SIMULTANEOUS SPECTROPHOTOMETRIC AND CHEMOMETRIC DETERMINATION OF CHOLESTEROL AND MONO-/POLYUNSATURATED FATTY ACIDS

By

GERARD G. DUMANCAS

Bachelor of Science in Chemistry
University of the Philippines
Miag-ao, Iloilo, Philippines
2005

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May 2012

SIMULTANEOUS SPECTROPHOTOMETRIC

AND CHEMOMETRIC DETERMINATION

OF CHOLESTEROL AND

MONO-/POLYUNSATURATED

FATTY ACIDS

Dissertation Approved:

Dr. Neil Purdie

_____

Dissertation Adviser

Dr. Kevin Ausman

_____

Dr. Kenneth Darrell Berlin

_____

Dr. Frank Blum

_____

Dr. Edralin Lucas

_____

Outside Committee Member

Dr. Sheryl A. Tucker

_____

Dean of the Graduate College

PREFACE

The Centers for Disease Control and Prevention (CDCP) has identified heart diseases, cancer, stroke (cerebrovascular diseases), alzheimer's, and diabetes as the major causes of deaths in the United States for 2007. It has been believed that such diseases are known to be caused by major risk factors including cholesterol and polyunsaturated fatty acids (PUFAs). With the emergence of such major risk factors, scientists are coming up with detection methods to efficiently identify and quantify such analytes in various types of biological media, including human serum. Current methodologies of identifying cholesterol and PUFAs in various types of biological media constitute techniques such as gas chromatography (GC) which requires significant time in order to carry out the analysis. This study is aimed in completing the analysis of the seven most abundant lipids in human serum using chemometric algorithms and consequent validation with gas chromatography-mass spectrometry (GC-MS). An extension of the development and application of the assay to different types of biological media including vegetable oils was also done. The results of these studies were published in the journals mentioned in this dissertation. As part of unpublished results, pattern recognition was also performed that allowed the discrimination of the various food samples from one another. The last part of this study is aimed in the inclusion of a monounsaturated fatty acid *oleic acid* in human serum as well as the spiking of human serum samples to build a new calibration model. Throughout the course of the study, a variety of chemometric algorithms were applied for the simultaneous determination of cholesterol and

polyunsaturated/monounsaturated fatty acids (PUFA/MUFA) in various biological samples.

A lot of effort was put in carrying out this study. First, I sincerely thank the Lord God guiding me always. My dissertation would have not also been completed without the support of my thesis advisor, Dr. Neil Purdie, who has helped me reach this milestone in my career. He has been very supportive of me of whatever endeavors I have in life. His guidance and determination in completing the validation of the assay and its application to various types of biological samples has been very admiring. To my labmate, Mary Kimani, who has also helped me in the validation of the assay, I truly appreciate your efforts.

I sincerely appreciate the help of Dr. Lisa Reilly, whose suggestions helped me a lot in continuing this research. Dr. Rockley has also helped me a lot especially when I first started understanding the basics of Moore-Penrose Pseudoinverse Algorithms. The suggestions, contributions, and support of my committee members, Dr. Ausman, Dr. Berlin, and Dr. Blum, are greatly appreciated as well.

Dr. Lucas who is also another one of my committee members and also part of the Filipino community, has been a consistent supporter of mine. I really appreciate her efforts both intellectually and socially. I would also like to extend my thanks to all my Filipino family in Stillwater for always constantly supporting and praying for my endeavors. Also, the friendship of my colleagues especially those coming from Dr. Ausman's lab is appreciated. Lastly, this study would have not been completed without my family. To my family, thank you for your constant prayers and guidance.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF SYMBOLS AND ABBREVIATIONS

γ-Linolenic acid	GLNA
α-Linolenic acid	ALNA
2-Nitrophenylhydrazine	2-NPH
3-Hydroxy-3-methylglutaryl-CoA	3-HMG-CoA
4-Bromomethyl-7-methoxy-coumarin	Br-Mmc
6,7-Dimethoxy-1-methyl-2(1H)-quinoxalinone -3-propionylcarboxylic acid hydrazide	DMEQ-hydrazide
Acetyl chloride	AC
Acyl coenzyme A	Acyl CoA
Adenosine monophosphate	ADM
Adenosine triphosphate	ATP
Adrenoleukodystrophy	ALD
Alzheimer's disease	AD
American Heart Association	AHA
American Heart Association	AHA
Amyloid β-peptide	Aβ
Apolipoprotein E	ApoE
Arachidonic acid	AA
Artificial neural network	ANN
Body mass index	BMI
Cardiovascular disease	CVD
Centers for Disease Control and Prevention	CDCP
Central nervous system	CNS
Coronary heart disease	CHD
Direct calibration	DC
DMEQ-hydrazide	HCPI
Docosahexaenoic acid	DHA
Docosapentaenoic acid	DPA
Eicosapentaenoic acid	EPA
Eicosatrienoic acid	ETA
Evaporative light-scattering detector	ELSD
Familial hypercholesterolemi	FH
Fatty acid methyl ester	FAME
Free fatty acid	FFA
Gas chromatography	GC
Gas chromatography/mass spectrometry	GC-MS
Gas liquid chromatography	GLC

Genetic algorithm	GA
Genetic algorithm partial least squares	GAPLS
High performance liquid chromatography	HPLC
High-density lipoprotein-cholesterol	HDL-C
High-density lipoproteins	HDL
Hillcrest Medical Center	HMC
Intermediate-density lipoproteins	IDL
Intracerebroventricular	ICV
K-matrix	KM
Leukotriene B_4	LTB_4
Linoleic acid	LA
Linolenic acid	LNA
Liquid chromatography/mass spectrometry	LC-MS
Liquid column chromatography	LCC
Low density lipoprotein	LDL
Low density lipoprotein-cholesterol	LDL-C
Mass spectrometry	MS
Monounsaturated fatty acid	MUFA
Non-negative least squares	NNLS
Nuclear magnetic resonance	NMR
Oleic acid	OA
Ordinary least squares	OLS
Paper chromatography	PC
Partial least squares	PLS
Partial least squares 1	PLS1
Partial least squares 2	PLS2
Partial least squares-genetic algorithm	PLS-GA
P-matrix	PM
Polyunsaturated fatty acids	PUFAs
Principal component	PC
Principal component analysis	PCA
Principal component regression	PCR
Refractive index	RI
Ridge regression	RR
Saturated fatty acid	SFA
Solid-phase extraction	SPE
Supercritical fluid extraction	SFE
Thin layer chromatography	TLC
Thromboxane A_2	TXA_2
Triacylglycerols	TAGs
Triglyceride	TG
Tumor necrosis factor	TNF
Ultraviolet visible	UV
United States Department of Agriculture	USDA
Very-low-density lipoproteins	VLDL

Perchloric acid	PA
Predicted residual error sum of squares	PRESS
Root mean square error	RMSE
Root mean square error of cross validation	RMSECV
Root mean square error of prediction	RMSEP

CHAPTER 1

INTRODUCTION

Part of this chapter has been published in the *Journal of Biotech Research*, the *Lecture Notes in Engineering and Computer Science*, and the *Lipid Technology Journal* and appears in this dissertation with the journals' permission.[1-4]

In the modern era, biomedical research plays a very critical role in human health. Within the biomedical research area, scientists are searching for new biomarkers that would serve to identify the causes of obesity, coronary heart disease, diabetes, hypercholesterolemia, and cancer among others. The Centers for Disease Control and Prevention (CDCP) reported that the diseases of the heart, cancer, stroke, Alzheimer's disease, and diabetes are among the top leading causes of deaths in the US for 2007.[5] The above mentioned diseases are all high risk conditions that are top priority in research laboratories. Cholesterol and polyunsaturated fatty acids (PUFAs) are among the biomarkers associated with the previously mentioned diseases.

The role of cholesterol and lipids in atherosclerosis has been studied for decades, and many of the cellular and molecular mechanisms have been worked out in considerable detail.[6] Across cultures, cholesterol is linearly related to coronary heart

disease (CHD) mortality, and the relative increase in CHD mortality rates with a given cholesterol increase is the same.[7] PUFAs in the diet, on the other hand, have long been considered essential to the growth and proper nutrition of humans and animals. On the contrary, they have also exhibited negative effects.[8] PUFAs exist in two major kinds, the ω-6 and the ω-3 forms. The ω-6 fatty acid esters such as the linoleic, conjugated linoleic, and arachidonic acids are known to enhance formation of cholesterol gallstones, a stimulus to carcinogenesis, increased vitamin E requirements, promotion of obesity, increased uptake of plant sterols, and increased cholesterol absorption.[8-10] The ω-3 esters of the fatty acids such as α-linolenic, eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), on the other hand, have effects on diverse physiological processes impacting normal health and chronic disease, such as the regulation of plasma lipid levels, cardiovascular and immune function, insulin action, and neuronal development and visual function.[11-21] Several studies have already shown the dietary effects of ω-3 and ω-6 fatty acids. A study by Cave showed that diets containing high levels of the ω-6 PUFAs have routinely enhanced tumorigenesis in lipid sensitive carcinogen-induced and tumor transplant tumor models, whereas diets with equivalent levels of ω-3 PUFAs have diminished tumorigenesis.[22]

The ω-6/ω-3 ratio is an important indicator of human health. There is evidence that a 4:1 ratio is required for maximum benefit for cardiovascular disease and less than 2:1 to have any effect on cancer.[23] Genetically speaking, human beings today live in a nutritional environment wherein major changes in our diet have taken place, particularly in the type and the amount of essential fatty acids and in the antioxidant content of foods.[24-28] Comparing the hunter-gatherer with the western diet and lifestyle, the ω-6 to

ω-3 ratio has shifted considerably from low to high.[26] Excessive amounts of ω-6 PUFA and a very high ω -6/ω -3 ratio, as is found in today's Western diets, promote the pathogenesis of many diseases, including cardiovascular disease, cancer, and inflammatory and autoimmune diseases, whereas increased levels of ω -3 PUFA (a lower ω -6/ ω -3 ratio), exert suppressive effects.[29]

Several methods now exist for the determination of cholesterol and PUFAs levels in human serum. Current methods of quantifying cholesterol and PUFAs in human serum are done separately from each other. For cholesterol, several methods exist for quantifying the analyte in human serum, which are mostly based on enzymatic tests. Since the introduction of enzymatic assays for total cholesterol in serum (i.e., the sum of free and fatty-acid-esterifled cholesterol) nearly about four decades ago, the methods have largely replaced the more laborious, interference susceptible, and less-specific routine nonenyzmatic methods.[30, 31] For PUFAs, on the other hand, gas chromatography (GC), thin layer chromatography (TLC), and high performance liquid chromatography (HPLC) are the methods commonly used for quantifying such analytes in human serum.[32] However, these methods are complicated, quite laborious, and suffer from the difficulty of obtaining meaningful concentrations.

