Sickle cell disease associated lipid changes and their relevance towards the disease pathogenesis And Lipid Biomarkers and Embryo Quality in In Vitro Fertilization; Pregnancy Success Differentially Expressed by Body Weight

Suad Alshammari

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/6851

This Thesis is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
Sickle cell disease associated lipid changes and their relevance towards the disease pathogenesis

And

Lipid Biomarkers and Embryo Quality in In Vitro Fertilization; Pregnancy Success Differentially Expressed by Body Weight

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

By

Suad Alshammari

Director:

Dayanjan S Wijesinghe

Virginia Commonwealth University

Richmond, Virginia

12, 2021
Acknowledgment

I would like to express my special thanks to my principal supervisor Dr. Dayanjan S Wijesinghe who guided me through my education and gave me the opportunities to achieve these projects. Also, I would thank my colleagues for their help and supports. A special thanks to my family for their supports through my journey.
TABLE OF CONTENTS

Contents

Table of Figures: .............................................................................................................................. 5

Chapter 1: ...................................................................................................................................... 6

Sickle cell disease associated lipid changes and their relevance towards the disease
pathogenesis: Literature Review ................................................................................................... 6

Introduction ....................................................................................................................................... 8

1 Background ...................................................................................................................................... 8

2 Pathophysiology ............................................................................................................................. 9

3 SCD Genotypes differentiation ........................................................................................................ 10

4 Global burden of SCD: .................................................................................................................. 15

5 SCD cell Disease in Saudi Arabia and the U.S ............................................................................ 16

6 Clinical manifestation and complication of SCD ........................................................................... 17

7 Method: .......................................................................................................................................... 18

8 Clinical manifestation and changes on lipidome in SCD ............................................................... 22

8.1 Pain and VOC ............................................................................................................................... 22

8.2 Acute chest syndrome (ACS) ...................................................................................................... 25

8.3 Hyper-coagulopathy and SCD ..................................................................................................... 26

9 Treatments ...................................................................................................................................... 29

9.1 Hydroxyurea ............................................................................................................................... 29

9.2 L-glutamine .................................................................................................................................. 30

9.3 Bone marrow transplant .............................................................................................................. 31

9.4 Emergency management .............................................................................................................. 31

9.4.1 VOC management: ................................................................................................................ 31

9.4.2 ACS management: ................................................................................................................... 32

10 Discussion and future directions ................................................................................................ 32

Chapter 2: .......................................................................................................................................... 35

Lipid Biomarkers and Embryo Quality in In Vitro Fertilization; Pregnancy Success
Differentially Expressed by Body Weight ........................................................................................ 35
Lipid Biomarkers and Embryo Quality in In Vitro Fertilization; Pregnancy Success Differentially Expressed by Body Weight

Abstract

Introduction

Materials and Methods

Results

Discussion

Bibliography

Supplementary appendix:
Table of Tables:

Table 1: SCD genotype, severity of complication, region, effected ethnicities, and percent of normal hemoglobin for each type: ..........................................................12
Table 2: Annual estimation of SCD patient according to WHO: ..............................................15
Table 3: Research terms used for the literature review: ...............................................................18
Table 4: Complications of SCD and a summary of the main characteristics..............................33
Table 1. Demographic and Baseline characteristics shows the similarities between normal weight women (not obese) and obese women.........................................................49

Table of Figures:

Figure 1 – Pathophysiology of SCD.....................................................................................10
Figure 2: Probability of gene inheritance in SCD.................................................................11
Figure 3: Number of newborns with SCA in 2010..............................................................16
Figure: 4 PRISMA Flow Diagram.........................................................................................21
Figure 1: Statistical analysis workflow..................................................................................48
Figure 2: Metabolites model for 2PN acquired from PPP sample......................................53
Figure 3: PL model for 2PN acquired from FF sample.......................................................54
Figure 4: rLDA of 6-weeks of pregnancy and EHR. ............................................................57
Figure 5: rLDA of 6-weeks of pregnancy and metabolites acquired from PPP. ..............59
Figure 6: rLDA of 6-weeks of pregnancy and lipids without GL species acquired from FF sample. ......................................................................................................................60
Figure 7: Final model for BMI from EHR data and PPP small molecules.........................64
Figure 8: Final model for BMI from EHR data and FF small molecules............................65
Chapter 1:

Sickle cell disease associated lipid changes and their relevance towards the disease pathogenesis: Literature Review
Sickle cell disease associated lipid changes and their relevance towards the disease pathogenesis: Literature Review

Suad Alshammari¹,², Monther Alsultan¹,³, Joshua Morriss¹, Silas Contaifer¹, Daniel Contaifer¹, Wally R. Smith, Donald F. Brophy and Dayanjan S Wijesinghe¹,⁴,⁵

¹ Department of Pharmacotherapy and Outcomes Sciences, School of Pharmacy, Virginia Commonwealth University, Richmond, Virginia, USA.
² Faculty of Pharmacy, Northern Border University, Saudi Arabia.
³ College of Clinical Pharmacy, King Faisal University, Saudi Arabia
⁴ Da Vinci Center, Virginia Commonwealth University, Richmond, Virginia, USA
⁵ Institute for Structural Biology Drug Discovery and Development (ISB3D), School of Pharmacy, Virginia Commonwealth University, Richmond, Virginia, USA

Abstract

Sickle cell disease (SCD) is a group of genetic disorder that occurs due to genetic mutation of a beta-globin gene that lead to production of pathogenic hemoglobin S (HB S). Genotypes of SCD include Hb SS (sickle cell anemia) which is the most common and severe form of SCD affects about 20 to 25 million people worldwide, HbSC, Hb Sβ+- thalassemia, and Hb Sβ0-thalassemia. SCD is characterized by multiorgan complications that, in turn, affect lipids composition. During hypoxia a sequence of changes will take place such as HbS polymerization, erythrocyte rigidity and stickiness, and oxidative stress. The combination of these changes will affect lipids components such as polyunsaturated fatty acids (PUFA), which are substrates for a
significant number of bioactive lipids such as the eicosanoids and some of the endocannabinoids. For example, vaso-occlusion crisis, the most common cause of SCD hospitalization, is found to be accompanied by changes in PUFA components of RBCs cell membrane encompassing Omega-3 and Omega-6. This comprehensive review outlines lipid changes that accompany SCD and also identify the gaps in our knowledge. This review will also allow us to devise better treatment options to manage the different pathophysiology and complications of SCD.

**Introduction**

1 Background

Sickle cell disease (SCD) is a group of genetic disorders. It is caused by genetic mutation in the beta-globin gene which results in the production of abnormal hemoglobin S (HbS), or the HbS trait. According to the World Health Organization (WHO), 20-25 million cases of global SCD are of the homozygous HbSS subtype. The vast majority of these cases are located in sub-Saharan Africa, accounting for approximately 12-15 million people, with India as a close second of nearly 5-10 million cases.

The biochemical variations that accompany the SCD mutation are numerous. Among such variation are expected difference in the lipidome. Lipidomics is a field concerned with investigating the roles lipids have in various diseases; the primary functions of lipids include them serving as energy reservoirs, signaling molecules, and to compartmentalize (via membrane formation) various biochemical reactions. Understanding the structures and functions of these molecules can allow for a better understanding of pathways that underly any given disease or injury. For example, the high oxidative stress caused by sickle cell disease may result in a high degree of oxidation of polyunsaturated fatty acids (PUFAs), which are substrates for a significant number of bioactive lipids such as the eicosanoids. Despite more apparent metabolic modulations, the lipidome has not been studied in detail with respect to SCD. Here we outline what is known to
date with regards to the lipid and metabolic changes that accompany SCD, and also identify the
gaps in the current literature with respect to these topics. A better understanding of the lipid and
metabolic changes accompanying SCD will allow us to devise better treatment options to manage
the different pathophysiology of SCD.

2 Pathophysiology

At the molecular level, normal hemoglobin (Hb) consists of two α and two β chains
genetically coded by genes on chromosomes 16 and 11. When a mutation occurs on
chromosome 11, the nitrogen base adenine codon GAG is substituted by thymine (GTG) leading
to replacement of glutamic acid (hydrophilic) to valine (hydrophobic) in position 6 of the globin’s
β chain n-terminal end. The ultimate result of this process is the formation of pathogenic HbS.

During hypoxia, the deoxygenated HbS upregulates a long chain of insoluble polymers, causing
rigidity and damage to the red blood cells (RBCs), resulting in the characteristic “sickle” shape.
Furthermore, HbS gives RBCs a “sticky” property due to the secretion of adhesion molecules
such as Integrin α4β1 which interact with endothelial cells’ membranes, altering the blood supply.
Consequently, this affects the half-life of RBCs which is reduced to 10-20 days instead of 120
days as is the half-life of normal RBCs.

A series of changes occur due to the chromosome 11 mutation (Figure 1). During sickling,
hemolysis occurs releasing free hemoglobin into the serum, which causes calcium pump
dysfunction as RBCs obtain excess Na+ and Ca2+ ions with a loss of K+ ions and water.
Additionally, nitric oxide (NO) production is inhibited due to hypoxia, enabling sickle cells to stick
more readily to the endothelium. Moreover, the extracellular hemoglobin interacts with NO leading
to decreased nitric oxide concentration in plasma and endothelial tissue, inducing vascular
complications such as vasoconstriction.
Figure 1 – Pathophysiology of SCD: a: The β globin chain mutation encoded in chromosome 11 results in substitution of adenine (GAG) by thymine (GTG), enabling the “sickled” erythrocyte. b: Under certain conditions such as those from environmental factors and hypoxia, Hbs is deoxygenated, causing long-chain polymerization which activates the sickling of erythrocyte.

3 SCD Genotypes differentiation

When a single copy HbS inherited from each of the parents leads to homozygous HbSS also known as sickle cell anemia which is a very common SCD disorder\textsuperscript{11,12}. Coinheritance of HbS with other pathogenic hemoglobin forms results in heterozygous classes of SCD \textsuperscript{12}. Heterozygous genotypes include hemoglobin SC (HbSC) HbSβ\textsuperscript{+}- thalassemia, HbSβ\textsuperscript{0}-thalassemia and other rare types as can be seen in Figure 2 \textsuperscript{4,8}. 
Figure 2: Probability of gene inheritance in SCD: If each of the parents carries a single copy of pathogenic hemoglobin (HbAS, HbAC, HbSβ\(^0\), HbSβ\(^+\)), the probability for children to inherit the disease as follows: 25% are affected (HbSS, HbSC, HbSβ0, HbSβ+), 50% are carriers (HbAS, HbAC, HbSβ0, HbSβ+), and 25% are normal (HbAA).

One of the main consequences of SCD is the onset of anemia. Anemia is caused by reduction in hemoglobin (Hb) concentration level. The short lifespan of RBCs in SCD, again approximately 10-to-20 days in contrast to 120 days for healthy individuals, causes a decreased number of RBCs and prompts a reduction in oxygen delivery to tissues (anemia). Hb in healthy individuals have either HbA, HbA2 and HbF genotypes, usually maintaining oxygen homeostasis.
throughout the body. However, sickle cell anemia (SCA) incurred by the inherited HbS gene is associated with different sub-genotypes: HbSS(Sickle Cell Anemia), HbSC, HbSβ0 thalassemia, HbSβ+ thalassemia, HbSα thalassemia, HbSD and HBAS (Sickle cell trait). The different genotypes, their respective phenotypes and prevalence are shown in Table 1.

Table 1: SCD genotype, severity of complication, region, effected ethnicities, and percent of normal hemoglobin for each type:

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Definition1</th>
<th>Severity of clinical manifestations1,2</th>
<th>Ethnicities or regions affected1,3</th>
<th>Prevalence among SCD1,5</th>
<th>% of normal vs abnormal Hb4,6</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbSS</td>
<td>Sickle Cell Anemia homozygous for S globin1,4</td>
<td>Severe or moderately severe and associated with shortest survival duration</td>
<td>70% of African ancestry</td>
<td>74-76%</td>
<td>HbA2:&lt;3.5% HbF:5-15% or higher in rare cases</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>HbSC</td>
<td>Double heterozygous for HbS and HbC</td>
<td>Moderate severity</td>
<td>25%-30% of African ancestry</td>
<td>18-24%</td>
<td>HbA2:&lt;3.5% HbF:1-5% or higher in rare cases</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>HbSB0 Thalassemia</td>
<td>Double heterozygous for HbS and b0 Thalassemia</td>
<td>Severe and indistinguishable from SCA</td>
<td>Most common among Middle eastern and Indian ancestry</td>
<td>1-6%</td>
<td>HbA2:&lt;3.5% HbF:2-15% HbS:80-92%</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>Hemoglobin Syndrome</td>
<td>Description</td>
<td>Symptoms</td>
<td>Prevalence</td>
<td>Hemoglobin Values</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
<td>----------</td>
<td>------------</td>
<td>------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>HbSb+ Thalassemia</td>
<td>Double heterozygous for HbS and b+ Thalassemia</td>
<td>Mild to moderate and varies among different ethnicities</td>
<td>Most common among those of Middle eastern and Indian ancestry</td>
<td>HbA:3-30%</td>
<td>1,2,3,4</td>
<td></td>
</tr>
<tr>
<td>HbSa Thalassemia</td>
<td>Double heterozygous for HbS and a Thalassemia</td>
<td>Mild to moderate and varies based on type of thalassemia inherited</td>
<td>30% in African region 50% in Middle Eastern 50% of Indian ancestry</td>
<td>N/A</td>
<td>1,2,3,4,5</td>
<td></td>
</tr>
<tr>
<td>HbSD</td>
<td>Double heterozygous for HbS and HbD</td>
<td>Severe symptoms</td>
<td>Most common in those of Northern Indian ancestry</td>
<td>HbA2: 3.8%</td>
<td>1,2,3,6</td>
<td></td>
</tr>
<tr>
<td>HbS/HPFP</td>
<td>HbS and hereditary persistence of fetal Hb</td>
<td>Very mild or not symptomatic due to high HbH</td>
<td>Reported in different ethnicities and multiple regions worldwide</td>
<td>N/A</td>
<td>1,2,3,4,5</td>
<td></td>
</tr>
<tr>
<td>HbS/HbE syndrome</td>
<td>Double heterozygosity for HbS and HbE</td>
<td>Very mild symptoms or sometimes similar to HbSb+ Thalassemia</td>
<td>Southeast Asian ancestry</td>
<td>HbE:&lt;30%</td>
<td>1,2,3,4,5</td>
<td></td>
</tr>
<tr>
<td><strong>HbAS</strong></td>
<td>Sickle cell trait A combination of a single HbS with normal Hb</td>
<td>No symptoms</td>
<td>AA-7 to 10% Sub-Saharan African ancestry 20% Latino ancestry: 0.2-6.3% Indian ancestry: 13% Middle Eastern ancestry: 0.2-27% Grecian ancestry: 4-10%</td>
<td>Not considered a form of SCD</td>
<td>HbA: 50-60% HbA2: &lt;3-30% HbF: &lt;2% HbS: 36-45%</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Other rare phenotypes combinations</strong></td>
<td>Abnormal combination of HbS with HbD Los Angeles, G-Philadelphia or HbO Arab</td>
<td>Very rare combinations and clinical symptoms and may develop with severe hypoxia</td>
<td>Very rare</td>
<td>N/A</td>
<td>1,2,3,4,5</td>
<td></td>
</tr>
</tbody>
</table>


4 Global burden of SCD:

Globally, SCD is common in countries where malaria is widespread. Annually, 275,000 to 300,000 children are born with SCD about 75% to 85% of them occur in Africa (Table2). However, these numbers are expected to reach 400,000 births/year by 2050. Regarding the survival rate, in developed countries, children with SCD survive to adolescence. Whereas, in developing countries, the mortality rate among children with SCD is higher. For example, in Africa, 50% to 80% of children with SCD die before five years old of age and 16.5% in West Africa (Aygun & Odame, 2012).

