

In addition to protein or peptide biomarkers, several potential biochemical markers were likely lipids based on their elution in the hydrophobic region of the cLC chromatogram. These candidates were also submitted to MS fragmentation studies. The fragmentation data were reviewed for peaks characteristic of some lipid classes, including the frequently seen peak at m/z 184.07 known to represent a phosphocholine head-group [22-23].

Accurate mass determination of lipids was compared to the LIPID MAPS database (http://www.lipidmaps.org/) to find archived lipids having very similar masses. LIPID MAPS MS prediction tool (http://www.lipidmaps.org/tools/index.html) was also used to predict possible fragments representing Sn1 and Sn2 acyl losses, which was then compared to actual peak masses present in MS/MS spectra to identify possible fatty acid constituents. Additionally, previously

reported MS/MS fragment assignments for glycerophosphocholines (PC) were used to confirm identities of some candidate PC biomarkers based on comparable MS/MS spectra**.**

2.3.8 Evaluating Possible Lipid Dimer Formation of During ESI

Some observed marker peaks were consistent with the presence of lipid homodimers (dimers of the same two PC species presenting as $2M+H^+$) or heterodimers (dimers of two different PC species presenting as $M_1+M_2+H^+$). To assess whether lipid dimers were formed using our instrument, a lyso PC standard, 1-oleoyl-sn-glycero-3-phosphocholine, having a MW of 521.67, was used (Sigma Aldrich). It was diluted with acetonitrile to a concentration of 12 μM followed by addition of 88% formic acid (20 μL per 1 mL sample) and was directly infused into the ESI-TOF MS at a flow rate of 10 μL/min.

2.4 Results

The question addressed by this study was whether there were biomarkers observable in the serum of pregnant women predictive of PE development later in the same pregnancy. To assess this, a novel serum proteomic approach was used to interrogate the low abundance, LMW proteins, peptides and other biomolecules. The approach found more than 60 candidate biomarkers that were statistically or near statistically different in pregnant women sampled at 12-14 weeks gestation who developed PE weeks to months later compared with pregnant women sampled at the same time who had uncomplicated, term pregnancies. Enough specimen was available to conduct MS/MS fragmentation studies on 39 of these. These candidate biomarkers are summarized in Table 2.2.

Table 2.2 Statistically significant potential biomarkers

S No.	m/z	P value	Higher in	AUC
1	942.5	0.017	Controls	0.611
2.	571.3	0.005	Controls	0.585
3.	593.3	0.034	Controls	0.583
4.	619.8	0.03	Controls	0.566
5.	1238.5	0.03	Controls	0.562
6.	539.6	0.05	Cases	0.177
7.	676.7	0.02	Cases	0.477
8.	568.8	0.05	Cases	0.505
9.	1071.4	0.05	Cases	0.242
10.	601.3	0.012	Controls	0.647
11.	649.3	0.00003	Cases	0.636
12.	509.3	0.001	Controls	0.693
13.	569.3	0.004	Controls	0.67
14.	621.4	0.03	Controls	0.568
15.	723.47	0.02	Controls	0.519
16.	767.5	0.03	Controls	0.502
17.	503.3	0.02	Controls	0.573
18.	508.3	0.0002	Controls	0.748
19.	513.3	0.003	Controls	0.674
20.	553.3	0.0001	Controls	0.753
21.	594.3	0.02	Cases	0.589
22.	665.4	0.006	Controls	0.628
23.	697.4	0.007	Controls	0.637
24.	739.4	0.0003	Controls	0.809
25.	975.6	0.0009	Controls	0.698
26.	1015.6	0.01	Controls	0.021
27.	1026.6	0.0001`	Controls	0.769
28.	1069.7	0.05	Cases	0.514
29.	1074.6	0.0005	Controls	0.7
30.	1111.7	0.08	Controls	0.5
31.	639.38	0.0001	Controls	0.79
32.	634.39	0.003	Controls	0.675
33.	634.39	0.002	Controls	0.707
34.	756.5	0.01	Controls	0.613
35.	1540.1	0.03	Controls	0.574
36.	1516.1	0.01	Controls	0.629
37.	718.8	0.001	Cases	0.591
38.	719.2	0.05	Controls	0.612
39.	734.8	0.000003	Cases	0.705

2.4.1 MS Structural Characterization of PE Peptide Biomarker Candidates

Fragmentation studies on the 39 peaks suggested that 9 candidates were peptides based on higher charge states, earlier elution times and fragmentation patterns consistent with a peptide, e.g. presence of an immonium ion. Of these, 6 were identified using MS-MS fragmentation data, which provided complete or near complete b- and y-series fragment ions and were identified by MASCOT. Some were not identified by MASCOT because of high charge state ($z \geq 3+$), but were identified by *de novo* sequencing. Among the 6 peaks identified, two sets of peaks represented the same peptide found in different charge states. Hence, 4 unique peptides were successfully sequenced.

The peptide with m/z 942.6 was identified as being a fragment of complement C3. The peptides m/z 1070.8 and m/z 676.7 were found to be fragments of inter-alpha-trypsin inhibitor heavy chain H4 isoform 2. Furthermore, peptides m/z 676.7 and 1014.5, were identical fragments of the same protein with different charge states, where 676.7 was in the +3 form and 1014.5 was the +2 form. Finally, the peptide with m/z 1238.5 was identified as being a fragment of β-fibrinogen with a pyro-glutamic acid modification on its N terminus representing a glutamine or glutamate. These are summarized in Table 2.3.

As indicated, 3 biomarkers that were likely to be peptides were not identified. This was due to some combination of the following: high charge state, the inability to identify y1 or b1 ions, higher MW (e.g. 40+ amino acids) coupled with low abundance leading to incomplete b- and y-ion series and in all instances inadequate specimen to carry out fragmentation studies at multiple collision energies.

Table 2.3 Amino acid sequence of peptide biomarkers and elemental composition and class for lipid biomarkers successfully fragmented

m/z	charge	MW	sequence/structure	parent protein/ compound/class	Elemental composition
942.5	1	941.5	hwesasll	Complement C3	
1070.8	4	4279.25	nvhsagaagsrmnfrpgvlss rqlglpgppdvpdhaayhpf	Inter-alpha-trypsin inhibitor heavy chain H4	
676.7	3	2026.98	qlglpgppdvpdhaayhpf	Inter-alpha-trypsin inhibitor heavy chain H4	
1014.5	2	2026.98	qlglpgppdvpdhaayhpf	Inter-alpha-trypsin inhibitor heavy chain H4 isoform	
619.8	2	1237.5	pyro-egvndneegff	Beta-fibrinogen	
1238.5	1	1237.5	pyro-egvndneegff	Beta-fibrinogen	
756.6	1	755.6	PC(18:2/16:1)	1-(9Z,12Z-octadecadienoyl)-2-(9Z-hexadecenoyl)- glycero-3-phosphocholine	$[C_{42}H_{78}NO_8P+H^+]$
508.3	1	507.3	PC(O-18:1)	1-(11Z-octadecenyl)-sn-glycero-3-phosphocholine	$[C_{26}H_{54}NO_6P+H^+]$
594.3	1	593.3	PC(16:0/5:0(CHO))	1-hexadecanoyl-2-(5-oxovaleroyl)-sn-glycero-3-phosphocholine	$[C_{29}H_{56}NO_9P+H^{+]}$
1069.7	1	1068.7 (m/z 524.35 + m/z 546.35)	Dimer of LPC (18:0) + LPC (20:3)	Glycerophosphocholines	$[C_{26}H_{54}NO_7P+C_{28}H_{52}NO_7P] + H^+$
1516.1	1	1515.1 (m/z 758.6 + m/z 758.6)	Dimer of PC(34:2) + PC(34:2)	Glycerophosphocholines	$[C_{42}H_{80}NO_8P+C_{42}H_{80}NO_8P] + H^+$

1540.1	1	1539.1 (m/z 758.57 + m/z 782.57)	Dimer of PC(34:2) + PC (36:4)	Glycerophosphocholines	$[C_{42}H_{80}NO_8P + C_{44}H_{80}NO_8P] + H^+$
1111.7	1	1110.7 (m/z 544.3 + m/z 568.34)	Dimer of LPC (20:4) + LPC (22:6)	Glycerophosphocholines	$[C_{28}H_{50}NO_7P + C_{30}H_{50}NO_7P] + H^+$
1015.7	1	1014.7 (m/z 496.35 + m/z 520.35)	Dimer of LPC (16:0) + LPC (18:2)	Glycerophosphocholines	$[C_{24}H_{50}NO_7P + C_{26}H_{50}NO_7P] + H^+$
975.7	1	974.7 (m/z 480.35 + m/z 496.35)	Dimer of LPC (O-16:1) + LPC (16:0)	Glycerophosphocholines	$[C_{24}H_{50}NO_6P + C_{24}H_{50}NO_7P] + H^+$
634.4	1	633.4		Glycerophosphocholines	$[C_{33}H_{64}NO_8P] + H^+$
1026.7	1	1025.7		Glycerophosphocholines	$[C_{51}H_{98}N_2O_{16}P] + H^+$
513.3	1	512.3		Sphingolipids	$[C_{24}H_{53}N_2O_7P] + H^+$
553.3	1	552.3		Sphingolipids	$[C_{25}H_{49}N_2O_9P] + H^+$
509.3	1	508.3		Sphingolipids	
569.33	1	568.3		Sphingolipids	$[C_{29}H_{49}N_2O_7P] + H^+$
593.3	1	592.3		Sphingolipids	

Note: for candidate m/z 509.3 it was likely that this was a unique and unusual distinct sphingomyelin whose fatty acid constituent could not be identified based on its fragmentation spectrum. Also, for candidate m/z 593.3 the elemental composition could not be determined because exact mass studies could not be performed due to its very low abundance in the samples and limited samples.

This group included the following additional markers (represented by their m/z ratio): 718.8, 719.2, and 734.8.

2.4.2 MS Structural Characterization of PE Lipid Biomarker Candidates

Several of the candidates did not appear to be peptides based on elution time, m/z and fragmentation. This pattern included no evidence of immonium ions, no fragments with intervals consistent with amino acids or modified amino acids but fragment features consistent with known lipid classes, e.g. a peak at m/z 184.07 consistent with a phosphocholine head group.

2.4.2.1 Identified PE Lipid Biomarkers

Several of the lipid biomarkers were successfully categorized as phospholipids. The markers with m/z 756.5, 594.3 and 508.3 were identified respectively to be glycerophosphocholine (PC) with fatty acyl groups represent a 16 carbon fatty acid with one double bond (16:1) and an eighteen carbon fatty acid with two double bonds (18:2), a lyso-PC (18:1) and an oxidized PC (16:0/5:0(CHO)). Several species, m/z 1069.7, m/z 1026.7, 1516.1, 1540.1, 1111.7, 1015.7 and 975.7 were found to be dimers containing two copies of a single PC or coupling of two different PC species (Table 2.3).

The MS-MS fragmentation spectrum for a number of other candidate markers, i.e. those having m/z ratio of 634.4, m/z 513.3, m/z 553.3, m/z 509.3, m/z 569.3 and m/z 593.3 displayed a prominent peak at m/z 184.07 indicating the presence of a phosphocholine moiety. Their elemental composition was determined based on exact mass studies. However, their complete structure could

not be determined. According to the nitrogen rule, the $M+H^+$ even m/z (having odd neutral mass) should correspond to the presence of odd number of nitrogen atoms.

Likewise, the odd $M+H^+$ m/z (having even neutral mass) should represent even numbered nitrogen atom containing species. Therefore the markers with m/z 634.4, 1026.7 are likely glycerophosphocholines, while the markers with m/z 513.3, m/z 553.3, m/z 509.3, m/z 569.3 and m/z 593.3 are likely to be sphingomyelins (Table 2.3).

A number of the other biomarker candidates were likely to be lipids, but did not belong to identifiable classes of lipids, i.e. they lacked identifiable head groups or other recognizable constituents and the fragmentation pattern, however detailed, remained uninterpretable using the latest data bases and software. These included those species having m/z values of 639.4, 739.4, 503.3, 601.3, 649.3, 621.4, 665.5, 697.4, 571.3, 568.8, 539.6, 723.5, 767.5 and 509.3.

2.4.3 Biomarker Panel Development

As shown in Figure 2.1, the best single marker having m/z=739.4 provided an area under the curve (AUC) of 0.809, and a sensitivity of 79% at a specificity of 67%. Statistical analyses were further performed on potential markers to determine multi-marker models with better predictive values. Only those 39 species submitted to fragmentation and having some chemical characterization were used in multi-marker modeling. Several combinations were formed having an AUC > 0.9 as summarized in Table 2.4. Furthermore, statistical analysis limited to just those biomarkers with clear chemical identities or classifications yielded three combinations of

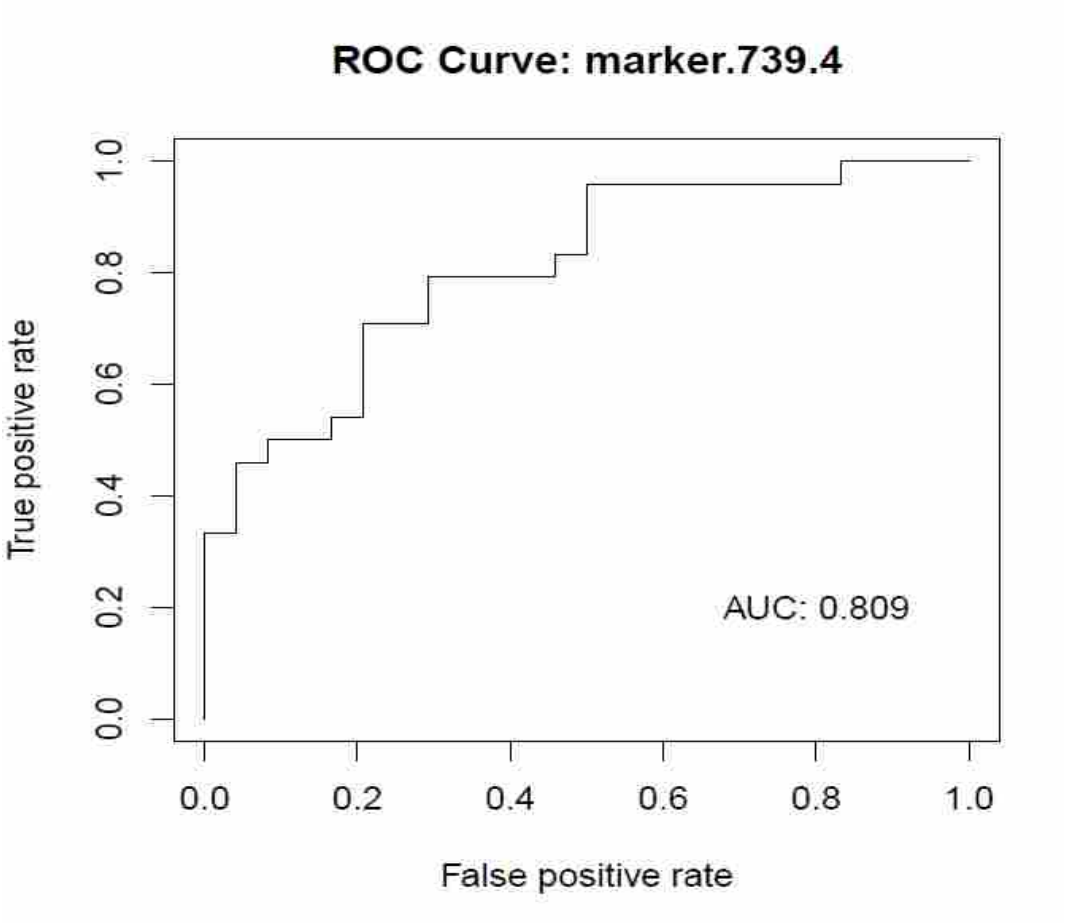


Figure 2.1 ROC curve for the single best serum PE marker having an m/z 739.4.

Table 2.4 Multi-marker panels constructed from combinations of the 39 biomarkers for which we have fragmentation and classification data. Only the 6 top sets with their respective ROC curve AUC values are shown

m/z	AUC
619.8, 739.4, 649.3, 1111, 756.5, 723.47	1
734.8, 639.38, 975.6, 756.5, 571.3, 568.8	1
785.5, 739.4, 649.3, 503.3, 639.38, 756.5, 676.7	1
639.38, 734.8, 975.6, 756.5, 571.3, 568.8	1
568.8, 734.8, 719, 718, 593.2, 1026.6, 503.3	0.99
1111, 639.8, 649.3, 509.3, 1516.1, 665.4,	0.98

markers with AUC > 0.8 (Table 2.5). The ROC curve for the panel of identified markers (m/z 508.3, 594.3 and 756.5) having an AUC of 0.832 is shown in Figure 2.2.

2.4.4 Dimer Formation for Glycerophosphocholines

When a lyso-PC standard was directly introduced by ESI onto the MS system a small amount of dimer was formed (See Figure 2.3). The peak at m/z 522.35 (ion count 270000) represented the monoisotopic peak ($M+H^+$) for the standard and the peak at m/z 1043.7 (ion count 32000, 11.8%) corresponded to the homodimer of the same ($2M+H^+$, m/z 1516.1). Table 2.6 represents the intensities of the PC and its dimer in 5 samples.

2.5 Discussion

Preeclampsia is a serious, life-threatening complication of pregnancy whose cause is incompletely understood. Currently, there is no effective way to predict who will likely develop PE during a given pregnancy. Yet, the very possibility of developing this complication has contributed substantially to the almost universal surveillance of pregnant women over the second half of their pregnancies. Several biochemical and hemodynamic abnormalities accompany established PE, but none have been shown to be useful in its prediction. The inability to accurately predict the disease has made trials of pharmacologic agents to prevent or reduce the incidence of PE large and expensive. This study sought to find predictive serum biomarkers of PE.

This approach found over 60 novel molecular species that were statistically or near statistically different between cases and controls but only 39 of these could be characterized due to limited

Table 2.5 Multi-marker panels limited to just the 7 identified markers. Lipid dimers and other lipids where the fatty acid constituents could not be unequivocally identified were excluded. Only those combinations that resulted in ROC curves having an AUC > 0.8.

m/z of panel of markers	AUC	Sensitivity	Specificity
508.3, 594.3, 756.5	0.832	88%	70%
508.3, 594.3, 942.5	0.812	67%	79%
508.3, 594.3, 942.5, 1014.5	0.810	80%	70%

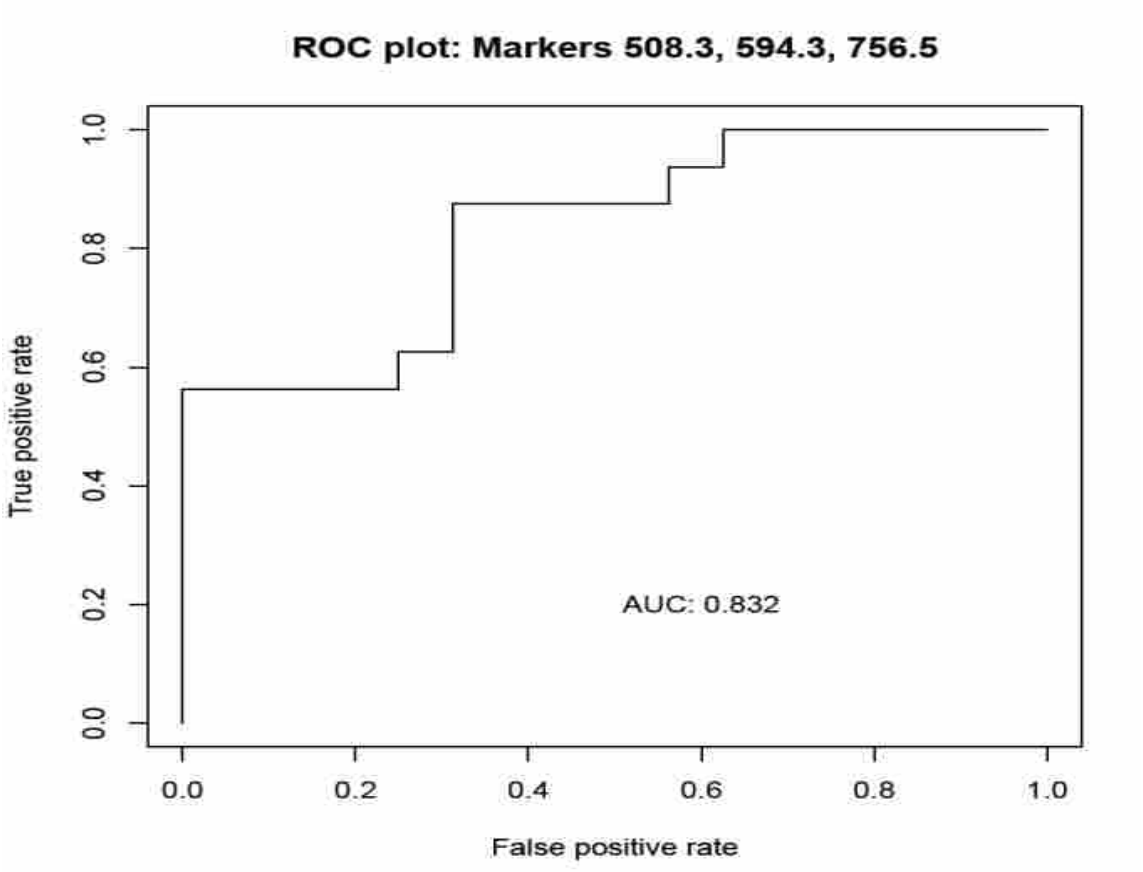


Figure 2.2 ROC curve for the best multi-marker panel found using just the 7 serum PE biomarkers that have been effectively identified.