The Purdie Assay was established to enable the simultaneous quantification of cholesterol and PUFAs in synthetic mixtures and human serum without the need for analytical separations.[2] The experimental method is accomplished in less than twenty minutes without the need for sophisticated analytical instrumentations like GC or HPLC. The assay originated with the Liebermann-Burchard reaction that was once the current gold standard for cholesterol and was later based upon a reaction attributed to Chugaev

and Gastev. The assay reagent had the extra selectivity of acylation of the α- over the ß-position at the C-17 carbon that enabled the differentiation of anabolic steroids. In that and a following study, it was also determined that if multiple unsaturated lipids were present, the resulting compounded spectrum was the simple addition of the weighted spectra for each of the components, assuming no extraneous interferences were present.[33]

The ultimate goal of this research project was to develop a simple, direct alternative method for the simultaneous quantitative determination of cholesterol and PUFAs (linoleic acid (LA), linolenic acid (LNA), arachidonic acid (AA), eicosapentaenoic acid (EPA), docosahexaenoic acid (DHA), and conjugated linoleic acid (CLA)) in human serum by exploitation of various chemometric algorithms and consequent validation with the gas chromatography-mass spectrometry (GC-MS) method. An inclusion of oleic acid (OA), in addition to the above-mentioned analytes, was also performed. In this study, the performance of the various chemometric algorithms in all seven and eight components was compared. The study started with multiple linear regression techniques [K-matrix (KM), P-matrix (PM), ridge regression (RR), K-matrix ordinary least squares (OLS), and K-matrix non-negative least squares (NNLS)] and is extended to factor-based techniques (Principal component regression (PCR) and partial least squares (PLS1 and PLS2)) and finally to artificial neural network (ANN). Previous study has compared the performance of some of the above mentioned algorithms in four components. In this dissertation, the performance of the above mentioned algorithms is compared with as much as eight components present in human serum. In addition, GC-MS validation is also performed to determine the best algorithm that would be suited for a typical clinical analysis. In addition, extensions of the

application of the assay to various types of samples, including vegetable oils and various food and biological samples, were also done.

For vegetable oil samples, an independent study was done with the objective of simultaneously quantitating oleic, linoleic, and linolenic fatty acids in vegetable oil samples using the Purdie Assay. HPLC, GC, or hyphenated methods such as HPLC/MS, and GC/MS are classical methods of determining the fatty acid composition and levels in vegetable oils.[34, 35] For the determination of the fatty acid composition, the triacylglycerols (TAGs) are transesterified to give the methyl esters prior to analysis because the esters are less polar than the corresponding fatty acids, and, thus, are more compatible with the various chromatographic systems.[36] Though these procedures have been successfully used in various chemical analyses, they, however, do suffer from the disadvantages of being time, labor, and resources consuming. In this study, various chemometric algorithms consisting of KM, NN, PCR, PLS1, and PLS2 were utilized for the deconvolution of the spectrophotometric data using the Purdie Assay. The most robust chemometric models were then compared against each and their regression coefficients applied for the molar concentration determination of oleic, linoleic, and linolenic fatty acids in olive and sunflower oil samples. The obtained concentrations were then validated with the existing United States Department of Agriculture (USDA) database concentrations.[37]

In addition to the above mentioned researches, pattern recognition was also performed in the above-mentioned samples to determine the discrimination capability of the Purdie Assay.

CHAPTER 2

POLYUNSATURATED FATTY ACIDS AND CHOLESTEROL

2.1 CHOLESTEROL

2.1.1 STRUCTURE

Cholesterol is a soft, waxy substance found among the lipids (fats) in the bloodstream and in all our body's cells. It's an important part of a healthy body because it's used to form cell membranes, some hormones, vitamin D, and bile acids and is needed for other functions.[38] Because of its hydrocarbon content (Figure 1 illustrates the chemical structure), cholesterol is not soluble in water.

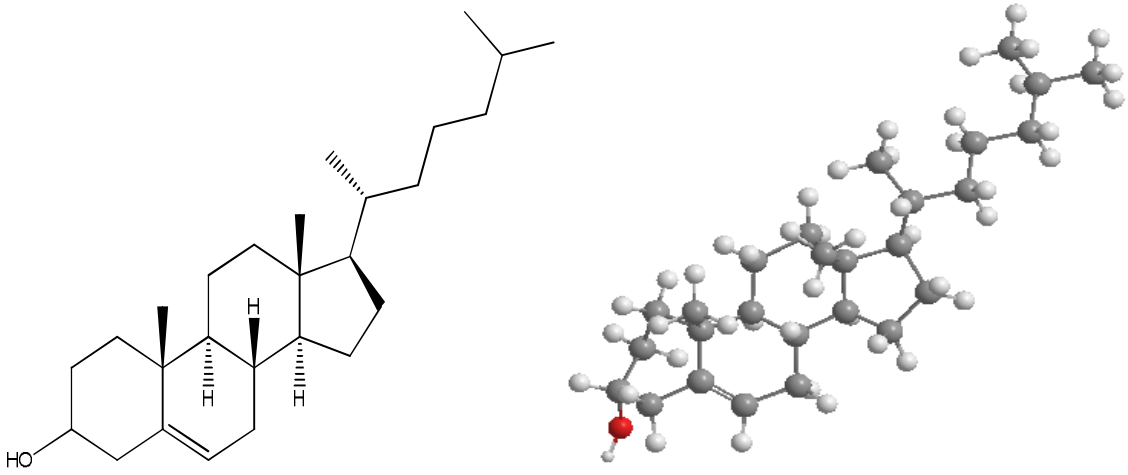


Figure 1. The chemical structures of cholesterol in 2D (left) and 3D.

The exact empirical formula of cholesterol was accurately established in 1888 by Austrian botanist Friedrich Reinitzer, who worked at the Imperial Institute of Plant Physiology at the German University in Prague. Interested in the biologic roles of cholesterol in plants, Reinitzer initially studied cholesterol isolated from the carrot root. However, its cholesterol content was so minute that Reinitzer resorted to purchasing cholesterol from a factory. After purifying the sample by treatment with alcoholic sodium hydroxide, Reinitzer treated cholesterol with bromine and obtained a compound that "precipitates out as *splendid crystals.*" Using a rudimentary but a reliable method called elemental analysis involving combustion of the compound and then analysis of the carbon and hydrogen contents, he deduced the precise molecular formula. In his publication in the prestigious chemistry journal *Monatshefte fur Chemie (Chemical Monthly*) in 1888, Reinitzer was very confident: "The formula of cholesterol *must read* C_{27} H_{46} O.[39] It has a molecular mass of 386.5 g/mol.

2.1.2 FUNCTIONS

Cholesterol has many vital functions in our body. In membrane function, it forms part of each cell in the body. Because it is hard fat, it gives the membrane rigidity and stability. Cholesterol is also important in the synthesis of steroid hormones estrogen, progesterone, and testosterones, as well as in the synthesis of adrenal hormones like aldosterone which regulates water and sodium balance in our body, and cortisol which regulates metabolism, suppresses inflammation, and is produced in response to stress. When we are under chronic stress, our bodies produce a great deal of cholesterol.[40] In general, steroids derived from cholesterol in animals include five families of hormones (the androgens, estrogens, progestins, glucocorticoids, and mineralocorticoids) and bile

acids (Figure 2). Androgens, such as testosterone and estrogens such as estradiol, mediate the development of sexual characteristics and sexual functions in animals. The progestins such as progesterone participate in control of the menstrual cycle and pregnancy. Glucocorticoids (cortisol, for example) participate in the control of carbohydrate, protein, and lipid metabolism, whereas mineralocorticoids regulate salt (Na^+, K^+, Cl^- Na^+) balances in tissues. The bile acids (including cholic and deoxycholic acid) are detergent molecules secreted in bile from the gallbladder that assist in the absorption of dietary lipids in the intestine.[41]