Table 2: Annual estimation of SCD patient according to WHO:

<table>
<thead>
<tr>
<th>Region</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>African region</td>
<td>233,289</td>
</tr>
<tr>
<td>American region</td>
<td>9,047</td>
</tr>
<tr>
<td>Eastern Mediterranean region</td>
<td>6,491</td>
</tr>
<tr>
<td>European region</td>
<td>1,292</td>
</tr>
</tbody>
</table>
### Figure 3: Number of newborns with SCA in 2010

The figure is reproduced using data by R software version (3.6.3).

### 5 SCD cell Disease in Saudi Arabia and the U.S

SCD is one of the most prevalent genetic disorders in Saudi Arabia. It was detected for the first time in the Eastern province in the 1960s. This discovery initiated multiple screening...
investigations across different regions of Saudi Arabia to determine the impact of SCD. The prevalence of SCD among the Saudi population is still not fully understood. However, the Premarital Screening Program (PSP) from 2004 to 2005 reported that nearly 4.2% of screened adults carried the HbS gene, or exhibited the SCD trait, and 0.26% of the total population had SCD. For thalassemia, 3.22% carried the trait and 0.07% were affected. Another study conducted in 2011 suggested that somewhere between 2% to 27% of the Saudi population carry SCD trait. Nonetheless, newborn screening for SCD is absent in Saudi Arabia and, therefore, the true prevalence is likely underestimated.

In the U.S, the sickle cell trait (SCT) differs greatly across states, between different ethnicities and races. It is likely that many SCT carriers in the United States are not aware of their medical sickle cell condition due to the mildness of symptoms. Thus, the exact number of cases in the U.S is unspecified. However, it has been estimated that roughly 100,000 American are affected with SCD. Additionally, approximately 15.5 out of every 1000 American births have been speculated to have SCD, of which the group with the highest prevalence were African Americans representing 73.1 cases per 1000 births. In 2004, the average hospitalization with underlying SCD-related complications reached 83,149 with a cost of roughly $488 million.

6 Clinical manifestation and complication of SCD

SCD is pervasive by its multisystem effects that are a result of vaso-occlusion (VOC). These effects include acute pain crises and chronic complications. VOC occurs as a consequence of sickled RBCs and the short life cycle of erythrocytes. Damaged erythrocytes block the blood supply in small vessels leading to VOC and organ damage. Although VOC’s acute pain crisis is the most common SCD complication that results in hospitalization, chronic complications often come from VOC including cardiovascular, pulmonary, nervous, gastrointestinal, musculoskeletal, urogenital, spleen and systemic inflammation. These complications can vary across each
patient, and could range from mild to severe problems. Even though some complications can be improved with current therapeutic interventions, such as pain crisis and leg ulcers, there is little evidence supporting the improvement in other more chronic, end-stage complications like renal dysfunction.

The SCD complications’ severity can vary among haplotypes and coinheritance of fetal hemoglobin (HbF) and α-thalassemia. HbF, found in newborns, has a higher affinity to carry oxygen than normal adult Hb α2β2; thus, SCD-related symptoms start to emerge after six months postpartum, when HbF levels decline. It has been demonstrated that individuals with homozygous SCD in the Saudi-Indian haplotype have a high level of HbF, exhibiting reduction in SCD complications as a result of the decrease in HbS polymerization rate. Another genetic modifier that impacts the manifestation of SCD complications is the coinheritance of α-thalassemia. Coinheritance of α-thalassemia incurs a reduction in the HbS polymerization rate, which leads to an increase in Hb concentration and the RBCs’ lifespan. Also, it reduces hemolysis of RBCs leading to decrease in some complications like stroke and leg ulcers. Despite these advantages, α-thalassemia increases the progression of other complications such as pain crisis and acute chest syndrome.

7 Method:

We acquired the studies by searching in PubMed. In the beginning, we searched using two terms (sickle cell and lipidome), and we get only 9 publications. After that, we tried to expand our research terms by using the disease name, lipid types, and the complications of our interest. The detailed terms are represented in table 3.
<table>
<thead>
<tr>
<th>Database</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>PubMed</td>
<td>(((&quot;sickle&quot;[All Fields] OR &quot;sickled&quot;[All Fields] OR &quot;sickles&quot;[All Fields] OR &quot;sickling&quot;[All Fields]) AND (&quot;cells&quot;[MeSH Terms] OR &quot;cells&quot;[All Fields] OR &quot;cell&quot;[All Fields]) AND (&quot;fatty acids&quot;[MeSH Terms] OR (&quot;fatty&quot;[All Fields] AND &quot;acids&quot;[All Fields]) OR &quot;fatty acids&quot;[All Fields]) OR (&quot;glycerolipid&quot;[All Fields] OR &quot;glycerolipids&quot;[All Fields]) OR (&quot;glycerophospholipids&quot;[MeSH Terms] OR &quot;glycerophospholipids&quot;[All Fields] OR (&quot;glycerophospholipid&quot;[All Fields]) OR (&quot;sphingolipids&quot;[MeSH Terms] OR (&quot;sphingolipid&quot;[All Fields]) OR (&quot;sterols&quot;[MeSH Terms] OR &quot;sterols&quot;[All Fields] OR &quot;sterol&quot;[All Fields]) OR (&quot;prenol&quot;[Supplementary Concept] OR &quot;prenols&quot;[All Fields] OR &quot;sphingolipids&quot;[All Fields]) OR (&quot;polyketides&quot;[MeSH Terms] OR &quot;polyketides&quot;[All Fields]) OR (&quot;saccharolipids&quot;[All Fields]) AND (&quot;acute chest syndrome&quot;[MeSH Terms] OR (&quot;acute&quot;[All Fields] AND &quot;chest&quot;[All Fields] AND &quot;syndrome&quot;[All Fields]) OR (&quot;vaso-occlusive&quot;[All Fields] OR (&quot;vasooclusion&quot;[All Fields] OR (&quot;blood coagulation&quot;[MeSH Terms] OR (&quot;blood&quot;[All Fields]) AND (&quot;thrombin&quot;[MeSH Terms] OR &quot;thrombin&quot;[All Fields]) OR (&quot;hypercoagulable&quot;[All Fields]) OR (&quot;hypercoagulability&quot;[All Fields]) OR (&quot;hypercoagulation&quot;[All Fields]) OR (&quot;hypercoagulative&quot;[All Fields]) OR (&quot;thrombophilic&quot;[All Fields]) OR (&quot;thrombophilia&quot;[MeSH Terms] OR (&quot;hypercoagulant&quot;[All Fields]) OR (&quot;hypercoagulative&quot;[All Fields])) AND (&quot;proteomic&quot;[All Fields] OR &quot;proteomic&quot;[MeSH Terms] OR (&quot;proteomics&quot;[MeSH Terms] OR &quot;proteomics&quot;[All Fields]))) NOT (&quot;covid 19&quot;[All Fields] OR &quot;covid 19&quot;[MeSH Terms] OR &quot;covid 19&quot;[All Fields]) OR (&quot;covid 19&quot;[All Fields])) OR &quot;covid 19&quot;[All Fields]) OR (&quot;covid 19&quot;[MeSH Terms] OR &quot;covid 19&quot;[All Fields])) OR &quot;covid 19&quot;[All Fields]) OR (&quot;covid 19&quot;[MeSH Terms] OR &quot;covid 19&quot;[All Fields])))</td>
</tr>
</tbody>
</table>
In PubMed, the search generated 33 articles. The duplicated and irrelevant articles were excluded. After that, the titles and abstracts were reviewed. The article that was investigated unapproved medications, evaluated types of oxidative stress compounds, and investigated a novel agent were excluded. We also exclude articles that do not include lipid changes and evaluate unrelated complications. We searched for additional articles manually using “google scholar”. This search method is limited to lipids changes in sickle cell disease (not include the introduction, background, and disease burden sections).

Figure: 4 PRISMA Flow Diagram.
8 Clinical manifestation and changes on lipidome in SCD

8.1 Pain and VOC

VOC acute pain is the most common cause of SCD hospitalization. Pain episodes vary from one patient to another according to stress, physical, and physiological conditions. In SCD, the mechanism of pain is more complicated than the pain mechanisms in other diseases due to an involvement of multiple factors. These factors include HbS polymerization and adhesion molecules; HbS polymerization leads to RBC sickling, the release of inflammatory mediators, mast cell activation, and the release of endothelial factors; adhesion molecules, from both red and white blood cells. During the VOC pain crisis, the accumulation of inflammatory mediators leads to nociceptor activation. For example, increased interleukin-1 (IL-1) is associated with elevated prostaglandins E2 (PGE2) and I2 (PGI2) synthesis which, in turn, causes nociceptor activation, inducing nerve fiber damage and pain. Prostaglandins have also been found to affect signaling pathways that cause hyperalgesia and increase pain sensation. The enzyme cyclooxygenase-2 (COX-2) converts fatty acid arachidonic acid (AA) to prostaglandin E2 (PGE2), and oxidizes endocannabinoids 2-arachidonoylglycerol (2-AG) into prostaglandin glycerol esters (PG-Gs) and anandamide into prostaglandin ethanolamides (prostamides), respectively.

HbS polymerization affects the RBC membrane. In SCD, there is an elevation in phosphatidylserine (PS) expression both extra-and-intra-cellularly. According to van Tits et al. (2009), there is a greater increase in PS levels and coagulation factors during pain crises. PS contributes to VOC episodes through the activation of endothelial cells and causes adhesion of damaged (sickled) RBCs. In addition, secretory phospholipase A2 causes PS hydrolysis, which in turn leads to thromboxanes and prostaglandins generation, subsequently incurring the onset of acute chest syndrome—the leading cause of SCD deaths. PS has a role in shortening the lifespan of RBCs. Besides PS, there are also changes in the polyunsaturated fatty acid (PUFA)
components of RBCs’ cell membranes. Altered PUFAs include increase AA 20:4n-6, adrenic acid (22:4n-6), and osbond acid (22:5n-6) with decreased linoleic acid (LA, 18:2n-6), EPA (20:5n-3), and DHA (22:6n-3)27–29. These changes may participate in VOC by facilitating RBCs-endothelial interaction. Daak et al. demonstrated the efficacy of omega-3 fatty acid supplementation in HbSS patients, as they found a reduction of hospitalizations pertaining to SCD pain and VOC in patients that were administrered EPA and DHA supplements. In addition, during VOC, there is an elevation in thromboxane B2 and 6-keto-prostaglandin F1α (6kPGF1α). This finding reflects the imbalance between vasoconstrictr and vasodilator in SCD patients.

Opioids are the most commonly used analgesics for the management of both acute and chronic pain in SCD. Although morphine is the first-choice opioid for pain management in SCD, it causes mast cell activation, and this leads to hyperalgesia. In addition, morphine causes tolerance, respiratory depression, and nephropathy. In pain conditions not pertaining to SCD, persistent long-term opioid use leads to opioid receptors adaptation, efficacy reduction, and pronociceptive neurotransmitters activation (pronociceptive neurotransmitters have a contrary action of opioids analgesic effects). The neural adaptation to opioids is extremely complex and encompasses multiple pathways, such as activation of calcitonin gene-related peptide (CGRP) and neuropeptide substance P (sP), as well as downstream messengers derived from the aforementioned AA metabolism through cyclooxygenase (COX), lipoxygenase (LOX) and endocannabinoid pathways. During acute opioids treatment, sensory transmitters are inhibited, causing partial analgesia. In chronic opioids use, the secretion of CGRP and sP are increased; also, there is postsynaptic activation of CGRP, NK-1, N-methyl-d-aspartate (NMDA) receptors, coinciding with the synthesis of PGs through cyclooxygenase pathways, and lipoxygenase metabolites (LT) through lipoxygenase pathways.

New strategies have been considered to reduce these side effects and increase efficacy by using adjuvant drugs or nutrients. One of these strategies involves combining morphine with omega-3 fatty acids. Omega-3 fatty acids such as EPA and DHA have anti-inflammatory and
antinociceptive effects. This antinociceptive effect provides a new strategy to prevent and treat chronic pain in several diseases, and may improve the efficacy and reduce its side effects. A study conducted on male Wistar rats demonstrated analgesic and antinociceptive effects of omega-3 after 16 days of administration. In addition, using omega-3 fatty acid in combination with morphine induced the antinociceptive effect in both acute and chronic morphine use as well as reduction of tolerance in chronic morphine treatment. Furthermore, when used with omega-3 fatty acids, morphine’s analgesic effects were compounded despite administration with a sub-therapeutic dose. The mechanism of omega-3 fatty acid supplementation producing antinociceptive is still poorly understood. However, it could be through (i) inhibition of the AA cascade, which leads to blocking the production of pro-inflammatory eicosanoids and cytokines, or (ii) through the production of E-series (RvE1, RvE2) and D-series (RvD1, RvD2) resolvins from EPA and DHA as precursors that have potent analgesic actions. This combinational therapy to our knowledge has not been tested in humans to treat SCD-linked pain and VOC. However, contrary to previous findings, Matte et al. (2019) evinced that using DHA in HbSS mice do not increase RvD1 level in plasma, eliciting the need for further study in the effects of omega 3 fatty acids.

The endocannabinoids system are another group of lipids that participate in pain regulation. Some studies have suggested that increased concentrations of endocannabinoids in the peripheral nervous system leads to reduced hyperalgesia severity in chronic pain. A study conducted on HbSS-BERK sickle mice using inhibitor of fatty acid amid hydrolase (FAAH), URB597, lead to increased anandamide in the periphery by suppressing its degradation through inhibition of FAAH, the enzyme responsible for endogenous anandamide breakdown. That lead to decreased hyperalgesia and nociceptor sensitization via anandamide binding to CB1 receptor on the nociceptor terminal. Anandamide stimulates eryptosis at concentrations of 2.5 μM; eryptosis is inherently enhanced in SCD mouse models. In SCD, eryptosis stimulation may lead to anemia.
17R-resolvin D1 (17R-RvD1) is a novel promising therapeutic used to suppress inflammation related to SCD. A study conducted on humanized mice model assessed the impairment of specialized pro-resolving lipid mediators involving resolving-D1 (RvD1) and 17R-RvD1 during hypoxia/reoxygenation. In this in vivo model, coupled to an ex vivo model using SCD human blood, blood treated with RvD1 and 17R-RvD1 saw a reduction in the adhesion of leukocytes to tumor necrosis factor-a (TNF-a)–activated endothelial cells. This finding was validated in vivo by the HbSS mice model, which revealed the efficacy of both RvD1 and 17R-RvD1 in diminishing leukocyte adhesion to the endothelial surface. This study also demonstrated the efficacy of administrating 17R-RvD1 in the reduction of neutrophil adhesion and transmigration in HbSS mice model. Since neutrophil adhesion preludes thrombus formation, which is considered the primary stage of VOC, prevention of neutrophil adhesion and transmigration will improve vaso-occlusion outcomes.