Table 2.6 Abundances for a dimer marker (m/z 1516.1) and its monomer (m/z 785.5) in some of the samples

Sample	Abundances for m/z 758.57	Abundances for m/z 1516.1	Ratio
1	8617	3296	2.6
2	9314	4258	2.1
3	9463	5236	1.8
4	9521	4008	2.2
5	9627	5753	1.7

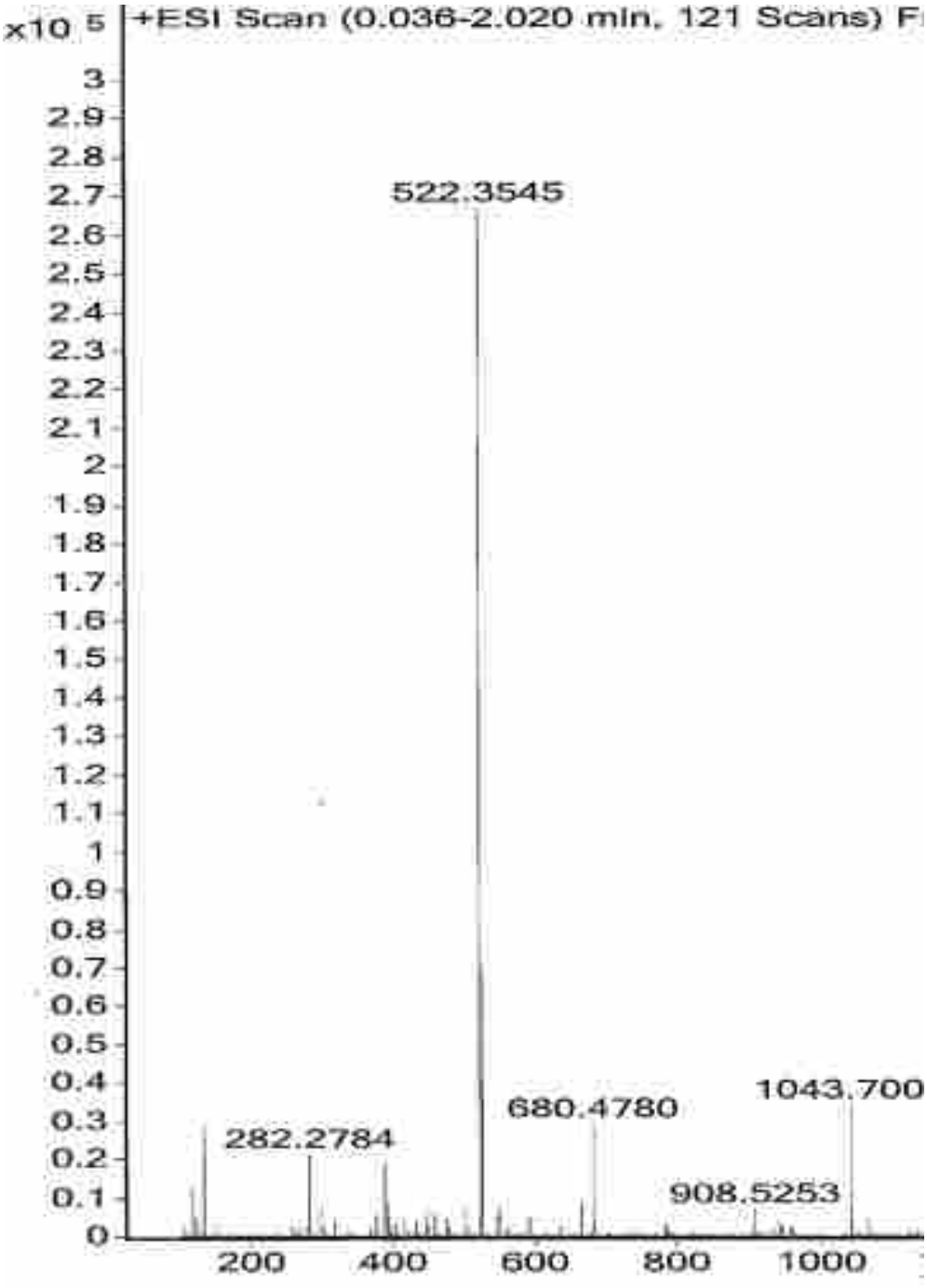


Figure 2.3 Direct infusion spectrum for the lyso-PC standard m/z 522.3 and the appearance of an m/z 1043.7 representing its dimer
.This represents ~12% conversion of this lipid to its dimer under in vitro conditions described in Methods.

specimen availability. Modeling experiments found 14 multi-marker panels with AUCs of 0.90 or more that appeared to be effective in identifying most pregnant women who would develop PE in the same pregnancy. It is recognized that these statistics were developed on a small number of samples and will work less well in a larger, unselected population of pregnant women. Nevertheless, the results suggest that the method is robust, with its ability to survey routinely ~7000-8000 low molecular weight species and able to find candidate predictive biomarkers at 12-14 weeks gestation. Clearly, other follow-up studies will be needed to verify these findings.

This approach has the additional advantage of allowing for the seamless transition from discovering potential biomarkers to their further chemical characterization as part of tandem MS-MS fragmentation studies. Identification methods for peptides and small proteins are much more highly developed than approaches for characterizing lipids. Not surprisingly, most peptides were identified. Two peptides (m/z 718.8, m/z 719.2) that were not sequenced had high charge states (z = +6) making fragmentation data search engines useless and with their low abundances in MS^1, longer length (40+ amino acids), and very limited sample made *de novo* sequencing incomplete without a clear identification. A third peptide was unsuccessfully characterized (m/z 734.4, z = +2). This was a result of having no additional sera with sufficient levels of the peptide to complete all of the several fragmentation studies that would have been necessary.

Among the 4 peptides identified, peptides m/z 942.6 and m/z 1238.5 were found to be higher in controls and are fragments of the parent proteins complement C3 and beta-fibrinogen respectively. Decreased levels of fibrinogen in PE pregnancies due to extensive platelet activation and consumption as compared to uncomplicated pregnancies have been reported by others [24-25].

The biological activity or role of peptide fragments of fibrinogen in PE, if any, has not been investigated. Likewise, activation of the complement system, resulting in depressed levels of complement C3 in PE patients, has also been documented, but, again the possible activity or biology implications of this fragment has not been explored [26-27]. Both the peptides with m/z 676.7 and 1070.8 were found to be elevated in PE cases and belonged to the same parent protein, inter-alpha trypsin inhibitor heavy chain 4 (ITIH4), a kallikrein-sensitive acute phase reactant found to be elevated in inflammatory states [28]. Peptide m/z 676.7 has earlier been shown to be depressed in sera of women at 24 weeks of pregnancy who suffered from early labor and subsequent spontaneous preterm birth[29].

As part of this study, several candidate biomarkers were found to be lipids. Marker m/z 508.3 was identified to be a lyso-phosphatidylcholine (lyso-PC). It was found to be lyso-PC (18:1) and was higher in controls. Some studies suggest that phospholipases, especially phospholipase A2, may be increased in the setting of PE [30]. However, most of these phospholipases are cytosolic enzymes and less likely to affect extracellular lipid levels. Consequently, there is not a particular pathway known to be altered in PE for circulating lipids or the production of lyso-phospholipids specifically. The peak having m/z 756.5 was shown to be a glycerophosphatidylcholine (PC) (16:1/18:2) and was higher in women having uncomplicated pregnancies. The lipid having m/z 594.3 was found to be higher in PE cases and appears to be an oxidized glycerophosphatidylcholine, in which one of the fatty acyl chains is oxidized to an aldehyde group. The compound appears to be 1-hexadecanoyl-2-(5-oxovaleroyl)-sn-glycero-3- phosphocholine. It has been previously demonstrated that PC species if exposed to reactive oxygen species (ROS) can undergo an aldehyde modification [31]. There have been several reports of increased ROS in PE

and this association might explain the increased levels of this species in PE cases as compared to controls [32-33].

Several of the lipids were characterized and found to be lipid dimers composed of two glycerophosphatidylcholines. The two PCs were not simply associated with each other but were covalently linked. Those that could be more completely characterized as lipid dimers included lipids with m/z = 1069.7, 1015.7, 975.7, 1111.7, 1516.1, 1540.1, 1026.7 and 1074.7. The source of these lipid dimers remains unclear. There are some previously publications suggesting that these may be formed as a result of the ionization process in the ESI source [34-35]. The extent of dimerization by ESI under the conditions used for these studies was probed by submitting a PC standard of known composition to direct injection on the ESI-TOF-MS system, albeit at a higher concentration than observed for any of the lipids found endogenously. The dimer (m/z 1516.1) of the standard was observed (Fig 2.3) but dimerization, even with high concentrations was ~12% of the standard added. In contrast, the abundance of lipid dimer observed at m/z 1516.1 in several actual serum samples was found to be ~2 times greater than that of the monomer at m/z 758.5 (Table 2.6). This suggests that the great majority of this dimer and presumably others is produced in vivo, perhaps through free radical exposure. However, the possibility of non-biological production of dimers during the ionization cannot be ruled out.

While no other group has carried a global (or shotgun) proteomic study of lower molecular weight PE biomarkers, many other groups have worked to develop useful biomarkers for PE focused on a single target or a few molecular species (including some MS approaches), and a few appear to have good AUC values and high detection rates in initial studies [13]. For example, an

AUC value of 0.7 was reported for pregnancy associated plasma protein-A (PAPP-A) in predicting PE when measured in the first trimester [36]. Nevertheless, the AUC and PE detection rates observed in our study exceed previously reported values and at a minimum represent complementary tests that could enhance early PE testing. Relying on a single marker to predict all PE may not be possible and not surprisingly, there are no currently clinically accepted prognostic biomarkers for PE. We believe that a more global approach is needed to overcome these limitations. It is likely that a combination of markers that detect PE of different etiology, severity, genetic factors will prove better for the purpose of pre-symptomatic detection of pregnant women at risk for PE. Even in what seems to be a somewhat homogenous population, PE still presents a spectrum of disease severity, timing, complications and in different populations, e.g. obese versus lean women, there are likely to be other biochemical differences leading to different proteomic profiles. Hence, more inclusive multi-maker panels are likely to be more predictive of PE generally, reflecting possible differences in etiology or involving different pathologic pathways that produce the diagnostic PE phenotype. As such, biomarkers representing biologically diverse pathways would likely then provide better overall coverage.

2.6 Future Perspective

These results appear promising with many novel candidate serum biomarkers discovered and many effective combinations developed providing potentially valuable risk assessment of PE. Future studies will need to retest and validate these candidate markers in independent specimens from different populations with laboratory analysis carried out absent knowledge of case or control status. This will require accumulation of or access to similarly timed specimens, processed expeditiously to allow for similar results and not submitted to freeze-thaw cycles. If the markers

are shown to be valid over time and considered useful in the screening of pregnant women, the MS method could be scaled up (as was done for the preterm labor marker fetal fibronectin [37]) to service high numbers of specimens. Alternatively, individual high throughput assays, e.g. immunoassays, could be developed for several of the best markers as is done for other clinical assay panels (lipids, cardiovascular risk factors, etc). Also, these markers themselves suggest biochemical changes occurring early in the pregnancies of women who will develop PE many weeks later. Further studies of these suggested pathways or processes, e.g. ROS, would be helpful in more fully explaining or verifying very early pathologic changes leading to PE. Having useful, verified biomarkers that predict later PE would allow for far smaller clinical studies, including evaluation of pharmacologic or other medical treatments for this potentially fatal disease.

2.7 References

1. Practice, A. C. o. O., Practice bulletin# 33: diagnosis and management of preeclampsia and eclampsia. *Obstetr. Gynecol.* **2002,** *99* (1), 159-167.

2. Huppertz, B. In *The feto–maternal interface: setting the stage for potential immune interactions*, Semin. Immunopathol., Springer: 2007; pp 83-94.

3. Huppertz, B., Placental origins of preeclampsia challenging the current hypothesis. *Hypertension* **2008,** *51* (4), 970-975.

4. Roberts, J.; Cooper, D., Pathogenesis and genetics of pre-eclampsia. *The Lancet* **2001,** *357* (9249), 53-56.

5. Cnossen, J. S.; Morris, R. K.; ter Riet, G.; Mol, B. W.; van der Post, J. A.; Coomarasamy, A.; Zwinderman, A. H.; Robson, S. C.; Bindels, P. J.; Kleijnen, J., Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis. *Can. Med. Assoc. J.* **2008,** *178* (6), 701-711.

6. Hung, T.-H.; Burton, G. J., Hypoxia and reoxygenation: a possible mechanism for placental oxidative stress in preeclampsia. *Taiwanese J. Obstetr. Gynecol.* **2006,** *45* (3), 189-200.

7. Maynard, S. E.; Min, J.-Y.; Merchan, J.; Lim, K.-H.; Li, J.; Mondal, S.; Libermann, T. A.; Morgan, J. P.; Sellke, F. W.; Stillman, I. E., Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J. Clin. Invest.* **2003,** *111* (5), 649-658.

8. Ruano, R.; Fontes, R. S.; Zugaib, M., Prevention of preeclampsia with low-dose aspirin: a systematic review and meta-analysis of the main randomized controlled trials. *Clinics* **2005,** *60* (5), 407-414.

9. Villar, J.; Purwar, M.; Merialdi, M.; Zavaleta, N.; Anthony, J.; De Greeff, A.; Poston, L.; Shennan, A., World Health Organisation multicentre randomised trial of supplementation with vitamins C and E among pregnant women at high risk for pre-eclampsia in populations of low nutritional status from developing countries. *BJOG.* **2009,** *116* (6), 780-788.

10. Kumar, A.; Devi, S. G.; Batra, S.; Singh, C.; Shukla, D. K., Calcium supplementation for the prevention of pre-eclampsia. *Int. J. Gynecol. Obstetr.* **2009,** *104* (1), 32-36.

11. Lam, C.; Lim, K.-H.; Karumanchi, S. A., Circulating angiogenic factors in the pathogenesis and prediction of preeclampsia. *Hypertension* **2005,** *46* (5), 1077-1085.

12. Venkatesha, S.; Toporsian, M.; Lam, C.; Hanai, J.-i.; Mammoto, T.; Kim, Y. M.; Bdolah, Y.; Lim, K.-H.; Yuan, H.-T.; Libermann, T. A., Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nat. Med.* **2006,** *12* (6), 642-649.

13. Kleinrouweler, C.; Wiegerinck, M.; Ris-Stalpers, C.; Bossuyt, P.; van der Post, J.; Von Dadelszen, P.; Mol, B.; Pajkrt, E., Accuracy of circulating placental growth factor, vascular endothelial growth factor, soluble fms-like tyrosine kinase 1 and soluble endoglin in the prediction of pre-eclampsia: a systematic review and meta-analysis. *BJOG,* **2012,** *119* (7), 778-787.

14. Sunderji, S.; Gaziano, E.; Wothe, D.; Rogers, L. C.; Sibai, B.; Karumanchi, S. A.; Hodges-Savola, C., Automated assays for sVEGF R1 and PlGF as an aid in the diagnosis of preterm preeclampsia: a prospective clinical study. *Am. J. Obstetr. Gynecol.* **2010,** *202* (1), 40. e1-40. e7.

15. Powers, R.; Roberts, J.; Cooper, K.; Gallaher, M.; Frank, M.; Harger, G.; Ness, R., Maternal serum soluble fms-like tyrosine kinase 1 concentrations are not increased in early pregnancy and decrease more slowly postpartum in women who develop preeclampsia. *Am. J. Obstetr. Gynecol.* **2005,** *193* (1), 185-191.

16. Chaiworapongsa, T.; Romero, R.; Kim, Y. M.; Kim, G. J.; Kim, M. R.; Espinoza, J.; Bujold, E.; Gonçalves, L.; Gomez, R.; Edwin, S., Plasma soluble vascular endothelial growth factor receptor-1 concentration is elevated prior to the clinical diagnosis of pre-eclampsia. *J. Maternal-Fetal and Neonatal Med.* **2005,** *17* (1), 3-18.

17. Stepan, H.; Kramer, T.; Faber, R., Maternal plasma concentrations of soluble endoglin in pregnancies with intrauterine growth restriction. *J. Clin. Endocrinol. Metabol.* **2007,** *92* (7), 2831-2834.

18. Merrell, K.; Southwick, K.; Graves, S. W.; Esplin, M. S.; Lewis, N. E.; Thulin, C. D., Analysis of low-abundance, low-molecular-weight serum proteins using mass spectrometry. *J. Biomol. Tech.* **2004,** *15* (4), 238.

19. Merrell, K.; Thulin, C. D.; Esplin, M. S.; Graves, S. W., Systematic Internal Standard Selection for Capillary Liquid Chromatography–Mass Spectrometry Time Normalization to Facilitate Serum Proteomics. *J. Biomol. Tech.* **2008,** *19* (5), 320.

20. Agresti, A., *An introduction to categorical data analysis*. Wiley New York: 1996; Vol. 135.

21. Devijver, P. A.; Kittler, J., *Pattern recognition: A statistical approach*. Prentice-Hall London: 1982; Vol. 761.

22. Ivanova, P. T.; Milne, S. B.; Myers, D. S.; Brown, H. A., Lipidomics: a mass spectrometry based systems level analysis of cellular lipids. *Curr. Opin. Chem. Biol.* **2009,** *13* (5), 526-531.

23. Ellis, S. R.; Wu, C.; Deeley, J. M.; Zhu, X.; Truscott, R. J.; Cooks, R. G.; Mitchell, T. W.; Blanksby, S. J., Imaging of human lens lipids by desorption electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2010,** *21* (12), 2095-2104.

24. Açmaz, G.; Tayyar, A.; Tayyar, M., Assessment of the Role of Fibrinogen in Preeclampsia. *Erciyes Med. J./Erciyes Tip Dergisi* **2008,** *30* (3).

25. Konijnenberg, A.; Stokkers, E. W.; van der Post, J. A.; Schaapb, M. C.; Boer, K.; Bleker, O. P.; Sturk, A., Extensive platelet activation in preeclampsia compared with normal pregnancy: enhanced expression of cell adhesion molecules. *Am. J. Obstetr. Gynecol.* **1997,** *176* (2), 461-469.

26. Denny, K. J.; Woodruff, T. M.; Taylor, S. M.; Callaway, L. K., Complement in pregnancy: a delicate balance. *Am. J. Reprod. Immunol.* **2013,** *69* (1), 3-11.

27. Derzsy, Z.; Prohászka, Z.; Rigó Jr, J.; Füst, G.; Molvarec, A., Activation of the complement system in normal pregnancy and preeclampsia. *Mol. Immunol.* **2010,** *47* (7), 1500-1506.

28. Pineiro, M.; Andres, M.; Iturralde, M.; Carmona, S.; Hirvonen, J.; Pyörälä, S.; Heegaard, P. M.; Tjørnehøj, K.; Lampreave, F.; Pineiro, A., ITIH4 (inter-alpha-trypsin inhibitor heavy chain 4) is a new acute-phase protein isolated from cattle during experimental infection. *Infect. Immun.* **2004,** *72* (7), 3777-3782.

29. Esplin, M. S.; Merrell, K.; Goldenberg, R.; Lai, Y.; Iams, J. D.; Mercer, B.; Spong, C. Y.; Miodovnik, M.; Simhan, H. N.; van Dorsten, P., Proteomic identification of serum peptides predicting subsequent spontaneous preterm birth. *Am. J. Obstetr. Gynecol.* **2011,** *204* (5), 391. e1-391. e8.

30. Bowen, R. S.; Zhang, Y.; Gu, Y.; Lewis, D. F.; Wang, Y., Increased phospholipase A< sub> 2</sub> and thromboxane but not prostacyclin production by placental trophoblast cells from normal and preeclamptic pregnancies cultured under hypoxia condition. *Placenta* **2005,** *26* (5), 402-409.

31. Wu, J.; Teuber, K.; Eibisch, M.; Fuchs, B.; Schiller, J., Chlorinated and brominated phosphatidylcholines are generated under the influence of the Fenton reagent at low pH—a MALDI-TOF MS study. *Chem. Phys. Lipids* **2011,** *164* (1), 1-8.

32. Packer, C. S., Reactive oxygen species and pre-eclampsia. *J. hypertension* **2003,** *21* (2), 263-264.

33. Touyz, R. M., Reactive Oxygen Species, Vascular Oxidative Stress, and Redox Signaling in Hypertension What Is the Clinical Significance? *Hypertension* **2004,** *44* (3), 248-252.

34. Eibisch, M.; Zellmer, S.; Gebhardt, R.; Süß, R.; Fuchs, B.; Schiller, J., Phosphatidylcholine dimers can be easily misinterpreted as cardiolipins in complex lipid mixtures: a matrix-assisted laser desorption/ionization time-of-flight mass spectrometric study of lipids from hepatocytes. *Rapid Commun. Mass Spectrom.* **2011,** *25* (18), 2619-2626.

35. Brady, J. J.; Judge, E. J.; Levis, R. J., Analysis of amphiphilic lipids and hydrophobic proteins using nonresonant femtosecond laser vaporization with electrospray post-ionization. *J. Am. Soc. Mass Spectrom.* **2011,** *22* (4), 762-772.

36. Goetzinger, K. R.; Singla, A.; Gerkowicz, S.; Dicke, J. M.; Gray, D. L.; Odibo, A. O., Predicting the risk of pre-eclampsia between 11 and 13 weeks' gestation by combining maternal characteristics and serum analytes, PAPP-A and free β-hCG. *Prenat. Diagn.* **2010,** *30* (12-13), 1138-1142.

37. Lockwood, C. J.; Senyei, A. E.; Dische, M. R.; Casal, D.; Shah, K. D.; Thung, S. N.; Jones, L.; Deligdisgh, L.; Garite, T. J., Fetal fibronectin in cervical and vaginal secretions as a predictor of preterm delivery. *N. Eng. J. Med.* **1991,** *325* (10), 669-674.

Chapter 3 Detection and Confirmation of Serum Lipid Biomarkers for Preeclampsia Using Direct Infusion Mass Spectrometry

Disclaimer- This chapter is mainly reproduced from a manuscript officially accepted by Journal of Lipid Research. Swati Anand, Sydney Young, Sean M. Esplin, Benjamin Peaden, H. Dennis Tolley, T. Flint Porter, Michael W. Varner, Mary E. D'Alton, Bruce Jackson, Steven W. Graves

3.1 Abstract

Despite substantial research, the early diagnosis of preeclampsia remains elusive. Lipids are now recognized to be involved in regulation and pathophysiology of some disease. Shotgun lipidomic studies were undertaken to determine if serum lipid biomarkers exist that predict preeclampsia later in the same in pregnancy. A discovery study was performed using sera collected at 12-14 weeks pregnancy from 27 controls with uncomplicated pregnancies and 29 cases that later developed preeclampsia. Lipids were extracted and analyzed by direct infusion into a time-of-flight mass spectrometer. MS signals, demonstrating apparent differences were selected, their abundances determined and statistical differences tested. Statistically significant lipid markers were reevaluated in a second confirmatory study having 43 controls and 37 preeclampsia cases. Multi-marker combinations were developed using those lipid biomarkers confirmed in the second study. The initial study detected 45 potential preeclampsia markers. Of these, 23 markers continued to be statistically significant in the second confirmatory set. Most of these markers representing several lipid classes were chemically characterized typically providing lipid class and potential molecular components using tandem MS. Several multi-marker panels with AUC>0.85 and high predictive values were developed. Developed panels of serum lipidomic biomarkers

appear able to identify most women at risk for preeclampsia in a given pregnancy at 12-14 weeks gestation.

3.2 Introduction

Preeclampsia is a potentially life-threatening disorder of pregnancy characterized by new-onset hypertension and proteinuria after 20 weeks gestation. It constitutes a leading cause of maternal and perinatal mortality and morbidity[1-2]. Estimates are that as many as 75,000 women worldwide die each year from preeclampsia[3]. Treatment options are very limited and frequently require termination of the pregnancy, regardless of gestational age, accounting for ~20% of all preterm births[4]. Furthermore, infants born to preeclamptic mothers may be at an increased risk of hypertension, heart disease and diabetes beyond their being premature[5].

While there are known risk factors for preeclampsia, it is not yet possible to precisely distinguish pregnancies destined to develop preeclampsia from those that will not. This has made prospective clinical studies large and expensive. Also, it has to some degree limited studies of early changes that may lead to preeclampsia. In the absence of a complete animal model of preeclampsia, the actual cause or causes of this disorder remains unknown. The pathogenesis is acknowledged to be complex and although incompletely understood, it is generally believed that one or more very early events in the pregnancy contribute to preeclampsia. One long held theory to explain this disease involves an incomplete remodeling of maternal spiral arteries by invasive extravillous placental trophoblast cells resulting in inadequate perfusion of the fetal-placental unit with attendant ischemia[6-7]. Abnormal waveform patterns and an increased pulsatility index observed in uterine Doppler ultrasound studies support the concept of an underperfused fetus prior

to clinically apparent preeclampsia[6-7]. There is also evidence to suggest that biochemical changes may be seen in women who, later in the same pregnancy, develop this disease. Several other biochemical changes appear to precede by a few weeks clinically evident preeclampsia, including hypoxia-reoxygenation[8], abnormal expression of angiogenic and anti-angiogenic factors in the maternal circulation and endothelial dysfunction with a pro-inflammatory response[9-10]. Collectively, these data suggest that there are biochemical abnormalities that occur prior to clinically evident preeclampsia signs and symptoms. However, to date there are still no accepted, predictive biomarkers for this disease[11-13] despite some initial promise for a number of proposed candidates. Also currently, the specific causes (as opposed to consequences) of preeclampsia are still debated and have yet to be established.