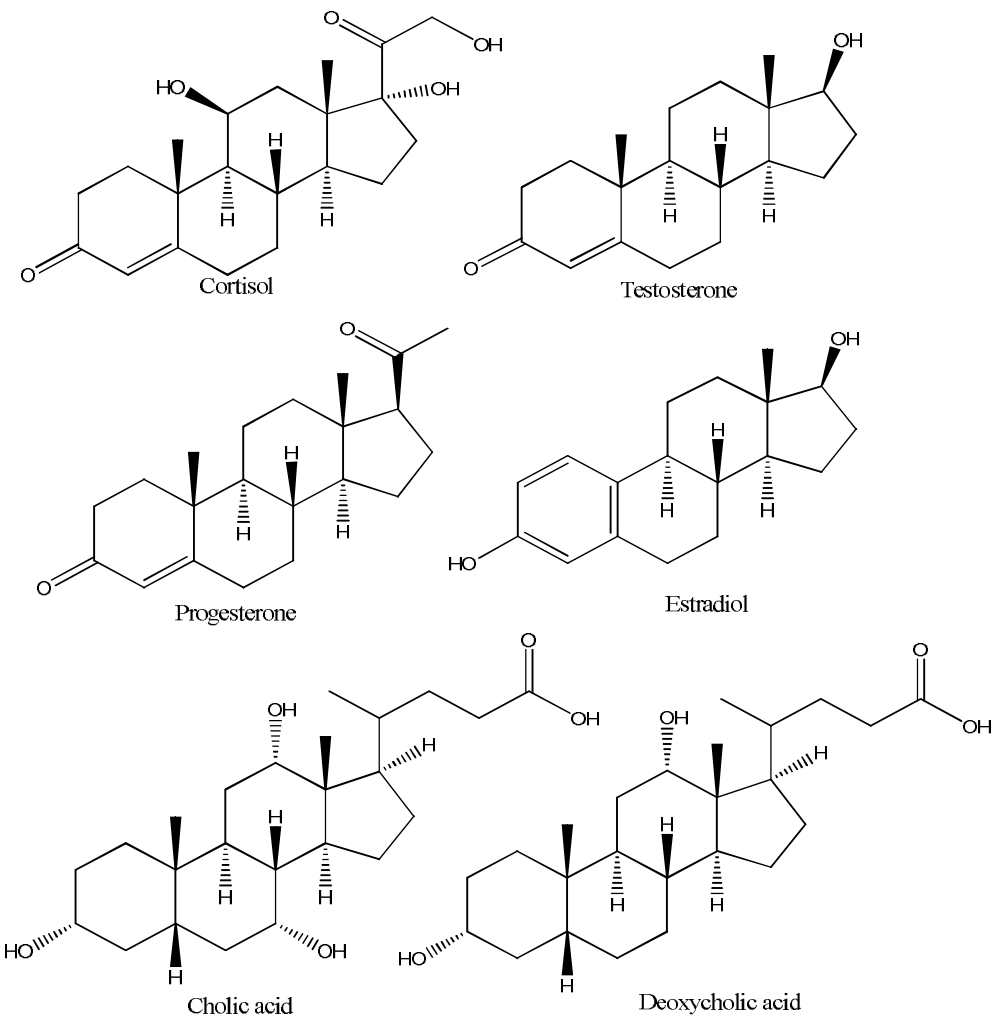
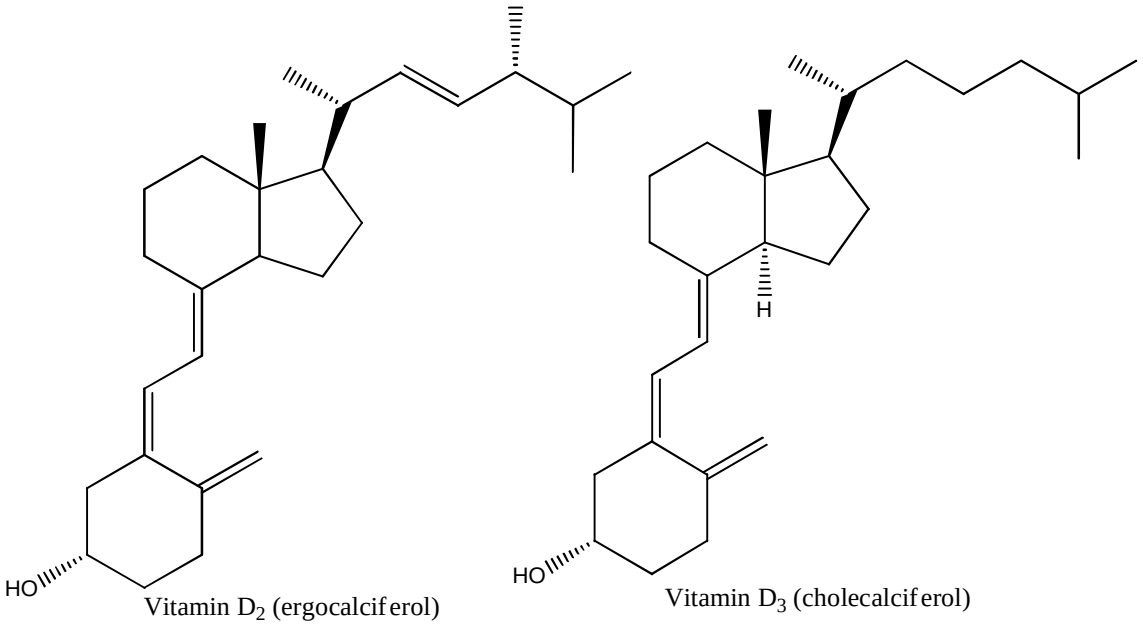


Figure 2. The structures of several important sterols derived from cholesterol.

About 80% of the cholesterol in the body is used by the liver to produce bile salts. Bile is stored in the gallbladder and is used to help in the digestion and absorption of dietary fats and fat soluble vitamins. This is the major route of exit of cholesterol from the body. Bile is secreted into the intestines and leaves the body in bowel movements.[40]

Cholesterol is also important in vitamin D synthesis. The sunlight hitting the human skin converts cholesterol into vitamin D, which is needed to keep the bones strong. Vitamin D (Figure 3) has other important functions in the body like boosting the immune system and helping keep the blood pressure normal. Getting a bit of sunlight on our skin most days of the week can help lower the cholesterol level by facilitating its conversion to vitamin D.[40]



Vitamin D_2 (ergocalciferol) Vitamin D_3 (cholecalciferol)

Figure 3. The two major forms of vitamin D.

In skin protection, cholesterol is secreted into the skin, where it covers and protects us from dehydration, cracking, and the drying effects of the elements. It helps to

keep the skin looking plump and wrinkle-free. Cholesterol plays an important role in healing, as high amounts of it are found in scar tissue.[40]

In serotonin function, cholesterol is necessary for the function of serotonin receptors in the brain. Serotonin (Figure 4) is a chemical that helps to protect us from depression. On the other hand, cholesterol is the main fat present in the myelin sheath, which coats the nerve cells and enables electrical impulses to occur in the brain and spinal cord. A healthy myelin sheet is needed for good concentration and memory.[40]

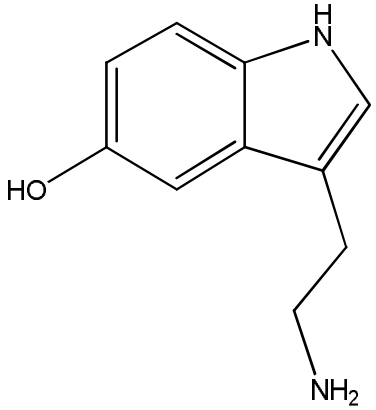


Figure 4. Chemical structure of serotonin (5-hydroxytryptamine).

Lastly, in antioxidant function, cholesterol helps to transform fat soluble antioxidants around our body, such as vitamins E and A (Figure 5), and several antioxidant enzymes.[40]

2.1.3 DIETARY SOURCES AND BIOSYNTHESIS

Cholesterol is a structural lipid that is not required in the diet, since ample amounts are synthesized in the body.[42] It is found in all animal tissues, so that some are present in all foods of animal origin, but eggs are the only common foods rich in

cholesterol. Principal dietary sources include meats (liver, 370 mg; veal, lamb, and beef, 80-85 mg/3 oz serving), eggs (one large, 252 mg), shellfish (shrimp, 128 mg/3 oz serving), poultry (chicken, 74 mg/3 oz serving), fish (45-60 mg/3 oz serving), and dairy products (whole milk, 34 mg/8 oz; ice cream, 54-98 mg/cup; American cheese, 28 mg/oz). It is virtually absent in foods of plant origin, which, do however, contain other sterols.[43]

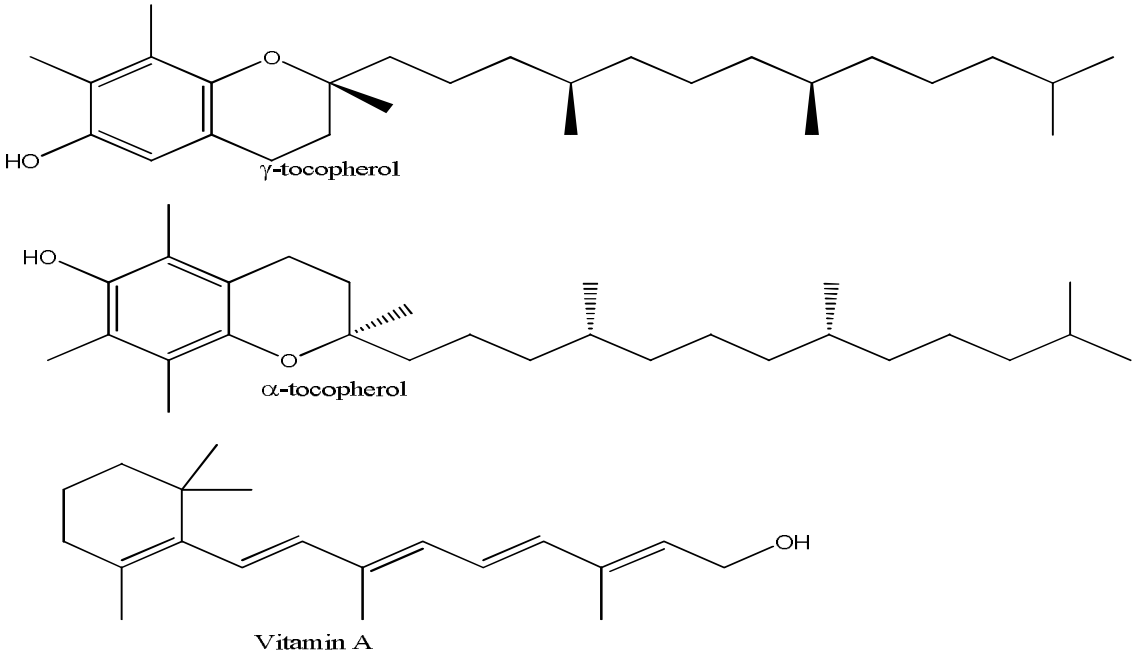


Figure 5. The two major forms of vitamin E, and vitamin A.