8.2 Acute chest syndrome (ACS)

Acute chest syndrome (ACS) is a life-threatening complication characterized by chest pain, fever, breathing difficulties, and pulmonary infiltrate which is apparent in pulmonary radiographs. It is the second most common cause of SCD patients’ hospitalization, and one of leading cause of patient’s death. ACS affects approximately 40% of SCD patients and is the underlying cause for nearly 25% of SCD-related deaths. The main three causes of ACS are pulmonary infection, pulmonary embolization, and pulmonary infarction. Indeed, pulmonary infection is considered the predominant cause of ACS. The National Acute Chest Syndrome Study assessed 671 ACS episodes for 538 SCD patients through the period of March 1993 to March 1997. The study found chlamydia infections is the most common cause for ACS episode in all age groups. However, Mycoplasma and viral pneumonia were most prevalent for ACS in young children. Besides infection, pulmonary embolism can induce ACS through infarcted bone marrow. Fat emboli is produced from infarcted bone marrow and travel throughout bloodstream...
to the lung and causes inflammation and ACS \(^{43}\). In addition, pulmonary infarction participates in ACS development through adhesion of sickled RBCs to endothelial cells that leads to VOC and hypoxia \(^{43}\).

During fat embolism, secretory phospholipase A2 (sPLA2), a potent inflammatory mediator, is upregulated followed by inflammation \(^{45}\). When the upregulated sPLA2 hydrolyses phospholipids into FFA and lysophospholipids, it generates inflammatory mediators including AA. As a consequence, AA produces farther inflammatory products that encompass thromboxane, leukotrienes, and prostaglandins \(^{45,46}\). Therefore, the remarkable elevation of sPLA2 before 24-48 hours of ACS could predict the episodes earlier in SCD patients \(^{45}\).

Lipid peroxidation also play a key role in ACS mechanism. AA undergo peroxidation through noncyclooxygenase pathway to produce F2 isoprostane. A study was conducted on HbSS and HbSC patients to assess the level of F2 isoprostanes in patients’ plasma during baseline, VOC and ACS. The study found a dramatic elevation (9-fold) of F2 isoprostanes concentration compared to normal volunteer and 6-fold increase compared to VOC \(^{41}\). In addition, F2 isoprostanes level declined after blood transfusion by 50-85%. However, transfusion leads to the incidence of bone marrow transplant rejection which is the only treatment for SCD \(^{47}\).

8.3 Hyper-coagulopathy and SCD

Coagulation activation is a common phenomenon in SCD. Markers of platelet activation can be found in SCD patients, suggesting a constant hypercoagulable state. In a study of SCD patients with PHT, Ataga et al found evidence of thrombin generation by detecting elevation of markers such as high fibrin degradation product, D-dimer, Thrombin–antithrombin complex (TAT) and prothrombin fragment 1.2 (F1+2) levels, that were higher in SCD patients than control, although not statistically significant. In addition, it was found that there were high levels of markers
for platelet and endothelial activation, such as sCD40L and sVCAM-1, respectively. These results indicate the intrinsic association of coagulation activation and endothelial dysfunction in SCD \(^{48}\). Elevated plasma levels of fibrinogen and whole-blood viscosity in SCD patients are also suggestive of a state of hypercoagulability. These markers have been associated with the pain crisis, suggesting a role of increased fibrinogen-mediated platelet aggregation in blood viscosity during pain crises in SCD patients \(^{49}\). This relationship was also reported by Tomer et al illustrating that pain episodes were associated with an elevation of plasma levels of F1+2, TAT, and D-dimer. They also demonstrated that frequency of pain episodes is correlated platelet procoagulant activity and plasma fibrinolytic activity \(^{50}\).

One of the main phospholipids involved in activation of coagulation is platelet activating factor (PAF). It is an acetyl-glyceryl-ether-phosphorylcholine that participates in the activation of many cells such as platelets, endothelial cells, neutrophils, monocytes, and macrophages. In normal conditions, it is continuously produced in low quantities by the activity of PAF acetylhydrolases \(^{51}\). The involvement of PAF in SCD has been demonstrated in several studies. For example, Oh et al. found high level of endogenous PAF in the plasma of patients with SCD when compared with normal volunteers, and suggested that higher circulating level of PAF in patients with SCD during steady state may play a role in maintaining a state of inflammation and circulatory stress in microcirculation. In this study, it was also demonstrated that high levels of PAF were not caused by attenuation of PAF catabolism. Interestingly, they found that platelets are not the major source of circulating PAF and suggested that the higher plasma level of PAF results from the abnormal interaction of neutrophils and vascular endothelium with sickle RBCs \(^{52}\).

There is evidence that the synthesis of PAF by human polymorphonuclear cells (PMN) and endothelial cells could be modulated by nitric oxide (NO) levels. Mariano et al demonstrated that blockade of NO synthesis triggers a spontaneous production of PAF by endothelium and PMN cells, enhancing PAF\(^*\) synthesis induced by TNF-\(\alpha\) and LPS \(^{53}\). In general, synthesis
of NO promotes vasodilatation and inhibits platelet activation, aggregation and endothelial adhesion. However, NO inactivation by cell-free plasma Hb produced by hemolysis can block the NO effect on platelets. Therefore, the interaction between hemolysis, impaired NO bioavailability, and platelet activation might contribute to thrombosis and PHT in SCD. It has been suggested that SCD patients have decreased NO reserves, probably due to low plasma levels of L-arginine (the precursor to NO), particularly during VOC, and these levels are inversely correlated with pain symptoms. The authors propose that the three major mechanisms of impaired NO bioavailability are decreased plasma L-arginine, consumption of NO by cell-free plasma hemoglobin, and depletion by reactive oxygen species.

A literature review on patients with SCD reported that chronically elevated PMN counts in SCD are associated with higher risk of sickle cells complication and mortality, suggesting a constant pro-inflammatory state even in the absence of an acute crisis. Moreover, the review also reported a higher than normal presence of PMNs in the pulmonary microcirculation at the capillary–tissue interface in SCD patients. The importance of leukocyte-intermediated inflammation in SCD has also been reported. A study in isolated perfused lungs evinced that the release of PAF and leukotriene (LTB4) from the activated PMN increased retention/adherence of sickled RBC. The authors argue that both PAF and LTB4 can influence vascular permeability, cell infiltration, and PMN adhesion to vascular endothelium by causing the release of other inflammatory mediators.

Other mechanisms occurring in the membrane of sickled RBC could have the potential to exacerbate the pro-coagulation state in the disease pathology. As covered previously, studies assessing the roles of erythrocyte adhesion markers in patients with SCD demonstrate that erythrocyte PS exposure in the outer side of the cell membrane is a major determinant of RBC–endothelial adhesion. The authors suggest that this loss of phospholipid asymmetry could participate in SCD vaso-occlusion. Therefore, the phenomenon of PS exposure in sickled red cells could aggravate the sickle cell anemia due to enhanced activation of coagulation. This PS
exposure could also provide a negatively charged surface in RBCs facilitating the clotting cascade components contact. A study of blood from sickle cell patients demonstrated that PS positive RBC are positively correlated with plasma levels of F1+2, D-dimers, and plasmin-antiplasmin complex, suggesting that sickle RBCs are responsible to the hypercoagulable state in SCD 59.

An important pro-coagulant phenomenon in SCD is the osmotic RBC shrinkage due to changes on the Hb effects on osmotic pressure. The kinetic loss of osmotic pressure of sickle-cell Hb is caused by deoxygenation that dehydrate the erythrocyte and promote Hb aggregation 60. Osmotic erythrocyte shrinkage can activate cation channels with subsequent Ca2+ influx and stimulation of sphingomyelininase to increase production of ceramides. Ca2+ and ceramide in sickled RBC are involved in change of phosphatidylserine asymmetry of the cell membrane. These mechanisms are also linked to the observation that PAF is released from erythrocytes triggered by hyperosmotic cell shrinkage. Lang et al showed that PAF in erythrocytes binds to PAFr and stimulates the breakdown of sphingomyelin with consequent production of ceramide, and that PAF activates cell shrinkage and PS exposure in the external membrane 61.

An indirect consequence of high PAF production in SCD patients is the over-expression of platelet-activating factor receptor (PAFr), and its relationship with an increased chance of bacterial infection. PAFr is a G protein–coupled receptor of the lipid chemokine PAF. The role that PAFr plays in pneumococcal invasion has been established in vitro and in mice’ models of SCD. The proposed mechanism is that the phosphorylcholine on the pneumococcal cell wall mimics PAF and binds the bacteria to the endothelial cell membrane and also serves to insert the bacteria into the cell through PAFr uptake pathway 62.

9 Treatments

9.1 Hydroxyurea
The first discovery of Hydroxyurea was in 1869 by Dresler and Stein. In 1967, it was approved by the FDA as an antineoplastic agent to treat several types of tumors and leukemias. Decades later, it showed efficacy in the production of fetal Hb in SCD patients. For that, the FDA approved it as the first medication for SCD treatment in 1998. Hydroxyurea, a ribonucleotide reductase inhibitor, leads to limiting the transformation of ribonucleosides into deoxyribonucleosides (the DNA building unit) which causes less DNA production which in turn impacts erythropoiesis process. Consequently, stress hematopoiesis is prompted leading to recruitment of early erythroid progenitor cells that have ability to produce HbF; however this mechanism remains to be fully understood. In the term of disease outcomes, Hydroxyurea reduces the frequency of SCD complications such as VOC pain by about 50% as well as ACS and stroke. Besides, it inhibits WBC and reticulocytes' production that acts as inflammatory mediators and reduces the frequency of blood transfusion by 50%. In children with SCD, it shows clinical efficacy in minimizing the painful crisis, organ damage and stroke. Also, it enhances the growth and development of SCD children. The side effects of hydroxyurea are mild, and it mainly include nausea, headache, reticulocytopenia and neutropenia. Severe side effects are rare and may include cytopenia.

9.2 L-glutamine

SCD patients demonstrate impairment in nicotinamide adenine dinucleotide (NAD) metabolism with decrease of the reduced NAD (NADH) to NAD ratio leading to NAD redox potential depletion. In 2017, FDA approved L-glutamine for children and adults with SCA. It improves the NADH:NAD ratio resulting in enhancement NAD redox potential. Thus, it prevents erythrocytes sickling by the inhibition of oxidative stress, namely by its protective properties against free radicals. Also, it has been demonstrated to prevent other complications including VOC pain and reduce patient’s hospitalization. The therapy’s side effects include
gastrointestinal effects accompanied by headache. However, hepatic and renal functions need to be further studied during medication course.

### 9.3 Bone marrow transplant

Bone marrow transplant is currently the only cure for SCD. Although the cure rate is more than 80%, this procedure is limited among patients due to the lack of appropriate donors, and the severe side effects brought on by the procedure itself. Hematopoietic cell transplantation (HCT) outcomes are influenced by patients’ condition and SCD complications. In young patient with B-thalassemia and low risk disease the survival rate shown to be 93% and the disease-free survival rate is 91%. These percentages are worsened in high risk patients and could reach 79% for survival rate and 58% for disease-free survival in sever risk patients.

However, HCT outcomes are influenced by several factors. Mainly, graft-versus-host disease (GVHD) causes HCT failure which promotes organ damage and death. Besides, some cases reported a delay in sexual organ development after HCT but the evidence was insufficient.

### 9.4 Emergency management

#### 9.4.1 VOC management:

The new clinical evidence of the complication management cause elevation in survival rate of patients with acute complications. Multiple institutions published a guideline algorithms to facilitate management of acute pain due to VOC such as National Heart, Lung and Blood Institution (NHLBI). The complex pathophysiology of VOC causes a complexity of the management of the crisis. Basically, opioids are widely used in acute VOC pain management. Guidelines recommended intravenous (IV) or subcutaneous morphine (dose 0.1-0.15 mg/kg) within 15 to 20 minutes of patients arrival in emergency department. According to NHLBI
recommendations, administration of adjuvants medications beside opiates such as sedative and antihistamine. Also, NHLBI and other guidelines recommended further treatment for patients with inadequate pain relief. Oral pain control medications with comparable potency of IV medications must be prescribed to discharged patients. Guidelines also recommended the administration of oxygen to maintain oxygen saturation more than 92%.

9.4.2 ACS management:

Before treatment of ACS, patient assessment is required. X-ray should be obtained to assess the presence of new infiltrate. Usually the upper or middle lungs lobe affected in pediatric patients while the lower lobe is impacted in adult. The etiology of ACS includes infection (viral or bacterial) and embolism. The recommended management of ACS involves administration of broad-spectrum antibiotics and oxygen (O₂ saturation more than 95%). Also, blood transfusion may be required if the Hb level less than baseline level by 1.0 g/dl to improve O₂ capacity.

10 Discussion and future directions

SCD is characterized by severe complications that could be life-threatening such as VOC and hemolysis. A summary of the SCD complications is described in Table 3. The severity of complications ranges from patient to patient with regard to the disease phenotype and environmental factors. This review discussed the lipidomic changes of SCD for some complication as they are presently understood. However, there is a limited number of studies regarding lipidomics modulations in SCD, and the majority of them discuss only the HbSS type. In addition, there are limited studies that examine the lipid changes pre and post medications use. Morphine is widely used to manage VOC pain. However, due to its associated complications, new strategies have been used to reduce its side effects. Concomitant use of omega-3 fatty acids provides antinociceptive effects and increases morphine efficacy even in sub-therapeutic dosing.
Resolvins and endocannabinoids systems are also being explored as therapeutics to target and manage pain in SCD.

Table 4: Complications of SCD and a summary of the main characteristics.