It is now possible to survey hundreds to thousands of molecules in biological specimens in an unbiased way using mass spectrometry (MS). Proteomics is by far and away the most common of these approaches and has been employed to study several diseases, but its use to explore diagnostic or predictive biomarkers in serum is difficult due to ~30 highly abundant serum proteins that mask the vast majority of lower abundance species in serum due to ion suppression. Other methods have targeted peptides and lipids, although much less frequently in serum or plasma.

Lipids are increasingly recognized as having important biological roles or representing important biochemical correlates of clinical changes. A wide variety of human diseases are associated with aberrant lipid metabolism including Alzheimer's disease, diabetes and atherosclerosis[14-16]. Alterations in lipids may represent by-products of underlying pathophysiology but could also represent primary disease mediators. For example, arachidonic acid is a precursor

for eicosanoids that have a significant role in inflammatory processes. There is evidence of changes in downstream products of arachidonic acid in preeclampsia[17]. As another example, oxidized lipid species can reflect increased reactive oxygen species (ROS) which are produced as a part of several diseases[18-20]. Therefore, there is ample reason to study lipid profiles in pregnant women at an early stage seeking predictive lipidomic biomarkers to identify patients at significant risk for preeclampsia. Consequently, a 'global' serum lipidomic approach, involving lipid extraction followed by direct injection, time-of-flight mass spectrometry was used to identify and chemically characterize predictive serum preeclamptic lipid biomarkers. We hypothesized that this approach would find individual lipids and sets of lipids that would allow for the prediction of a substantial portion of women who would later develop this disease in the same pregnancy and that the changes would provide insights into the mechanisms early in the development of preeclampsia.

3.3 Methods

3.3.1 Study population

Serum specimens used for both the discovery and confirmatory studies were obtained from the Department of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, UT and from Intermountain Health Care (Intermountain), Murray, UT. All samples were banked sera, obtained from a previously completed clinical study, "First- And Second-Trimester Evaluation of Risk"[21]. Institutional Review Board (IRB) approval for these studies was obtained from the University of Utah, Intermountain and Brigham Young University prior to our initiating these experiments. Sera had been collected from pregnant women at 12-14 weeks gestation

followed through the completion of their pregnancies. Specimens were analyzed without clinical identifiers.

A discovery cohort involved sera from 27 controls, having term, uncomplicated pregnancies and 29 cases, who developed preeclampsia later during the index pregnancy. The second confirmatory study of the promising potential lipid biomarkers from the discovery cohort involved serum from 37 cases and 43 controls, also collected at 12-14 weeks gestation. Demographic characteristics are summarized in Tables 3.1 and 3.2.

3.3.2 Materials

Glycerophosphocholine lipid standards: PC (14:0/16:0) and PC (18:0/18:2) were purchased from Avanti Lipids (Alabaster, AL, USA). A triacylglycerol standard: TG (16:0/18:0/16:0) was purchased from Sigma Aldrich (St. Louis, MO, USA).

3.3.3 Sample preparation

All specimens had been stored at -80^{0}C until our receiving them on dry ice and were maintained at -80^{o}C before and after sample processing. Serum lipid extraction efficiency of five different organic extraction solutions were performed as follows: for a 200 µL sample 1) 2.5 mL of 4:1 chloroform:methanol; 2) 2.5 mL of 4:1methyl-tert-butyl ether:methanol; 3) 1.8 mL of 3:2 hexane:isopropanol; 4) 2.3 mL of 3.6:1hexane:methanol; or 5) 2.0 mL of 3:1benzene:methanol. Of these the procedure producing both the highest number of MS features and best signal to noise was chosen for use here. This was a modified, previously described extraction protocol that involved a solvent mixture of hexane: isopropanol (3:2)[22]. To 200 µL of serum, 1.8 mL hexane:

Table 3.1 Demographics for the First, Discovery Study

	Maternal age (yrs)	Gestational age (wks)	BMI (kg/m^2)	SBP	DBP	Birthweight (gm)	Race
Cases	30 ± 4.6	36.4 ± 3.07	30.8± 5.42	149.5± 17.4	90.1± 12.3	2755 ± 801	28/29 Cau
Controls	27 ± 4.8	39 ± 1.23	30.4 ± 5.6	138.6± 14.7	80 ± 12.5	3459 ± 349	27/27 Cau

Note: BMI = body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; Cau = Caucasian

Table 3.2 Demographics for the Second, Confirmatory Study

	Maternal age (yrs)	Gestational age (wks)	BMI (kg/m^2)	SBP	DBP	Birthweight (gm)	Race
Cases	29 ± 4.7	36.3 ± 3.2	31.7 ± 5.89	147 ± 20.8	90 ± 15.1	2817 ± 860.9	37/37 Cau
Controls	26 ± 4.7	38.7 ± 1.6	27.9 ± 4.7	117 ± 14.4	70.4 ± 8.2	3295 ± 535.2	43/43 Cau

Note: BMI = body mass index; SBP = systolic blood pressure; DBP = diastolic blood pressure; Cau = Caucasian

isopropanol (3:2) and 300 µL of 0.5 M KH_2PO_4 were added in a glass tube followed by vigorous vortexing. Samples were further agitated on a shaking platform for 1 h at room temperature at the speed of 80 rpm. To complete the extraction, 150 µL of water were added, mixed and centrifuged at 2000 rpm for 12 min. The upper, organic phase containing lipids was collected and dried completely under nitrogen. Dried lipid extracts were redissolved in 200 µL of chloroform: methanol (3:1) and stored at -80 ^{0}C.

3.3.4 Mass spectrometric analysis of the lipid extract

To a 20 µL aliquot of the sample extract, 23 µL of chloroform, 46 µL of methanol and 14 µL of 12 µM ammonium acetate were added. The samples were directly injected into the mass spectrometer (6230 TOF LC/MS Agilent Technologies) through an electrospray ionization (ESI) source operated in the positive ion mode. A syringe pump was utilized to inject samples at the flow rate of 2 µL/min. The ESI source employed a microspray needle having an i.d. of 50 µm. The capillary voltage was set at 3500 V. MS data was collected over mass to charge ratios (*m/z*) of 100-3000 with an acquisition rate of 1 spectrum/sec. Nebulizer gas and dry gas parameters were optimized to obtain a stable flow. The dry gas was set to 5 L/min at 325^{0}C with a nebulization gas pressure of 1.03 bar. Mass Hunter-Qualitative software [Agilent Technologies] was used for data analysis. Each specimen generated a mass spectrum from which the total ion chromatogram (TIC the sum of all ion counts) from *m/z* 100 to 3000 was determined. A peak list having *m/z* values for all peaks with their abundances was generated from the mass spectrum.

To reduce analytical variability all MS peaks were normalized. For normalization, 7 abundant peaks representing different classes of lipids but showing similar abundance in both case

and control sera ($p>0.40$) were chosen as a reference set and their abundances averaged. The *m/z* values and classes for these peaks are as follows: *m/z* 203 (fatty acid), 369 (sterol), 666 (cholesterol ester), 758 (phosphatidyl choline (PC)), 782 (PC), 810 (PC) and 848 (triacylglycerol). These peaks were consistently seen in all the samples. The ratio of the average intensity of these 7 peaks was comparable to or more consistent than the use of TIC counts for peak normalization, but avoided the occasional high TIC for a few MS runs with a high background. Therefore, the average of the combined intensity of these 7 peaks was used to normalize all peaks across the each individual mass spectrum.

For the second confirmatory study, all the samples were processed using the same method and were directly injected onto the same MS instrument at a flow rate of 10 μL/min, using a standard sprayer having an i.d. of 120 μm. The microbore needle had previously been blocked on several occasions. The larger bore needle used here eliminated that problem. The increased flow rate and shorter run times kept the TIC, peak list and abundances comparable to the initial study. All the other parameters were kept identical. All the potential significant markers from the initial, discovery set were reanalyzed for their performance in the second confirmation set using Student's T-test.

3.3.5 Chemical characterization of the replicating lipid biomarkers

After candidate biomarkers were found and confirmed significant by ESI-TOFMS, tandem MS (MS^2) was performed to fragment and characterize all validated biomarkers. Lipid markers were fragmented using a QSTAR Pulsar 1 quadrupole (Sciex, Framingham, MA) and as needed an Agilent 6530 Q-TOF MS (Agilent, Santa Clara, CA) in the positive ion mode. For the QSTAR,

samples were injected directly at a flow rate of 2 μL/min. The capillary voltage was set to 4200 V. The selection of the mass range targeted for fragmentation depended upon the *m/z* of the precursor ion with a spectral acquisition rate of 1 spectrum/sec. Declustering and focusing potentials were set to 65 V and 290 V respectively. Nitrogen and/or argon gas was used for collisionally-induced fragmentation. Multiple fragmentation energies were employed to obtain as complete fragmentation as possible. MS2 spectra were collected for 2 min and the multi-channel analyzer (MCA) function was turned on resulting in summation of all 120 MS2 spectra together thus increasing signal to noise ratio.

For the Agilent 6530, sample injection was carried out at the flow rate of 10 μL /min. The capillary voltage was set to 3500 V. The drying gas flow rate and the temperature were 5 L/min and 300^0C. MS2 spectra were collected from *m/z* 50-3000 and the spectral acquisition rate was 3 spectra /sec. Collision energies were optimized depending upon the precursor to obtain maximum fragmentation coverage. Individual scans were summed using the add feature of the Mass Hunter program to obtain greater signal to noise. The targeted MS2 mode was used to isolate and fragment the precursor ion.

Exact mass studies were done while using a set of internal standards. Lipidmaps.org was used to tentatively determine the lipid class from the different possible classes for a particular precursor ion. Furthermore, product ions, neutral losses, fragmentation patterns of standards or fragmentation patterns found in the literature, in conjunction with predicted fragments from the databases and predicted elemental compositions based on exact mass studies were compared with

the actual observed fragments to determine the class and to evaluate the possible or probable components of the individual candidate markers.

3.3.6 Statistical Analyses

For comparisons between the cases and controls, the normalized intensities of all the peaks (abundance > 200 ion counts) in the first study were subjected to a two-tailed Student's t-test as an initial statistical comparison. A p-value <0.05 for any peak, was used to designate a potential biomarker. It is recognized that since the t-test statistic was performed for multiple peaks, separately, the true p-values are much larger than 0.05. Consequently, any candidate biomarker potentially may not be statistically significant between the cases and controls. However, using a one-at-a-time p-value calculation with a minimum threshold reduces the list of potential biomarkers. We augment this list by adding peaks that had low p-values close to 0.05 but greater than this threshold, resulting in a list of 45 potential biomarkers. Those peaks that continued to show differential expression, as measured by a p-value of less than 0.05 as determined from a one-at-a-time t-test in the second, independent set of specimens were considered highly likely to be biomarkers. Given the goal of generating panels of biomarkers, candidates with a p-value between 0.05 and 0.10 were also included because it is recognized that some markers may be selective for a subgroup within the broader diagnosis of preeclampsia and hence may be complementary to other markers. For this reason, they were considered in potential biomarker panels and were modeled to provide better sensitivities and specificities. Biomarker panel development employed a forward selection; leave one out, logistic regression analysis approach[23-24]. In modeling the MS data, there were a few peaks in a small number of specimens for which intensities could not be determined due either to ion suppression by other nearby peaks or because of a sub-threshold

abundance of a specific peak in that sample. To account for these missing values, the values for these peaks were estimated using multivariate imputation by chained equations. The method of chained equations used to predict the missing values was through predictive mean matching. This uses the values of the other samples to impute missing values, allowing for all peaks to still be used[25]. Using logistic regression analysis receiver operator characteristic curves were generated for these panels to allow for the determination of sensitivity (true positive) and specificity (1-true negative or false positive rate).

3.4 Results

3.4.1 Serum lipid preeclamptic biomarkers in a discovery study

The hypothesis of this study was that there would be serum lipid biomarkers present at 12-14 wks pregnancy predictive of preeclampsia later in the same pregnancy. Using a global, serum lipidomic approach, the initial study of 27 controls and 29 cases found 45 candidate markers that were statistically or near statistically different in women who developed the disease compared with women who had uncomplicated pregnancies. The candidate markers are listed in table 3.3.

3.4.2 Replicating serum lipid biomarkers established in the second confirmatory study

A second confirmation study of the 45 candidate biomarkers was performed to evaluate their performance in a second set of specimens processed and analyzed comparably. This set, also collected at 12-14 wks pregnancy, involved 43 controls having uncomplicated term pregnancies and 37 cases having preeclampsia later in the same pregnancy.

Table 3.3 Candidate Preeclamptic Lipid Biomarkers

No.	*m/z*	P value	Higher in
1.	211.0	0.1	Cases
2.	228.2	0.07	Cases
3.	239.1	0.06	Cases
4.	256.2	0.08	Cases
5.	257.1	0.07	Cases
6.	263.2	0.08	Cases
7.`	280.9	0.04	Controls
8.	282.2	0.08	Cases
9.	301.2	7.1 X 10^-6	Controls
10.	383.3	2.5 X 10^-8	Cases
11.	425.1	0.08	Controls
12.	430.2	0.05	Cases
13.	462.3	0.05	Cases
14.	445.4	3.4 X 10^-4	Controls
15.	520.3	0.08	Cases
16.	531.4	0.02	Cases
17.	548.4	0.05	Cases
18.	642.6	0.03	Cases
19.	645.5	0.06	Controls
20.	663.5	2.5 X 10^-5	Cases
21.	714.6	0.009	Controls
22.	734.6	0.07	Cases
23.	760.6	0.08	Cases
24.	774.5	0.04	Cases
25.	784.6	0.05	Cases
26.	894.7	0.02	Cases
27.	788.6	0.06	Cases
28.	796.6	0.09	Cases
29.	798.6	0.03	Cases
30.	810.6	0.1	Cases
31.	836.6	0.1	Cases
32.	838.6	0.07	Cases
33.	895.7	0.02	Controls
34.	896.7	0.02	Controls
35.	898.8	0.02	Controls
36.	904.8	0.08	Controls
37.	916.8	0.01	Controls
38.	920.7	0.009	Controls
39.	922.7	0.001	Controls
40.	924.7	0.003	Controls
41.	926.8	0.005	Controls
42.	928.8	0.006	Controls
43.	954.8	0.01	Controls
44.	956.8	0.004	Controls
45.	958.8	0.005	Controls

Of the 45 potential biomarkers, 23 continued to show a statistically significant or near significant p-values when considered one at a time. These markers are listed table 3.4.

3.4.3 Multi-marker panel development

Statistical modeling, using logistic regression analysis, was performed on the 23 replicating markers to develop multi-marker panels with higher predictive values. Several panels having combinations of 3-6 markers were obtained demonstrating areas under the curve (AUC) >0.85 for receiver operator characteristic (ROC) curves as illustrated in table 3.5. The ROC curves for two multi-marker combinations having AUC of 0.89 (sensitivity ~91% at a specificity of ~82% and the second ~86% and specificity ~81%) are shown in Figures 3.1 and 3.12. A plot showing the classification of cases and controls by means of a panel of 6 markers, m/z 263.3, 383.3, 462.3, 645.5, 784.6 and 920.7 is shown in Figure 3.3. The overall AUC of this set is 0.88.

3.4.4 Chemical characterization of the validated serum lipid preeclamptic biomarkers

While absolute chemical structures of most lipids are not possible using MS, substantial chemical characterization is often provided by tandem MS fragmentation studies. Tandem MS studies were performed on all 23 confirmed serum lipid biomarkers. This provided lipid classes, exact mass studies suggested possible or probable elemental composition and fragments representing neutral losses suggested possible or probable structural components for most of the confirmed lipid biomarkers. The molecular features most consistent with the data are described in Table 3.6.

Table 3.4 Candidate Biomarkers that Continued to be Statistically Significant in the Second Confirmation Study

S. No.	*m/z*	P value	Higher in
1.	263.2	0.07	Cases
2.	301.2	0.005	Controls
3.	383.3	0.05	Cases
4.	425.1	7.8 X 10^-5	Controls
5.	462.3	0.04	Cases
6.	445.4	0.03	Controls
7.	645.5	0.0001	Controls
8.	714.6	0.008	Controls
9.	734.6	0.1	Cases
10.	760.6	0.06	Cases
11.	784.6	0.005	Cases
12.	788.6	0.08	Controls
13.	796.6	0.06	Cases
14.	798.6	0.06	Cases
15.	810.6	0.007	Cases
16.	836.6	0.06	Cases
17.	895.7	0.009	Controls
18.	916.8	0.003	Controls
19.	920.8	0.003	Controls
20.	928.8	0.05	Controls
21.	954.8	0.003	Controls
22.	956.8	0.003	Controls
23.	958.8	0.0008	Controls

Table 3.5 Multi-Marker Panels with AUC > 0.85

AUC	Modeled Marker Combinations
0.89	263, 383, 445, 645, 784, 916
0.89	383, 462, 445, 645, 784, 810
0.89	383, 445, 784, 796, 798, 920
0.88	263, 383, 645, 784, 836
0.86	263, 383, 645, 784

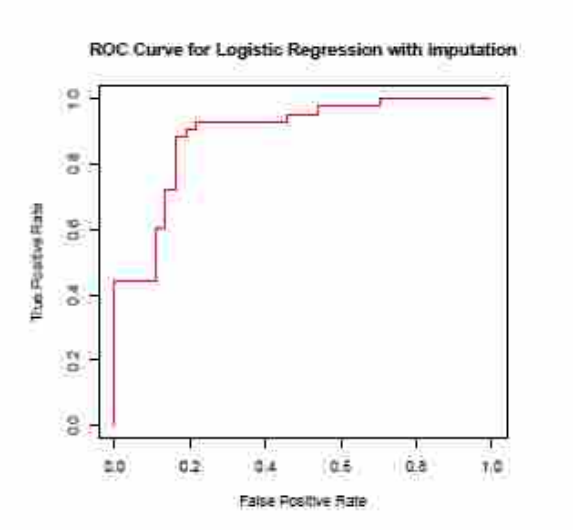


Figure 3.1 Resulting receiver operator characteristic curve generated by logistic regression analysis that included 6 lipid biomarkers having mass to charge ratios of 383, 445, 784, 796, 798 and 920. The AUC was 0.89 with a sensitivity of ~91% at a specificity of ~82%.

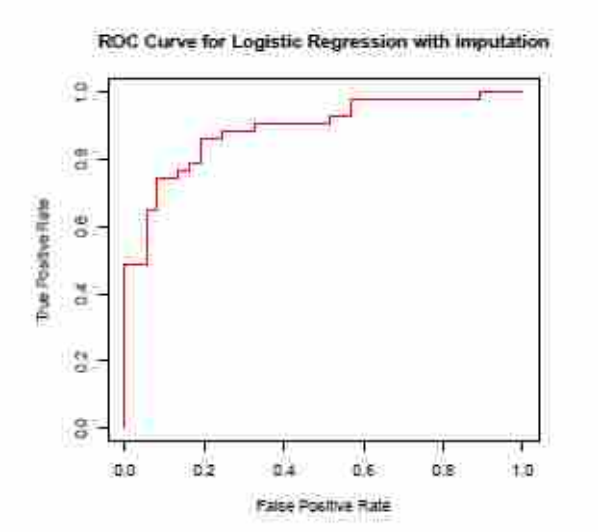


Figure 3.2 Resulting receiver operator characteristic curve generated by logistic regression analysis of another combination of 6 lipid biomarkers having mass to charge ratios of 263, 383, 445, 645, 784 and 916. Some peaks are common to the two panels. The AUC of this set was also 0.89 with a sensitivity of ~86% at a specificity of ~81%.

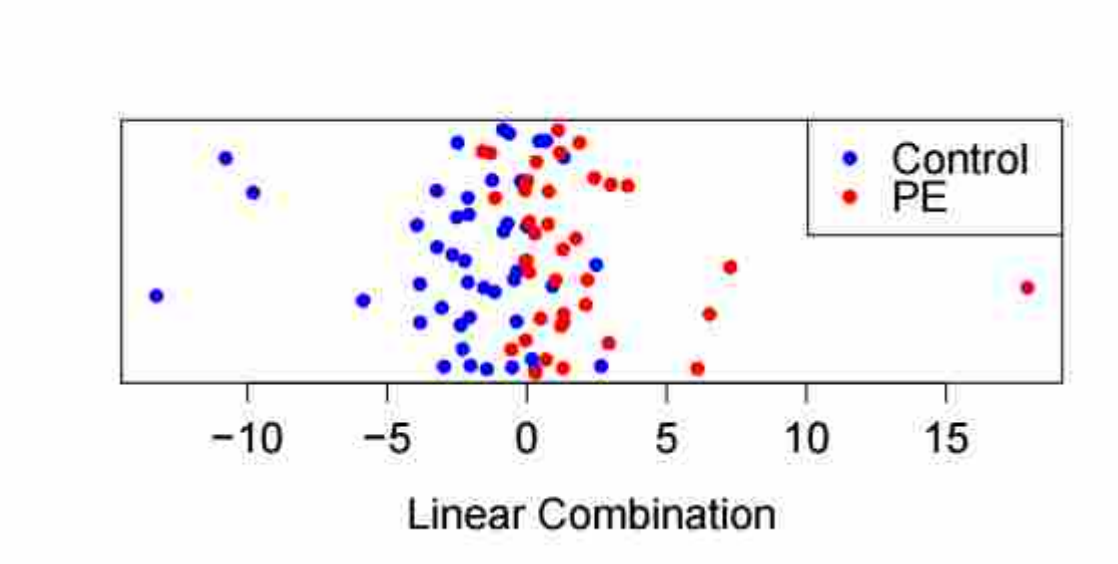


Figure 3.3 A plot showing the ability of a panel of 6 biomarkers, m/z 263.3, 383.3, 462.3, 645.5, 784.6, 920.7, to correctly classify cases and controls.
Controls are shown in blue and cases are displayed in red. Each point represents a different patient. The overall AUC is 0.88 for this set of markers

Table 3.6 Chemical Characterization of Preeclamptic lipid Biomarkers

	m/z	Adduct	Class	Possible identity	
1	263.2	$M+H^+$	Fatty alcohol and aldehydes		
2	301.2	$M+H^+$	Fatty acids and conjugates		
3	383.3	M^+	Oxidized cholesterol		
4	425.1	------	----------	----------	
5	445.4	$M+H^+$	Cholesterol derivatives		
6	462.3	-------	----------	----------	
7	645.5	$M+H^+$	Cholesteryl esters	C18:4 cholesteryl ester	
8	714.6	$M+NH_4^+$	Cholesteryl esters	C22:6 cholesteryl ester	
9	734.6	$M+H^+$	Glycerophosphocholine	PC-32:0	16:0 and 16:0
10	760.6	$M+H^+$	Glycerophosphocholine	PC-34:1	16:0 and 18:1
11	784.6	$M+H^+$	Glycerophosphocholine	PC-36:3	18:1 and 18:2
12	788.6	$M+H^+$	Glycerophosphocholine	PC-36:1	18:0 and 18:1
13	796.6	$M+H^+$	Glycerophosphocholine	PC-O-38:4	O-18:0 and 20:4
14	798.6	$M+H^+$	Oxidized Glycerophosphocholine	PC-36:4+OH	16:0 and 20:4 +OH
15	810.6	$M+H^+$	Glycerophosphocholine	PC-38:4	18:0 and 20:4
16	836.6	$M+H^+$	Glycerophosphocholine	PC-40:5	18:0 and 22:5

17	895.7	$M+H^+$	Oxidized sphingomyelin		
18	916.8	$M+NH_4^+$	Oxidized triacylglycerol		OOOepoxide
19	920.7	$M+NH_4^+$	Triacylglycerol	TG-56:8	16:0, 18:2 and 22:6
20	928.8	$M+NH_4^+$	Triacylglycerol		LLL-mono-hydroperoxide
21	954.8	$M+NH_4^+$	Triacylglycerol		
22	956.8	$M+NH_4^+$	Triacylglycerol		
23	958.8	$M+NH_4^+$	Triacylglycerol		

Fragmentation results with any chemical structural information are summarized briefly for the validated markers:

The biomarker *m/z* 263.2, when fragmented, demonstrated several hydrocarbon fragments in the low *m/z* region consistent with its having an alkane region. This marker's fragmentation pattern was most consistent with that of a fatty alcohol or aldehyde. Its elemental composition of $C_{18}H_{30}O$, as determined on LIPID MAPS indicates a fatty alcohol or aldehyde. This compound could also correspond to an $[acyl]^+$ ion for linoleic acid as reported previously[26]. A fragment ion at *m/z* 245.2, predicted to result from the $(linoleoyl\text{-}H_2O)^+$ ion, was also observed[26] in our MS^2 studies of the precursor *m/z* 263.2.The ions at *m/z* 263.2 and *m/z* 245.2 have also been observed in the MS^2 spectrum of linoleic acid (*m/z* 298 for $M+NH_4^+$) in previous studies[26]. However, the occurrence of *m/z* 263.2 corresponding to the $(linoleoyl)^+$ ion was not observed in the MS^1 spectrum in that report indicating that *m/z* 263.2 is more likely to be a protonated fatty alcohol or aldehyde and not the $(linoleoyl)^+$ ion The marker with *m/z* 301.2 was predicted to have an elemental composition of $C_{20}H_{28}O_2$ consistent with its being a fatty acid or fatty acid conjugate.