Cholesterol is derived about equally from the diet and biosynthesis. The synthesis of cholesterol occurs in the smooth endoplasmic reticulum. In a long reaction chain, starting with C_2-units, the C_{27}-sterol is formed.[44]

Acetate → Mevalonate → [Isoprene] → Squalene → Cholesterol
C_2 C_6 C_5 C_{30} C_{27}

Acetyl-CoA is the source of all carbon atoms in cholesterol. In particular, the acetyl group of acetyl-CoA is the ultimate precursor of all the carbon atoms in cholesterol

and in the other steroids that are derived from cholesterol.[44] There are many steps in the biosynthesis of steroids. The synthesis proceeds with the formation of acetoacetyl-CoA, 3-hydroxy-3-methylglutaryl-CoA (3-HMG-CoA) and mevalonate. Mevalonate, consisting of six carbon atoms, is formed by the condensation of three acetyl groups of acetyl-CoA. Decarboxylation of mevalonate produces the five-carbon isoprene units frequently encountered in the structure of lipids. The involvement of isoprene units is a key point in the biosynthesis of steroids and of many other compounds that have the generic name *terpenes*. Vitamins A, K, and E come from reactions involving terpenes that humans cannot carry out. That is why we must consume these vitamins in our diet. After isoprene formation, the six isoprene units condense to form squalene, which contains 30 carbon atoms. Finally squalene is converted to cholesterol, which contains 27 carbon atoms (Figure 6). This is made possible via lanosterol formation.[44]

The key enzyme of cholesterol synthesis is HMG-CoA reductase which catalyzes the formation of mevalonate from 3-HMG-CoA. Free cholesterol and glucagon inhibit the activity of HMG-CoA reductase. Insulin and thyroxin stimulate the enzyme. Decreased intracellular cholesterol concentration and pharmacologic inhibition of HMG-CoA reductase (statins) lead to an increase of low density lipoprotein (LDL)-receptors in the hepatocyte membrane.[45]

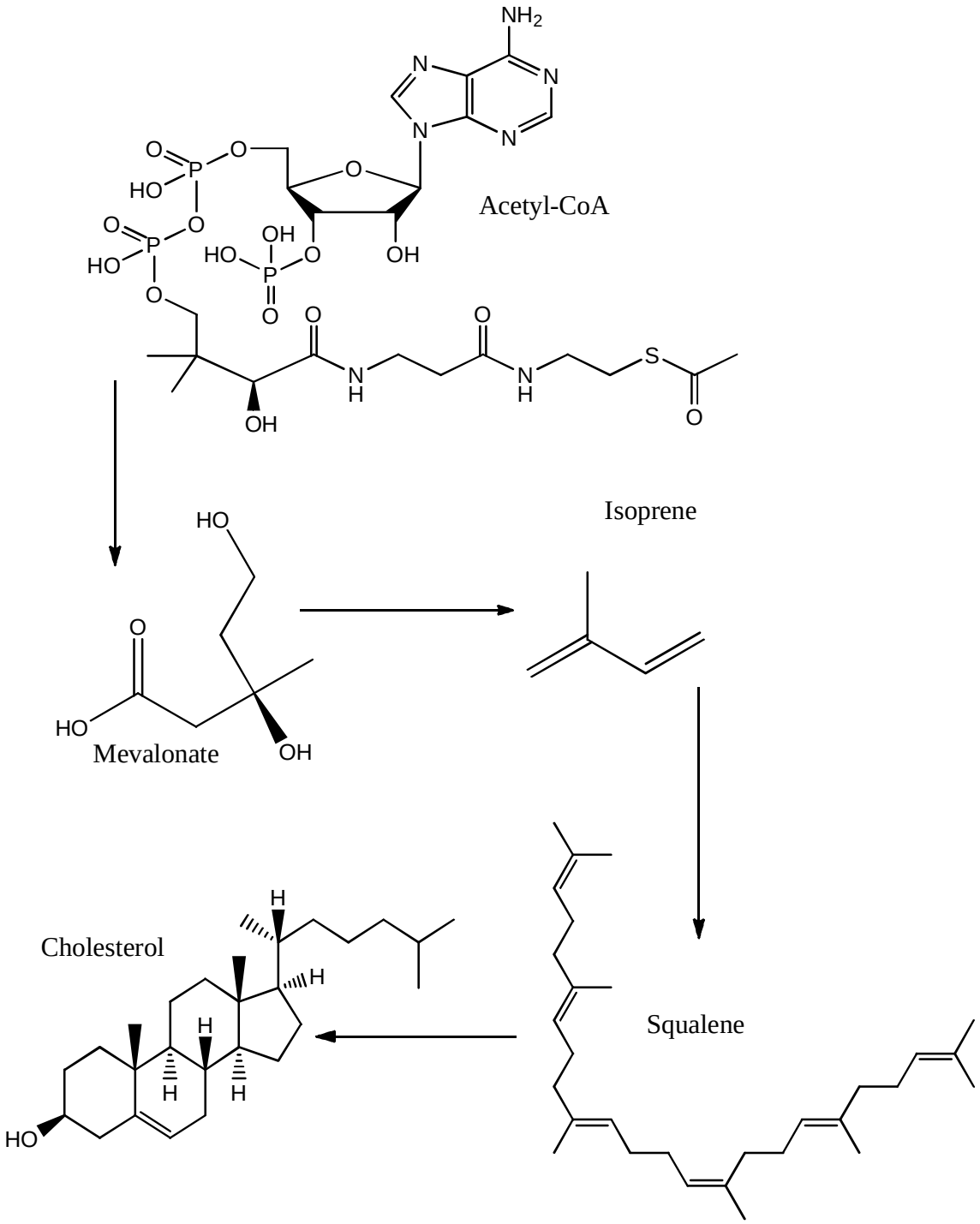


Figure 6. Outline of the biosynthesis of cholesterol.

2.2 POLYUNSATURATED FATTY ACIDS (PUFAS)

Fatty acids are linear hydrocarbon chains with a methyl (-CH_3, also called the ω-end) and a carboxyl (-COOH) end. They vary in their number of carbon atoms and double bonds and can be classified as saturated fatty acids (SFAs), monounsaturated fatty acids (MUFAs), and polyunsaturated fatty acids (PUFAs).[46]

SFAs have no double bond in their hydrocarbon chain. MUFAs on the other hand, contain a double bond in their hydrocarbon chain. They have both systematic and common names. In the case of unsaturated fatty acids, isomerism is possible. The double bond exists either in *cis* (Z) or *trans* (E) configuration, and the double bond may be located at different carbon atoms (Table 1). Natural unsaturated fatty acids are mostly of the *cis*-configuration. Traces of *trans*-unsaturated fatty acids (0.04-0.05%) are detected even in cold-pressed edible oils.[47] The position of double bonds is usually counted from the carboxyl carbon, but in texts dealing with nutrition, calculation is often from the final methyl group. In such case, they are defined, e.g., as an n-6 or ω-6, if the double bond is located at the 6^{th} carbon atom from the final methyl group. The most common MUFA is oleic acid (an ω-9 fatty acid), which is present in nearly all lipids at least in small amounts.[48]

Unsaturated fatty acids have lower melting points than the respective saturated fatty acids, and are better soluble in organic solvents. *Trans* unsaturated fatty acids have higher melting points than the respective *cis* acids (Table 1).[48]

Table 1. The most important monounsaturated fatty acids (MUFAs).[48]

Systematic name	Common name	Number of carbon atoms	Molecular weight	Melting point [°C]
9-*cis*-Hexadecenoic	Palmitoleic	16	254.4	34.5
6-*cis*-Octadecenoic	Petroselinic	18	282.5	31.5
9-*cis*-Octadecenoic	Oleic	18	282.5	16
9-*trans*-Octadecenoic	Elaidic	18	282.5	45.5
11-*trans*-Octadecenoic	Vaccenic	18	282.5	39
13-*cis*-Docosenoic	Erucic	22	338.6	33.5
13-*trans*-Docosenoic	Brassidic	22	338.6	60.6

Polyunsaturated fatty acids (PUFAs), which are the major analytes of concern in this study, contain 2-6 double bonds. Many isomers are possible, but only a few are really found in nature. The majority among them belong to essential fatty acids as they cannot be synthesized in the human body.[49] Most natural PUFAs have common names, which are more widely used than the systematic names (Table 2). The allylic (pentadienoic) configuration is most frequent, -CH=CH-CH_2-CH=CH-, where the double bonds are separated by a methylene group. Both ω-6 and ω-3 PUFAs belong to the essential fatty acids. They are enzymatically transformed into eicosanoids.[50]

Conjugated unsaturated acids, -CH=CH-CH=CH- are rather rarely found, e.g., in milk fat or on some nonedible oils. In the most common fatty acids, the first double bond is located at the 9^{th} or 6^{th} carbon atoms (from the carboxyl group) or at the 3^{rd} or 6^{th} carbon atom (from the methyl group).[48]

The structures of the most common mono- and polyunsaturated fatty acids are shown in Figure 7.