<table>
<thead>
<tr>
<th>Complications</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute pain</td>
<td>• The most common reason for hospitalization.</td>
</tr>
<tr>
<td></td>
<td>• Affects teenagers more than adults.</td>
</tr>
<tr>
<td></td>
<td>• May cause early death due to pain severity.</td>
</tr>
<tr>
<td></td>
<td>• Occurs in HbSS more than HbSC</td>
</tr>
<tr>
<td></td>
<td>• Its frequency decreases with high Hb F concentration.</td>
</tr>
<tr>
<td></td>
<td>• Managed by using opiates.</td>
</tr>
<tr>
<td>Infection</td>
<td>• Occurs mainly in children.</td>
</tr>
<tr>
<td></td>
<td>• It is due to spleen impairment, micronutrients deficiency and complement activation deficiency.</td>
</tr>
<tr>
<td></td>
<td>• <em>S. pneumoniae, H. influenza</em>, and non-typhi <em>Salmonella</em> species are the most frequent infections.</td>
</tr>
<tr>
<td></td>
<td>• Prevented by using penicillin prophylaxis and immunization.</td>
</tr>
<tr>
<td>Neurological complications</td>
<td>• Occurs due to vasculopathy.</td>
</tr>
<tr>
<td></td>
<td>• The most common cause of stroke in pediatric is HbSS.</td>
</tr>
<tr>
<td></td>
<td>• Stroke occurs as a result of blood flow dysregulation.</td>
</tr>
<tr>
<td></td>
<td>• 90% of stroke risk could be reduced by blood transfusion.</td>
</tr>
<tr>
<td></td>
<td>• Neurocognitive problems may be seen in children.</td>
</tr>
<tr>
<td></td>
<td>• Intracranial bleeding is more frequent in adults.</td>
</tr>
<tr>
<td>Condition</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Acute chest syndrome       | • The second leading cause of hospitalization.  
                                • Cause acute lungs injury.  
                                • 13% of cases need mechanical ventilation.  
                                • Managed by using broad-spectrum antibiotics, bronchodilators, and oxygen. |
| Heart disease              | • Involve both diastolic and systolic dysfunctions.  
                                • Increase mortality rate.  
                                • 13% of SCD patients have left-side heart disease. |
| Renal complications        | • In SCD patients, partial oxygen and pH decrease with an increase in osmolality in the kidneys. That leads to an increase in HbS polymerization.  
                                • Renal infarction occurs due to vaso-occlusion.  
                                • Early age renal dysfunction and albuminuria may occur. |
| Hemolysis                  | • Most severe in HbSS phenotype.  
                                • Occurs because NO reduction.  
                                • Leads to other complications such as chronic anemia and PHT. |
| Hyper-coagulopathy         | • Elevation of coagulation markers in SCD patients even in steady state.  
                                • Leads to other complications such as venous thromboembolism, ACS, PHT and stroke. |
Chapter 2:

Lipid Biomarkers and Embryo Quality in In Vitro Fertilization;
Pregnancy Success Differentially Expressed by Body Weight
Lipid Biomarkers and Embryo Quality in In Vitro Fertilization;  

Pregnancy Success Differentially Expressed by Body Weight

1,2 Suad Alshammari; 1,3 Monther Alsultan; 1 Joshua Morriss; 1 Silas Contaifer; 1 Lama Basalelah; 1 Daniel Contaifer; Kai I. “Annie” Cheang; R. Scott Lucidi; Erin Mooney; Michelle Rhea; 1,4,5 Dayanjan S. Wijesinghe

1 Department of Pharmacotherapy and Outcome Sciences, School of Pharmacy, Virginia Commonwealth University, Richmond VA 23298

2 Faculty of Pharmacy, Northern Border University, Saudi Arabia.

3 College of Clinical Pharmacy, King Faisal University, Saudi Arabia.

4 Da Vinci Center, Virginia Commonwealth University, Richmond VA 23284

5 Institute for Structural Biology Drug Discovery and Development (ISB3D), School of Pharmacy, Virginia Commonwealth University, Richmond VA 23298

Please send correspondences to:

Dayanjan S. Wijesinghe, Ph.D.
Department of Pharmacotherapy and Outcome Sciences
School of Pharmacy, Virginia Commonwealth University,
PO Box 980533
Robert Blackwell Smith Bldg. Rm. 652A
Richmond, VA 23298-5048
USA
Abstract:

Introduction: A common risk factor for infertility is obesity. The global rise of obesity accompanied with infertility has led to widespread adoption of assisted reproductive technologies such as in vitro fertilization (IVF) to achieve pregnancy. However, pregnancy outcomes such as embryo quality vary after IVF, possibly due to disruptions in metabolism. Previous metabolomic studies investigating embryo quality were limited to characterizing broadly lipid classes, or a few molecular lipid species. Here, we sought to determine specific circulating lipids and metabolites with matrix-specific effects that could serve as putative biomarkers of embryo quality, correlated with BMI, and predicted clinical pregnancy subsequent IVF.

Methods: Electronic health record (EHR) data, as well as lipids and metabolites obtained from follicular fluid (FF) and platelet poor plasma (PPP), were collected from women (n = 26) undergoing IVF. Lipids and metabolites were acquired via untargeted mass spectrometry. For embryo quality and BMI, we performed multiple linear regression analysis to find correlates. For 6 weeks pregnancy, we applied a linear discriminant analysis to select lipids and metabolites that allowed for group determination.

Results: Several lipids and metabolites were selected from both matrices (FF and PPP) that either outperformed models containing only EHR or added value to EHR models. In predicting embryo quality, glycerophospholipids obtained from PPP produced the best fit model. The predicted values include (LPC) 22:6, phosphatidylcholine (PC) 16:1/22:6, and phosphatidylethanolamine-plasmalogen (PE-P) 16:0/22:6 were negatively
correlated with 2PN while Phosphatidylethanolamine (PE) 18:0/20:3, lysophosphatidylethanolamine (LPE) 18:1, PC 14:0/16:1, and PE-P 16:0/20:5 were positively correlated with 2PN (R adjusted = 0.730, RMSE = 0.329). For rLDA of 6-weeks of pregnancy, the best model was the metabolite model obtained from Platelet Poor Plasma (misclassification = 3.85%, Entropy R-squared = 0.809).

The BMI multicomponent domain model obtained from FF, LPC 18:1, PC 16:1/22:6, and malic acid were negatively associated with BMI while Fasting insulin and PC 16:0/22:4 were positively correlated with BMI values (R-square adjusted = 0.819, RMSE = 0.127). However, the combined data model for FF has the best prediction of BMI values. In this model, PE-P 16:0/22:6, aspartic acid, and fasting insulin as positively correlated variables with BMI values, whereas indole-3-propionic acid was negatively correlated with BMI (R-squared adjusted = 0.856, RMSE = 0.113).

**Conclusion:** High BMI is known to have detrimental effects on the development and success of pregnancy. It is also known to disrupt the endocrine pathways involved in the pregnancy process. This study provides information about metabolic signatures associated with BMI and pregnancy outcomes. Also, it shows the effect of 2PN value on embryo quality and the success of pregnancy.

**Introduction**

Infertility is a medical condition characterized by the inability of a couple to achieve clinical pregnancy after 12 months of regular, unprotected sexual intercourse. The prevalence of infertility for women of reproductive age is estimated to be 1 in every 7 couples, and 1 in every 4 couples, for those within Western countries and economically
developing countries respectively. Although the strongest negative predictor of fertility is increasing age for women, it is believed that other risk factors such as lifestyle and environmental factors contribute significantly towards the development of infertility.

A common risk factor for infertility is obesity, a disorder that is estimated to affect 42.1% of women, within the United States alone, who are 20 years of age or older. Obesity is defined as a body mass index (BMI) ≥30 kg/m², characterized by excessive adiposity. Not only does obesity increase time to conception, it increases the relative risk of chronic anovulation, decreases the chance of spontaneous conception, and is linked with greater odds of miscarriage. A global rise in the prevalence of obesity in tandem with infertility has led to an increasing number of women to utilize assisted reproductive technologies such as in vitro fertilization (IVF) to achieve pregnancy. However, women who are obese exhibit low responses to controlled ovarian stimulation via IVF. Cellular and subcellular disruptions co-occurring with obesity may account for these poor reproductive outcomes subsequent to IVF.

The impact of obesity on reproductive function can be understood through aberrant endocrine signaling pathways that dysregulate the hypothalamic-pituitary-ovarian (HPO) axis, the system that drives reproduction by cyclic production of gonadotropic and steroid hormones. Notably, obesity exerts an inhibitory effect on gonadotropic production, and women who are obese require higher doses of administered gonadotropins such as follicle stimulating hormone (FSH) during IVF in order to achieve an ovulatory response, comparable to women of normal weights. Women who are obese additionally have markedly lower pulsatile luteinizing hormone (LH) amplitudes and progesterone
metabolite secretion, further implicating the HPO axis as a central mechanism for poor IVF outcomes. Bridging obesity with the dysregulation of the HPO axis in women who are obese is insulin resistance (IR), which is exacerbated by lipotoxicity.

Lipotoxicity in ovarian tissues can exacerbate IR, is a potential mechanism for oocyte organelle damage, and has the potential to induce hyperinsulinemia, which can promote obesity. Under euglycemic, hyperinsulinemia conditions, elevated non-esterified fatty acid (FFA) content concomitant high insulin levels can selectively impair LH and FSH, suggesting that the endocrine disruption in women who are obese may be mediated by direct lipotoxic effects on the HPO axis. During IVF, concentrations of FFA, triglyceride (TAG), and long-chain acylcarnitine species within follicular fluid (FF) of women who are obese are positively correlated with gonadotropin-stimulated androgen levels. Additionally, in the FF of women with normal BMI who failed to respond to IVF, there is a significant shift in the accumulation of monoacylglycerols (MAG) and glycerophospholipids like phosphatidylethanolamine (PE) towards triacylglycerols (TAG) and cholesteryl esters. Lastly, circulating levels of TAG as a total lipid class is negatively correlated with embryo quality, as determined by the presence of two clearly distinct pronuclei and two polar bodies (2PN). Taken together, a diversity of lipids and metabolites dysregulated in an obese state promote a number of deleterious pathophysiological processes, and exert modulatory effects on fertility.

Yet which lipid species and metabolites are clinically relevant towards the predicting the outcomes of IVF in women who are obese remains unknown. Current clinical evidence for IVF success or failure have been methodologically limited to only a few lipid species studied, or do not associate specific lipid species with embryo quality or
fertility, only assessing lipids by their classes (i.e. total TAG). Therefore, we sought to determine specific circulating lipids and metabolites, within the plasma and FF, that could serve as putative biomarkers of 2PN, and which play a role in reduced IVF pregnancy rates associated with obesity.

**Materials and Methods**

1) *Patients*

Follicular fluid (FF), platelet poor plasma (PPP), and electronic health record (EHR) data were collected from women (n = 26) undergoing IVF. Body composition (percentage of body fat mass) was obtained by bioelectrical impedance\(^{98}\). Women’s weights were recorded as either normal, overweight, or obese. The study design has been previously reported (VCU IRB#: HM20004988), but in brief; starting on menstrual cycle day 3, patients self-administered gonadotropins and underwent transvaginal ultrasounds every 2 days. Blood was drawn prior to ovarian stimulation via subcutaneous injection with human chorionic gonadotropin (hCG, 250 µg), and processed into plasma, once adequate mature follicles were detected (approximately cycle day 10-12). A follow-up visit coincided 34-36 hours post hCG, where blood was drawn and FF was obtained during egg retrieval. Patients were asked to fast after midnight of the night before the baseline visit, and were provided an optional 2-hour oral glucose tolerance test to determine if they were diabetic.
(GDM threshold was ≥140 Mg/dl). Additionally, patients were asked to recall what they had eaten within the previous 24 hours prior to their follow-up visit. Whether or not pregnancy was achieved 6 weeks after IVF was recorded. Two weeks after oocyte retrieval, serum hCG was used to evaluate pregnancy and, if serum hCG was positive, an additional visit 6-weeks post hCG occurred to confirm absence or presence of pregnancy.

2) Sample Preparation and Storage

During egg retrieval, discarded FF was obtained and transferred to Heathrow 2.0 mL Screw-Top Tubes with O-Ring caps. Samples of FF were then stored at -80°C until needed for lipidomic and metabolomic analysis.

The following protocol was used to prepare PPP: whole blood (7.4 mL) was collected in K2EDTA tubes upon the patient’s visit. Whole blood was then centrifuged at 1300 g for 10 minutes using an Eppendorf centrifuge 5810 R to separate the platelet rich plasma (PRP) layer. The upper PRP layer was then collected, avoiding contamination of the buffy coat. To obtain PPP, PRP was centrifuged at 2000 g for 15 minutes. Samples of PPP were then aliquoted at 0.2mL into Eppendorf tubes and stored at -80°C until needed for lipidomic and metabolomic analysis.

3) Lipidomic and Metabolomic Extraction

3a) Extraction of lipids and metabolites for liquid chromatography (LC) lipidomic and metabolomic analysis

Either 20 µL aliquot of FF or PPP had their lipids, along with water-soluble metabolites, extracted by a biphasic solvent system. In detail, sample was introduced
to 225 µL of cold methanol in a 1.5 mL polypropylene tube, mixed with odd chain and
deuterated lipid internal standards [lysophosphatidylethanolamine (lysoPE 17:1),
lysophosphatidylcholine (lysoPC 17:0), phosphatidylethanolamine (PE 17:0/17:0),
phosphatidylcholine (PC 12:0/13:0), phosphatidylglycerol (PG 17:0/17:0), sphingosine
(d17:1), d7 cholesterol, sphingomyelin (SM 17:0), C17-ceramide (d18:1/17:0), d3-palmitic
acid (16:0/16:0/16:0), 1-margaroyl-glycerol (MAG 17:0/0/0/0), 1-oleoyl-2-acetyl-sn-
glycerol (DAG 18:1/2:0/0:0), 1,2-diacyl-sn-glycerol (DAG 12:0/12:0/0:0), and 1,3(d5)-
diheptadecanoyl-2-(10Z-heptadecenoyl)-glycerol (d5-TAG 17:0/17:1/17:0)]. Thereafter
750 µL of cold methyl tert-butyl ether (MTBE) was added and vortexed, and phase
separation induced by 188 µL of liquid chromatography mass spectrometry (LC-MS)-
grade water. Extracts were centrifuged at 12,300 rpm for 2 minutes, with the upper
organic phase (300µL) collected for analysis of lipids, and evaporated under vacuum
conditions. Resuspension occurred with 110µL of methanol/toluene (9:1, v/v) containing
12-[(cyclohexylamino) carbonyl]amino]-dodecanoic acid (CUDA, 50 ng/ml; internal
standard for quality control of injection), vortexed, and centrifuged at 800 rpm for 5
minutes. The resultant supernatant (100µL) underwent lipidomic analysis. An aliquot
(125µL) of the lower phase was evaporated under vacuum conditions and resuspended
in pure acetonitrile (ACN). These extracted metabolites were then subjected to
metabolomic acquisition via hydrophilic interaction (HILIC) LC-MS/MS101.

3b) Extraction of metabolites for gas chromatography (GC) metabolomic analysis

Small metabolites (<650 Daltons) were acquired from 30µl aliquots of FF or PPP
added to 1.0 mL of cold extraction solution consisting of acetonitrile, isopropanol (IPA)
and water (3:3:2, v/v/v)102. Samples were vortexed for 5 minutes at 4°C and centrifuged
for 2 minutes at 14,000 relative centrifugal force (rcf). A total of 450μL of the supernatant was dried under a cold trap concentrator and then reconstituted with 450μL ACN:water (50:50 v/v) solution, centrifuged for 2 minutes at 14,000 rcf, and then subjected to drying under a cold trap. Derivatization occurred with the addition of 10μL of 40 mg/mL methoxyamine hydrochloride solution to the dried samples and quality controls, including N-methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA) and fatty acid methyl esters (FAME). Samples were shaken at 30°C for 1.5 hours, then mixed with 91μL of MSTFA and FAME. After shaking at maximum speed at 37°C, for 30 minutes, the contents were transferred for metabolomic acquisition via gas chromatography mass spectrometry (GCMS)²⁷.