Another marker having *m/z* 383.3 produced a fragmentation pattern consistent with it being an oxidized (keto or epoxy) cholesterol[27-28]. The lipid ion peak at *m/z* 383.3 has been identified to be dehydrated 7-ketocholestrol in a number of published studies involving human blood samples[29-31], although, we recognize that oxidation of cholesterol at some other position is possible and would result in similar fragmentation.

The species at *m/z* 445.4 was likely to be a cholesterol derivative as indicated by its fragmentation pattern. Fragmentation studies for two other markers having *m/z* 645.5 and *m/z*

714.6 resulted in a fragment ion at *m/z* 369.4, characteristic of cholesteryl esters, formed with octadecatetraenoic acid (C18:4) and docosahexaenoic acid (22:6) respectively.

To improve interpretation of fragmentation data for glycerophosphocholine (PC) lipids, studies were performed on a standard, PC (14:0/16:0). Figure 3.4a shows the MS^2 spectrum of protonated PC (14:0/16:0) having *m/z* 706.5 displaying a highly abundant signal at *m/z* 184.1 arising from the protonated phosphocholine moiety, characteristic of fragmentation of lysophosphatidylcholines (LPC), PC and SM species[32]. Figure 3.4b displays the fragment ion at *m/z* 468.3 corresponding to protonated LPC (14:0/0:0) and a fragment ion at *m/z* 496.3 corresponding to protonated LPC (16:0/0:0) resulting from the neutral losses of the fatty acyl chains. The peak at *m/z* 523.4 $(M+H-183)^+$ results from the loss of the entire phosphocholine head group from the precursor at *m/z* 706.5. These are all predicted fragments. No other fragments were seen for unpredicted truncations, combinations or any other unanticipated collision product. We also performed MS^2 studies on a second standard, PC (18:0/18:2) showing a protonated peak at *m/z* 786.6. Its fragmentation spectrum displayed fragment ions at *m/z* 184, 520.3 and 524.3 (Fig 3.5a and 3.5b). The signal at *m/z* 184 corresponds to protonated phosphocholine whereas *m/z* 520.3 and 524.3 correspond to LPC (18:2/0:0) and LPC (18:0/0:0) produced by neutral losses of the fatty acyl chains. Again, all fragments were exactly as predicted and no other fragments were observed. Studies involving identification of PC's based on the neutral losses of the fatty acyl chains are commonly observed in literature[33-35].

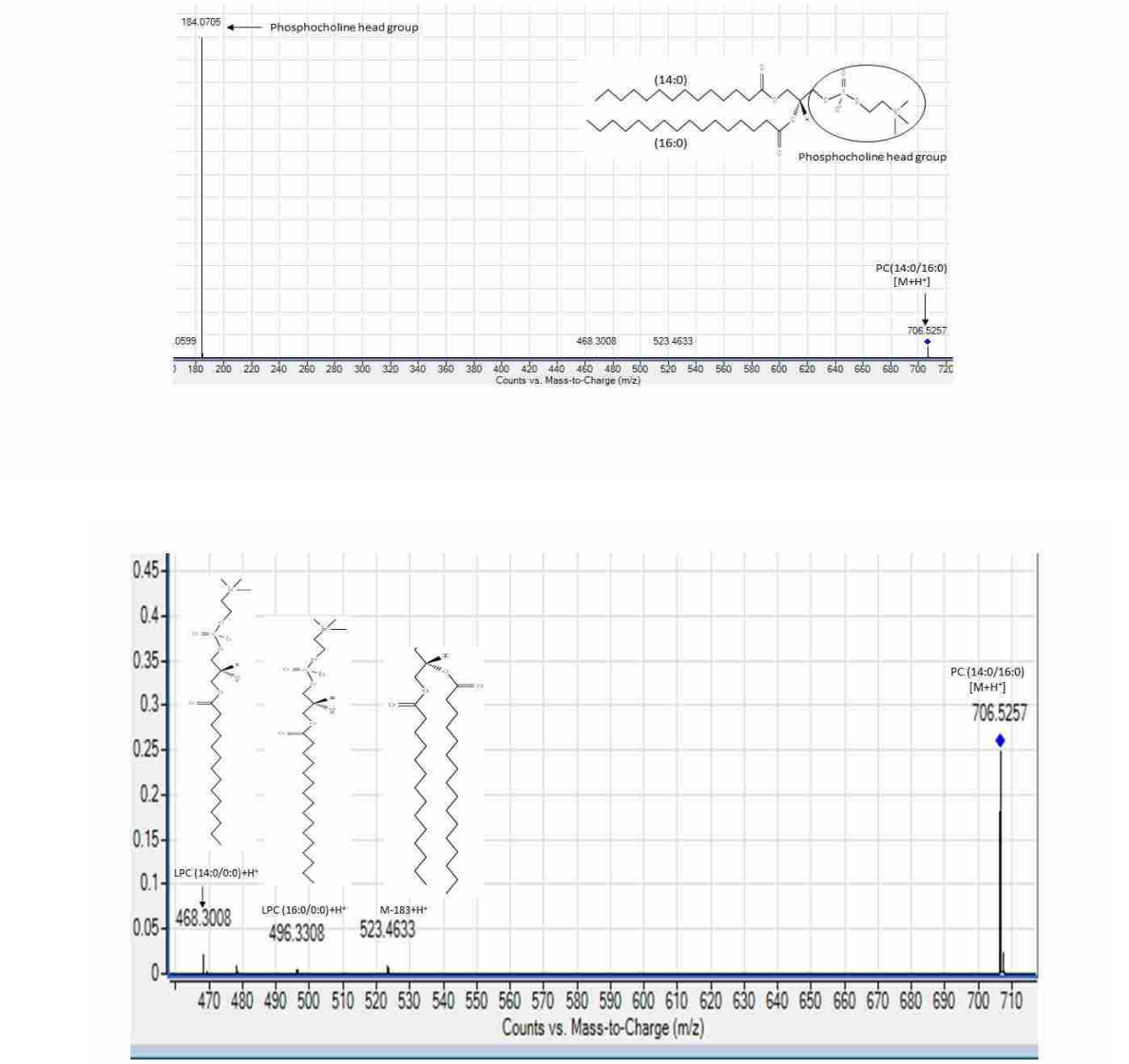


Figure 3.4 MS2 spectrum for a glycerophosphocholine standard PC (14:0/16:0) with *m/z* 706.5 ($M+H^+$). a) The abundant fragment ion at *m/z* 184 corresponds to the protonated phosphocholine head group.b) The fragment at *m/z* 468.3 corresponds to protonated LPC (14:0/0:0). Similarly the fragment at *m/z* 496.3 corresponds to protonated LPC (16:0/0:0). Neutral loss of the entire phosphocholine head group leads to the formation of the ion at $(M+H-183)^+$ at *m/z* 523.

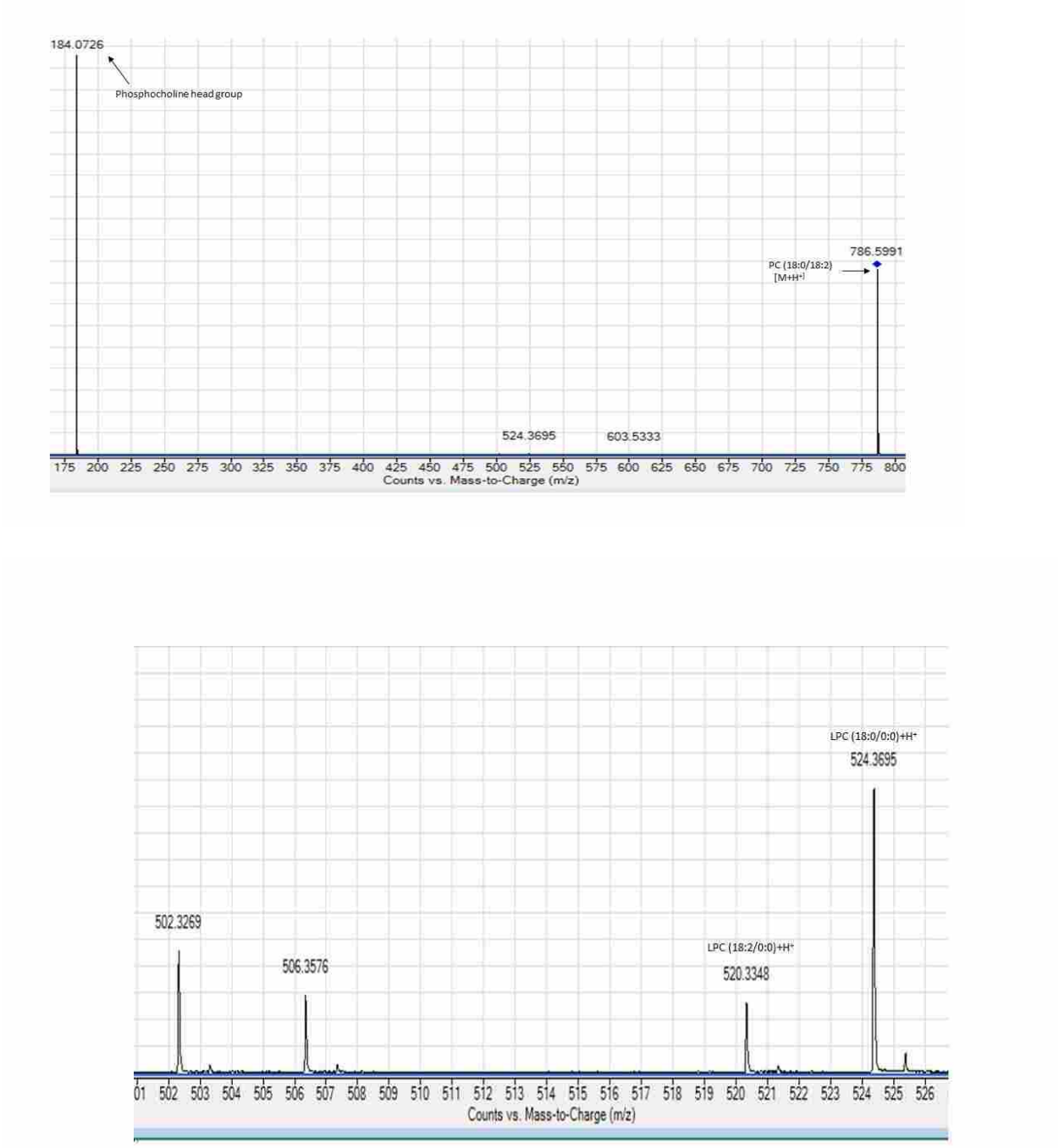


Figure 3.5 MS2 spectrum for a glycerophosphocholine standard PC (18:0/18:2) with *m/z* 786.6 ($M+H^+$).a) The abundant fragment ion at *m/z* 184 corresponds to protonated phosphocholine head group. Neutral loss of the entire phosphocholine head group leads to the formation of the ion at $(M+H-183)^+$ at *m/z* 603.5. b) The fragment at *m/z* 520.3 corresponds to protonated LPC (18:2/0:0). Similarly the fragment at *m/z* 524.3 corresponds to protonated LPC (18:0/0:0).

Collectively, our study and published literature suggest strongly that fragmentation patterns observed with PC lipid species follow predictable fragmentations.

Based on their fragmentation patterns, several of the lipid markers were categorized as PC lipids. The MS^2 fragmentation spectrum for markers with *m/z* values of 734.6, 760.6, 784.6, 788.6, 796.6, 798.6, 810.6, 836.6 and 895.7 displayed a prominent peak at *m/z* 184.07 corresponding to a phosphocholine moiety. Their elemental compositions were predicted using exact mass studies and the fatty acyl constituents were assigned based on the neutral losses from the precursor ion with help from the lipid MS predictor feature from LIPID MAPS and following the fragment patterns observed with PC standards.

The fragmentation spectrum of the marker with *m/z* 734.6 also displayed low abundance fragment ion with *m/z* 496.3 suggesting the presence of 16:0 fatty acyl chain[33] (Fig 3.6). Therefore, putative structure for *m/z* 734.6 is likely to be PC (32:0) $+H^+$ having two 16:0 fatty acyl chains. The MS^2 spectrum of the marker with *m/z* 760.6 displayed fragment ion peaks at *m/z* 496.3 and *m/z* 504.3 consistent with the presence of 16:0 and 18:1 fatty acyl chains[34]. Therefore, this lipid *m/z* 760.6 is likely to be PC (34:1)$+H^+$. Similarly, markers with *m/z* 784.6, 788.6, 810.6 and 836.6 were predicted to be PC (36:3)$+H^+$, PC (36:1)$+H^+$, PC (38:4)$+H^+$ and PC (40:5)$+H^+$ [34]. Based on the fragmentation patterns, the fatty acid components for these same markers *m/z* 784.6, 788.6, 810.6 and 836.6 were predicted to be (16:0 and 18:1), (18:0 and 18:1), (18:0 and 20:4) and (18:0 and 22:5) respectively[34].

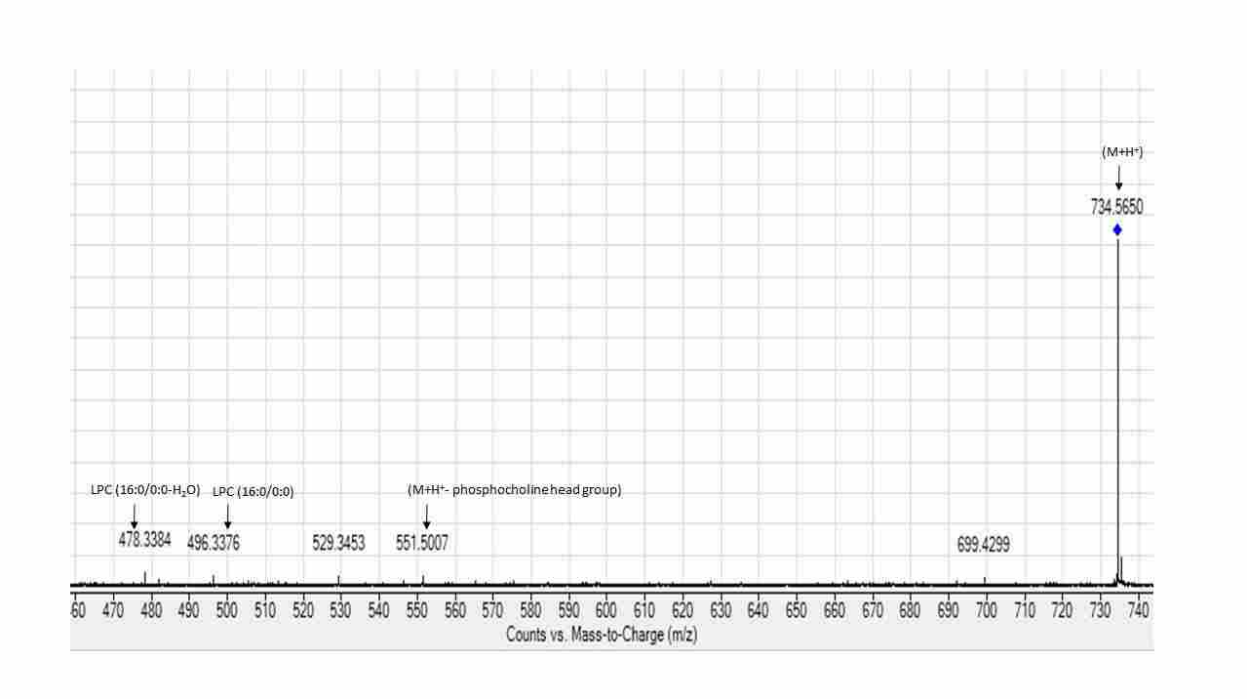


Figure 3.6 MS2 spectrum for a lipid marker with m/z 734.5.
Fragmentation suggests that it is likely to be a protonated adduct of PC -32:0. The fragment ion at m/z 496.3 is likely due to protonated LPC (16:0/0:0) resulting from the neutral loss of another 16:0 from the precursor ion.

The MS^2 spectrum of marker *m/z* 796.6 displayed a fragment ion at *m/z* 510.3 previously reported to indicate the presence of an O-18:0 fatty acyl chain[36]. Based on the several fragments observed, it is predicted to be PC(O-38:4) having a combination of O-18:0 and 20:4 fatty acyl chains consistent with the previously published literature[36].

The fragmentation spectrum of marker *m/z* 798.6 displayed a prominent fragment ion peak at *m/z* 780.6 likely to result from water loss from the precursor ion and is not otherwise commonly observed in the MS^2 spectrum of PC's. This major neutral water loss peak observed for this precursor, suggests strongly an easily removable hydroxyl group[37]. A fragment ion at *m/z* 496.3, most likely corresponding to 16:0 fatty acid, was also observed. This fragmentation pattern most likely represents a hydroxylated PC (16:0/20:4) and this agrees with previously published literature[37]. Therefore, marker with *m/z* 798.6 was predicted to be hydroxylated PC (36:4) $+H^+$.

Due to the presence of a fragment ion at *m/z* 184.07 and a peak indicating water loss from the precursor, the marker with *m/z* 895.7 likely represents an oxidized sphingomyelin (SM) (odd $M+H^+$ having even neutral mass). Oxidized SM's have not been commonly studied using mass spectrometry[38]. Therefore, the structure for *m/z* 895.7 could not be predicted with certainty.

Tandem MS studies were also performed on a triacylglycerol standard, TG (16:0/18:1/16:0) that displayed an ammoniated precursor peak at *m/z* 850.7. As shown in Figure 3.7, in the fragmentation spectrum of precursor *m/z* 850.7 ($M+NH_4^+$) a fragment ion at *m/z* 551.5 corresponding to diacylglycerol (DG) $16:0/16:0^+$ and a fragment ion at *m/z* 577.5 corresponding

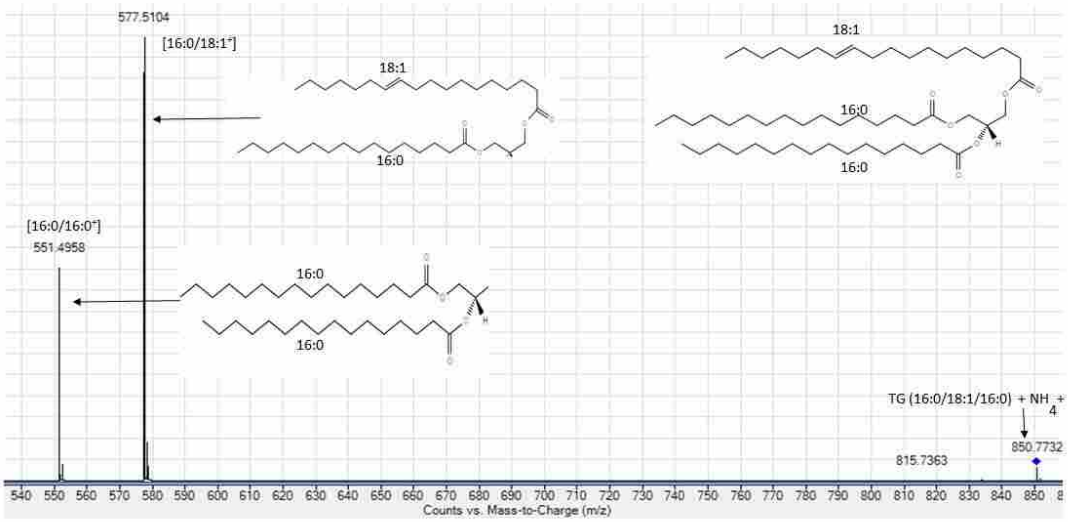


Figure 3.7 MS2 spectrum for a triacylglycerol standard TG (16:0/18:1/16:0) having *m/z* 850.7 ($M+NH_4^+$). The fragment ion at *m/z* 551.5 corresponds to diacylglycerol 16:0/16:0$^+$ resulting from the neutral loss of 18:1 from the precursor. Similarly, the fragment ion at *m/z* 577.5 corresponds to diacylglycerol 16:0/18:1$^+$ resulting from the neutral loss of 16:0 from the precursor ion.

to diacylglycerol (DG) 16:0/18:1^+ were observed. These ions result from the neutral losses of 18:1 and 16:0 from the precursor ion. These fragments were predicted and no other fragments due to truncations, recombinations, or other transformations were observed. Based on MS^2 fragmentation results, the markers having *m/z* values of 916.8, 920.7, 928.8, 954.8, 956.8 and 958.8 represented triacylglycerols. MS^2 spectrum of the marker *m/z* 916.8 (Fig 3.8) displayed a fragment at *m/z* 881.7 likely to be due to neutral loss of water and ammonia. This suggested that *m/z* 916.8 is an ammoniated adduct ($M+NH_4^+$) of an oxidized triacylglycerol. The peak was predicted to be epoxide of OOO_{ep} based on published studies of such compounds and identified by the presence of oxidized DAG fragments[39-40]. For our biomarker, a fragment at *m/z* 617.5 corresponding possibly to epoxy diacylglycerol of OO_{ep}^{+}[39] and a fragment at *m/z* 603.5 corresponding possibly to OO^+ [41] were observed in the MS^2.