Table 2. The most important polyunsaturated fatty acids (PUFAs).[48]

Systematic name	Common name	Number of carbon atoms	Number of double bonds	Molecular weight
9-cis, 12-cis-Octadecadienoic	Linoleic	18	2	280.4
9-cis, 12-cis, 15-cis-Octadecatrienoic	Linolenic	18	3	278.4
6-cis, 9-cis, 12-cis-Octadecatrienoic	γ-Linolenic	18	3	278.4
9-trans, 11-trans, 13-trans-Octadecatrienoic	Eleostearic	18	3	278.4
5, 8, 11, 14-all-cis-Eicosatetraenoic	Arachidonic	20	4	304.6
5, 8, 11, 14, 17-all-cis-Eicosapentaenoic	EPA	20	5	302.4
4, 7, 10, 13, 16, 19-all-cis-Docoshexaenoic	DHA	22	6	328.5

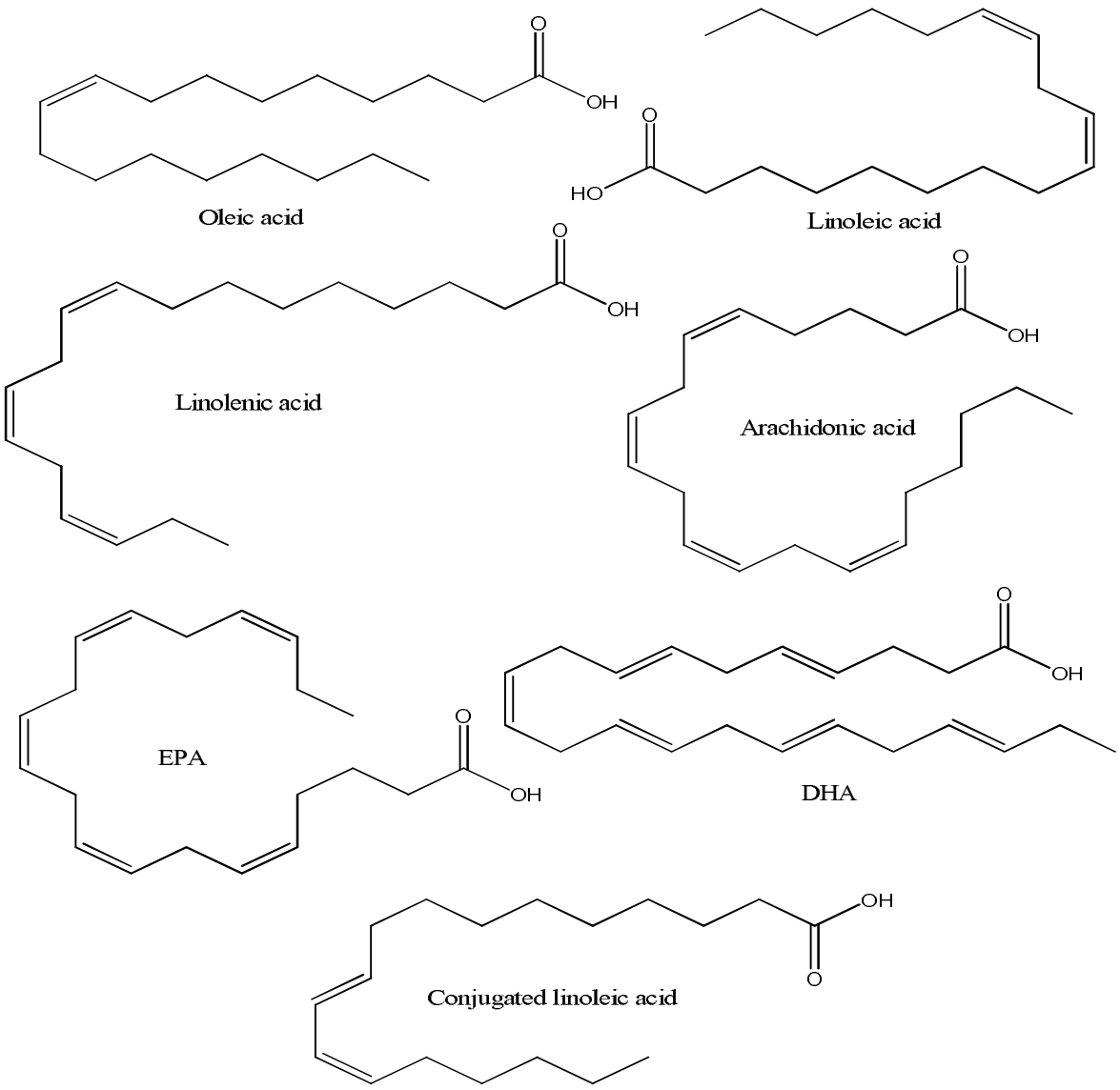


Figure 7. The structures of the mono and polyunsaturated fatty acids of interest in this study.

Polyunsaturated fatty acids (PUFAs) can be classified into major types—the ω-6 and the ω-3 PUFAs. ω-6 PUFAs have the first double bond at carbon number 6 counting from the methyl end. The major ω-6 PUFAs in the diet are linoleic acid (LA), γ-linolenic acid (GLNA), and arachidonic acid (AA). ω-3 PUFAs, on the other hand, have the first double bond at carbon number 3 counting from the methyl end. The major ω-3 PUFAs in

the diet are α-linolenic acid (ALNA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA).

2.3 THE OMEGA-6/OMEGA-3 PUFAS RATIOS AND ITS IMPORTANCE IN HEALTH RELATED ISSUES

In the Western diet, 20-25 fold more ω-6 fats than ω-3 fats are consumed.[24] It has been known that the predominance of ω-6 fat is due to the abundance of LA (18:2ω-6), which is present in high concentrations in certain vegetable oils such as soy, corn, safflower, and sunflower oils. By contrast, there is a low intake of the ω-3 homologue of LA, ALNA (18:3ω-3), which is present in leafy green vegetables and in flaxseed and canola oils. Compared with LA, there is little dietary intake of AA and EPA, which are present in meat and fish, respectively.[51]

LA and ALNA are necessary for complete health and cannot be synthesized in vertebrates; therefore, they are essential fatty acids. As a consequence, the relative dietary amounts of ω-6 and ω-3 fatty acids are determinants of the relative cellular amounts of LA and ALA (figure 8).[51]

On the other hand, unlike the 18-carbon ω-3 fatty acid ALNA, oleic acid (OA) (18:1ω-9), is consumed in substantial amounts in the typical Western diet and is not an essential fatty acid. There is little eicosatrienoic acid (ETA; 20:3ω-9) in cell membranes, however, probably this is because of the overwhelming competition from dietary LA for the relevant desaturase and elongase enzymes (Figure 8).[52]

Human beings evolved on a diet in which the ratio of ω-6/ω-3 essential fatty acids was about 1, whereas in the Western diets, the ratio is 15/1 to 16.7/1. Such evidence comes from studies on the evolutionary aspects of diet, modern day hunter-gatherers, and traditional diets (Figure 8). Many of the chronic conditions, cardiovascular disease, diabetes, cancer, obesity, autoimmune diseases, rheumatoid arthritis, asthma, and depression are associated with increased production of thromboxane A_2 (TXA_2), leukotriene B_4 (LTB_4), IL-1β, IL-6, tumor necrosis factor (TNF), and C-reactive protein. All these factors increase by increases in ω-6 fatty acid intake and decrease by increases in ω-3 fatty acid intake, either ALNA or EPA and DHA.[53]

Fatty acid designation	**ω-9**	**ω-6**	**ω-3**
18-carbon fatty acids	**Oleic acid (OA)**	**Linoleic acid (LA)** (dietary essential fatty acids	**α-linolenic acid (ALNA)** (dietary essential fatty acids
Dietary sources	Olive oil, sunola oil, meat	Soy, corn, sunflower, safflower oils	Flaxseed, canola, soy oils
Dietary intake	Large intake (8-15% dietary energy)	Large intake (7-8% dietary energy)	Minor intake (0.3-0.4% dietary energy)
	↓	↓	↓
20-carbon fatty acids	**Eicosatrienoic acid**	**Arachidonic acid (AA)**	**Eicosapentaenoic acid (EPA)**
Sources	Potentially synthesized from OA	Mainly synthesized from ingested LA, relative small amounts in diet (meat, offal)	Synthesized from ingested ALNA or ingested as fish or fish oil
Leucocyte content (% of total fatty acids)	*0.1%*	*10-16%*	*0.1-0.3%*

Figure 8. Dietary fatty acids intake and their metabolism after ingestion. Sunola oil; Meadow Lea Foods Ltd, Sydney, Australia.[51]

A study was shown to summarize the current evidence on the consumption of ω-6 PUFAs, particularly LA, and CHD risk. Aggregate data from randomized trials, case-control and cohort studies, and long-term animal feeding experiments indicate that the consumption of at least 5% to 10% of energy from ω-6 PUFAs reduces the risk of CHD relative to lower intakes. The data from the study also suggest that higher intakes appear to be safe and may be even more beneficial (as part of a low-saturated-fat, low cholesterol diet). In summary, the American Heart Association (AHA) supports an ω-6 PUFA intake of at least 5% to 10% of energy in the context of other AHA lifestyle and dietary recommendations. To reduce ω-6 PUFA intakes from their current levels would be more likely to increase than to decrease risk for CHD.[54] However, as mentioned earlier, chronic excessive production of ω-6 eicosanoids is associated with heart attacks, thrombotic stroke, arrhythmia, arthritis, osteoporosis, inflammation, mood disorders, obesity, and cancer.[55]