4) Lipidomic and Metabolomic Data Acquisition

4a) LC-MS Lipidomic data acquisition

Reverse phase, untargeted analysis was performed with a Sciex TripleTOF 6600 coupled to an Agilent 1290 LC. Lipids were separated on an Acquity ultra-performance liquid chromatography (UPLC) CSH C18 column (100 × 2.1 mm; 1.7μm) (Waters, Milford, MA, USA) as previously reported by our group²⁸. The column was maintained at a temperature of 65 °C with a constant flow-rate of 0.6mL/min. The mobile phases were Mobile Phase A at 60:40 (v/v) ACN:water with 10 mM ammonium acetate, and Mobile Phase B at 90:10 (v/v) IPA:ACN with 10 mM ammonium acetate. Lipid separation was conducted following a stepwise gradient program as follows: 0-2 minutes (Mobile Phase B: 15-30%), 2-2.5 minutes (Mobile Phase B: 48%), 2.5-11 minutes (Mobile Phase B: 82%), 11-11.5 minutes (Mobile Phase B: 99%), 11.5-12 minutes (Mobile Phase B: 99%), 12-12.1 minutes (Mobile Phase B: 15%), 12-14 minutes (Mobile Phase B: 15%). Negative
and positive electrospray ionization (ESI) modes were applied with nitrogen serving as the desolvation gas and the collision gas. The MS instrument parameters were as follows: curtain gas: 35; collision-induced dissociation (CAD): high; ion spray voltage: 4500V; source temperature: 350°C; gas 1: 60; gas 2: 60; declustering potential: ±80V, and a collision energy of ±10^3.

4b) LC-MS Metabolomic data acquisition

Detection of water-soluble PPP and FF metabolites underwent separation via normal phase LC on a Sciex TripleTOF 6600 coupled to Agilent 1290 LC. Metabolites were separated by an Acquity UPLC BEH Amide (130Å, 1.7 μm, 2.1 mm X 150 mm) column (Waters, Milford, MA, USA). Column temperature was maintained at 45°C at a constant flow-rate of 0.4 mL/min. Two mobile phases were utilized, consisting of Mobile Phase A, which was LC-MS grade water with 10 mM of ammonium formate and 0.125% formic acid, and Mobile Phase B at 90:10 (v/v) ACN:water with 10 mM of ammonium formate and 0.125% formic acid. Analytes were separated using a stepwise gradient program as follows: 0–2 minutes (Mobile Phase B: 100%), 2–7.7 minutes (Mobile Phase B: 70%), 7.7–9.5 minutes (Mobile Phase B: 40%), 9.5–10.25 minutes (Mobile Phase B: 30%), 10.25–12.75 minutes (Mobile Phase B: 100%), 12.75–16.75 minutes (Mobile Phase B: 100%). Injection volumes were 1μL and 3μL for the positive and negative modes, respectively. Positive and negative ESI modes utilized nitrogen serving as the desolvation gas and the collision gas, and the MS instrument parameters were: curtain gas: 35; CAD: high; ion spray voltage: 4500V; source temperature: 300°C; gas 1: 60; gas 2: 60; declustering potential: ± 80V, and collision energies ± 10^3.

4c) GC-MS Metabolomic data acquisition
Metabolites were acquired on a Leco Pegasus IV time of flight mass spectrometer, coupled with an Agilent 6890 GC that was equipped with a Gerstel automatic liner exchange system (ALEX) that had a multipurpose sample (MPS2) dual rail, and a Gerstel CIS cold injection system (Gerstel, Muehlheim, Germany). The transfer line was maintained at 280°C, and separation was achieved on a 30 m long, 0.25 mm i.d. Rtx-5Sil MS column (0.25 μm; 95% dimethyl 5% diphenyl polysiloxane film) with a built-in, integrated guard column (10m; Restek, Bellefonte PA) with a constant flow (1 mL/min) of helium (99.999%; Airgas, Radnor, PA, U.S.A.) as the carrier gas. The oven temperature was held constant at 50°C for 1 minute, and then ramped at 20°C/min to 330°C, which was held constant for 5 minutes. The GC temperature program was set as follows: 50°C to 275°C final temperature at a rate of 12°C/s and held for 3 minutes. The injection volume was 1 μL in splitless mode at 250°C. Electron impact ionization at 70 eV was employed with an ion source temperature of 250°C. The scan mass ranged from 85- to 500 Da with an acquisition rate of 17 spectra/second.

5) Statistical Analysis

Baseline characteristics and demographic parameters were assessed for any significant (p<0.05) differences that could confound our findings between women with obesity and healthy-weight women undergoing IVF. Lipidomic and metabolomic data were acquired as ratios of mass-to-charge, and were normalized by mTIC to enable the merger of these two datasets, obtained from different instruments and techniques. Metabolic data were filtered to exclude any detected exogenous metabolites identified as products of gut microbiota so that we obtained only endogenous metabolites, relevant to the physiological response of the patient undergoing IVF. Winsorization of outliers
occurred for any value outside of 1.5 times the interquartile range\textsuperscript{104}, and missing values for EHR data were imputed by the median of the variable. Any missing values for the outcome measures, 6-week pregnancy (a binary of “yes” or “no”) and 2PN (nominal) were excluded. Then the data were log-transformed and pareto-scaled to correct for technical variation and other factors not attributable to induced biological variation\textsuperscript{105}. Phospholipids were identified by their head group, parent-fragment pair, and by their sum fatty acid composition.

An overview of the statistical analytical workflow is presented within Figure 1, and this analysis was performed separately for FF and PPP analytes. Four domain models were used to ascertain the most important predictors belonging to lipids, metabolites, and EHR data for our chosen outcomes. These models consisted of either strictly metabolites, EHR data, or lipids, with two models containing either with glycerophospholipids (and lysophospholipids) or excluding glycerophospholipids. We performed stepwise multiple linear regression analysis (MLRA) with forward selection of candidate predictors among EHR, lipids, and metabolites that best fit BMI and 2PN. For our binary outcome of 6-weeks pregnancy, we conducted a stepwise, regularized linear discriminant analysis (rLDA) to discriminate between pregnant and non-pregnant women. Lastly, predictors from lipids, metabolite, and EHR models were combined into a multi-component model for the final discriminant analysis model. To control for confounders, univariate models assessed whether selected predictors were also associated with confounders, and if so, the final multivariable model was fit to identify discriminatory lipids and metabolites after adjusting for potential confounders. The statistical analysis described herein was performed with JMP14Pro, MetaboAnalyst 4.0, and with R-scripts written by us.
Figure 1: Statistical analysis workflow.

**Results**

Current clinical evidence for IVF success or failure have been methodologically limited, and this is problematic as lipids and cholesterol species in particular display high levels of compositional and functional heterogeneity. For example, cholesteryl esters, the single most abundant lipid class in human plasma and preferred form of cholesterol in lipoprotein particles, consist of several hundreds of molecular species. We sought a qualitative approach to best characterize specific lipid species and metabolites that could sufficiently determine embryo quality and predict pregnancy outcomes in women with obesity following IVF.

*Characteristics of the study cohort and adjustments by Fasting Insulin*
At the enrollment of this study, patients in either weight group had similar distributions of age, maximal serum estradiol, highest dosage of FSH used, total cumulative FSH used, and dietary scores (Table 1). T-tests at baseline determined no statistically significant (p>0.05) difference between demographic parameters for the treatment groups, except with BMI (p<0.0001), fasting insulin concentrations (p = 0.001), and the homeostatic model assessment for insulin resistance (HOMA-IR, p = 0.003). Because of this difference in insulin resistance, features selected for the final models of 2PN, BMI, and 6-weeks pregnancy outcomes were adjusted by fasting insulin concentrations. Interestingly, when considering only weight class, both obese and healthy-weight groups had similar 6-weeks pregnancy (p = 0.229) and 2PN (p = 0.208) outcomes following IVF (Table 1).

**Table 1.** Demographic and Baseline characteristics shows the similarities between normal weight women (not obese) and obese women

<table>
<thead>
<tr>
<th></th>
<th>Normal (n=10)</th>
<th>Obese (n=16)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female age, mean (SD)</strong></td>
<td>33.30(5.1)</td>
<td>33.99(3.2)</td>
<td>0.712</td>
</tr>
<tr>
<td><strong>BMI</strong></td>
<td>21.71(1.7)</td>
<td>34.42(9.7)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Fasting glucose</strong></td>
<td>84.7(4.42)</td>
<td>87.12(7.11)</td>
<td>0.29</td>
</tr>
<tr>
<td><strong>Fasting insulin</strong></td>
<td>4.63(2.06)</td>
<td>11.01(1.53)</td>
<td>0.001</td>
</tr>
<tr>
<td><strong>HOMA-IR</strong></td>
<td>0.97(0.466)</td>
<td>2.31(1.47)</td>
<td>0.003</td>
</tr>
<tr>
<td><strong>Mediterranean diet score</strong></td>
<td>25.1(7.35)</td>
<td>22.06(1.67)</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>Total score black food fat</strong></td>
<td>14.4(7.39)</td>
<td>17(5.93)</td>
<td>0.361</td>
</tr>
<tr>
<td><strong>screener</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total score block</strong></td>
<td>13(3.43)</td>
<td>11.06(4.73)</td>
<td>0.240</td>
</tr>
<tr>
<td><strong>questionnaire fruit</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>vegetable</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Average setting</strong></td>
<td>335.57(131.24)</td>
<td>311.51(169.62)</td>
<td>0.689</td>
</tr>
<tr>
<td><strong>Cycle plan, n(%)</strong></td>
<td>10(38.46)</td>
<td>15(57.69%)</td>
<td>0.318</td>
</tr>
<tr>
<td><strong>Antagonist downreg</strong></td>
<td>0(0%)</td>
<td>1(3.85%)</td>
<td></td>
</tr>
<tr>
<td><strong>Highest dosage of FSH</strong></td>
<td>397.5(50.62)</td>
<td>389.06(73.58)</td>
<td>0.732</td>
</tr>
<tr>
<td><strong>used</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total cumulative FSH</strong></td>
<td>4327.5(1392.86)</td>
<td>4397.65(1375.86)</td>
<td>0.901</td>
</tr>
<tr>
<td><strong>used</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maximal serum estradiol</strong></td>
<td>1312.78(739.94)</td>
<td>1511.425(952.54)</td>
<td>0.557</td>
</tr>
<tr>
<td><strong>6 weeks pregnancy</strong></td>
<td>0.5(0.527)</td>
<td>0.25(0.447)</td>
<td>0.229</td>
</tr>
<tr>
<td><strong>2PN</strong></td>
<td>17.1(10.80)</td>
<td>11.68(2.34)</td>
<td>0.208</td>
</tr>
</tbody>
</table>

**Note**, 2PN=2 polar bodies, SD= Standard Deviation, BMI= Body Mass Index, HOMA ISI= Homeostatic model assessment Insulin Sensitivity Index. Data presented as mean (SD).

**Domain Independent and multicomponent MLRA models predict 2PN values with lipids and organic compounds, differing across PPP and FF matrices**

The EHR model for predicting 2PN retrieved female age in years, fasting glucose, fasting insulin, Mediterranean diet score, the total score from the block food questionnaire fat screener (23 high), the total score from the block food questionnaire fruit and vegetable screener (11 low), average sitting time as total minutes per day (test statistic = 0.665,
Prob > |t| = 0.0221), and total cumulative FSH used (test statistic = -1.44, Prob > |t| = 0.0015) as covariates, while only the average sitting time (total of minutes per day) was negatively associated with an increase in 2PN values (R adjusted = 0.769, RMSE = 0.144). From the EHR data used to construct this model, only average sitting time and total cumulative FSH used were the clinical parameters that were significant after two-way hypothesis testing. Yet in constructing the MLRA model, predictors are covaried with another, thus those that were non-significant following a two-way hypothesis test were kept within this clinical parameter model and all remaining MLRA models.

- *Metabolites acquired from PPP allowed for the best fit model across other domain models for 2PN values*

Without exclusion of PPP lipids belonging to the glycerophospholipid (PL) family, we observed the following to best predict increasing 2PN values: PL that decreased in circulating with increasing 2PN numbers were lysophosphatidylcholine (LPC) 22:6 (test statistic = -1.01, Prob > |t| = 0.0231), phosphatidylcholine (PC) 16:1/22:6, and phosphatidylethanolamine-plasmalogen (PE-P) 16:0/22:6 (test statistic = -0.899, Prob > |t| = 0.0004). Phosphatidylethanolamine (PE) 18:0/20:3, lysophosphatidylethanolamine (LPE) 18:1 (test statistic = 0.928, Prob > |t| = 0.0005), PC 14:0/16:1 (test statistic = 0.4698, Prob > |t| = 0.0060), and PE-P 16:0/20:5 (test statistic = 0.483, Prob > |t| = 0.0013) were those PL that were either positively associated or correlated with 2PN (R adjusted = 0.730, RMSE = 0.329).

Exclusion of PL and inclusion of other lipid species from our PPP model derived a model with poorer fit for 2PN (R adjusted = 0.666, RMSE = 0.365). Nonetheless, we found
the following nonesterified fatty acids (FFA) and cholesteryl esters (CE) to be associated or correlated with 2PN: those that increased with 2PN values included CE 14:0, CE 20:5 (test statistic = 0.931, Prob > |t| = 0.0236), FFA 19:1 (test statistic = 1.01, Prob > |t| = 0.0288), and myristic acid. Inversely, those that decreased corresponding a rise in 2PN values were CE 16:0, CE 22:6 (test statistic = -1.555, Prob > |t| = 0.0007), FFA 14:1, FFA 18:2, FFA 20:5, and lauric acid (test statistic = -0.771, Prob > |t| = 0.0137).

Comparatively, the metabolic model performed better against both lipid domain models and the final model (R adjusted = 0.749, RMSE = 0.317), as depicted in Figure 2, and included the following small molecules: 2-ketoisovaleric acid (test statistic = -0.971, Prob > |t| = 0.0001), 3-aminoisobutyric acid (test statistic = -0.611, Prob > |t| = 0.0004), beta alanine (test statistic = -0.638, Prob > |t| = 0.0007), glutamic acid (test statistic = -0.887, Prob > |t| = 0.0002), and serine (test statistic = -0.897, Prob > |t| = 0.0033) were negatively correlated with 2PN. However, those metabolites that were positively correlated with 2PN values involved aspartic acid (test statistic = 0.500, Prob > |t| = 0.0236), ornithine (test statistic = 0.683, Prob > |t| = 0.0225), and phosphate (test statistic = 0.852, Prob > |t| = 0.0030).
When combining all data domains, the final model for PPP to predict 2PN values was constructed: positively correlated analytes included LPE 20:4 (test statistic = 0.765, Prob > |t| = 0.0023), PC 14:0/16:1 (test statistic = 0.255, Prob > |t| = 0.0329), and phosphate (test statistic = 1.03, Prob > |t| = 0.0003). Negatively correlated predictors were LPC 22:6 (test statistic = -0.652, Prob > |t| = 0.0057), CE 16:0 (test statistic = -0.595, Prob > |t| = 0.0056), and FA 20:5 (test statistic = -0.289, Prob > |t| = 0.0212). Although better fit than the lipid models (R adjusted = 0.727, RMSE = 0.331), the combined model performed worse than the PPP metabolites model for 2PN.