Fragmentation studies of marker *m/z* 920.8 resulted in the production of fragment ions at *m/z* 575.5, 623.5 and 647.5 corresponding quite possibly to DAG fragments of 16:0/18:2^+, 16:0/22:6^+ and 18:2/22:6. Therefore, it is predicted to be a TG (16:0/18:2/22:6^+) $+NH_4^+$. Fragmentation spectrum of the marker with *m/z* 928.8 displayed fragment ions at *m/z* 601.5, 613.5, 629.5 and 631.5 consistent with the fragmentation pattern of ammoniated adduct ($M+ NH_4^+$) of LLL-mono-hydroperoxide[42].

MS^2 studies of lipids with *m/z* 954.8, 956.8 and *m/z* 958.8 resulted in water losses from their ammoniated ($M+NH_4^+$) precursor ions suggesting oxidized lipids. Some other abundant fragment ion peaks represented neutral loss of fatty acids from TAG species indicating that markers with *m/z* 954.8, 956.8 and 958.8 belong to the class of TAG lipids.

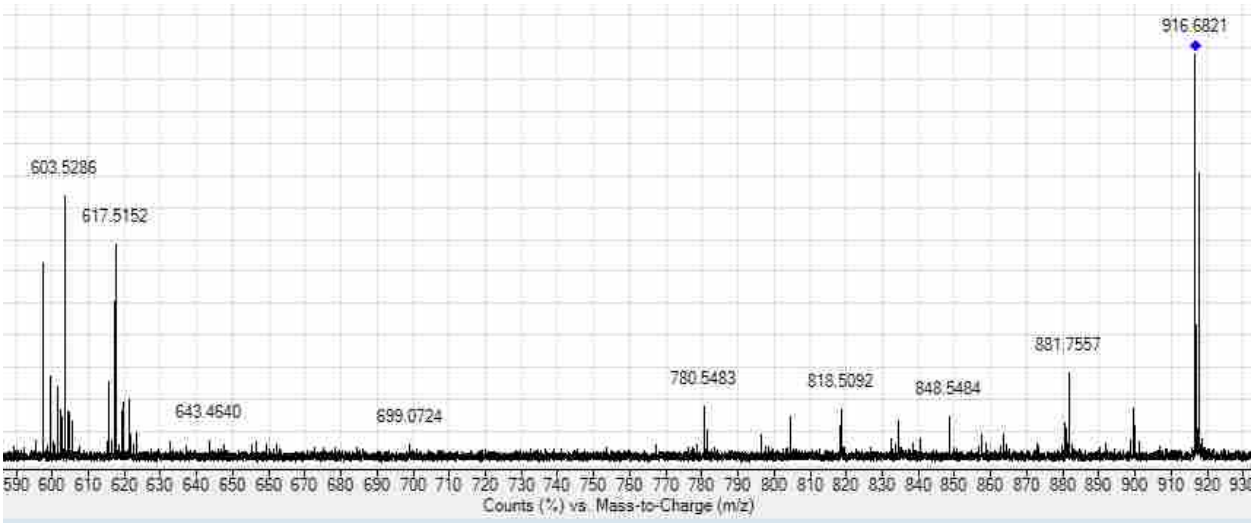


Figure 3.8 MS2 spectrum for a lipid marker with m/z 916.8. It was predicted to be an ammoniated adduct of OOO_{ep}. The fragment ions at m/z 617.5 and m/z 603.5 are likely to be OO_{ep}^{+} and OO^{+}.

Unambiguous identification could not be made for these TAG's because of their complicated spectra which may have contained a second overlapping isobaric or near isobaric species.
The markers with *m/z* 425 and *m/z* 462 are very likely lipids but with unknown identities due to fragmentation patterns without precedence in the literature or any of the databases.

3.5 Discussion

These studies tested the hypothesis that measurement of serum lipid biomarkers early in pregnancy would identify patients at risk for developing preeclampsia later in that pregnancy. Using a 'global' or in-depth or 'shotgun' serum lipidomics approach, these studies suggest that there are predictive preeclampsia biomarkers.

Global serum lipidomic approaches are relatively new and have been applied to only a few clinical indications, e.g. Alzheimer's disease[43-44]. To our knowledge they have not been previously applied to preeclampsia, although there have been a few reports of altered cholesterol and triglyceride profiles in preeclampsia[45-47].

Interest in lipids has increased as it has been recognized that lipids are more widely involved in cell regulatory pathways than previously thought. Lipids then may not only be altered in response to disease but it is possible that some circulating or cellular lipid species may mediate or contribute to aspects of disease. Given the uncertainties in the etiology and prediction of preeclampsia, there is interest in both developing better assessments of preeclampsia risk as well as better understanding the early changes that precede the fully manifest disorder.

Our approach led to the detection and confirmation of 23 unique lipid biomarkers for preeclampsia. It is entirely possible that these circulating factors reflect the consequence of early disease, but some may have a more direct biological role. In the absence of an accepted animal model for preeclampsia, early events leading to this disease are poorly understood. A number of lipidomic biomarkers based on their lipid class and other potential characterization could be related to pathological processes linked to preeclampsia. Thus, considering the changes in these specific biomarkers may be useful in providing biochemical information about early antecedents of this disease. Based on previous literature, some are at least suggestive. For example, among the several markers, the peak at *m/z* 383.3, was predicted to be an oxidized cholesterol (keto or epoxy). 7-ketocholesterol, which has the same mass, has been proposed to contribute to atherosclerosis[48] and vascular changes similar to atherosclerotic disease have been reported in preeclampsia[49]. This might explain the higher levels of this marker observed in women with later preeclampsia.

The candidate biomarkers having *m/z* values 734, 760, 784, 796, 810 and 836 were found to be higher in the serum of preeclamptic women and belong to the lipid class of glycerophosphocholines (PC). Placental ischemia and apoptosis have been reported in preeclampsia, resulting in cell lysis with release of membrane constituents, including likely phosphatidylcholines, into the circulation[50-51]. This might explain the higher levels of these markers in preeclamptic cases.

The marker with *m/z* 798 was predicted to be an oxidized (hydroxylated) PC (16:0/20:4) as indicated by the fragmentation pattern seen and consistent with other studies [37]. The exposure of PC species to reactive oxygen species (ROS) can result in oxidation of the species. There have

been several reports of increased production of ROS in preeclampsia, which might explain the increased production of this marker in preeclamptic cases[52].

Currently, there are no accepted biomarkers for predicting preeclampsia. A number of candidates have been proposed, especially pro-angiogenic and anti-angiogenic factors, and while they have repeatedly demonstrated changes in many women with established disease, they have shown poor sensitivity and specificity in predicting all forms of preeclampsia[11-13]. Therefore, there continues to be a need for predictive biomarkers of this disorder with more reliable sensitivities and specificities. The AUC values of 0.89 for two panels of markers detected in our study are substantially better than the previously reported values in the literature. The utility of these markers across two studies suggests these they should be considered of interest, especially given their being found early in the pregnancy and predicting events months later.

Utilitzation of this direct lipidomic approach as used here could provide a high throughput method for analysis of individual lipid species from diverse classes without a chromatographic separation step. This method provides comparative quantitation of species and with standards could allow for absolute quantitation. MS methods are currently in use in the clinical laboratory, e.g. drugs of abuse analysis. In addition, our approach allows for chemical characterization of interesting lipids as well as analysis analogous to multiple reaction monitoring as used for peptides and proteins. This method can have a linear range above 1000, even over the low concentration range, making it efficient for studying low abundant lipid species[53].

The approach analyzed ~2,000 lipids but could also be carried out in the negative ion mode providing an even more comprehensive evaluation of serum lipids were it needed. This expansion of scope would likely provide additional useful lipid biomarkers for early diagnosis of preeclampsia. The current approach represents a versatile method requiring relatively simple specimen preparation and demonstrating adequate reproducibility. As such, it could be applied to the analysis of blood samples for potentially many other clinical targets.

While the current studies are suggestive, the specimens available were primarily from Caucasian subjects. It is recognized that additional studies are needed to determine whether the markers found are more universal in their application. Certainly, more serum lipidomic studies are needed to demonstrate clinical utility. Nevertheless, global serum lipidomics is a valuable and robust approach for the discovery and identification of lipid biomarkers for disease and may have provided useful predictive markers for preeclampsia.

3.6 References

1. Practice, A. C. o. O., ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002. American College of Obstetricians and Gynecologists. *Int. J. Gynaecol. Obstet.* **2002,** *77* (1), 67.

2. Higgins, J. R.; de Swiet, M., Blood-pressure measurement and classification in pregnancy. *The Lancet* **2001,** *357* (9250), 131-135.

3. Bellamy, L.; Casas, J.-P.; Hingorani, A. D.; Williams, D. J., Pre-eclampsia and risk of cardiovascular disease and cancer in later life: systematic review and meta-analysis. *Bmj* **2007**.

4. Goldenberg, R. L.; Rouse, D. J., Prevention of premature birth. *N. Engl. J. Med.* **1998,** *339* (5), 313-320.

5. Barker, D. J.; Godfrey, K. M.; Gluckman, P. D.; Harding, J. E.; Owens, J. A.; Robinson, J. S., Fetal nutrition and cardiovascular disease in adult life. *The Lancet* **1993,** *341* (8850), 938-941.

6. Espinoza, J.; Romero, R.; Nien, J. K.; Gomez, R.; Kusanovic, J. P.; Gonçalves, L. F.; Medina, L.; Edwin, S.; Hassan, S.; Carstens, M., Identification of patients at risk for early onset and/or severe preeclampsia with the use of uterine artery Doppler velocimetry and placental growth factor. *Am. J Obstet. Gynecol.* **2007,** *196* (4), 326. e1-326. e13.

7. Cnossen, J. S.; Morris, R. K.; ter Riet, G.; Mol, B. W.; van der Post, J. A.; Coomarasamy, A.; Zwinderman, A. H.; Robson, S. C.; Bindels, P. J.; Kleijnen, J., Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis. *Can. Med.Assoc.J.* **2008,** *178* (6), 701-711.

8. Hung, T.-H.; Burton, G. J., Hypoxia and reoxygenation: a possible mechanism for placental oxidative stress in preeclampsia. *Taiwanese J. Obstet. Gynecol.* **2006,** *45* (3), 189-200.

9. Maynard, S. E.; Min, J.-Y.; Merchan, J.; Lim, K.-H.; Li, J.; Mondal, S.; Libermann, T. A.; Morgan, J. P.; Sellke, F. W.; Stillman, I. E., Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J. Clin. Invest.* **2003,** *111* (5), 649.

10. Roberts, J. M.; Taylor, R. N.; Musci, T. J.; Rodgers, G. M.; Hubel, C. A.; McLaughlin, M. K., Preeclampsia: an endothelial cell disorder. *Am. J. Obstet. Gynecol.* **1989,** *161* (5), 1200-1204.

11. Kleinrouweler, C. E.; Wiegerinck, M. M.; Ris-Stalpers, C.; Bossuyt, P. M.; van der Post, J. A.; von Dadelszen, P.; Mol, B. W. J.; Pajkrt, E., Accuracy of circulating placental growth factor, vascular endothelial growth factor, soluble fms-like tyrosine kinase 1 and soluble endoglin in the prediction of pre-eclampsia: a systematic review and meta-analysis. *BJOG:* **2012,** *119* (7), 778-787.

12. Powers, R.; Roberts, J.; Cooper, K.; Gallaher, M.; Frank, M.; Harger, G.; Ness, R., Maternal serum soluble fms-like tyrosine kinase 1 concentrations are not increased in early pregnancy and decrease more slowly postpartum in women who develop preeclampsia. *Am. J. Obstet. Gynecol.* **2005,** *193* (1), 185-191.
13. Chaiworapongsa, T.; Romero, R.; Kim, Y. M.; Kim, G. J.; Kim, M. R.; Espinoza, J.; Bujold, E.; Gonçalves, L.; Gomez, R.; Edwin, S., Plasma soluble vascular endothelial growth factor receptor-1 concentration is elevated prior to the clinical diagnosis of pre-eclampsia. *J. Matern. Fetal. Neonatal.Med.* **2005,** *17* (1), 3-18.

14. Wenk, M. R., The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **2005,** *4* (7), 594-610.

15. Watson, A. D., Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Lipidomics: a global approach to lipid analysis in biological systems. *J. Lipid Res.* **2006,** *47* (10), 2101-2111.

16. Steinberg, D., Thematic review series: the pathogenesis of atherosclerosis. An interpretive history of the cholesterol controversy, part V: the discovery of the statins and the end of the controversy. *J. Lipid Res.* **2006,** *47* (7), 1339-1351.

17. Balazy, M., Eicosanomics: targeted lipidomics of eicosanoids in biological systems. *Prostaglandins Other Lipid Mediat.* **2004,** *73* (3), 173-180.

18. Butterfield, D. A.; Lauderback, C. M., Lipid peroxidation and protein oxidation in Alzheimer's disease brain: potential causes and consequences involving amyloid β-peptide-associated free radical oxidative stress 1, 2. *Free Radic. Biol. Med.* **2002,** *32* (11), 1050-1060.

19. Suarna, C.; Dean, R. T.; May, J.; Stocker, R., Human atherosclerotic plaque contains both oxidized lipids and relatively large amounts of α-tocopherol and ascorbate. *Arterioscler. Thromb.Vasc. Biol.* **1995,** *15* (10), 1616-1624.

20. Berliner, J. A.; Navab, M.; Fogelman, A. M.; Frank, J. S.; Demer, L. L.; Edwards, P. A.; Watson, A. D.; Lusis, A. J., Atherosclerosis: basic mechanisms oxidation, inflammation, and genetics. *Circulation* **1995,** *91* (9), 2488-2496.

21. Malone, F. D.; Canick, J. A.; Ball, R. H.; Nyberg, D. A.; Comstock, C. H.; Bukowski, R.; Berkowitz, R. L.; Gross, S. J.; Dugoff, L.; Craigo, S. D., First-trimester or second-trimester screening, or both, for Down's syndrome. *N. Eng. J. Med.* **2005,** *353* (19), 2001-2011.

22. Hara, A.; Radin, N. S., Lipid extraction of tissues with a low-toxicity solvent. *Anal. Biochem.***1978,** *90* (1), 420-426.

23. Agresti, A., *An introduction to categorical data analysis*. Wiley New York: 1996; Vol. 135.

24. Devijver, P. A.; Kittler, J., *Pattern recognition: A statistical approach*. Prentice-Hall London: 1982; Vol. 761.

25. Buuren, S.; Groothuis-Oudshoorn, K., mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **2011,** *45* (3).

26. Rocha, J. M.; Kalo, P. J.; Ollilainen, V.; Malcata, F. X., Separation and identification of neutral cereal lipids by normal phase high-performance liquid chromatography, using evaporative light-scattering and electrospray mass spectrometry for detection. *J. Chromatogr. A* **2010,** *1217* (18), 3013-3025.

27. Kemmo, S.; Ollilainen, V.; Lampi, A.-M.; Piironen, V., Determination of stigmasterol and cholesterol oxides using atmospheric pressure chemical ionization liquid chromatography/mass spectrometry. *Food Chem.* **2007,** *101* (4), 1438-1445.

28. Souidi, M.; Dubrac, S.; Parquet, M.; Milliat, F.; Férézou, J.; Sérougne, C.; Loison, C.; Riottot, M.; Boudem, N.; Bécue, T., Effects of dietary 27-hydroxycholesterol on cholesterol metabolism and bile acid biosynthesis in the hamster. *Can. J. Physiol. Pharmacol.* **2003,** *81* (9), 854-863.

29. Jiang, X.; Sidhu, R.; Porter, F. D.; Yanjanin, N. M.; Speak, A. O.; te Vruchte, D. T.; Platt, F. M.; Fujiwara, H.; Scherrer, D. E.; Zhang, J., A sensitive and specific LC-MS/MS method for rapid diagnosis of Niemann-Pick C1 disease from human plasma. *J. Lipid Res.* **2011,** *52* (7), 1435-1445.

30. Lin, N.; Zhang, H.; Qiu, W.; Ye, J.; Han, L.; Wang, Y.; Gu, X., Determination of 7-ketocholesterol in plasma by liquid chromatography mass spectrometry for rapid diagnosis of acid sphingomyelinase deficient Niemann-Pick disease. *J. Lipid Res.* **2013**, jlr. D044024.

31. Helmschrodt, C.; Becker, S.; Schröter, J.; Hecht, M.; Aust, G.; Thiery, J.; Ceglarek, U., Fast LC–MS/MS analysis of free oxysterols derived from reactive oxygen species in human plasma and carotid plaque. *Clin. Chim.Acta* **2013,** *425*, 3-8.

32. Ivanova, P. T.; Cerda, B. A.; Horn, D. M.; Cohen, J. S.; McLafferty, F. W.; Brown, H. A., Electrospray ionization mass spectrometry analysis of changes in phospholipids in RBL-2H3 mastocytoma cells during degranulation. *Proc. Natl. Acad. Sci.***2001,** *98* (13), 7152-7157.

33. Chughtai, K.; Jiang, L.; Greenwood, T. R.; Glunde, K.; Heeren, R. M., Mass spectrometry images acylcarnitines, phosphatidylcholines, and sphingomyelin in MDA-MB-231 breast tumor models. *J. Lipid Res.* **2013,** *54* (2), 333-344.

34. Enomoto, H.; Sugiura, Y.; Setou, M.; Zaima, N., Visualization of phosphatidylcholine, lysophosphatidylcholine and sphingomyelin in mouse tongue body by matrix-assisted laser desorption/ionization imaging mass spectrometry. *Anal. Bioanal. Chem.* **2011,** *400* (7), 1913-1921.

35. Sudano, M. J.; Santos, V. G.; Tata, A.; Ferreira, C. R.; Paschoal, D. M.; Machado, R.; Buratini, J.; Eberlin, M. N.; Landim-Alvarenga, F. D., Phosphatidylcholine and sphingomyelin profiles vary in Bos taurus indicus and Bos taurus taurus in vitro-and in vivo-produced blastocysts. *Biol. Reprod.* **2012,** *87* (6), 130.

36. Stegemann, C.; Drozdov, I.; Shalhoub, J.; Humphries, J.; Ladroue, C.; Didangelos, A.; Baumert, M.; Allen, M.; Davies, A. H.; Monaco, C., Comparative lipidomics profiling of human atherosclerotic plaques. *Circulation: Cardiovasc. Gen.* **2011,** *4* (3), 232-242.

37. Reis, A.; Domingues, P.; Domingues, M., Structural motifs in primary oxidation products of palmitoyl-arachidonoyl-phosphatidylcholines by LC-MS/MS. *J. Mass Spectrom.* **2013,** *48* (11), 1207-1216.

38. Domingues, M. R. M.; Reis, A.; Domingues, P., Mass spectrometry analysis of oxidized phospholipids. *Chem. Phys. Lipids* **2008,** *156* (1), 1-12.

39. Zeb, A., Triacylglycerols composition, oxidation and oxidation compounds in camellia oil using liquid chromatography–mass spectrometry. *Chem. Phys. Lipids* **2012,** *165* (5), 608-614.

40. Byrdwell, W.; Neff, W. E., Dual parallel electrospray ionization and atmospheric pressure chemical ionization mass spectrometry (MS), MS/MS and MS/MS/MS for the analysis of triacylglycerols and triacylglycerol oxidation products. *Rapid Commun. Mass Spectr.* **2002,** *16* (4), 300-319.

41. Zeb, A.; Murkovic, M., Analysis of triacylglycerols in refined edible oils by isocratic HPLC-ESI-MS. *Eur. J. Lipid Sci Technology* **2010,** *112* (8), 844-851.

42. Zeb, A.; Murkovic, M., Determination of thermal oxidation and oxidation products of β-carotene in corn oil triacylglycerols. *Food Res. Int.***2013,** *50* (2), 534-544.

43. Han, X.; Rozen, S.; Boyle, S. H.; Hellegers, C.; Cheng, H.; Burke, J. R.; Welsh-Bohmer, K. A.; Doraiswamy, P. M.; Kaddurah-Daouk, R., Metabolomics in early Alzheimer's disease: identification of altered plasma sphingolipidome using shotgun lipidomics. *PloS one* **2011,** *6* (7), e21643.

44. Han, X., Potential mechanisms contributing to sulfatide depletion at the earliest clinically recognizable stage of Alzheimer's disease: a tale of shotgun lipidomics. *J. Neurochem.* **2007,** *103* (s1), 171-179.

45. Siddiqui, I., Maternal Serum Lipids in Women with Pre‐eclampsia. *Ann. Med. Health Sci. Res.* **2015,** *4* (4), 638-641.

46. Aziz, R.; Mahboob, T., Pre-eclampsia and lipid profile. *Pak. J. Med. Sci.* **2007,** *23* (5), 751.

47. Islam, N.; Chowdhury, M.; Kibria, G.; Akhter, S., Study of serum lipid profile in pre-eclampsia and eclampsia. *Faridpur Med. Coll. J.* **2010,** *5* (2), 56-59.

48. Rao, X.; Zhong, J.; Maiseyeu, A.; Gopalakrishnan, B.; Villamena, F. A.; Chen, L.-C.; Harkema, J. R.; Sun, Q.; Rajagopalan, S., CD36-dependent 7-ketocholesterol accumulation in macrophages mediates progression of atherosclerosis in response to chronic air pollution exposure. *Circulation Res.* **2014,** *115* (9), 770-780.

49. Staff, A. C.; Dechend, R.; Pijnenborg, R., Learning From the Placenta Acute Atherosis and Vascular Remodeling in Preeclampsia–Novel Aspects for Atherosclerosis and Future Cardiovascular Health. *Hypertens.* **2010,** *56* (6), 1026-1034.

50. Neale, D. M.; Mor, G., The role of Fas mediated apoptosis in preeclampsia. *J. Perinat. Med.* **2005,** *33* (6), 471-477.

51. Levy, R., The role of apoptosis in preeclampsia. *Isr. Med. Assoc. J.* **2005,** *7* (3), 178-81.

52. Packer, C. S., Reactive oxygen species and pre-eclampsia. *J. Hypertens.* **2003,** *21* (2), 263-264.

53. Han, X.; Cheng, H., Characterization and direct quantitation of cerebroside molecular species from lipid extracts by shotgun lipidomics. *J. Lipid Res.* **2005,** *46* (1), 163-175.

Chapter 4 Discovery and Confirmation of Diagnostic Serum Lipid Biomarkers for Alzheimer's Disease using Direct Infusion Mass spectrometry

4.1 Abstract

Alzheimers disease (AD) is a neurodegenerative disorder lacking early diagnosis and treatment. Lipids have been implicated as key regulators in a number of neurodegenerative disorders including Alzheimer's disease. A shotgun lipidomic approach was undertaken to determine if lipid biomarkers exist that can discriminate AD cases (all stages) from controls. The discovery study involved sera from 29 different stage AD cases and 32 controls. Lipid extraction was performed using organic solvent and the samples were directly infused into a time-of-flight mass spectrometer. Statistically significant lipid markers demonstrating differences between AD cases and controls were detected. These potential lipid markers were reevaluated in a second confirmatory study involving 27 cases and 30 controls. The initial study detected 89 potential AD markers. Of these, 35 markers continued to be statistically significant in the second confirmatory set. Tandem MS studies were performed and almost all of the confirmed markers were identified. Using the confirmed markers, several multi-marker panels with AUC > 0.87 were developed for any stage AD cases vs controls. Additionally, using confirmed biomarkers, models were developed for individual stages of AD. Multi-marker panels with AUCs > 0.90 were developed for each specific CDR vs controls, including the earliest stage of AD. These lipidomic biomarkers are likely to distinguish AD cases regardless of the stage from controls.