A high consumption of ω-6 PUFAs, which are found mainly in most types of vegetable oils, may increase the likelihood that postmenopausal women will develop breast cancer.[56] Similar effect was observed on prostate cancer.[57] ω-3 fatty acids, on the other hand, consist mostly of EPA and DHA and are found in fish and fish oils. Epidemiological studies have shown an inverse relation between the dietary consumption of fish containing EPA/DHA and CHD mortality. These relationships have been proven from blood measures of ω-3 fatty acids including DHA as a physiological biomarker for ω-3 fatty acid status. Controlled intervention trials with fish oil supplements enriched in EPA/DHA have shown their potential to reduce mortality in post-myocardial infarction patients with a substantial reduction in the risk of sudden cardiac death. The

cardioprotective effects of EPA/DHA are widespread, appear to act independently of blood cholesterol reduction, and are mediated by diverse mechanisms. Their overall effects include anti-arrhythmic, blood-triglyceride-lowering, anti-thrombotic, and endothelial relaxation among the few.[58]

2.4 THE MONOUNSATURATED OLEIC FATTY ACID

Oleic acid (OA) is known to be a pheromone released by bees when they die. This pheromone upon release initiates necrophoric behavior to other bees. Thus, when other bees detect the pheromone, the worker bees would grasp the corpse and move it a short distance towards the hive entrance. This behavior in bees serves as a hygienic procedure by keeping the dead bees away from the hive to avoid the spread of infection to the other bees.[59] Oleic acid plays a key role in plants. It is probably synthesized from short-chain fatty acids and can be dehydrogenated into more highly unsaturated fatty acids, be converted into substituted acids such as ricinoleic acid; or elongated into erucic acid by specific enzyme systems (Figure 9).[60]

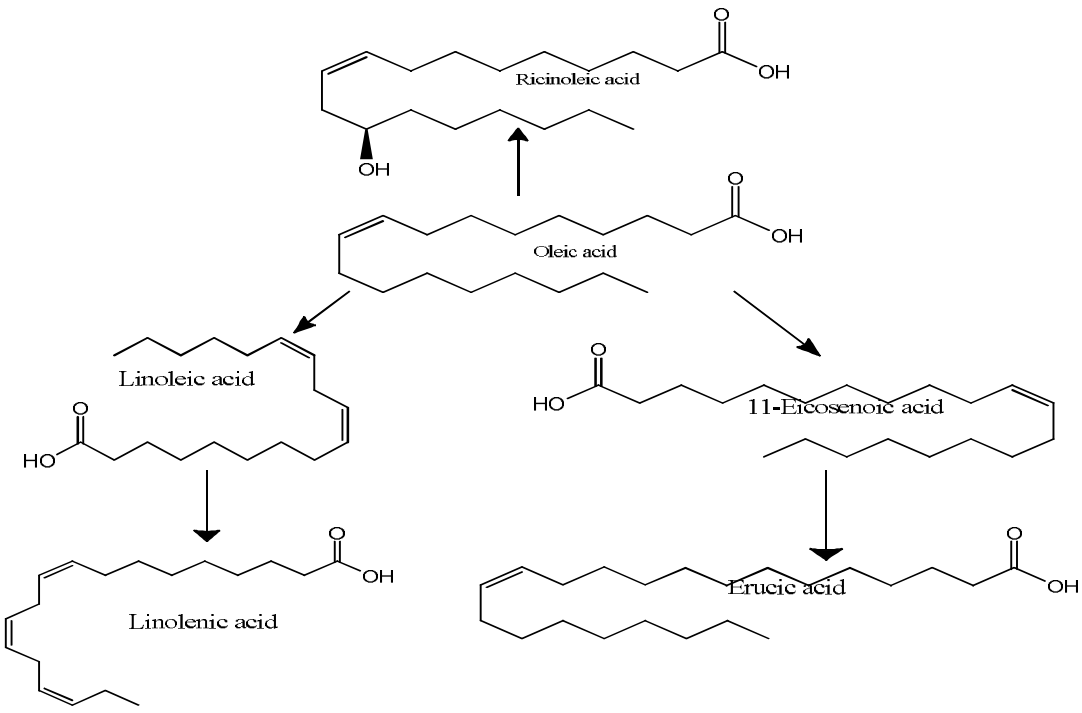


Figure 9. Unsaturated fatty acids derived from oleic acid.[60]

OA is known to be the present in various types of oils including olive oil, pecan oil (59-75%), peanut oil (36-67%), grape seed oil (15-20%), sesame oil (15-20%), and poppyseed oil (14%).[61-63] In animal fats, OA constitutes to about 37-56% in chicken and turkey.[64] In lard, OA is present at 44-47%. Moreover, the human adipose tissue consists mostly of OA.[65]

OA is known to slow down the progression of adrenoleukodystrophy (ALD), a rare, inherited disorder leading to progressive brain damage, failure of the adrenal glands and eventually death.[66] Studies by Pala and colleagues also show that OA and MUFA levels in red blood cells membranes are known to be associated with increased risk of breast cancer.[57] Moreover, OA in olive oil is also known to possess hypotensive properties.[67] Intracerebroventricular (ICV) administration of OA is known to inhibit

glucose production and food intake. This means that OA provides a signal of “nutrient abundance” to discrete areas within the central nervous system (CNS). This signal in turn activates a chain of neuronal events designed to promote a switch in fuel sources from carbohydrates to lipids and to limit the further entry of exogenous and endogenous nutrients in the circulation.[68]

CHAPTER 3

METHODS OF CHOLESTEROL AND MONO-/POLYUNSATURATED FATTY ACIDS DETERMINATION

Part of this chapter has been published in *Lipid Technology Journal* and appears in this dissertation with the journal's permission.[2]

3.1 REVIEW OF METHODS OF CHOLESTEROL AND MONO-/POLYUNSATURATED FATTY ACIDS (PUFAs) DETERMINATION IN HUMAN SERUM

The current methods for the direct determination of cholesterol and PUFAs in human serum are done independently of the other. This section divides the discussion into two parts; the first part focuses on cholesterol, and the second on PUFAs.

3.1.1 CLASSIFICATION OF METHODS FOR THE DETERMINATION OF CHOLESTEROL

Zak reviewed the different classification schemes for cholesterol determination methodologies and will be discussed in this section.[69]

3.1.1.1 DIRECT REACTIONS FOR THE DETERMINATION OF CHOLESTEROL

The methods involved in the direct reactions for cholesterol determination require no partial or full separation of cholesterol. The procedure usually involves mixing the reagents with sample and determining the endpoint by spectrophotometric methods.[70-92] It is also possible to perform electrochemical measurement of the reaction product.[93, 94] Direct reactions have the absence of phase separations. Specific examples of direct reactions include the ones based on the Liebermann-Burchard reaction such as the Wybenga methods using an Fe (III) reaction,[72] Pearson and colleagues' method,[70] or that of Huang.[73] Allain's procedure basing on a coupled enzyme system is another example of a direct reaction.[31] Direct procedures have the primary advantage of being simple.

3.1.1.2 PARTIAL PURIFICATION OF CHOLESTEROL WITH ORGANIC SOLVENTS

Liquid-solid or liquid-liquid extractions are the two methods of achieving cholesterol partial purification.[95-109] In liquid-solid extractions, the extracting fluid is soluble in water and serves to remove the cholesterols from their protein-binding sites into the liquid phase, where the measurement can be made either directly[110] or on a later solubilized residue of an evaporated extract.[111] It is also possible to dry the serum into one of several solid materials, from which it can be eluted into a separating liquid.[112-114] Biphasic liquid-liquid extractions with[115] or without[99] saponification of cholesterol esters before the extraction have been used. Saponification method is required in some

procedures because the extracting fluid will not completely separate both free and esterified cholesterol.[115]

For the automated mode, there has been at least one report in an enzymatic test[116] of dialysis of product, peroxide, where the latter could be measured in a diffusate which seemed free of interferences. Any potential competition of bilirubin with the color reaction was avoided by a preliminary treatment of the sample with alkali to oxidize it when total cholesterol was determined, because peroxidase was generated on the same side of the membrane where the bilirubin was present. However, bilirubin reaction with peroxide is minimized if no peroxidase is present to catalyze it.

Partial purification and elimination of most interferences should result in a more nearly pure end-point determination and more nearly accurate results. However, the possibility of errors increases with increasing number of steps in the reaction process. Moreover, as the complexity of the process increases, the potential for automation also increases.[69]

3.1.1.3 COMPLETE ISOLATION OF CHOLESTEROL

Studies were done isolating free cholesterol after extraction and saponification, by precipitation using different precipitating agents,[103, 117-124] and this purified derivative has then been subjected to endpoint analyses.[106, 107, 117, 125] Aluminum hydroxide[118] or aluminum chloride[119] can be used to hasten precipitation.

Chromatographic processes such as gas chromatography (GC)[124-132], liquid column chromatography (LCC)[133, 134], paper chromatography (PC), or thin layer

chromatography (TLC)[135-141] after a preliminary partial purification by extraction, can also be used to separate cholesterol or its esters.

Purifying an analyte before it is determined offers the advantage of serving as a reference procedure for the more common procedures used routinely.[69] However, high instrumentation costs and prolonged analyses times are the major disadvantages.

3.1.1.4 MISCELLANEOUS METHODS

Several other miscellaneous procedures exist for cholesterol determination. One good example is one in which cholesterol and its esters have been extracted from the protein zones known to contain them after electrophoresis of serum proteins was carried out.[142-145] Another example is that lipoproteins containing cholesterol have been selectively precipitated out with dextran sulfate before colorimetric determination by either the Liebermann-Burchard or Fe (III).[146, 147] A selective partition procedure has been described for isolating free cholesterol.[148] Any procedure devised in the future that does not fit into I, II, or III would fall in IV.