- **Including Glycerophospholipids produced the best fit FF model for predicting 2PN values against other Lipid, Metabolic, and Combined domain models**
Inclusion of PL from FF generated a model that included two acylcarnitine species as well various PLs. Acylcarnitine C16:0 (test statistic = 0.417, Prob > |t| = 0.0033), containing palmitic acid, was positively correlated with 2PN values, yet acylcarnitine C18:1 (test statistic = -0.542, Prob > |t| = 0.0358), or an L-carnitine esterified with oleic acid, was negatively correlated. Retained PL from FF that decreased with 2PN values were LPC 16:1 (test statistic = -0.853, Prob > |t| = 0.0047), lysophosphatidylinositol (LPI) 18:2 (test statistic = -1.00, Prob > |t| = 0.0023), and PC-Plasmalogen (PC-P) 18:0/22:6 (test statistic = -0.799, Prob > |t| = 0.0013). Those PL that were positively correlated with 2PN numbers were LPC 22:4 (test statistic = 0.603, Prob > |t| = 0.0028), LPE 20:4 (test statistic = 0.927, Prob > |t| = 0.0042), and LPI 18:1 (test statistic = 1.16, Prob > |t| = 0.0048). Of the FF models, this model performed the best for predicting 2PN values (R adjusted = 0.755, RMSE = 0.312), the results of which can be viewed in Figure 3.

![Figure 3: PL model for 2PN acquired from FF sample.](image-url)
From our FF lipid model (R adjusted = 0.726, RMSE = 0.331) excluding PL, lipids that positively associated or correlated with 2PN involved CE 16:1, CE 18:1 (test statistic = 1.26, Prob > |t| = 0.0009), and arachidic acid. Among the negatively associated or correlated lipids, when excluding PL, included CE 18:3, CE 22:6 (test statistic = -0.914, Prob > |t| = 0.0003), FA 14:0 (test statistic = -0.439, Prob > |t| = 0.0410), FA 16:0 (test statistic = -0.585, Prob > |t| = 0.0410), FA 22:3 (test statistic = -0.894, Prob > |t| = 0.0003).

Metabolites acquired from FF were all negatively correlated with 2PN values, with the exception of 1-monoolein, which bore a positive association with increasing 2PN. Nonetheless, the metabolites from our model included 2-aminobutyric acid (test statistic = -0.646, Prob > |t| = 0.0021), alanine (test statistic = -1.09, Prob > |t| = <0.0003), hexose (test statistic = -0.591, Prob > |t| = <0.0010), phenylalanine (test statistic = -0.908, Prob > |t| = 0.0108), and phenylethylamine (test statistic = -0.975, Prob > |t| = 0.0002). Despite the metabolites model performing better than the FF lipid model excluding PL (R adjusted = 0.7299, RMSE = 0.329), metabolites alone failed to outperform the FF lipid model including PL.

A final model with all 4 data domains gave a model without any EHR parameters (R-squared adjusted = 0.748, RMSE = 0.317). Lipids and metabolites selected were negatively correlated or associated with 2PN values except arachidonic acid containing LPE 20:4 (test statistic = 0.854, Prob > |t| = 0.0040). Remaining molecules were PC-P 18:0/22:6 (test statistic = -0.622, RMSE = 0.0115), 2-aminobutyric acid, alanine (test statistic = -0.771, RMSE = 0.0062), phenylethylamine, and FA 22:3 (test statistic = -0.429, RMSE = 0.0438).
**Regularized Linear Discriminant analysis (rLDA) of 6-weeks of pregnancy**

Pregnancy was discretized as a binary, categorical outcome, or as either “yes” (1) or “no” (0), 6-weeks after IVF. Ergo, the objective of these stepwise, regularized linear discriminant analytical (rLDA) models was to classify patients along this binary. EHR data alone surprisingly gave the highest rate of misclassification (19.23%) and the poorest entropy R-squared, a measure of classification quality or goodness of separation, across all of the component models (Entropy R-squared = 0.356). EHR parameters included in this model were female age, fasting glucose, Mediterranean diet score, total score block food fat screener (23 high), total score block food fruit and vegetable screener (11 low), average sitting time (total of minutes per day), total cumulative FSH used, and maximal serum estradiol value. A 95% confidence level ellipse was plotted for each canonical variables mean, pregnant or non-pregnant, as well as an ellipse denoting a 50% contour for each group, depicting a region that contains approximately 50% of the observations. The set of rays in the biplot represents the covariates, whose coefficients in the linear combination were canonical weights. Thus, the larger a covariate’s canonical weight, the greater its association with the canonical variable. The length and direction of each ray in the biplot indicates the degree of association of the corresponding covariate with the canonical variable, 6-weeks pregnant or failed pregnancy (Figure 4).
High goodness of separation without misclassification for 6-weeks pregnancy outcomes was achieved with the Metabolite model obtained from Platelet Poor Plasma.

In classifying patients into their pregnancy status following IVF, lipids obtained from PPP that did not include the glycerophospholipid (GL) class resulted in one false positive (misclassification = 3.85%, Entropy R-squared = 0.809). Five lipid species were associated with pregnancy failure, those involving CE 17:1 (standardized scoring coefficient = -1.64), CE 22:5 (standardized scoring coefficient = -3.67), FA 20:4 (standardized scoring coefficient = -3.08), capric acid (standardized scoring coefficient = -1.47), and lauric acid (standardized scoring coefficient = -1.72). Successful pregnancies were associated with higher concentrations of CE 18:3 (standardized scoring coefficient = 3.69), CE 20:4 (standardized scoring coefficient = 1.99), CE 20:5 (standardized scoring coefficient = 1.49), FA 20:2 (standardized scoring coefficient = 2.51), arachidic acid (standardized scoring coefficient = 2.26), and myristic acid (standardized scoring coefficient = 2.17).
No change was observed in the percent misclassification (3.85%), nor the number of false positives, when GL species were included as covariates of the lipid model (Entropy R-squared = 0.813). Of note was that circulating total cholesterol (standardized scoring coefficient = 0.350) was positively associated with 6-weeks pregnancy following IVF. Other positively associated covariates included LPE 20:4 (standardized scoring coefficient = 2.60), PC 16:1/18:2 (standardized scoring coefficient = 0.880), PC 17:0/20:3 (standardized scoring coefficient = 1.28), and PI 16:0/18:1 (standardized scoring coefficient = 0.964).

With regards to the metabolite model, none of the patients were misclassified into their pregnancy groups, and relatively high goodness of separation was attained (Entropy R-squared = 0.939). Metabolites, with the exception of 2-ketoisovaleric acid (standardized scoring coefficient = -1.02), glucose-1-phosphate (standardized scoring coefficient = -0.945), and ornithine (standardized scoring coefficient = -1.19), were positively associated with pregnancy and can be viewed in Figure 5. These metabolites were 2-hydroxyglutaric acid (standardized scoring coefficient = 1.71), serine (standardized scoring coefficient = 1.15), threonine (standardized scoring coefficient = 1.42), and uric acid (standardized scoring coefficient = 1.47).
A final model was constructed without clinical parameters, containing only lipids and metabolites. Total cholesterol (standardized scoring coefficient = 1.30) was positively associated with pregnancy outcomes, as well as LPE 20:4 (standardized scoring coefficient = 1.21), uric acid (standardized scoring coefficient = 1.08), and arachidic acid (standardized scoring coefficient = 0.753). Negatively associated lipids and metabolites, or those that were higher with negative pregnancy outcomes, were LPC 20:4 (standardized scoring coefficient = -1.77), 2-ketoisovaleric acid (standardized scoring coefficient = -0.818), and lauric acid (standardized scoring coefficient = -0.452).

- **Lipids excluding Glycerophospholipids obtained from Follicular Fluid resulted in No Misclassification and the Best Separation for 6-weeks pregnancy groups**

Classification of the 6-week pregnancy groups was achieved with 0% misclassification and high goodness of separation (Entropy R-squared = 0.977) when using a model of lipids, excluding from selection all GL species. Covariates that were weighted heavily...
towards pregnancy failure included CE 20:5 (standardized scoring coefficient = -0.959), capric acid (standardized scoring coefficient = -0.870), and stearic acid (standardized scoring coefficient = -1.08). Pregnancy success was associated with higher levels of CE 18:3 (standardized scoring coefficient = 1.31), FA 18:3 (standardized scoring coefficient = 0.867), heptadecanoic acid (standardized scoring coefficient = 1.71), and palmitoleic acid (standardized scoring coefficient = 0.874). Comparatively, this lipid model performed the best against all other FF models for discrimination of the pregnancy classes (**Figure 6**).

**Figure 6**: rLDA of 6-weeks of pregnancy and lipids without GL species acquired from FF sample.

Coupling FF obtained lipid species with GL for a model of 6-week pregnancy was achieved with modest misclassification (7.69%) and admissible goodness of separation (Entropy R-squared = 0.750). Lysophospholipids and phospholipids predominated as the covariates of this model where IVF proved efficacious were more closely associated with LPC 14:0 (standardized scoring coefficient = 1.34), LPI 18:1 (standardized scoring coefficient = 0.830), LPI 20:4 (standardized scoring coefficient = 1.47), and PC 18:1/20:3.
(standardized scoring coefficient = 1.27). Patients who were classified as having not achieved pregnancy saw higher concentrations of LPC 20:2 (standardized scoring coefficient = -2.07), LPE 18:2 (standardized scoring coefficient = -2.07), and PC 16:0/20:4 (standardized scoring coefficient = -1.06).

Analogous to the 6-week pregnancy lipid model excluding glycerophospholipids, the metabolite model for FF had no misclassification (0%) and high goodness of separation (Entropy R-squared = 0.943). Only alpha-ketoglutarate (standardized scoring coefficient = -1.26) was associated with failed pregnancy. The remaining covariates of fumaric acid (standardized scoring coefficient = 0.793), hexose (standardized scoring coefficient = 0.825), indole-3-propionic acid (standardized scoring coefficient = 0.566), and uric acid (standardized scoring coefficient = 1.35) were associated positively associated with pregnancy.

The final model of FF analytes and clinical parameters for 6-weeks pregnancy status resulted in one false negative (misclassification = 3.85%, Entropy R-squared = 0.889). Despite including all variables across the data domains, only one lipid and metabolites were selected, of which were FA 18:3 (standardized scoring coefficient = 0.616), fumaric acid (standardized scoring coefficient = 1.08), hexose (standardized scoring coefficient = 0.730), and uric acid (standardized scoring coefficient = 1.06) that were positively associated with pregnancy. A singular metabolite, alpha-ketoglutarate (standardized scoring coefficient = -1.29), was elevated in FF for those that did not achieve pregnancy.
Domain Independent and domain-integrated MLRA models predict BMI values with phospholipids and fatty acid derivatives, varied across PPP and FF matrices

Our EHR model for predicting BMI values retrieved fasting glucose, fasting insulin, and the total score from the block food questionnaire fat screener (23 high) as positively associated covariates, while only the average sitting time (total of minutes per day) was negatively associated (R adjusted = 0.769, RMSE = 0.144). Fasting insulin (test statistic = 0.473, Prob > |t| = <0.0001) and total score block food fat screener (23 high; test statistic = 0.167, Prob > |t| = 0.0054) were clinical parameters that were significant after two-way hypothesis testing. As mentioned prior, in constructing the MLRA model, predictors are covaried with another, thus those that were non-significant following the two-way hypothesis test were kept within this clinical parameter model.

- A multicomponent domain model for PPP outperformed data-domain independent Lipid, Metabolic, and EHR domain models for BMI

With regards to our PPP PL model, LPC 18:1 (test statistic = -0.261, Prob > |t| = 0.0133), LPC-P 16:0 (test statistic = -0.233, Prob > |t| = 0.0021), PC 16:1/22:6 (test statistic = -0.243, Prob > |t| = 0.0002), and PI 18:0/20:4 (test statistic = -0.268, Prob > |t| = 0.0422) were negatively correlated with an increase in BMI values, while, interestingly, recovered LPC 22:6 (test statistic = 0.305, Prob > |t| = 0.0127) and PC 16:0/22:4 (test statistic = 0.185, Prob > |t| = 0.0027) were positively correlated predictors of BMI values (R adjusted = 0.786, RMSE = 0.138).

From our lipid model excluding PL, lipids that were positively correlated with BMI were CE 18:0 (test statistic = 0.467, Prob > |t| = 0.0014), CE 20:3 (test statistic = 0.364,
Prob > |t| = 0.0015), FA 17:1 (test statistic = 0.691, Prob > |t| = 0.0005), and heptadecanoic acid (test statistic = 0.413, Prob > |t| = 0.0005), while CE 22:4 (test statistic = -0.869, Prob > |t| = <0.0001), FA 14:1 (test statistic = -0.904, Prob > |t| = <0.0001) were negatively correlated (R adjusted = 0.726, RMSE = 0.156).

The metabolic model using PPP metabolites demonstrated that 2-hydroxyglutaric acid (test statistic = -0.203, Prob > |t| = 0.0189), malic acid (test statistic = -0.225, Prob > |t| = 0.0349), serine (test statistic = -0.620, Prob > |t| = <0.0001), and alanine (test statistic = -0.315, Prob > |t| = 0.0144) were negatively correlated with BMI, yet tyrosine (test statistic = 0.276, Prob > |t| = 0.0157), glucose-1-phosphate (test statistic = 0.307, Prob > |t| = 0.0015) were positively correlated (R adjusted = 0.739, RMSE = 0.152).

Inclusion of predictors from all domain models (EHR, PL exclusive, excluding PL, and metabolites) presented the final model with negatively correlated LPC 18:1 (test statistic = -0.210, Prob > |t| = 0.0244), PC 16:1/22:6 (test statistic = -0.090, Prob > |t| = 0.0594), and malic acid (test statistic = -0.207, Prob > |t| = 0.0239) where only PC 16:1/22:6 was not significant following a two-way t-test. Fasting insulin (test statistic = 0.269, Prob > |t| = 0.0018) and PC 16:0/22:4 (test statistic = 0.163, Prob > |t| = 0.0044) were positively correlated with BMI values (R-square adjusted = 0.819, RMSE = 0.127). A visualization of this final model for BMI from EHR data and PPP small molecules can be found in Figure 7.
- **Combined data domains for FF enabled the best prediction of BMI values against the other Lipid, Metabolic, and EHR domain models**

As determined from FF, the PL model selected fewer lipid species such as LPC 18:2 (test statistic = -0.390, Prob > |t| = 0.0061), and LPC-(11Z-octadeceneyl) (LPC-O) 18:1 (test statistic = -0.466, Prob > |t| = 0.0006) which were negatively correlated with BMI values. Those PL that were positively correlated with BMI included LPC 20:2 (test statistic = 0.385, Prob > |t| = 0.0280), PC 20:3/20:4 (test statistic = 0.193, Prob > |t| = 0.0003) and PE-P 16:0/22:6 (test statistic = 0.478, Prob > |t| = <0.0001). With fewer lipids, the FF PL model achieved a greater fit with a lower RMSE (R adjusted = 0.797, RMSE = 0.134).