4.2 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the most common cause of age-related dementia[1-2]. Currently, as many as 25 million people may have AD with more than 4 million new cases reported every year[3]. AD is characterized by declining cognitive function including memory, language, motor control, spatial ability and executive function and in many individuals an eventual marked change in behavior. General risk factors for AD include age, female sex and certain genetic backdrops. However, the actual risk of an individual for AD cannot be determined currently. This in turn has limited drug trials.

Several hypotheses exist for AD etiology. The most common theory proposes that increased amyloid β peptide (Aβ) deposits, resulting from the cleavage of amyloid precursor protein by β-secretase and γ-secretase, lead to amyloid plaque formation[4]. Another hypothesis proposes the formation of neurofibrillary tangles of (NFT) of abnormally phosphorylated tau protein[5]. These are proposed to be the primary pathological events leading to loss of synapses and neurons producing the clinical sequelae.

Mild cognitive impairment (MCI) is often a precursor of AD but may be linked to other types of dementia. The definite diagnosis of AD can only be made by postmortem histopathological analysis. Currently, the clinical diagnosis of AD is still subjective and based on medical records, lab tests, neuroimaging and neuropsychological evaluation. Early diagnosis is challenging, although clinical centers specializing in AD appear better able to predict which subjects with MCI will progress to AD, but generally accuracies vary from 20-100%[6].

More quantitative methods of AD diagnosis exist but are not routinely used. A wide variety of imaging techniques has been used in the past for AD diagnosis. Magnetic resonance imaging (MRI) measuring hippocampal volume as a measure of neuron loss has been widely reviewed in literature[7]. Positron emission tomography (PET) using ligands such as Pittsburgh compound B (^{11}C-PIB) that bind Aβ plaques[8-9] have also been extensively studied. A novel ligand (S)-[^{18}F]THK5117 that binds to tau aggregates has also been explored recently[10]. Fluoro-deoxyglucose positron emission tomography (FDG-PET) has been used in past studies involving mild AD patients that measures reduced glucose metabolism due to synaptic loss in brain[11]. These techniques are often sensitive but are not broadly used, due to their being very expensive, and hence are more often employed for clinical research[12]. Studies of cerebrospinal fluid Aβ and phosphorylated tau have found them to be reasonably accurate for mid to late stage AD but of limited use in early or very early stage AD[6]. Blood levels of tau protein are below the detection limit for many assays[13]. Serum or plasma Aβ peptide measured revealed poor diagnostic value for AD in a number of studies[14-15]. Also, some studies indicated that there is no correlation between Aβ peptide and the severity of dementia[16]. While there is an increase in AD risk when an individual has the apolipoprotein E ε4 (ApoE4) allele or alleles, yet, it likewise cannot be used as an AD biomarker because of its low predictive value[17].

Currently then, there are as yet no accepted serologic or radiologic tests to diagnose AD, most especially early stage AD. Imaging using radio-ligands is promising and shows good results for many with AD but it is expensive, time consuming and invasive. Consequently, there is an unmet need to find serum biomarkers for AD, especially for very early disease. Moreover, other

assessments such future risk of AD, rates of AD progression, responsiveness to therapies may yet be possible with as yet undiscovered serum biomarkers.

Lipids are increasingly recognized as participants in disease, frequently modified in consequence of disease but also occasionally as mediators of disease[18]. In AD, neurodegeneration, a key feature of AD, leads to breakdown of cell membrane, resulting in abnormal metabolism of membrane lipids[19]. Also, there is a strong evidence of oxidative stress in brains of AD patients [20-21]. There are reports that continuous oxidative stress led to accumulation of oxidized lipid species namely, isoprostanes[22] or malinoaldehyde[23]. Some studies have demonstrated alterations in ceramides with different stages of AD[24-26]. A recent study indicated breakdown of phosphatidylcholine species was an indicator of neurodegeneration[19].

Therefore, it is reasonable that global characterization of alterations in the serum lipidome between AD cases and controls would likely identify biomarkers for diagnosis of AD, perhaps even early AD. The research reported here tested this hypothesis by applying an untargeted, label free, global lipidomic approach that employs direct electrospray ionization, time-of-flight mass spectrometry (ESI-TOFMS) to serum from AD patients and controls. Diagnostically useful AD lipid biomarker panels were sought and the results suggest that such lipid biomarkers exist and represent biologically relevant changes in consequence of AD.

4.3 Methods

4.3.1 Study Population

The serum specimens were obtained from the Knight Alzheimer's Disease Research Center (knight ADRC) at the Washington University School of Medicine, St. Louis, MO. The study was approved by Institutional Review Boards (IRB) at both the Knight ADRC and Brigham Young University (BYU) for specimen analysis.

In an initial discovery set, sera from 29 cases and 32 controls were analyzed. AD cases included patients having different stages of AD as indicated by differing clinical dementia ratings (CDR). Cases included 10 subjects with mild cognitive impairment (MCI) or very mild AD (CDR 0.5), 10 with mild AD (CDR 1) and 9 with moderate AD (CDR 2). The confirmatory set included 27 cases and 30 controls. AD cases included 9 patients with CDR 0.5, 9 with CDR 1and 9 with CDR 2. The demographics for the patients whose samples were used for the discovery and confirmatory set are summarized in Tables 4.1 and 4.2. Cases and controls excluded patients with co-morbidities including thyroid disease, diabetes, etc. Blood specimens after collection, were allowed to clot at room temperature for 30 min followed by separation of serum, removal of aliquots with immediate freezing. Serum specimens were maintained frozen at $< -80^0$C until used in these studies.

4.3.2 Sample Preparation

Lipid extraction from serum was accomplished using a modified published method (31). To 200 µL of serum, 1.8 mL of a mixture of hexane and isopropanol (3:2, v: v) and 300 µL of 0.5 M KH_2PO_4 were added to a glass tube followed by vigorous vortexing. The stoppered tubes were further mixed on a motorize shaker at room temperature for 1 hr at 80 rpm. To enhance phase separation, 150 µL of water was added and the samples were then centrifuged at 2000 rpm for 12 min.

Table 4.1 AD demographics for the first discovery set of samples

	No.		Age (mean)	Apo E status
CASES	29	18/29 F, 11/29 M	79.7 ± 6.6	Allele 4 positive - 20
				Allele 4 negative - 9
CONTROLS	32	18/32 F, 14/32 M	78.1 ± 6.3	Allele 4 positive - 7

Table 4.2 AD demographics for the second confirmatory set of samples

	No.		Age (mean)	Apo E status
CASES	27	16/27 F, 11/27 M	79.6 ± 8.5	Allele 4 positive - 14
				Allele 4 negative - 13
CONTROLS	30	12/30 F, 18/30 M	77.2 ± 6.2	Allele 4 positive - 12

The upper organic phase containing the lipids was collected and dried under nitrogen.

Dried lipid samples were redissolved in 200 µL of chloroform: methanol (3:1) and stored at -80 ^{0}C until MS analysis.

4.3.3 Mass Spectrometric Analysis of the Lipid Extract

To a 20 µL aliquot of the sample extract, 23 µL of chloroform, 46 µL of methanol and 14 µL of 12 µM ammonium acetate were added and mixed. The samples were directly injected into the mass spectrometer (6530 LC/MS ESI-QTOF Agilent Technologies) through an electrospray ionization (ESI) source operated in the positive ion mode. A syringe pump was utilized to inject samples at the flow rate of 10µL/min for 4 min. The ESI source used a standard spray needle having an i.d. of 120 µm. The capillary voltage was set at 3500 V. MS data was collected from mass to charge ratios (m/z) of 100-3000 with an acquisition rate of 1 spectrum/sec. The nebulizer gas and drying gas parameters were optimized to obtain a stable flow. The drying gas was set to 5 L/min at 325^{0}C with a nebulization gas pressure of 1.03 bar. Agilent Mass Hunter-Qualitative software was used for data analysis. Each specimen generated a mass spectrum from m/z 100 to 3000 from the total ion chromatogram (TIC the sum of all ion counts). A peak list having m/z values for all peaks with their abundances was generated from the mass spectrum.

To reduce analytical variability all MS peaks were normalized. For normalization, 7 abundant peaks representing different classes of lipids were chosen as a reference set. The sum of the intensities of these 7 peaks was used for normalizing all the peaks across the mass spectrum.

For the second confirmatory study, all the samples were processed and directly infused into the instrument using the same method at a flow rate of 10 µL/min for 5 min, the change being due to

a somewhat larger diameter spray needle, employed to reduce blockage of the ESI needle. All other parameters were kept identical. All the candidate markers significant in the initial, discovery set were reanalyzed for their performance in the second confirmation set.

4.3.4 Chemical Characterization of the Replicating Lipid Biomarkers by Tandem MS

Tandem MS studies were performed on all 35 replicating lipid AD biomarkers. Fragmentation was accomplished using collision-induced dissociation (CID) on an ESI-quadrupole-TOFMS. Both a QSTAR Pulsar 1 quadrupole and an Agilent 6530 Q-TOF MS were employed using the positive ion mode.

For the QSTAR, the capillary voltage set to 4200 V. Samples were injected directly at a flow rate of 2 μL/min and a specific precursor ion targeted in the quadrupole. A spectral acquisition rate of 1 spectrum/sec was used. Focusing and declustering potentials were set to 290 V and 65 V. CID was performed using nitrogen or argon gas. In order to obtain more complete fragmentation coverage, multiple fragmentation energies were used. MS/MS spectra were collected for 2 min. The (multi-channel analyzer function was used to sum all 120 MS/MS spectra and this resulted in higher signal to noise.

For the Agilent 6530, samples were injected at the flow rate of 10 μL/min with the capillary voltage set to 3500 V. The drying gas flow rate and the temperature were 5 L/min and 300^0C. MS/MS spectra were collected from m/z 50-3000 at a rate was 1 spectrum /sec. Collision energies were optimized to obtain maximum fragmentation coverage. Individual scans were summed using the add feature of the Mass Hunter program to improve the signal to noise. The targeted MS/MS mode

was used to select and fragment the parent ion of interest. The resolution was set to narrow with 1.3 m/z range.

Exact mass studies were accomplished by reference to a set of internal standards added to specimen. Lipidmaps.org was used to tentatively determine all the possible lipid species from different classes for a particular parent peak. Furthermore, product ions, neutral fragment losses, fragmentation patterns and exact masses of parent and daughter ions were evaluated and where possible compared with experimentally observed fragmentation patterns or with in silico predicted fragmentations to clarify and confirm the classes and composition of the markers.

4.3.5 Statistical Analyses

The normalized intensities of all monoisotopic peaks (abundance>150) were subjected to a two-tailed Student's t-test. Peaks having a p-value < 0.05 in the first study were considered for further evaluation in a second, independent study. The markers that continued to show p-values < 0.05 in the second confirmatory set were considered candidates and used for biomarker panel development. In addition to peaks having p-values <0.05 being considered, species with p-values < 0.10 were also considered for multi-marker modeling recognizing that they might identify more specifically subgroups of patients within the broader diagnosis and hence be complementary. Models were developed to improve diagnostic performance, using a series of forward stepwise variable selection procedures based on the AUC (corrected for optimism using the traditional bootstrap[27]) or overall error rate (based on leave-one-out cross validation or the optimism-corrected .632+ bootstrap method[28]) were performed. To find the combination of variables resulting in the best discriminatory ability, variable addition ended when the inclusion of any other variables did not result in improved performance. Furthermore, with the goal of identifying the

most parsimonious models, the forward stepwise variable selection procedure was sometimes constrained to stop after the inclusion of 3 or 4 variables. Models were then compared (in terms of estimated AUC, corrected AUC, and corrected error rate). Due to the binary nature of the data (presence or absence of Alzheimer's disease), all models employed logistic regression.

4.4 Results

The question addressed by this work was whether one or more serum lipids could identify patients with AD. A 'global' serum lipidomics approach was used to determine if lipid biomarkers exist that identifies patients clinically diagnosed with AD, including those with early stage AD. The approach was used in two separate studies and the answer appears to be yes.

4.4.1 Discovery of candidate serum lipid AD diagnostic biomarkers

The first study of 29 cases and 32 controls was used to discover potential biomarkers for any stage AD. The approach yielded 89 lipids of interest as part of this discovery set. These are summarized in Table 4.3.

4.4.2 Confirmation of candidate AD biomarkers from the first study

A second confirmatory study was conducted to evaluate the performance of the 89 previously discovered lipid markers in a new set of samples. This new study involved serum from 27 AD cases and 30 controls. Of the 89 markers, 35 continued to show statistically significant or near significant p-values. These diagnostic candidates are listed in Table 4.4.

Table 4.3 Candidate PE lipid biomarkers

No.	m/z	P value	Higher in
1.	228.2	0.06	Controls
2.	229.1	0.1	Controls
3.	243.2	0.004	Controls
4.	281.1	$2.6X10^{-5}$	Controls
5.	282.2	$8.8x10^{-5}$	Controls
6.	280.9	0.04	Controls
7.	295.1	$6.7X10^{-5}$	Controls
8.	371.35	0.001	Controls
9.	387.17	0.04	Cases
10.	391.17	0.009	Controls
11.	415.2	$2.3X10^{-5}$	Controls
12.	430.3	0.03	Cases
13.	432.2	$2.6X\ 10^{-5}$	Controls
14.	437.1	0.01	Controls
15.	447.3	$2.2X10^{-10}$	Cases
16.	453.1	0.07	Cases
17.	463.3	0.02	Cases
18.	488.3	0.004	Controls
19.	496.3	$1.1X\ 10^{-6}$	Cases
20.	514.4	0.01	Cases
21.	520.3	0.0004	Cases
22.	522.3	$5.4X10^{-6}$	Cases
23.	524.3	$2.7X10^{-5}$	Cases
24.	577.5	0.008	Cases
25.	579.5	0.001	Cases
26.	601.5	0.04	Cases
27.	602.4	0.02	Cases
28.	603.5	0.01	Cases
29.	610.5	0.05	Cases
30.	611.3	0.1	Cases
31.	614.5	0.07	Controls
32.	616.4	0.08	Controls
33.	620.4	0.004	Controls
34.	627.4	0.07	Cases
35.	630.5	0.03	Cases
36.	639.4	0.03	Cases
37.	640.6	0.0001	Controls
38.	642.6	$5.3X10^{-9}$	Controls
39.	654.3	$7.4X10^{-7}$	Controls
40.	664.6	0.1	Controls
41.	666.9	0.01	Controls

42.	669.6	$4.4X10^{-6}$	Controls
43.	671.6	0.005	Controls
44.	688.6	0.01	Controls
45.	701.5	0.008	Controls
46.	703.5	0.02	Controls
47.	714.6	0.07	Controls
48.	717.5	0.06	Controls
49.	724.5	0.0007	Controls
50.	727.5	$2.3X10^{-10}$	Cases
51.	729.6	$3.6X10^{-8}$	Cases
52.	731.6	0.04	Controls
53.	742.6	0.07	Controls
54.	744.6	0.04	Controls
55.	750.5	$4X10^{-5}$	Controls
56.	752.5	0.001	Controls
57.	754.5	0.03	Controls
58.	759.1	0.007	Controls
59.	760.8	0.09	Controls
60.	760.9	0.08	Controls
61.	761.1	0.08	Controls
62.	766.6	0.02	Controls
63.	768.6	0.0008	Controls
64.	778.5	0.005	Controls
65.	786.6	0.02	Controls
66.	792.6	0.005	Controls
67.	794.6	0.02	Controls
68.	799.6	0.1	Controls
69.	810.8	0.02	Controls
70.	812.6	0.07	Controls
71.	820.7	0.07	Cases
72.	822.7	0.03	Cases
73.	824.6	0.02	Cases
74.	842.6	0.02	Cases
75.	844.7	0.06	Controls
76.	846.7	0.01	Cases
77.	848.7	0.02	Cases
78.	850.7	0.03	Cases
79.	860.7	0.01	Cases
80.	862.7	0.007	Cases
81.	864.6	0.003	Cases
82.	881.7	0.0005	Cases
83.	886.7	0.02	Cases
84.	888.8	0.01	Cases
85.	890.8	0.01	Cases
86.	907.8	0.009	Cases

87.	912.7	0.005	Cases
88.	915.8	$2.3X10^{-5}$	Cases
89.	930.8	0.01	Cases

Table 4.4 Potential markers continued to be statistically significant in the second confirmatory set

S. No	Exact mass	P value in first discovery	P value in confirmatory	Higher in
1.	229.14	0.1	0.003	Controls
2.	430.38	0.03	0.001	Cases
3.	488.39	0.004	0.09	Controls
4.	496.33	10^{-5}	0.06	Cases
5.	514.38	0.01	0.01	Cases
6.	522.35	10^{-6}	0.02	Cases
7.	577.51	0.008	0.009	Cases
8.	601.52	0.04	0.02	Cases
9.	603.53	0.02	0.003	Cases
10.	602.44	0.02	0.006	Cases
11.	610.53	0.05	0.001	Cases
12.	620.42	0.004	0.0003	Controls
13.	630.47	0.03	0.007	Cases
14.	703.56	0.02	0.1	Controls
15.	714.62	0.07	0.002	Controls
16.	724.52	0.0007	0.04	Controls
17.	778.54	0.005	0.09	Controls
18.	766.57	0.02	0.07	Controls
19.	799.66	0.1	0.05	Controls
20.	820.73	0.07	0.02	Cases
21.	822.75	0.03	0.01	Cases
22.	824.62	0.02	0.01	Cases
23.	842.61	0.02	0.007	Cases
24.	846.75	0.01	0.002	Cases
25.	848.77	0.02	0.004	Cases
26.	850.78	0.03	0.01	Cases
27.	862.78	0.007	0.007	Cases
28.	860.77	0.01	0.006	Cases
29.	864.61	0.001	0.08	Cases
30.	881.74	0.0005	0.1	Cases
31.	886.78	0.02	0.01	Cases
32.	888.80	0.01	0.008	Cases
33.	890.81	0.005	0.003	Cases
34.	912.76	0.005	0.05	Cases
35.	930.84	0.01	0.01	Cases

4.4.3 Chemical characterization of the validated AD diagnostic lipid biomarkers

Tandem MS studies were performed on all the 35 confirmed lipid biomarkers to chemically characterize them. The exact mass, elemental composition, lipid class and molecular components were successfully determined for most of the confirmed lipid biomarkers. The assigned class and identified features of these markers are provided in Table 4.5. The rationale for the structural chemical assignments of the validated markers are also explained in Table 4.5.

Our fragmentation studies on biomarker m/z 430.4 produced fragments at m/z 165 and m/z 205 consistent with the published fragmentation pattern of vitamin E[29]. Based additionally on additional peaks in the broader fragmentation spectrum and on its putative elemental composition ($C_{29}H_{50}O_2^+$) as determined by exact mass studies, this lipid with m/z 430.4 was completely coinsistent with its being vitamin E[29].

Fragmentation spectra of the markers with m/z 496.3 and 522.3 displayed a prominent peak at m/z 184.07 corresponding to a phosphocholine moiety. Fragmentation patterns indicating neutral losses from the parent precursor ion suggest fatty acids present in the molecule. The fatty acid constituents were determined by comparing the theoretically produced fragments generated by the Lipid MS predictor feature of Lipid Maps with the experimentally produced fragments from biomarker.

Elemental compositions were also predicted by exact mass studies. These peaks were classified as lysophosphatidylcholines (LPC). The markers with m/z 496.3 and m/z 522.3 were most likely to be LPC-16:0 and LPC-18:1, consistent with the previously published literature[30].

The elemental composition of marker m/z 577.5 was predicted to be $C_{37}H_{69}O_4^+$. This species could arise from the neutral loss of the head group from a parent phospholipid (i.e. $M+H^+$- head group), most likely a glycerophosphoserine, glycerophosphoethanolamine or glycerophosphoglycerol, and represent the remaining two fatty acids attached to the glycerol backbone. However, it could also arise from the neutral loss of a fatty acid from a triacylglycerol [31]. We also observed the product ion peaks at m/z 265 and m/z 239. In the published literature, these were shown to be acylium ions (RCO^+) for 18:1 and 16:0 fatty acid substituents[31].

Similarly, the elemental compositions of markers m/z 601.5 and m/z 603.5 were proposed to be $C_{39}H_{69}O_4^+$ and $C_{39}H_{71}O_4^+$. These are also predicted to arise from the neutral loss of the head group from a glycerophosphoserine, glycerophosphoethanolamine, glycerophosphoglycerol, glycerophosphatidylcholine or through the neutral loss of a fatty acid from a triacylglycerol and appears to represent the remaining two fatty acids attached to a glycerol backbone. Based on the published literature, the acylium ions we observed in the MS^2 spectrum of the precursor ion m/z 601.5 were consistent with 18:1 (m/z 265) and 18:2 (m/z 263)[31]. Likewise, the fragmentation spectrum of the marker with m/z 603.5 presented peaks consistent with acylium ions from 18:1 (m/z 265) and 18:1 (m/z 265) fatty acids[31].

The markers with m/z 229.1, 602.4, 610.5, 620.4 and 630.4 were very likely lipids but with fragments and fragment patterns not previously described. Their predicted elemental compositions based on exact mass studies are summarized in Table 4.5.

A dominant peak at m/z 184.06 was also displayed in the MS^2 spectra of markers with m/z 703.5 and 799.6 indicating phosphocholine containing species. Therefore, the fragmentation spectra of the markers with m/z 703.5 and m/z 799.5 were most consistent with and were predicted to be sphingomyelins (SM) (SM (d18:1/16:0) or SM (d16:1/18:0)) [32]and (SM (d18:2/23:0) or SM (d17:1/24:1)) .

Fragmentation studies applied to the marker m/z 714.6 resulted in a product fragment ion m/z 369.4. This same fragment has been shown to be characteristic of cholesterol esters in the studies of others[33]. The fragmentation results were consistent with and suggest it to be a cholesterol ester of docosahexaenoic acid (22:6)[34].

The markers with m/z 724.5 and m/z 778.5 were likely glycerophosphoethanolamines (GPETn) as indicated by neutral loss of 141 from the parent ion with fragmentation. MS/MS spectra of m/z 724.5 displayed product ions at m/z 364 and m/z 361 consistent with previous studies as being a 16:0p fatty acid group at the *sn*-1 position and a 20:4 fatty acid at *sn*-2 position[35]. Its putative identification (16:0p/20:4) GPETn was also consistent with its elemental composition of $C_{41}H_{74}NO_7P+H^+$. Similarly, the fragmentation of the marker m/z 778.5 produced product ions at m/z 392 and m/z 387 potentially, indicating the presence of an 18:0p fatty acyl group at *sn*-1 position and a 22:5 acyl group at *sn*-2 position as observed in the previously published literature[35]. Based on prior studies, m/z 778.5 was predicted to be (18:0p/22:5) GPETn with an elemental composition of $C_{45}H_{80}NO_7P +H^+$.