3.1.1.5 SCREENING, DEFINITIVE, AND REFERENCE PROCEDURES

A well-established method for the determination of cholesterol has not been described yet. However, at least three procedures, although unofficial, have proven worthy as reference methods[115, 125, 149] while all the rest have been applied for routine circumstances. Several studies can be categorized as screening tests[70, 72, 73, 150] since their primary purpose is to find abnormal values quickly, although perhaps not providing nearly the accuracy of the reference procedures when the latter are carefully carried out.[151] The previously mentioned screening procedures have led to an era where

cholesterol assay became relatively inexpensive. The introduction of the enzymatic procedures, which require simple handling, as compared to the complicated methods,[115, 125] led to superior precision of cholesterol routine determinations.

3.1.2 POLYUNSATURATED FATTY ACIDS (PUFAs) DETERMINATION IN HUMAN SERUM

Most polyunsaturated fatty acids (PUFAs) analyses are done for qualitative purposes only. Determination of PUFAs in human serum can be both qualitative and quantitative. This section will review the various methods of determining PUFAs both qualitatively and quantitatively.

In general, determination of PUFAs in human serum or any type of biological media can be divided into two major categories—those requiring lipid extraction and separations, and those requiring a little or none of those. Although the discussion in this section will focus on the qualitative determination of PUFAs, a particular emphasis will be on the quantitative means.

3.1.2.1 DIRECT DETERMINATION OF PUFAs IN HUMAN SERUM

The direct methods constitute those procedures which require little or no extraction and separation techniques. Most of the direct methods for the determination of PUFAs are usually done by enzymatic and spectrophotometric procedures.[152-156]

MacGee[152] described a simple and rapid enzymatic procedure for the quantitative estimation of total *cis*-methylene-interrupted polyenoic acids. Linoleic, linolenic, and arachidonic fatty acids were used to calibrate the method. In his procedure, the potassium

salts of the fatty acids are oxidized by atmospheric oxygen in the presence of the enzyme lipoxidase, and the absorption of the conjugated diene hydroperoxide is measured at 234 mμ. As little as 5 γ of linoleic acid can be quantitatively measured with good accuracy and precision. The % total content of PUFAs containing the *cis*-methylene-interrupted diene structure of fats, oils, hydrogenated oils, fatty acids, esters, blood plasma, microorganisms, and plant seeds has been measured directly by this method.[152]

A spectrophotometric micromethod for determining PUFAs was described by Herb and Riemenschneider.[153] The method is considered to be an improvement over the usual spectrophotometric methods for determining PUFAs in fats and oils. Although, there were prior improvements in spectrophotometric methods for determining PUFAs in fats and oils requiring only 100 mg of sample[155, 156], there was considered to be a great deal of interest in a micromethod that would require only a one to ten mg of fat. Herb and Riemenschneider's procedure[153] involves blanketing a 21% potassium hydroxide-glycol reagent with oxygen-free nitrogen in a reaction tube and heating in the bath at 180°C for 15 minutes. One to ten mg of sample fat was then added to the tube followed by a 30-second interval shaking until the contents appeared clear and homogenous. The total reaction time from the point wherein the sample was added to the reaction tube should be exactly 15 minutes. After rapid cooling of the sample in cold water, the isomerized mixture was diluted to known volumes with absolute methanol until suitable spectral densities were reached. Appropriate readings were made in a Model DU Beckman spectrophotometer. The specific extinction coefficient of a fat or oil at a selected wavelength is equal to the sum of the specific extinction coefficients of the components, each multiplied by its proportion in the mixture. Therefore, simultaneous equations may

be set up for the system and solved for each component with the results expressed in terms of % of linoleic, linolenic, arachidonic, and pentaenoic acids. Overall, the data obtained were reproducible results.[153]

On the other hand, Miles and colleagues[154] described a rapid and precise microfluorometric method for the determination of free fatty acid concentrations in 2-5μL of plasma. The assay is performed directly on plasma, eliminating the need for extraction with organic solvents, and is based on the quantitation of adenosine monophosphate (AMP) generated from the formation of acyl coenzyme A (acyl CoA) in the presence of adenosine triphosphate (ATP) and acyl Co-enzyme A (CoA) synthetase.[154] Although this method is considered to be rapid, accurate, and relatively inexpensive, it is only limited to total free fatty acid molar concentration quantitation.

3.1.2.2 DETERMINATION OF PUFAs IN HUMAN SERUM REQUIRING LIPID EXTRACTIONS AND SEPARATIONS

The primary disadvantage of determining PUFAs in human serum requiring lipid extractions and separations is that it requires significant amount of time to carry out the analyses. Moreover, with the rising cost of enormous instrumentations, though offering great accuracy and reproducibility, suffer from the disadvantage of not being ideal in an ordinary clinical setting. Lastly, methods requiring lipid extractions and separations require in some cases toxic solvents and also expertise in handling the sophisticated instrumentations and consequent interpretations.

Large numbers of literature citations, which cannot all be discussed in this section, were already published in many scientific publications regarding the

determination of PUFAs in human serum requiring lipid extractions and separations. Lin and McKeon[157] described various methodologies of analyzing acyl lipids using high performance liquid chromatography (HPLC). HPLC of free fatty acids (FFA) and fatty acid methyl ester (FAME) has been reviewed.[158-160] HPLC of FA was reported in the early developing stage as isocratic reversed-phase C_{18} HPLC of FAME using a refractive index (RI) detector or ultraviolet visible (UV) detector.[161] The RI detector, however, was not sufficiently sensitive. To improve detection sensitivity, FA derivatives were formed that could be detected at low concentrations by UV or fluorescence. A large number of derivatization reagents for the determination of FA by HPLC have been developed and applied to study numerous biologically important processes. For UV detection, 2-Nitrophenylhydrazine (2-NPH) is useful for the simultaneous determination of various FA. It is also possible, on the other hand to achieve a highly sensitive detection at the femtomole, or much lower levels, using proper fluorescent derivatization reagents (e.g. 4-Bromomethyl-7-methoxy-coumarin (Br-Mmc), 6,7-Dimethoxy-1-methyl-2(1H)-quinoxalinone -3-propionylcarboxylic acid hydrazide (DMEQ-hydrazide), or 2-(4-Hydrazinocarbonylphenyl)-4,5-diphenylimidazole (HCPI).[162] The separation of derivatives usually involves a complicated gradient elution system, since it is based on the differences of alkyl chain lengths of FA. Moreover, some problems involved in FA analyses that need to be overcome include increasing throughput and the development of new derivatization reagents for simultaneous determination of FA ranging from short to very long chains in a single analytical run with high sensitivity, simplicity, and rapidity.[162]

In the derivatization of FA, the bulky and somewhat polar chromophores attached to the FA molecules diminish the selectivity related to substituents of the FA and can complicate their subsequent identification by gas chromatography/mass spectrometry (GC-MS) and liquid chromatography/mass spectrometry (LC/MS). The evaporative light-scattering detector (ELSD) is universal and sensitive. It usually eliminates the need to derivatize FA above the detection limit (about 1 μg).[163] The disadvantage of ELSD is that there is a difficulty to recover FA and any other compounds after HPLC. Moreover, the relationship between mass and peak area is not linear. Some have described this relationship as sigmoidal, exponential, or "nearly linear".[164]

Different tissues often have different requirements for sample handling and preparation. Prior to fractionation of total lipids, it is necessary to extract them. Basically, total lipids are extracted following the method of Folch[165] utilizing a solvent mixture of chloroform/methanol (2:1, v/v). A lipoprotein sampling process prior to GC analysis of FAME which circumvents the lengthy and cumbersome Folch extraction method is also available.[166] Consequent fractionation of total lipids after extraction includes liquid-liquid extraction, low-pressure solid-liquid column chromatography, thin-layer chromatography (TLC), solid-phase extraction (SPE), preparative HPLC, or supercritical fluid extraction (SFE).[32]

Besides HPLC, it is also possible to analyze FA by GC although there are many difficulties in the quantitative analysis of free FA by GC.[167] Lagerstedt and colleagues[168] developed a capillary gas-chromatography-electron-capture negative-ion mass-spectrometry (GC/MS) for the quantitative determination of C8-C26 total fatty acids in

plasma. Following hydrolysis, hexane extraction, and derivatization with pentafluorobenzyl bromide, fatty acid esters are analyzed in two steps: a splitless injection and a second split injection (1:100) for the quantitation of the more abundant long-chain species.

It is also possible to perform absolute quantitation of human lipoproteins and their lipid contents directly from plasma using ^{1}H nuclear magnetic resonance spectroscopy (NMR). Though the method enables complete lipoprotein lipid profiles to be obtained in a total time of less than one hour, it is only limited in determining the absolute concentrations of triglycerides, phospholipids, total cholesterol, free cholesterol, esterified cholesterol, total proteins, and total masses for the very low density lipoprotein (VLDL) and low density lipoprotein (LDL) fractions.[169] It doesn't allow the absolute concentration determination of the individual omega-6 and omega-3 fatty acids which provides vital information for determining lipid disease states. Quantitative metabolomics was first described in 1991, as a novel approach to perform serum lipid and lipoprotein particle profiling.[170] A current protocol, now called NMR LipoProfile, utilizes a reference spectral library of different lipoprotein subclasses for the identification and quantification of 15 different lipoprotein subcategories of VLDL, LDL, and HDL particles.[171] A number of variations of this method have been described in the literature, with some methods utilizing only 11 lipoprotein categories or others using more sophisticated wavelet-deconvolution algorithms.[172, 173] Although NMR-based lipid quantitative methods are considered to be non-destructive, non-biased, easily quantifiable, and requires little or no separation, they suffer from the disadvantage that they require relatively large sample sizes (~500 µL).[174] Moreover, the instruments are expensive and the method is

inappropriate in a typical clinical setting. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry can also be used for the analysis of FA,[175] but similar to NMR, the instrument itself is very expensive and inappropriate for a typical clinical setting.