From our FF lipid model excluding PL, lipids that positively correlated with BMI involved CE 20:3 (test statistic = 0.855, Prob > |t| = <0.0001), FA 19:0 (test statistic = 0.259, Prob > |t| = 0.0028), and FA 22:2 (test statistic = 0.356, Prob > |t| = 0.0239) while CE 16:0 (test statistic = -0.389, Prob > |t| = 0.0236), CE 18:3 (test statistic = -0.548, Prob > |t| = 0.0004), FA 18:1 (test statistic = -0.648, Prob > |t| = 0.0009) were lipids negatively correlated with BMI (R adjusted = 0.706, RMSE = 0.162).

Few FF metabolites were required to predict BMI values with good fit, those being aspartic acid (test statistic = 0.413, Prob > |t| = 0.0003) and tyrosine (test statistic = 0.583, Prob > |t| = <0.0001) that were positively correlated, and histidine (test statistic = -0.377, Prob > |t| = 0.0285) and indole-3-propionic acid (test statistic = -0.241, Prob > |t| = 0.0008) negatively correlated with BMI (R adjusted = 0.754, RMSE = 0.148).
A combination of all 4 components yielded a final model with PE-P 16:0/22:6 (test statistic = 0.206, Prob > |t| = 0.0040), aspartic acid (test statistic = 0.207, Prob > |t| = 0.0141), and fasting insulin (test statistic = 0.269, Prob > |t| = <0.0001) as positively correlated variables with BMI values, whereas indole-3-propionic acid (test statistic = -0.225, Prob > |t| = <0.0001) was negatively correlated with BMI (R-squared adjusted = 0.856, RMSE = 0.113), the results of which can be viewed in Figure 8.

![Figure 8: Final model for BMI from EHR data and FF small molecules.](image)

**Discussion**

This study was able to determine the metabolite and lipid profiles that were specific to changes in body weight, associated with pregnancy outcomes, and correlated to embryo quality following in vitro fertilization (IVF). By measuring two distinct matrices of 26 women, platelet poor plasma (PPP) and follicular fluid (FF), we delineated what matrix interactions were present on the composition of these significant metabolites and lipids and found inter-matrix fluctuations in the levels of these analytes that corresponded with our chosen pregnancy outcomes. Several studies have explored the utility of metabolomics to assess pregnancy outcomes in women of varying weight class.
subsequent IVF, yet were restrained in their findings by the level of specificity in their analyses or by not considering matrix effects$^{94-97}$.

Maternal central carbon metabolism undergoes substantial alterations early in gestation. One such example includes 2-hydroxyglutaric acid (2-HG), a derivative of Krebs cycle intermediate 2-oxoglutarate that, when obtained from PPP, was positively associated with pregnancy. Respiratory alkalosis is a common acid-base disorder that is observed in pregnancies, and the L-enantiomer of 2-HG accumulates in the context of a hypoxic ventilatory response to alleviate reductive stress$^{108-112}$. With regards to pregnancy, the metabolic pathway of L-2-HG in humans is poorly characterized and requires further study. Carbohydrates and lipid metabolites have also served as putative markers of embryo quality$^{97,113,114}$. Glucose-1-phosphate (G-1-P) obtained from PPP (Figure 5), at high concentrations, was negatively associated with pregnancy; notably, pregnancy alters glucose-6-phosphate dehydrogenase activity resulting in enzymatic deficiency$^{115}$. Furthermore, plasma concentrations of glycogen phosphorylase, which catalyzes the rate-limiting step of glycogen to G-1-P, are elevated in pregnant women experiencing preterm preeclampsia$^{116}$. We found that in PPP, the higher the concentration of phosphate, the greater 2PN values. Phosphate serves as an intracellular buffer and regulates the death of hypertrophic cells via caspase-9-mediated pathways$^{117}$. Mechanistic studies on oocyte progression to the blastocyst stage however have demonstrated that media devoid of glucose and phosphate, but enriched in glutamine, improves the development of embryos$^{118,119}$, complicating the interaction between phosphate and oocyte maturation.
Amino acid metabolism is also altered throughout the course of pregnancy; relative to the first trimester, 2-ketoisovaleric acid, a keto-acid (BCKA) intermediate within branched-chain amino acid catabolism (BCAA), has been shown to decrease in plasma\textsuperscript{112,120}. Elevated 2-ketoisovaleric acid (KV) is indicative of impaired BCAA, where an accumulation of toxic BCKA can promote the development of type 2 diabetes mellitus\textsuperscript{120}. Our results confirm the deleterious effects of accumulated BCKA like KV, when obtained from PPP, as higher KV was correlated with lower 2PN values and was negatively associated with clinical pregnancies. From our study, plasma concentrations of 3-aminoisobutyric acid (BAIBA) and glutamic acid decreased with 2PN values. Degradation of BAIBA, an endogenous protective myokine, regulates adipose tissue browning and improves insulin sensitivity\textsuperscript{121}. A reduction in the turnover rate, or a higher concentration, of amino acids like glutamic acid track with meiotic phase progression. An increase in glutamic acid corresponds with oocyte degradation\textsuperscript{122}. Pregnancy is a period of extreme physiological demand for protein synthesis. The anionic form of glutamic acid, glutamate, plays a vital role in protein synthesis and persists at low concentrations throughout gestation so that proteins may be retained to meet the metabolic demand; hence, elevated glutamic acid may indicate a breakdown in protein synthesis\textsuperscript{123}. Obtained from PPP, ornithine was elevated in response to higher 2PN values (Figure 2), however curiously higher ornithine concentrations corresponded with unsuccessful pregnancy following IVF. Ornithine can act as a substrate for putrescine synthesis, a conjugate base that positively influences oocyte maturation; however, elevation of ornithine may also correspond with an inborn error in the urea cycle\textsuperscript{124,125}. Future studies that parse out the exact role plasma ornithine has towards pregnancy outcomes.
When eliciting the effect of fatty acids (FA), it is of great import to determine where particular FAs are sequestered within the lipidome. Our analysis revealed free fatty acid and glycerophospholipid species consisting of capric acid (FA 10:0), palmitic acid (FA 16:0), palmitoleic acid (FA 16:1), heptadecanoic acid (FA 17:0), stearic acid (FA 18:0), oleic acid (FA 18:1), linoleic acid (FA 18:2), α-linoleic acid (FA 18:3), arachidonic acid (FA 20:4), docosatetraenoic acid (FA 22:4), and docosahexaenoic acid (FA 22:6) to elicit differential effects on embryo quality, body mass composition, and clinical pregnancy outcomes, dependent upon matrix. Accumulation of nonesterified FA 10:0 and FA 18:0 in FF was found to be negatively associated with 6-weeks clinical pregnancy, while greater concentrations of FA 17:0 and FA 18:3 were positively associated with achieving clinical pregnancy post IVF. Capric acid (FA 10:0) regulates inflammation by functioning as a modulating ligand for peroxisome proliferator-activated receptors, and by interfering with nuclear factor kappa beta activation\textsuperscript{126,127}. Stearic acid (FA 18:0) along with palmitic acid (FA 16:0), which are the predominant saturated free fatty acids in FF\textsuperscript{128}, activate inflammatory signaling pathways and can induce endoplasmic reticulum stress\textsuperscript{129,130}. Previous studies found that embryos which released greater concentrations of FA 10:0 and FA 18:0 into the microenvironment emulsion saw a greater reduction in birth\textsuperscript{131}. Conversely, FA 18:3, an essential ω-3 polyunsaturated FA, is evinced to increase the number of mature oocytes in a dose-dependent fashion\textsuperscript{132}. Additionally, high concentrations of cholesteryl esters (CE) bearing FA 18:3 (CE 18:3) were positively associated with pregnancy, while CE 20:5 higher concentrations in FF were negatively associated with pregnancy. Cholesteryl esters can accumulate in steroidogenic cells as
well as in pools of lipid droplets; their main functions are to store and transport cholesterol as well as PUFA\textsuperscript{133}. These PUFA are then integrated into steroidogenesis, and can serve as precursors to prostanoids and lipoxygenase products as well as signaling molecules\textsuperscript{134}.

The number of two polar bodies (2PN) increased with higher concentrations of LPI 18:1, LPE 20:4, LPC 22:4 in FF, whereas LPC 16:1 and LPI 18:2 were negatively correlated with 2PN. Interestingly, absolute plasma concentrations of LPI 18:2 among other lysophospholipids (lysoPL) are significantly reduced with gestational diabetes mellitus (GDM), possibly due to an altered glucose metabolism; additionally, the ratio of saturated lysoPL - to - unsaturated lysoPL is higher during pregnancy with GDM, indicating that lysoPL acyl moieties may determine the effect of lysoPL species in lipid metabolic pathways\textsuperscript{135}. Platelet poor plasma LPC 18:1, a purported inhibitor of reactive oxygen species synthesis and neutrophil effector responses, was found here to negatively correlate with BMI values\textsuperscript{136}. Species such as LPC 20:4 and LPC 22:4 can potentiate anti-inflammatory effects, neutralizing LPCs that have been shown to transduce pro-inflammatory signals such as LPC 16:0, and these very species are known to be altered significantly in the serum of obese women\textsuperscript{137,138}. Lysophospholipids therefore may serve as markers of oxidative stress, lipotoxicity, and other inflammatory factors that regulate embryo quality and body composition\textsuperscript{109,138,139}. Yet the physiological role of lysophospholipids, especially by their native matrix, are nascently understood despite their apparent involvement in numerous inflammatory, proliferative, and cytoskeletal responses as signaling molecules\textsuperscript{138}. 
Stereoisomeric and structural, or “positional”, isomeric lipids can have markedly different biological ramifications for overall health and pregnancy outcomes. The position of a fatty acyl moiety affects its distribution in phospholipid species. Phosphatidylcholines (PC) such as PC 16:0/22:4 acquired from PPP, and PC 16:0/22:6 obtained from FF, were positively correlated with rising BMI values, while PC 16:1/22:6 was negatively correlated with BMI. Interestingly, PCs containing FA 22:6, an ω-3 polyunsaturated fatty acid that serves as a precursor for several pro-resolving lipid mediators[40], varied here in its association with BMI dependent upon other fatty acyl chains in the sn1 position. While FA 22:6 has also been linked to positively influence biochemical processes in fetal development[41,42], these data exemplify the importance of metabolomic analyses to characterize the lipidome with a greater level of specificity that can capture the physicochemical properties of lipids.

Palmitic acid is a saturated FA that at high concentrations induces a proinflammatory response, as determined by cytokine and adipokine profiling; furthermore, elevated nonesterified palmitate correlates with insulin resistance in pregnant women[29] and at higher plasma concentrations is associated with a reduction in the number of births[31]. Reesterification of LPC to PC, and de novo synthesis of PC species such as PC 16:0/22:4, have been found to be amplified by the metabolic dysregulations typical of an obese state[43]. The inverse is true for palmitoleic acid (FA 16:1) bearing PCs; additionally, we found nonesterified FA 16:1, when obtained from FF,
at higher concentrations to be positively associated with clinical pregnancy, 6 weeks subsequent IVF. Palmitoleic acid is a lipokine that exhibits insulin-sensitizing and anti-inflammatory properties. During maternal obesity, it has been reported that the synthesis of free form palmitoleic acid, and storage lipids bearing FA 16:1, is downregulated\textsuperscript{141}. Two plasmalogens obtained from FF also correlated with pregnancy outcomes, with an ethanolamine plasmalogen (PE-P 16:0/22:6) positively correlating with BMI, and a choline plasmalogen (PC-P 18:0/22:6) negatively correlating with 2PN values. Concentrations of these lipid species tend to fluctuate with the physiological processes such as aging and serve as putative biomarkers of ovarian reserve when assayed from FF\textsuperscript{144}. Although the functions of plasmalogens are not fully understood, the literature suggests that these lipids act as reservoirs of PUFA, such as FA 22:6, to be further metabolized into second messenger molecules involved in regulating inflammation\textsuperscript{144,145}.

Acylcarnitine species esterified to palmitate for example (acylcarnitine FA 16:0), when obtained from FF, increases with 2PN values. It has previously been reported that IVF may result in excess carnitine consumption and greater depletion of long-chain acylcarnitine pools by oocytes\textsuperscript{146}. Ergo, a positive correlation for acylcarnitine C16:0 and 2PN values may suggest metabolic processes, such as a complete tricarboxylic acid cycle, that follow the course of pregnancy. Additionally, oleic acid (FA 18:1) bearing acylcarnitines, when obtained from FF, decreased with 2PN values. Oocyte competence can be negatively impacted by saturated free fatty acids like stearic acid and palmitic acid. When not sequestered to β-oxidation, free form oleic acid and palmitoleic acid can stimulate the distribution of saturated FA towards neutral storage lipids and β-oxidation, away from apoptotic, pro-inflammatory pathways\textsuperscript{141(p),147,148}.
Uric acid was positively associated with clinical pregnancy, when it was obtained from PPP. Early pregnancy serum uric acid levels tend to fall, related to the uricosuric effects from estrogen. Uric acid is also positively associated with insulin resistance, hence this finding elicits further study as to determine why uric acid was positively associated with pregnancy. In general, uric acid concentration decreases during the first three months of normal pregnancy due to increased glomerular filtration rate (GFR). However, a previous study showed that the concentration of uric acid in the plasma of women undergoing IVF decreased less than the women who had a normal pregnancy\textsuperscript{149}. To our knowledge, there are limited studies that assess kidney function and GFR in IVF. Another possible explanation for our finding of uric acid increasing with pregnancy is that IVF may increase the chance of induction of preeclampsia\textsuperscript{150}. Conceptions achieved by IVF resulted in a greater risk to develop preeclampsia (2.7 times)\textsuperscript{151,152}.