A product ion at m/z 184.06 was also observed in the fragmentation spectra of markers with m/z 766.6, and 842.6 indicating they are glycerophosphocholines. Based on the fragmentation pattern, Lipid Maps and exact mass studies, our results were most consistent with the m/z 766.6 being PC (O-16:1/20:4).

The marker with m/z 842.6 displayed fragment ions at m/z 808.6 and m/z 824.6 corresponding to loss of hydrogen peroxide (H_2O_2) and water consistent with previous published literature[36]. An additional fragment ion at m/z 524.3 indicating presence of a C18:0 fatty acid was also observed. Therefore the marker m/z 842.6 is predicted to be a protonated oxidized PC having a hydroperoxide group attached to PC (18:0/20:4).

The markers with m/z 824.6 and m/z 864.6 had an overlapping peak adjacent to them. These could not be specifically isolated by the instrument and the fragmentation spectrum involved fragments from the overlapping peaks as well. Although their fragmentation spectrum displayed a product ion at m/z 184.06 indicating a phosphocholine group, peak classification can't be made as the peak might represent a fragment ion of the overlapping peak. This problem could be solved with an instrument with a high resolving power such as Orbitrap Elite (Thermo scientific).

A number of peaks, i.e. those with m/z 488.3, 820.7, 822.7, 846.7, 848.7, 850.7, 860.7, 862.7, 886.7, 888.8, 890.8, 912.7 and 930.8 were likely to be ammoniated adducts of triacyglycerols (TAGs). A product ion representing the neutral loss of 17 was found for each and represents ammonia loss, which is consistent with the TAGs readily forming ammonium adducts, that is the precursor ion was ammoniated ($M+NH_4^+$) and ammonia was lost upon collision[37]. The species

with m/z 881.7 was found to be the protonated form ($M+H^+$) of a TAG. The combinations of fatty acids were predicted by observing the neutral loss of acylium ions from the precursor ion but absolute assignment is not possible in the absence of standards. The predicted fatty acid components for all the TAG's are provided in Table 4.5.

The species with m/z 912.8 displayed fragment ions at m/z 599.5, m/z 601.6 and m/z 615.5 corresponding to diacylglycerol fragments of $(18:2/18:2)^+$, $(18:2/18:1)^+$ and $(18:1/18:2_{epoxide})^+$. Therefore, marker m/z 912.8 was predicted to be an epoxide of 18:1/18:2/18:2 (OLL_{ep})+NH_4^+ consistent with the previously published literature [38].

Table 4.5 Chemical characterization of the lipid biomarkers

No	m/z	Exact mass	Adduct	Elemental composition	Class	Possible identity
1.	229.1	229.14		$C_{12}H_{21}O_4^+$		
2.	430.4	430.38	M^+	$C_{29}H_{50}O_2^+$	Sterols	Vitamin e
3.	488.4	488.39	$M+NH_4^+$	$C_{27}H_{50}O_6+NH_4^+$	Triacylglycerols	TG-8:0/8:0/8:0
4.	496.3	496.33	$M+H^+$	$C_{24}H_{50}NO_7P+H^+$	Lysophosphocholines	LysoPC 16:0
5.	514.4	514.38				---
6.	522.3	522.35	$M+H^+$	$C_{26}H_{52}NO_7P+H^+$	Lysophosphocholines	LysoPC 18:1
7.	577.5	577.51	M^+	$C_{37}H_{69}O_4^+$		
8.	601.5	601.52	M^+	$C_{39}H_{69}O_4^+$		
9.	603.5	603.53	M^+	$C_{39}H_{71}O_4^+$		
10.	602.4	602.44		$C_{44}H_{58}O^+$		
11.	610.5	610.53		$C_{41}H_{70}O_3^+$		
12.	620.4	620.42		$C_{43}H_{56}O_3^+$ or $C_{47}H_{56}$		
13.	630.4	630.47		$C_{46}H_{62}O^+$		
14.	703.5	703.56	$M+H^+$	$C_{43}H_{76}N_2O_6+H^+$	Sphingomyelins	SM(d18:1/16:0) or SM(d16:1/18:0)

15.	714.6	714.61	$M+NH_4^+$	$C_{49}H_{76}O_2+NH_4^+$	Cholesterol esters	C22:6 cholesterol ester
16.	724.5	724.52	$M+H^+$	$C_{41}H_{74}NO_7P+H^+$	Glycerophosphatidyleth-anolamines	PE(16:0p/ 20:4)
17.	778.5	778.57	$M+H^+$	$C_{45}H_{80}NO_7P+H$	Glycerophosphatidyleth-anolamines	PE(18:0p/ 22:5)
18.	766.6	766.57	$M+H^+$	$C_{44}H_{80}NO_7P+H^+$	Glycerophosphocholines	PC(O-16:1/20:4)
19.	799.6	799.66	$M+H^+$	$C_{46}H_{91}N_2O_6P+H^+$	Sphingomyelins	SM(d17:1/24:1) or SM(d18:2/23:0)
20.	820.7	820.73	$M+NH_4^+$	$C_{51}H_{94}O_6+NH_4^+$	Triacylglycerols	TG-16:0/16:1/16:1
21.	822.7	822.75	$M+NH_4^+$	$C_{51}H_{96}O_6+NH_4^+$	Triacylglycerols	TG-16:0/16:0/16:1
22.	824.6	824.61	$M+H^+$			
23.	842.6	842.61	$M+H^+$	$C_{46}H_{84}NO_{10}P+H^+$	Glycerophosphocholines	PC(18:0/20:4)+OOH $+H^+$
24.	846.7	846.75	$M+NH_4^+$	$C_{53}H_{96}O_6+NH_4^+$	Triacylglycerol	TG-16:0/16:1/18:2
25.	848.7	848.77	$M+NH_4^+$	$C_{53}H_{98}O_6+NH_4^+$	Triacylglycerol	TG-16:0/16:1/18:1
26.	850.7	850.78	$M+NH_4^+$	$C_{53}H_{100}O_6+NH_4^+$	Triacylglycerol	TG16:0/16:0/18:1
27.	862.7	862.78	$M+NH_4^+$	$C_{54}H_{100}O_6+NH_4^+$	Triacylglycerol	TG16:0/17:0/18:2
28.	860.7	860.77	$M+NH_4^+$	$C_{54}H_{98}O_6+NH_4^+$	Triacylglycerol	TG16:1/17:1/18:1
29.	864.6	864.61	$M+H^+$			

30.	881.7	881.74	$M+H^+$	$C_{56}H_{100}O_6 + H^+$	Triacylglycerol	TG-16:0/18:0/20:5
31.	886.7	886.78	$M+NH_4^+$	$C_{56}H_{100}O_6 + NH_4^+$	Triacylglycerol	TG-17:0/18:1/18:3
32.	888.8	888.80	$M+NH_4^+$	$C_{56}H_{102}O_6 + NH_4^+$	Triacylglycerol	TG-17:0/18:1/18:2
33.	890.8	890.81	$M+NH_4^+$	$C_{56}H_{104}O_6 + NH_4^+$	Triacylglycerol	TG-17:0/18:0/18:2
34.	912.7	912.76	$M+NH_4^+$	$C_{57}H_{98}O_7 + NH_4^+$	Oxidized triacylglycerols	18:1/18:2/18:2 epoxide
35.	930.8	930.84	$M+NH_4^+$	$C_{59}H_{108}O_6 + NH_4^+$	Triacylglycerol	TG-18:1/18:2/20:0

4.4.4 Modeling of Multi Marker Panels Diagnostic of AD

Statistical modeling was performed on the 35 validated markers to develop multi-marker panels with higher predictive values. Several panels using combinations of 4-6 markers were found to have AUCs > 0.87 as summarized in Table 4.6. The ROC curves for one of the best multi-marker combinations [AUCs of 0.90 (sensitivity > 85% and specificity > 80%)] are shown in Figure 4.1.

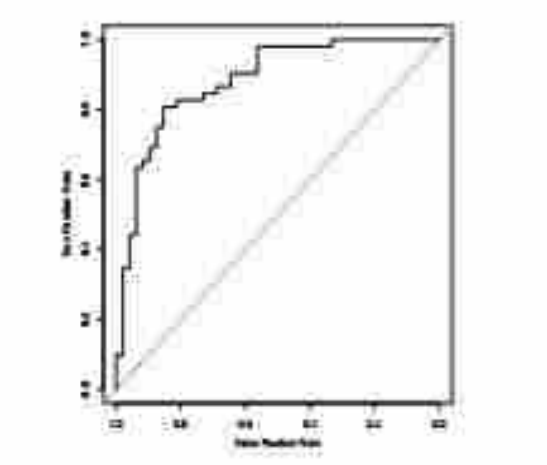
Likewise, multi-marker panels were also developed for individual stages of AD (CDR 0.5, 1 or 2) versus controls. Several panels formed from combinations of 3-5 markers had AUC values > 0.94 and are listed in Table 4.7. The ROC curves for the three best multi-marker combinations for each individual AD stage (CDR 0.5, 1 and 2) all having AUCs >0.93 and sensitivities >90% at specificities of at least 84% are shown in Fig 4.2, 4.3 and 4.4.

4.5 Discussion

The early diagnosis of AD remains challenging. With newer ligand imaging methods, earlier diagnosis appears possible for many patients with AD; however, the approach is expensive, invasive and involved and doesn't lend itself to routine screening or assessment. Still, many who investigate this disease believe that neurodegeneration in AD begins long before the appearance of clinical symptoms. These early changes then would involve molecular changes. As such, there is a reasonable presumption that very early AD can be identified and that it may be possible

Table 4.6 Multi-marker panels for any stage AD vs controls having an AUC > 0.87

AUC	Potential marker combinations
0.90	496.3, 714.6, 778.5, 864.6, 488.3, 602.3
0.90	522.3, 778.5, 864.6, 488.3, 610.5
0.88	864.6, 714.6, 620.4, 778.5
0.87	778.5, 864.6, 610.5, 488.3



ROC curve for m/z 522.3, 778.5, 864.6, 488.3, 610.5

Figure 4.1 Best modeled diagnostic serum biomarker set for AD.
This receiver operator characteristic curve (ROC), generated by logistic regression analysis, included 5 lipid biomarkers having mass to charge ratios of 522.3, 778.5, 864.6, 488.3 and 610.5. The AUC was 0.90 with a sensitivity of ~87% at a specificity of ~82%.

Table 4.7 Best multi-marker panels for individual stages of AD (CDR 0.5, 1 and 2) vs controls

CDR	AUC	Marker combinations in optimized set
0.5	0.95	620.4, 602.4, 778.5
1	0.93	864.6, 714.6, 778.5, 620.4
2	0.94	620.4, 602.4, 778.5, 488.3

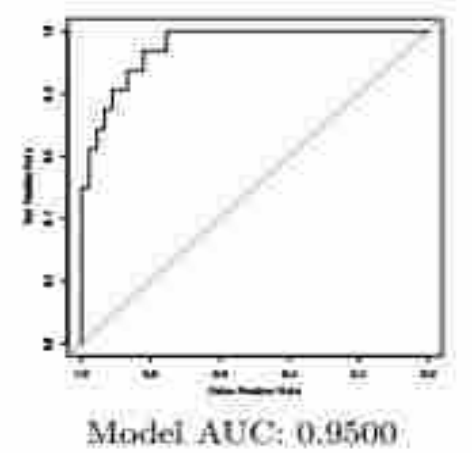


Figure 4.2 Best modeled diagnostic serum biomarker set for CDR 0.5 vs controls.
This ROC curve was generated by logistic regression analysis on 3 lipid biomarkers having mass to charge ratios of 620.4, 602.4 and 778.5. The AUC was 0.95 with a sensitivity of ~97% at a specificity of ~85%.

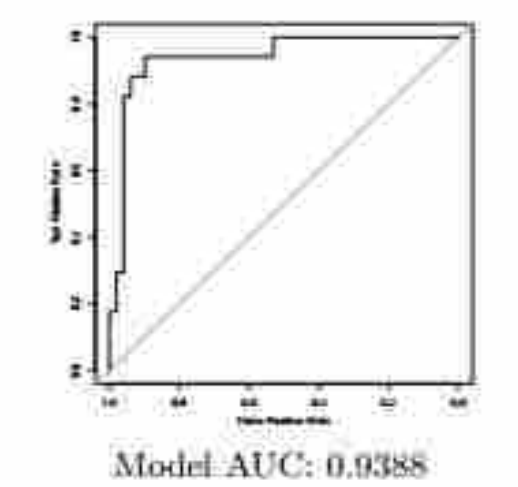


Figure 4.3 Best modeled diagnostic serum biomarker set for CDR 1 vs controls.
This ROC curve was generated by logistic regression analysis using 4 lipid biomarkers having mass to charge ratios of 864.6, 714.6, 778.5 and 620.4. The AUC was 0.93 with a sensitivity of ~97% at a specificity of ~88%.

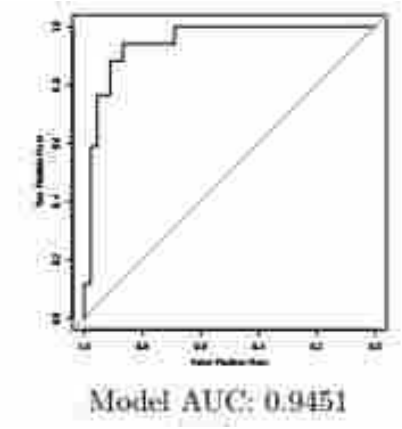


Figure 4.4 Best modeled diagnostic serum biomarker set for CDR 2 vs controls.
This ROC curve was generated using regression analysis of a set of 5 lipid biomarkers having mass to charge ratios of 864.6, 714.6, 778.5, 842.6 and 620.4. The AUC was 0.94 with a sensitivity of ~98% at a specificity of ~84%.

to observe and track many of these changes with a simple blood test. Indeed, other recent work suggests that peptides and lipids are altered early in the disease[39].

Lipids are an emerging area of interest in biomedical research[40]. Indeed, an important role for lipids in the pathogenesis of AD has been previously considered[5]. What were once thought of as largely structural molecules or passive fuel reserves have been more recently shown to reflect pathology and to participate in disease[41].

This is also one of a number of recent studies to employ a global lipidomic approach using direct infusion mass spectrometry to interrogate serum. Direct infusion mass spectrometry using ESI-MS has been shown to be quantitative and faster than other MS analyses and useful in the analysis of the lipids in the extracts of biological samples[42]. It allows for the screening of many individual lipids from diverse lipid classes. This study was conducted to determine if lipid biomarkers existed that discriminated AD cases from controls, especially patients with very early stage AD and to address the hypothesis that such markers would provide further insight into AD pathology.

Our approach led to the discovery and confirmation of 35 lipidomic markers differentiating any stage AD from control across two independent studies. We observed alterations in the levels of lipid species belonging to various lipid classes, including sterols, sphingomyelins (SM), glycerophosphocholines (PC), glycerophosphoethanolamines (PE), lysophosphocholines (LPC) and triacylglycerols (TAG). Combinations of these markers provided panels with ROC curves having AUCs greater than 0.87 for all stage AD. The best of these provided a sensitivity of 88% with a specificity of 82% for all stages of AD. The ability of these same markers to identify

patients with a particular stage of AD versus controls was also modeled. The best models provided sensitivities of ~97% for CDR 0.5, 1 or 2 individually with specificities of at least 85% for CDR 0.5 and ~87% for CDR 1 or CDR 2. While the number of patients in these subgroups was small, the results nonetheless suggest strongly that serum lipid biomarkers are present that diagnose AD, even very early stage AD.

Although a confirmation study was performed, it is well recognized replication of the results from this study needs to be performed in other large studies across many populations. It would also be of interest to monitor changes in serum biomarkers over time. Further, given the alterations in lipids in response to AD, it would be useful to more directly study possible biologic roles for some of the individual molecular species of lipids in the context of AD.

It was demonstrated that direct infusion mass spectrometry was reasonably effective in a broad, untargeted analysis of the serum lipidome. It allowed fast, uncomplicated, high throughput and semi quantitative analysis of >1000 molecular species of lipids. It resulted in reproducible discrimination of AD cases from controls.

The second goal of the research was to determine whether the observed markers provided insight into AD pathology. The classification and/or identification of the biomarkers was carried out with the intent of understanding biochemical changes present in AD, even early AD, asking whether there was biological plausibility for them.

One marker with m/z 430.4 was consistent with its being vitamin E. Its levels were found to be significantly increased in the serum samples of AD cases. This differs from previously reported results where vitamin E levels were found to be lower in the CSF and plasma of AD patients [43-44].

Two markers with m/z 496.3 and m/z 522.3 found to be higher in AD cases belonged to the class of lysophosphocholines (LPC). Neurodegeneration, a key feature of AD, results in the breakdown of cell membranes[19] potentially increasing the release of membrane molecules into the circulation. Also, published studies suggest that activation of phospholipase A_2 (PLA_2), responsible for hydrolysis of phosphocholines and leading to increased production of lysoPCs, was also increased in AD patients[19]. Furthermore, lysoPC produced through hydrolysis of PC by PLA_2 can be a precursor of platelet activating factor (PAF) which in turn acts as a mediator of inflammation[45]. There is clear evidence of inflammation in the brains of AD patients [46-47]. Alterations in levels of these markers then might be explained by and potentially represent known pathological events giving rise to higher levels of lyso-PCs in AD cases. Our results appear consistent with previous AD findings.

Two markers with m/z 703.5 and m/z 799.6, found to be lower in AD patients, were determined to be sphingomyelins (SM). This is consistent with a past study which also reported lower SM levels in AD patients[48]. SMs are metabolized to ceramides which can play important roles in cell differentiation, proliferation, and apoptosis[49]. Ceramides are also known to be involved in atherosclerosis[50]. All of these processes have been previously associated with the pathogenesis of AD[51-52].

Candidate biomarkers with m/z 724.5 and m/z 778.5 were decreased in AD cases. Both belonged to the class of PEs. Ethanolamine plasmalogen (PlsEtn), representing ~ 90% of all PEs, is highly enriched in white and grey matter of the human brain. PlsEtn provides a sheath for axons in white matter and acts as a source of lipid messengers participating in signal transduction in grey matter. There are several reports of gradual reduction in levels of plasmalogen with increasing severity of AD[53-54]. Its deficiency in white matter likely contributes to myelin sheath defects and axonal dysfunction leading to dementia in AD patients. Plasmalogen deficiency in cerebral grey matter could result in synapse loss due to membrane instability[55] which is another known early event in AD[56]. This same process, if somewhat more generalized, might explain the reduced levels of PE in AD patients in our study.

The candidate marker with m/z 842.6 was identified to be an oxidized PC species elemental an easily lost OH group. Therefore it's likely to be PC (18:0/20:4)+OOH $+H^+$ with an elemental composition of $C_{46}H_{84}NO_{10}P+H^+$. It was found to be higher in AD cases. The increased oxidative stress observed in neurodegenerative disorders [57] likely results in increased lipid oxidation and would possibly explain the higher levels of this marker among AD cases in our study. Similarly, the marker with m/z 912.7 was found to be an ammoniated adduct of an oxidized TAG and was higher in AD cases. AD is associated with increased levels of oxidative stress [58]. F2-isoprostane, produced from polyunsaturated fatty acids as a result of free radical mediated peroxidation has been reported to be higher in brain and CSF of AD patients. However, its increase in plasma remains controversial [59].

Several other candidate lipid markers, those with m/z 820.7, 822.7, 846.7, 848.7, 850.7, 860.7, 862.7, 881.7, 886.7, 888.8, 890.8 and 930.8, all belong to the class of triacylglycerols (TAG) and were found to be higher in AD cases. There is emerging evidence that elevated triglyceride levels precede amyloid deposition in AD mouse model[60], although the implications of this are unknown.

A previously published untargeted lipidomic and metabolomics approach revealed a blood based biomarker panel to identify normal individuals who will phenoconvert to MCI or AD within 2-3 years with high accuracy[61]. The number of subjects who phenocoverted was only 10. Currently, there are still no US-FDA approved biomarkers for AD diagnosis.

Our study detected novel, previously undescribed potential lipidomic biomarkers for AD. We also developed panels to predict CDR 0.5,1 and 2 individually. Our work then is complementary to the above mentioned study and is likely to contribute to improved clinical diagnosis of AD.

Collectively, the results of these studies suggest that useful serum lipid biomarkers allow for the accurate diagnosis of most patients with AD, regardless of stage. Moreover, the majority of these biomarkers were found to be relevant lipid species whose altered expression is consistent with known pathologies associated with and in some cases potentially contributing to AD pathology. These are novel, previously undescribed lipid biomarkers for AD and holdout diagnostic promise for patients with AD, including early stage disease. As such, they should help in identifying individuals with early AD events and provide a tool for easier and more rational selection of subjects for trials of potential AD therapies.

4.6 References

1. Schneider, J. A.; Arvanitakis, Z.; Bang, W.; Bennett, D. A., Mixed brain pathologies account for most dementia cases in community-dwelling older persons. *Neurology* **2007,** *69* (24), 2197-2204.

2. Schneider, J. A.; Aggarwal, N. T.; Barnes, L.; Boyle, P.; Bennett, D. A., The neuropathology of older persons with and without dementia from community versus clinic cohorts. *J.Alzheimer's Dis.* **2009,** *18* (3), 691.

3. Ferri, C. P.; Prince, M.; Brayne, C.; Brodaty, H.; Fratiglioni, L.; Ganguli, M.; Hall, K.; Hasegawa, K.; Hendrie, H.; Huang, Y., Global prevalence of dementia: a Delphi consensus study. *The lancet* **2006,** *366* (9503), 2112-2117.

4. Hardy, J., The amyloid hypothesis for Alzheimer's disease: a critical reappraisal. *J. Neurochem.* **2009,** *110* (4), 1129-1134.

5. Maccioni, R. B.; Muñoz, J. P.; Barbeito, L., The molecular bases of Alzheimer's disease and other neurodegenerative disorders. *Archives of medical research* **2001,** *32* (5), 367-381.

6. Jack, C.; Petersen, R. C.; Xu, Y. C.; O'Brien, P. C.; Smith, G. E.; Ivnik, R. J.; Boeve, B. F.; Waring, S. C.; Tangalos, E. G.; Kokmen, E., Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology* **1999,** *52* (7), 1397-1397.

7. Hampel, H.; Bürger, K.; Teipel, S. J.; Bokde, A. L.; Zetterberg, H.; Blennow, K., Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. *Alzheimer's & Demen.***2008,** *4* (1), 38-48.

8. Jack, C. R.; Lowe, V. J.; Senjem, M. L.; Weigand, S. D.; Kemp, B. J.; Shiung, M. M.; Knopman, D. S.; Boeve, B. F.; Klunk, W. E.; Mathis, C. A., 11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. *Brain* **2008,** *131* (3), 665-680.

9. Forsberg, A.; Engler, H.; Almkvist, O.; Blomquist, G.; Hagman, G.; Wall, A.; Ringheim, A.; Långström, B.; Nordberg, A., PET imaging of amyloid deposition in patients with mild cognitive impairment. *Neurobiol. Aging* **2008,** *29* (10), 1456-1465.

10. Wall, A.; Chiotis, K.; Saint-Aubert, L.; Wilking, H.; Sprycha, M.; Borg, B.; Thibblin, A.; Eriksson, J. P.; Sörensen, J.; Antoni, G., Tracer kinetic analysis of (S)-18F-THK5117 as a PET tracer for assessing tau pathology. *J. Nuclear Med.* **2016**, jnumed. 115.158519.