Indole-3-propionic acid (IPA), synthesized from tryptophan and a natural product derived from the gut microbiota, has been associated with insulin secretion and sensitivity, and negatively correlated with low-grade inflammation. This could be due to IPA being the most potent scavenger of the hydroxyl radicals, and elevated concentrations of IPA in plasma is associated with reduced risk for incident type 2 diabetes (T2D); yet this protective effect may be diminished or negated in patients who already have T2D. Obesity has had a documented signature on the metabolomic profile of FF, decreases IPA in FF as we determined here with higher BMI values. These data suggest that IPA obtained from FF may model antioxidant status, or a reflect the gut metabolism interacting with follicular development\textsuperscript{153}. IPA can also reduce the levels of various lipid peroxidation markers and DNA damage. In a previous study using rats
model, IPA was shown to improve the serum insulin resistance index, glucose, and insulin levels. Obesity has been studied in animals and humans. It causes alteration of bacterial species in the gut microbiome compared to controls. Considering this, Ruebel et al. hypothesized that obesity changes the profiles of bacteria species, reducing IPA production. Alternatively, the systemic inflammation caused by obesity could cause intestinal inflammation, which would modulate the gut microbiome and lower IPA levels. Future research in germ-free models will be needed to confirm this hypothesis as well as determine the role of FF IPA in oocyte maturation and development.\textsuperscript{153}

Malic acid from PPP is exogenous, obtained from a diet of citric fruits, and was also found to be negatively associated with BMI values (figure 7). Dietary choices and cooking methods may explain the decline for the obese group. The natural sources of malic acid ($p = 0.026$) are typically vegetables, berries, cherries, and citrus fruits; more recently, it has been incorporated into popular fruit vinegar drinks for its antioxidant properties and taste, suggesting that higher malic acid in PPP could result in a patient with lower systemic oxidative stress.\textsuperscript{6} (Figure 6). As the relation between obesity and oxidative stress is well characterized in previous research,\textsuperscript{154,155} an in vitro study shows that the decrease in malic acid concentration is an oxidative stress biomarker besides vitamin C.\textsuperscript{156}

Beta-alanine is a nonessential amino acid that was negatively associated with 2PN values as obtained from PPP (figure 2). During pregnancy, the development of gluconeogenesis is attenuated due to decreased alanine flux. This process is triggered by the intrahepatic mechanism which lowers the deamination of alanine.\textsuperscript{157} It has been hypothesized that the metabolic cooperation between oocytes and follicular cells leads to
the uptake of amino acids. In addition, these findings indicate that the cells do not have the capability to increase the level of leucine in mice. A liquid chromatography-tandem Mass spectrometry study revealed that the presence of aromatic amino acids in the fluid of polycystic ovary syndrome (PCOS) patients is associated with an increase in body mass index. These changes involve elevation of amino acids including glutamic acid, phenylalanine, alanine, and arginine. The study supports the notion that the lack of these nutrients can affect the quality of oocyte eggs. A previous study showed that evaluation of the non-invasive amino acid profiling of bovine oocytes with high-performance LC might be utilized to evaluate oocyte development in bovine. The results showed a significant decrease in glutamine and increase in alanine in the medium of oocytes that failed in the cleavage process after 3 days postfertilization

Aspartic acid is interestingly correlated with higher BMI values as well as a greater number of 2PN. A previous study showed that increased aspartic acid in PCOS women is related to increased chance of getting pregnant. In addition, D'Aniello et al. demonstrate a positive relation between aspartic acid elevation and egg quality in obese women undergoing IFV. Furthermore, d-aspartic acid improves the synthesis of some hormones such as LH and testosterone. The LH hormone released from the anterior pituitary gland is essential for pregnancy sustained through activation of corpus luteum which in turn causes production of progesterone. The mid-cycle luteinization surge is known to play a crucial role in the development of the oocyte and the follicularization process. However, recent studies indicate that the abnormal level of LH may be detrimental to the embryo development. In addition, the reduction in concentrations of human LH in the follicular fluid can cause decreased production of follicle-stimulating
hormones (FSH). This study supports the hypothesis that decreased concentrations of LH can affect the quality of oocytes\textsuperscript{159}.

Serine was positively associated with pregnancy along with threonine in platelet poor plasma. Valine, glycine, serine, and threonine concentrations in plasma were closely correlated with insulin resistance and obesity. In addition, a significant positive association of lactate and leucine concentrations with insulin resistance was observed regardless of obesity. Notably, significant decreases in serine and threonine concentration were observed in obese women with PCOS and patients with insulin resistance compared to non-obese PCOS patients and patients with normal insulin sensitivity, suggesting opposing effects of PCOS and their two common features on serine and threonine metabolism\textsuperscript{162}. An earlier study on obese mice supplemented with threonine for 10 weeks showed weight reduction and improvement of insulin resistance\textsuperscript{163}. In addition, a cross-sectional study evaluating the association of amino acids regarding insulin sensitivity in men showed a significant association between serine concentration and insulin sensitivity improvement\textsuperscript{164}.

In summation, our findings here of lipid species being correlated with 2PN values provides a basis for the detrimental effects of high BMI, and the lipid disruptions that accompany this physical state, on embryo quality and success of clinical pregnancy. Group II ovulatory disorders compose \textasciitilde85\% of all ovulation disorders and are caused by conditions such as abnormal BMI. Obesity is known to result in disruptions to endocrine and paracrine mechanisms such as insulin signaling, which is evinced to act on the HPO axis and can lead to adverse pregnancy outcomes.
Limitations with this analysis included a small cohort size, which carried the risk of overfitting our regression estimates. Standardization of the regression coefficients aided us in overcoming this challenge, and reliance on adjusted R-squared as a performance metric meant that our measures were adjusted to produce unbiased estimators of the population and effect. Relative quantification\textsuperscript{165} of circulating lipids and metabolites enabled us to determine the specific species that were correlated with the number of distinct pronuclei and two polar bodies (2PN) or embryo quality, molecules correlated with BMI, and putative markers for successful pregnancies following IVF.
Bibliography


38. Lang F, Qadri SM. Mechanisms and Significance of Eryptosis, the Suicidal Death of Erythrocytes. BPU. 2012;33(1-3):125-130. doi:10.1159/000334163


42. Gladwin MT. Pulmonary Complications of Sickle Cell Disease. *n engl j med*. Published online 2008:12.


46. Ren H, Ghebremeskel K, Okpala I, Ugochukwu CC, Crawford M, Ibegbulam O. Abnormality of erythrocyte membrane n-3 long chain polyunsaturated fatty acids in sickle cell haemoglobin C (HbSC) disease is not as remarkable as in sickle cell anaemia (HbSS).


Supplementary appendix:

1. Plasma data:
   a. BMI

<p>| Variable selected                  | Prob&gt;|t| | Model R² and R² adj |
|-----------------------------------|-----|-----------------|
| <strong>BMI and lipids with glycerophospholipids</strong> |     |                 |
| LPC (18:1)                         | 0.0133 | R² = 0.839 |
| LPC(22-6)                          | 0.0127 | R² adj= 0.785 |
| LPC (p-16:0)                       | 0.0021 | Root mean square error=0.138 |
| PC(16:0-22:4)                      | 0.0027 |                 |
| PC(16:1-22:6)                      | 0.0002 |                 |
| PI(18:0-20:4)                      | 0.0422 |                 |
| <strong>BMI and lipids without glycerophospholipids</strong> |     |                 |
| CE (18:0)                          | 0.0014 | R² = 0.794 |
| CE (20:3)                          | 0.0015 | R² adj= 0.726 |
| CE (22:4)                          | &lt;.0001 | Root mean square error=0.156 |
| FA (14:1)                          | &lt;.0001 |                 |
| FA (17:1)                          | 0.0005 |                 |</p>
<table>
<thead>
<tr>
<th>Chemical</th>
<th>Value</th>
<th>$R^2$</th>
<th>$R^2$ adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heptadecanoic acid</td>
<td>0.0089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-hydroxyglutaric acid</td>
<td>0.0189</td>
<td>0.804</td>
<td>0.739</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.0144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose-1-phosphate</td>
<td>0.0015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malic acid</td>
<td>0.0349</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>0.0157</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting glucose</td>
<td>0.3374</td>
<td>0.807</td>
<td>0.768</td>
</tr>
<tr>
<td>Fasting Insulin</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score Block food</td>
<td>0.0054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Sitting Total</td>
<td>0.1554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPC (18:1)</td>
<td>0.0244</td>
<td>0.856</td>
<td>0.818</td>
</tr>
<tr>
<td>PC (16:0-22:4)</td>
<td>0.0044</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC (16:1-22:6)</td>
<td>0.0594</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable selected</td>
<td>Prob&gt;</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td><strong>2PN and Lipids with glycerophospholipids</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPC(22:6)</td>
<td>0.0213</td>
<td>$R^2 = 0.805$</td>
<td></td>
</tr>
<tr>
<td>LPE(20:4)</td>
<td>0.0005</td>
<td>$R^2$ adj= 0.730</td>
<td></td>
</tr>
<tr>
<td>PC(14:0-16:1)</td>
<td>0.0060</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(16:1-22:6)</td>
<td>0.2452</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE(18:0-20:3)</td>
<td>0.2569</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE-P(16:0-20:5)</td>
<td>0.0013</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE-P (16:0-22:6)</td>
<td>0.0004</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2PN and Lipids without glycerophospholipids</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE(22:6)</td>
<td>0.00066</td>
<td>$R^2 = 0.799$</td>
<td></td>
</tr>
<tr>
<td>Lauric acid</td>
<td>0.01373</td>
<td>$R^2$ adj= 0.666</td>
<td></td>
</tr>
<tr>
<td>CE (20:5)</td>
<td>0.02362</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA(19:1)</td>
<td>0.02885</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE(16:0)</td>
<td>0.06641</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myristic acid</td>
<td>0.06739</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE(14:0)</td>
<td>0.11267</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (20:5)</td>
<td>0.14690</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (18:2)</td>
<td>0.38866</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (14:1)</td>
<td>0.59931</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**2PN and metabolite model**

<table>
<thead>
<tr>
<th>2-ketoisovaleric acid</th>
<th>0.0001</th>
<th>R² = 0.829</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-aminoisobutyric acid</td>
<td>0.0004</td>
<td>R² adj = 0.749</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>0.0236</td>
<td></td>
</tr>
<tr>
<td>Beta alanine</td>
<td>0.0007</td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>ornithine</td>
<td>0.0225</td>
<td></td>
</tr>
<tr>
<td>phosphate</td>
<td>0.0030</td>
<td></td>
</tr>
<tr>
<td>serine</td>
<td>0.0033</td>
<td></td>
</tr>
</tbody>
</table>

**2PN and clinical parameter model**

<table>
<thead>
<tr>
<th>Female age (yr)</th>
<th>0.4979</th>
<th>R² = 0.623</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Value</td>
<td>R² adj</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Fasting glucose</td>
<td>0.1282</td>
<td></td>
</tr>
<tr>
<td>Fasting insulin</td>
<td>0.9776</td>
<td></td>
</tr>
<tr>
<td>Mediterranean diet score</td>
<td>0.1929</td>
<td></td>
</tr>
<tr>
<td>Total score Block food Fat screener 23 high.</td>
<td>0.8744</td>
<td></td>
</tr>
<tr>
<td>Total score Block food questionnaire fruit vegetable screener 11 low.</td>
<td>0.0712</td>
<td></td>
</tr>
<tr>
<td>Average Sitting Total Minutes day</td>
<td>0.0221</td>
<td></td>
</tr>
<tr>
<td>Total cumulative FSH used</td>
<td>0.0015</td>
<td></td>
</tr>
</tbody>
</table>

**Final plasma 2PN model tested for clinical parameter (Exclude the clinical parameter)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>R²</th>
<th>R² adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPC (22:6)</td>
<td>0.0057</td>
<td>0.792</td>
<td>0.726</td>
</tr>
<tr>
<td>LPE (20:4)</td>
<td>0.0023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC(14:0-16:1)</td>
<td>0.0329</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE (16:0)</td>
<td>0.0056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (20:5)</td>
<td>0.0212</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phosphate</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Follicular fluid:

   a. BMI

| Variable selected | Prob>|t| | Model R² and R² adj |
|-------------------|-----|--------------------------|


<table>
<thead>
<tr>
<th><strong>BMI and with glycerophospholipids</strong></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>LPC(18:2)</td>
<td>0.0061</td>
<td></td>
<td>$R^2 = 0.839$</td>
</tr>
<tr>
<td>LPC (20:2.)</td>
<td>0.0280</td>
<td></td>
<td>$R^2$ adj= 0.797</td>
</tr>
<tr>
<td>LPC(o-18:1)</td>
<td>0.0006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC (20:3-20:4)</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PE-P(16:0-22:6)</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>BMI and without glycerophospholipids</strong></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CE (16:0)</td>
<td>0.0236</td>
<td>$R^2 = 0.779$</td>
<td></td>
</tr>
<tr>
<td>CE (18:3)</td>
<td>0.0004</td>
<td>$R^2$ adj= 0.705</td>
<td></td>
</tr>
<tr>
<td>CE(20:3)</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (18:1)</td>
<td>0.0009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA(19:0)</td>
<td>0.0028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA(22:2)</td>
<td>0.0239</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>BMI and metabolite model</strong></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspartic acid</td>
<td>0.0003</td>
<td>$R^2 = 0.795$</td>
<td></td>
</tr>
<tr>
<td>histidine</td>
<td>0.0285</td>
<td>$R^2$ adj= 0.754</td>
<td></td>
</tr>
<tr>
<td>Indole-3-propionic acid</td>
<td>0.0008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tyrosine</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Final BMI model

| Variable                        | Prob > |t| | Model R² and R² adj |
|--------------------------------|--------|---|---------------------|
| PE-P(16:0-22:6)                | 0.0040 |   | R² = 0.880          |
|                                |        |   | R² adj = 0.856      |
| Aspartic acid                  | 0.0141 |   |                     |
| Indole-3-propionic acid        | <.0001 |   |                     |
| Fasting insulin                | <.0001 |   |                     |

## b. 2PN:

<p>| Variable selected                          | Prob &gt; |t| | Model R² and R² adj |
|--------------------------------------------|--------|---|---------------------|
| 2PN and with glycerophospholipids          |        |   |                     |
| Acylcarnitine (C16:0)                      | 0.0033 |   | R² = 0.833          |
|                                            |        |   | R² adj = 0.755      |
| Acylcarnitine C18:1                        | 0.0358 |   |                     |
| LPC (16:1)                                  | 0.0047 |   |                     |
| LPC (22:4)                                  | 0.0028 |   |                     |
| LPE (20:4)                                  | 0.0042 |   |                     |
| LPI (18:1)                                  | 0.0048 |   |                     |
| LPI (18:2)                                  | 0.0023 |   |                     |
| PC-P(18:0-22:6)                             | 0.0013 |   |                     |
| 2PN and without glycerophospholipids        |        |   |                     |</p>
<table>
<thead>
<tr>
<th>Component</th>
<th>Value</th>
<th>R²</th>
<th>R² adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE (16:1)</td>
<td>0.0517</td>
<td>0.813</td>
<td>0.725</td>
</tr>
<tr>
<td>CE (18:1)</td>
<td>0.0009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE (18:3)</td>
<td>0.0732</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE (22:6)</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (14:0)</td>
<td>0.0410</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (16:0)</td>
<td>0.0404</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**2PN and metabolite model**

<table>
<thead>
<tr>
<th>Component</th>
<th>Value</th>
<th>R²</th>
<th>R² adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-monoolein</td>
<td>0.0834</td>
<td>0.794</td>
<td>0.729</td>
</tr>
<tr>
<td>2-aminobutyric acid</td>
<td>0.0021</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alanine</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hexose</td>
<td>0.0010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phenylalanine</td>
<td>0.0108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phenylethylamine</td>
<td>0.0002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Final 2PN model tested for clinical parameter**

<table>
<thead>
<tr>
<th>Component</th>
<th>Value</th>
<th>R²</th>
<th>R² adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPE (20:4)</td>
<td>0.0040</td>
<td>0.808</td>
<td>0.747</td>
</tr>
<tr>
<td>PC-P(18:0-22:6)</td>
<td>0.0115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-aminobutyric acid</td>
<td>0.0934</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alanine</td>
<td>0.0062</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compound</td>
<td>Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phenylethylamine</td>
<td>0.0564</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FA (22:3)</td>
<td>0.0438</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>