11. Samuraki, M.; Matsunari, I.; Chen, W.-P.; Yajima, K.; Yanase, D.; Fujikawa, A.; Takeda, N.; Nishimura, S.; Matsuda, H.; Yamada, M., Partial volume effect-corrected FDG PET and grey matter volume loss in patients with mild Alzheimer's disease. *Eur. J. Nucl. Med. Mol. Imaging* **2007,** *34* (10), 1658-1669.

12. Panegyres, P. K.; Rogers, J. M.; McCarthy, M.; Campbell, A.; Wu, J. S., Fluorodeoxyglucose-positron emission tomography in the differential diagnosis of early-onset dementia: a prospective, community-based study. *BMC Neurol.* **2009,** *9* (1), 41.

13. Tang, B. L.; Kumar, R., Biomarkers of mild cognitive impairment and Alzheimer's disease. *Ann. Acad. Med. Singapore* **2008,** *37* (5), 406-5.
14. Mayeux, R.; Honig, L.; Tang, M.-X.; Manly, J.; Stern, Y.; Schupf, N.; Mehta, P., Plasma Aβ40 and Aβ42 and Alzheimer's disease Relation to age, mortality, and risk. *Neurology* **2003,** *61* (9), 1185-1190.

15. Fukumoto, H.; Tennis, M.; Locascio, J. J.; Hyman, B. T.; Growdon, J. H.; Irizarry, M. C., Age but not diagnosis is the main predictor of plasma amyloid β-protein levels. *Archives of neurology* **2003,** *60* (7), 958-964.

16. Irizarry, M. C., Biomarkers of Alzheimer disease in plasma. *NeuroRx* **2004,** *1* (2), 226-234.

17. Patterson, C.; Feightner, J. W.; Garcia, A.; Hsiung, G.-Y. R.; MacKnight, C.; Sadovnick, A. D., Diagnosis and treatment of dementia: 1. Risk assessment and primary prevention of Alzheimer disease. *Can. Med. Assoc. J.* **2008,** *178* (5), 548-556.

18. Watson, A. D., Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Lipidomics: a global approach to lipid analysis in biological systems. *J. Lipid Res.* **2006,** *47* (10), 2101-2111.

19. Klein, J., Membrane breakdown in acute and chronic neurodegeneration: focus on choline-containing phospholipids. *J. Neural Transm.* **2000,** *107* (8-9), 1027-1063.

20. Huang, X.; Moir, R. D.; Tanzi, R. E.; Bush, A. I.; Rogers, J. T., Redox-Active Metals, Oxidative Stress, and Alzheimer's Disease Pathology. *Ann. N.Y.Acad. Sci.* **2004,** *1012* (1), 153-163.

21. Butterfield, D. A.; Perluigi, M.; Sultana, R., Oxidative stress in Alzheimer's disease brain: new insights from redox proteomics. *Eur. J.Pharmacol.* **2006,** *545* (1), 39-50.

22. Praticò, D., The neurobiology of isoprostanes and Alzheimer's disease. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2010,** *1801* (8), 930-933.

23. Bet, L.; Calabrese, E.; Bava, L.; Magni, E.; Rapuzzi, S.; Pezzoli, G.; Mariani, C., Malondialdheyde production features in platelets of Parkinson and Alzheimer patients and in physiological aging. *Neurosci. Res. Commun.* **1999,** *25* (1), 33-41.

24. Svennerholm, L.; Gottfries, C. G., Membrane Lipids, Selectively Diminished in Alzheimer Brains, Suggest Synapse Loss as a Primary Event in Early-Onset Form (Type I) and Demyelination in Late-Onset Form (Type II). *J. Neurochem.* **1994,** *62* (3), 1039-1047.

25. Han, X.; M Holtzman, D.; W McKeel, D.; Kelley, J.; Morris, J. C., Substantial sulfatide deficiency and ceramide elevation in very early Alzheimer's disease: potential role in disease pathogenesis. *J. Neurochem.* **2002,** *82* (4), 809-818.

26. Cutler, R. G.; Kelly, J.; Storie, K.; Pedersen, W. A.; Tammara, A.; Hatanpaa, K.; Troncoso, J. C.; Mattson, M. P., Involvement of oxidative stress-induced abnormalities in ceramide and cholesterol metabolism in brain aging and Alzheimer's disease. *Proc. Natl. Acad. Sci. USA* **2004,** *101* (7), 2070-2075.

27. Harrell, F. E.; Lee, K. L.; Mark, D. B., Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **1996,** *15*, 361-387.

28. Efron, B.; Tibshirani, R., Improvements on cross-validation: the 632+ bootstrap method. *JASA* **1997,** *92* (438), 548-560.

29. Touboul, D.; Roy, S.; Germain, D. P.; Chaminade, P.; Brunelle, A.; Laprévote, O., MALDI-TOF and cluster-TOF-SIMS imaging of Fabry disease biomarkers. *Int. J. Mass Spectr.* **2007,** *260* (2), 158-165.

30. Chughtai, K.; Jiang, L.; Greenwood, T. R.; Glunde, K.; Heeren, R. M., Mass spectrometry images acylcarnitines, phosphatidylcholines, and sphingomyelin in MDA-MB-231 breast tumor models. *J. Lipid Res.* **2013,** *54* (2), 333-344.

31. Hsu, F.-F.; Turk, J., Electrospray ionization with low-energy collisionally activated dissociation tandem mass spectrometry of glycerophospholipids: mechanisms of fragmentation and structural characterization. *J. Chromatogr. B* **2009,** *877* (26), 2673-2695.

32. Enomoto, H.; Sugiura, Y.; Setou, M.; Zaima, N., Visualization of phosphatidylcholine, lysophosphatidylcholine and sphingomyelin in mouse tongue body by matrix-assisted laser desorption/ionization imaging mass spectrometry. *Anal. Bioanal. Chem.* **2011,** *400* (7), 1913-1921.

33. Liebisch, G.; Binder, M.; Schifferer, R.; Langmann, T.; Schulz, B.; Schmitz, G., High throughput quantification of cholesterol and cholesteryl ester by electrospray ionization tandem mass spectrometry (ESI-MS/MS). *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids* **2006,** *1761* (1), 121-128.

34. Yu, S.; Dong, J.; Zhou, W.; Yang, R.; Li, H.; Zhao, H.; Zhang, T.; Guo, H.; Wang, S.; Zhang, C., A rapid and precise method for quantification of fatty acids in human serum cholesteryl esters by liquid chromatography and tandem mass spectrometry. *J. Chromatogr. B* **2014,** *960*, 222-229.

35. Berry, K. A. Z.; Murphy, R. C., Electrospray ionization tandem mass spectrometry of glycerophosphoethanolamine plasmalogen phospholipids. *J. Am. Soc. Mass Spectr.* **2004,** *15* (10), 1499-1508.

36. Uchikata, T.; Matsubara, A.; Nishiumi, S.; Yoshida, M.; Fukusaki, E.; Bamba, T., Development of oxidized phosphatidylcholine isomer profiling method using supercritical fluid chromatography/tandem mass spectrometry. *J. Chromatogr. A* **2012,** *1250*, 205-211.

37. Byrdwell, W.; Neff, W. E., Dual parallel electrospray ionization and atmospheric pressure chemical ionization mass spectrometry (MS), MS/MS and MS/MS/MS for the analysis of triacylglycerols and triacylglycerol oxidation products. *Rapid Commun. Mass Spectrom.* **2002,** *16* (4), 300-319.

38. Zeb, A., Triacylglycerols composition, oxidation and oxidation compounds in camellia oil using liquid chromatography–mass spectrometry. *Chem. Phys. Lipids* **2012,** *165* (5), 608-614.

39. Shah, D. J.; Rohlfing, F.; Anand, S.; Johnson, W. E.; Alvarez, M. T. B.; Cobell, J.; King, J.; Young, S. A.; Kauwe, J. S.; Graves, S. W., Discovery and confirmation of novel serum biomarkers diagnosing Alzheimer's disease.

40. Wenk, M. R., The emerging field of lipidomics. *Nat. Rev. Drug Discov.* **2005,** *4* (7), 594-610.

41. Hu, C.; van der Heijden, R.; Wang, M.; van der Greef, J.; Hankemeier, T.; Xu, G., Analytical strategies in lipidomics and applications in disease biomarker discovery. *J. Chromatogr.B* **2009,** *877* (26), 2836-2846.

42. Han, X.; Gross, R. W., Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude extracts of biological samples. *Mass Spectrom. Rev.* **2005,** *24* (3), 367-412.

43. Mangialasche, F.; Kivipelto, M.; Mecocci, P.; Rizzuto, D.; Palmer, K.; Winblad, B.; Fratiglioni, L., High plasma levels of vitamin E forms and reduced Alzheimer's disease risk in advanced age. *J. Alzheimer's Dis.***2010,** *20* (4), 1029.

44. KONTUSH, A.; SCHEKATOLINA, S., Vitamin E in neurodegenerative disorders: Alzheimer's disease. *Ann. N.Y. Ac.Sci.* **2004,** *1031* (1), 249-262.

45. BAZAN, N. G.; de TURCO, E. B. R.; ALLAN, G., Mediators of injury in neurotrauma: intracellular signal transduction and gene expression. *J. Neurotrauma* **1995,** *12* (5), 791-814.

46. Akiyama, H.; Barger, S.; Barnum, S.; Bradt, B.; Bauer, J.; Cole, G. M.; Cooper, N. R.; Eikelenboom, P.; Emmerling, M.; Fiebich, B. L., Inflammation and Alzheimer's disease. *Neurobiol. Aging* **2000,** *21* (3), 383-421.

47. Heneka, M. T., Inflammation in Alzheimer's disease. *Clin. Neurosci. Res.* **2006,** *6* (5), 247-260.

48. He, X.; Huang, Y.; Li, B.; Gong, C.-X.; Schuchman, E. H., Deregulation of sphingolipid metabolism in Alzheimer's disease. *Neurobiol. Aging* **2010,** *31* (3), 398-408.

49. Merrill, A.; Schmelz, E.; Dillehay, D.; Spiegel, S.; Shayman, J.; Schroeder, J.; Riley, R.; Voss, K.; Wang, E., Sphingolipids—the enigmatic lipid class: biochemistry, physiology, and pathophysiology. *Toxicol. Appl. Pharmacol.***1997,** *142* (1), 208-225.

50. Ichi, I.; Nakahara, K.; Miyashita, Y.; Hidaka, A.; Kutsukake, S.; Inoue, K.; Maruyama, T.; Miwa, Y.; Harada-Shiba, M.; Tsushima, M., Association of ceramides in human plasma with risk factors of atherosclerosis. *Lipids* **2006,** *41* (9), 859-863.

51. Rodríguez, J. J.; Jones, V. C.; Verkhratsky, A., Impaired cell proliferation in the subventricular zone in an Alzheimer's disease model. *Neuroreport* **2009,** *20* (10), 907-912.

52. Casserly, I.; Topol, E. J., Convergence of atherosclerosis and Alzheimer's disease: inflammation, cholesterol, and misfolded proteins. *The Lancet* **2004,** *363* (9415), 1139-1146.

53. Han, X.; Holtzman, D. M.; McKeel, D. W., Plasmalogen deficiency in early Alzheimer's disease subjects and in animal models: molecular characterization using electrospray ionization mass spectrometry. *J. Neurochem.* **2001,** *77* (4), 1168-1180.

54. Ginsberg, L.; Rafique, S.; Xuereb, J. H.; Rapoport, S. I.; Gershfeld, N. L., Disease and anatomic specificity of ethanolamine plasmalogen deficiency in Alzheimer's disease brain. *Brain Res.* **1995,** *698* (1), 223-226.

55. Ginsberg, L.; Xuereb, J. H.; Gershfeld, N. L., Membrane instability, plasmalogen content, and Alzheimer's disease. *J. Neurochem.***1998,** *70* (6), 2533-2538.

56. Terry, R. D.; Masliah, E.; Salmon, D. P.; Butters, N.; DeTeresa, R.; Hill, R.; Hansen, L. A.; Katzman, R., Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann.Neurol.* **1991,** *30* (4), 572-580.

57. Migliore, L.; Fontana, I.; Colognato, R.; Coppede, F.; Siciliano, G.; Murri, L., Searching for the role and the most suitable biomarkers of oxidative stress in Alzheimer's disease and in other neurodegenerative diseases. *Neurobiol. Aging* **2005,** *26* (5), 587-595.

58. Sultana, R.; Mecocci, P.; Mangialasche, F.; Cecchetti, R.; Baglioni, M.; Butterfield, D. A., Increased protein and lipid oxidative damage in mitochondria isolated from lymphocytes from patients with Alzheimer's disease: insights into the role of oxidative stress in Alzheimer's disease and initial investigations into a potential biomarker for this dementing disorder. *J. Alzheimer's Dis.: JAD* **2010,** *24* (1), 77-84.

59. Song, F.; Poljak, A.; Smythe, G. A.; Sachdev, P., Plasma biomarkers for mild cognitive impairment and Alzheimer's disease. *Brain Res. Rev.* **2009,** *61* (2), 69-80.

60. Burgess, B. L.; McIsaac, S. A.; Naus, K. E.; Chan, J. Y.; Tansley, G. H.; Yang, J.; Miao, F.; Ross, C. J.; van Eck, M.; Hayden, M. R., Elevated plasma triglyceride levels precede amyloid deposition in Alzheimer's disease mouse models with abundant Aβ in plasma. *Neurobiol. Dis.* **2006,** *24* (1), 114-127.

61. Mapstone, M.; Cheema, A. K.; Fiandaca, M. S.; Zhong, X.; Mhyre, T. R.; MacArthur, L. H.; Hall, W. J.; Fisher, S. G.; Peterson, D. R.; Haley, J. M., Plasma phospholipids identify antecedent memory impairment in older adults. *Nat. Med.* **2014,** *20* (4), 415-418.

Chapter 5 Concluding Remarks

5.1 Summary of Current Research Accomplishments

5.1.1 Summary of our PE proteomics research

A study involving sera from 24 controls, having term uncomplicated pregnancies and 24 cases, who developed PE later in the same pregnancy was conducted using a novel serum proteomic approach to detect low molecular weight biomarkers for PE. This study led to the detection of 66 species demonstrating statistical differences between the cases and the controls. Further statistical analysis on these biomarkers resulted in production of panels having combinations of markers with a higher sensitivity and specificity (AUC~0.99) as compared to individual markers. Chemical structure characterization MS/MS studies were successfully performed on 19 biomarkers. Six of these biomarkers were identified to be peptides based on the fragmentation pattern, while 13 were found to be lipids belonging to the class of glycerophosphocholines or sphingomyelins. Multi-marker panels were also developed using only the validated chemically characterized candidate markers that demonstrated AUC of around 0.8. Developed panels of serum biomarkers appeared effective in identifying pregnant women at 12–14 weeks gestation at risk of PE later in their pregnancy.

5.1.2 Summary of our PE lipidomics research

A discovery study was performed using sera collected at 12-14 weeks pregnancy from 27 controls with uncomplicated pregnancies and 29 cases that later developed PE in that pregnancy to find lipid biomarkers for PE. Lipids were extracted from the serum samples using organic solvents and were analyzed using direct infusion mass spectrometry. MS signals demonstrating statistical

differences were selected. This study led to detection of 45 potential lipid biomarkers. These lipid markers were reevaluated in a second confirmatory study having 43 controls and 37 preeclampsia cases. Of the 45 biomarkers from the first study, 23 markers continued to be statistically significant in the second confirmatory set. Most of these markers, representing several lipid classes, were chemically characterized typically providing lipid class and putative chemical structures using tandem MS. These preeclampsia biomarkers were found to belong to diverse lipid classes including fatty acids, cholesteryl esters, phosphatidylcholines, lysophosphocholines, triglycerides and sterols. Several multi-marker panels with AUC >0.85 and high predictive values were developed from the 23 validated biomarkers. Developed panels of serum lipidomic biomarkers appear able to identify most women at risk for preeclampsia in a given pregnancy at 12-14 weeks gestation.

5.1.3 Summary of AD lipidomic project

A shotgun lipidomic approach was undertaken to determine if lipid biomarkers exist that can discriminate AD cases (all stages) from controls. The discovery study involved sera from 29 any stage AD cases and 32 age and gender matched controls. Lipid extraction was performed using organic solvent and the samples were directly infused into a time-of-flight mass spectrometer. Statistically significant lipid markers demonstrating differences between AD cases and controls were detected. These potential lipid diagnostic biomarkers were reevaluated in a second confirmatory study involving 27 cases and 30 controls. The initial study detected 89 potential AD markers. Of these, 35 markers continued to be statistically significant in the second confirmatory set. Tandem MS studies were performed and almost all of the confirmed markers were chemically classified and characterized. Using the confirmed markers, several multi-marker panels with AUC

> 0.87 were developed for any stage AD cases vs controls. Additionally, using confirmed biomarkers, models were developed for individual stages of AD. Multi-marker panels with AUCs > 0.90 were developed for each specific CDR vs controls, including the earliest stage of AD. These lipidomic biomarkers are likely to distinguish AD cases from controls at any stage.

5.2 Limitations of our current research

5.2.1 Limitations of our PE proteomics research

Although this study provided many statistically significant potential biomarkers, some limitations need to be mentioned. This study was initially started with the purpose of detecting low M.W. peptides as biomarkers for PE. However, when tandem MS studies were performed on potential biomarkers, several markers were identified to belong to a lipid class of glycerophosphocholines. Several other markers remained unidentified probably belonging to some other class of metabolites. There were a number of markers which could not be fragmented due to lack of sample availability.

The micro capillary column used for peptide separation was packed with POROS R1 slurry having a particle size of 10 μm. This resulted in broader, less resolved and overlapping peaks complicating the quality of data and making the analysis at times difficult.

Although desirable, a validation study having a large set of PE cases and controls could not be carried out to re-evaluate the performance of detected potential biomarkers due to unavailability of serum samples. Such large prospective pregnancy studies are very few and none to our knowledge have also sampled women at 12-14 weeks gestation. Consequently, either later gestational time points might be studied in existing pregnancy if access to sera could be obtained

or the NIH would need to initiate a new prospective study that included these very early gestation age samples being collected.

5.2.2 Limitations of our PE lipidomics project

This study, involving direct infusion of lipid extracts into a TOF mass spectrometer provided several potential lipid biomarkers which in combination resulted in good predictive values. Some issues still need to be addressed.

Shotgun lipidomics using direct infusions mass spectrometry has an advantage of smaller run times leading to higher sample throughput and faster analysis. However, for comprehensive analysis of lipids, LC/MS is still preferred as it would result in 5 fold more MS variables or features being observable[1]. Also, separation of isobaric peaks is not always obtained in direct infusion mass spectrometry. Therefore, lipids having the same or very similar molecular masses belonging to either the same or different lipid classes are not always separated by direct infusion techniques using a TOF-MS.

Direct infusion also results in ion suppression due to so many different ionization efficiencies of lipids from diverse classes being introduced simultaneously. Although normalization strategies help, they do not eliminate the problem.

The reproducibility of the runs is also quite variable with directly infused samples as the sensitivity of the instrument varies for incompletely understood reasons (e.g. residual solvent) or the amount of specimen introduced varies due to clogging of the ESI needle and there are other sources of instrument or ionization variations.

Unlike proteomics studies, there are no biological assays to quantitate total lipids in human specimens. Similar to other lipid researchers, we also used the equal volume of different samples and infused it into mass spectrometer without knowing if total lipid concentration varied. This could lead to variability in peak height as the total lipid concentration may vary from person to person or from extraction to extraction.

5.2.3 Limitations of AD lipidomic project

Many of the analytical concerns here are identical to those described in the PE lipidomic study. Additionally, many AD cases detected with CDR of 0.5, 1 and 2 might be on some prescribed medications. This is typically an aged population with many co-morbities. These medications may have altered some biomarkers or produced artifacts in the study.

This study was carried out on a particular population group with limited numbers of cases and controls. It would be useful to carry out the study on populations from different parts of world as these biomarkers might be altered in other populations with or without disease.

The cases included small numbers of patients with CDR 0.5, 1 and 2. A greater set of patients each with CDR 0.5, 1 and 2 would allow us to compare the levels of these biomarkers with severity of the disease and have greater confidence in the markers found and may also allow us to discover and validate markers specific to a particular stage of disease.

5.3 Future objectives

5.3.1 Future research objectives of the PE proteomics study

This study appears promising with many novel candidate serum biomarkers discovered and many effective combinations developed providing potentially valuable risk assessment of PE. Future studies will need to retest and validate these candidate markers in independent specimens from different populations with laboratory analysis carried out without knowledge of case or control status. This will require accumulation of or access to similarly timed specimens, processed expeditiously to allow for similar results and not submitted to freeze thaw cycles, which degrade proteins. If the markers are shown to be valid over time and considered useful in the screening of pregnant women, the MS method could be scaled up (as was done for the preterm labor marker fetal fibronectin[2] to analyze high numbers of specimens at a reference laboratory. Alternatively, individual high throughput assays, for example, immunoassays, may be developed for several of the best markers as has been done for other clinical assay panels (lipids, cardiovascular risk factors, and so on). Also, these markers themselves suggest biochemical changes occurring early in the pregnancies of women who will develop PE many weeks later. Further studies of these suggested pathways or processes, for example, ROS, would be helpful in more fully explaining or verifying very early pathologic changes leading to PE. Having useful, verified biomarkers that predict later PE would allow for far smaller clinical studies, including evaluation of pharmacologic or other medical treatments for this potentially fatal disease.

5.3.2 Future research objective of PE lipidomics study

Although, we carried out an initial discovery study and a validation study which resulted in 23 out of 45 lipid markers being validated, several future studies could be proposed.

Numerous validation studies including larger sets of PE cases and controls from different parts of the world should be carried out before clinical assays could be developed for these markers.

This study was carried out in the positive ion mode. Similar studies in the future involving samples run in the negative ion mode will allow us to detect anionic and weakly anionic lipids.

Studies involving treatment of placental cells with synthetic standards of the individual lipid markers might provide some insights into the biological activity or roles of individual lipid biomarkers in PE.

Quantitative studies should be carried out in the future by using known concentrations of the internal standards of the validated lipid biomarkers to detect amounts of these markers in the diseased patients. This would dramatically reduce pre-analytical and analytical variability and allow the biological variability to be better established.

5.3.3 Future research objectives for our AD lipidomics study

In this study, the discovery set let to detection of 89 lipid biomarkers. In the confirmatory set 35 out of 89 markers were validated. Although, a confirmation study was performed, it is well recognized that replication of the results from this study needs to be performed in other larger and more diverse sets of populations. It would also be of interest to monitor changes in serum biomarkers over time in order to study the stability of the lipid markers. Further, given the alterations in lipids in response to AD, it would be useful to more directly study possible biologic roles for some of the individual molecular species of lipids in the context of AD.

5.4 References

1. Lin, L.; Yu, Q.; Yan, X.; Hang, W.; Zheng, J.; Xing, J.; Huang, B., Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabonomics? A serum metabonomic study of kidney cancer. *Analyst* **2010,** *135* (11), 2970-2978.

2. Lockwood, C. J.; Senyei, A. E.; Dische, M. R.; Casal, D.; Shah, K. D.; Thung, S. N.; Jones, L.; Deligdisgh, L.; Garite, T. J., Fetal fibronectin in cervical and vaginal secretions as a predictor of preterm delivery. *N. Eng. J. Med.* **1991,** *325* (10), 669-674.