3.2 REVIEW OF METHODS FOR POLYUNSATURATED FATTY ACIDS DETERMINATION IN VEGETABLE OILS AND FOOD/BIOLOGICAL SAMPLES

Vegetable oils primarily consist of oleic, linoleic, and linolenic fatty acids. These fatty acids are, thus, chosen in this study for analyses. Table 3 lists the fatty acid composition for the triacyglycerol oils of interest. Monounsaturated fatty acids such as palmitoleic (C16:1), gadoleic (C20:1), and erucic (C22:1) exist in traces[176] and, thus, are not taken into account in this study.

Table 3. Fatty acid composition ranges for sunflower,[176] soybean,[176] safflower,[176] corn,[177] and flaxseed oils[178] (% weight composition).

Fatty Acid	Sunflower	Soybean	Safflower	Corn	Flaxseed
18:1 (oleic)	14-65	19-30	8.4-21.3	19-49	19
18:2 (linoleic)	20-75	44-62	67.8-83.2	34-52	24.1
18:3 (linolenic)	<0.7	4.0-11	0-0.1	Trace	47.4

A number of published papers exist for the determination of fatty acids in vegetable oils. Fatty acid composition has commonly been determined by GC. The most commonly used methods for such analysis involve the conversion of lipids and/or oils into methyl esters before GC analysis.[179, 180] The total time required to carry out the GC data is approximately 30 minutes.[181] With the advent of NMR spectroscopy, several

studies have also been published regarding the analysis of oils by NMR.[182-189] Moreover, it is also possible to perform quantitative fatty acid analysis of vegetable oils by gas-liquid chromatography (GLC) which is limited in expressing fatty acid composition in terms of percentages[190]; MALDI-TOF mass spectrometry which provides quantitative measurement of fatty acid mixtures and/or soap formulations that contain saturated and unsaturated hydrocarbon moieties expressing the compositions in terms of fatty acid percentages[191]; HPLC via a flame ionization detector where quantitative analysis is based on a direct proportionality of peak areas[192]; silver-ion HPLC for the quantitative analysis (% composition) of monoenoic fatty acid.[193]

For most food and biological samples, the basic procedures of extraction prior to analyses are always performed. The analyses are then carried out by the previously mentioned methods for both human serum and vegetable oil samples, usually using sophisticated instrumentations.

3.3 THE PURDIE ASSAY FOR THE DIRECT DETERMINATION OF CHOLESTEROL AND MONO-/POLYUNSATURATED FATTY ACIDS DETERMINATION

The Purdie Assay originated with the Liebermann-Burchard reaction that was once the current gold standard for cholesterol and was later based upon a reaction attributed to Chugaev and Gastev. The reason for the change was the desire to lessen the high temperature and toxic experimental conditions that present personal discomfort and health risks to the users. The reagent was a 2:1 mixture of 20% w/v $ZnCl_2$ in glacial acetic acid combined with 98% acetyl chloride[194] and was used to measure cholesterol

with great accuracy by Hanel and Dam.[195] Refinements of the reagent system has led to a 25:1 mixture of acetyl chloride and 70% perchloric acid (PA), which gave the procedure it the capability to react with steroids and terpenes. This latter so-called Purdie reagent had the extra selectivity of acylation of the α- over the ß- position at the C-17 carbon that enabled the differentiation of anabolic steroids. In that and a following study, it was also determined that if multiple unsaturated lipids are present, the resulting compounded spectrum is the simple addition of the weighted spectra for each of the components, assuming no extraneous interferences are present.[33]

CHAPTER 4

MATERIALS AND METHODS

Part of this chapter has been published in the *Journal of Biotech Research*, the *Lecture Notes in Engineering and Computer Science*, and the *Lipid Technology Journal* and appears in this dissertation with the journals' permission.[1-4]

4.1 HUMAN SERUM

4.1.1 UNKNOWN HUMAN SERUM SAMPLES

The staff and volunteers at the Hillcrest Medical Center (HMC) in Tulsa, Oklahoma provided us with the human serum samples. Such anonymously named samples were from individuals who had requested a lipid profile and had given consent. No attempt was made to solicit samples nor was any extensive medical information derived from the samples. Prior to sample collection, the subjects fasted for at least 12 hours, and using a Vacutainer^{TM} red and grey capped separation tube, venous blood samples were collected from these individuals. After inversion of the tube five times to mix the blood and the components of the collection tube, the sample was centrifuged at 3400 RPM for 15 minutes. The collection tube contained a clotting activator which takes approximately 30 minutes to activate and a floating gel that separates the red blood cells from the serum during the centrifugation step. The serum, which was the top layer in the tube, was then transferred to a 10 mL glass vial with a screw cap. The

experimental assay was completed within three days of receiving the sample. Samples were stored in a refrigerator at 2-4°C and were allowed to return to room temperature prior to analyses. HMC samples were drawn from patients with normal to elevated cholesterol levels. For serum sample analysis, a 10 µL sample of serum was added to a 13 x 100 mm borosilicate disposable test tube. Then 1 mL of 98 % acetyl chloride (AC) (Acros) was added to the test tube. A 40 µL aliquot of perchloric acid (PA) (70% ACS reagent grade, GFS) was carefully added down the inside of the test tube and slowly introduced to the AC, sample solution. The reaction started on first contact with the perchloric acid. The solution was shaken by hand for twenty seconds to allow for the release of the small amount of gas from the reaction test tube. The test tube was then covered with a Teflon cap and placed into a centrifuge and spun for 3 minutes at 3400 RPM. After centrifugation, precipitated proteins were separated, and the reagent solution was transferred to a 10 mm pathlength optical glass cuvette that was fitted with a Teflon stopper for the remaining time. Absorbance spectra were measured after 15 minutes on an HP8452A Hewlett Packard spectrophotometer. A 5-second integration time and 2-nm spectral resolutions were used to collect the absorbance data over the range of 350-550 nm. This wavelength range was chosen for the reason that the lipid analytes exhibit spectral variations in this range. For the analysis, the visible spectrum obtained for a typical plasma sample turned out to be the linear sum of the weighted contributions from all seven analytes that – given the heterogeneity of blood samples – leads to a broad diversity in the spectral patterns.[2] The blank for each reaction was pure AC. The reagent mixture of AC, with PA, did produce a slight color at 15 minutes. The combination of

AC and PA was not used as a blank, due to the possible variability and small absorbance value out of such solution mixture.

4.1.2 SYNTHETIC HUMAN SERUM SETS

Methyl esters of ω-6 fatty acids (linoleic, conjugated linoleic, arachidonic), ω-3 fatty acids (α-linolenic, eicosapentaenoic, docosahexaenoic), oleic, and free cholesterol in chloroform solutions were all used to prepare synthetic mixtures to be used as training and prediction sets. The mentioned PUFAs were used in the study because they are the most abundant lipids present in human serum.[168] The training set was done using a full factorial design (n = 128), and the prediction set was done using D-optimal design (n = 16) using the SAS-JMP Software Package.[196] Tables 4 and 5 show the training and prediction sets, respectively. All of the standards were 90 to 99 % pure based on gas chromatographic analysis and were all purchased from Sigma-Aldrich. Stock solutions for each of the analytes with maximum total concentrations of 0.02 M and 0.04 M were prepared. The stock solutions were used to prepare mixtures to limit the maximum spectral response to ranges between 0.2 and 0.9 absorbance units. The inclusion of water was taken into account in this study. Serum normally consists of 97 % water.[197] With the sample size of serum being 10 µL, approximately 9.7 µL of water was added to the reagents in cases where synthetic mixtures are analyzed. The final experimental assay involved the addition of 10 µL of distilled water as the first step, followed by 1 mL AC, 10 µL chloroform mixture sample, and finally 40 µL PA. The final steps of the assay remained the same as in serum in order to maintain constancy during the 15-minute reaction period.

Table 4. Full factorial training set of synthetic human serum for seven components.

	Cholesterol	Linoleic	Linolenic	Arachidonic	EPA	DHA	Conjugated
Mixture1	7.78E-03	0	0	0	0	0	0
Mixture2	0	1.87E-02	0	0	0	0	0
Mixture3	7.78E-03	1.87E-02	0	0	0	0	0
Mixture4	0	0	0	0	0	0	0
Mixture5	7.78E-03	0	0	0	0	0	0
Mixture6	0	1.87E-02	1.87E-02	0	0	0	0
Mixture7	6.87E-03	1.65E-02	1.65E-02	0	0	0	0
Mixture8	0	0	0	1.92E-02	0	0	0
Mixture9	6.81E-03	0	0	9.58E-03	0	0	0
Mixture10	0	1.87E-02	0	1.92E-02	0	0	0
Mixture11	6.81E-03	9.35E-03	0	9.58E-03	0	0	0
Mixture12	0	0	1.87E-02	1.92E-02	0	0	0
Mixture13	6.81E-03	0	9.37E-03	9.58E-03	0	0	0
Mixture14	0	9.35E-03	9.37E-03	9.58E-03	0	0	0
Mixture15	6.81E-03	9.35E-03	9.37E-03	9.58E-03	0	0	0
Mixture16	0	0	0	0	1.61E-02	0	0
Mixture17	6.57E-03	0	0	0	8.03E-03	0	0
Mixture18	0	1.97E-02	0	0	1.61E-02	0	0
Mixture19	6.57E-03	9.85E-03	0	0	8.03E-03	0	0
Mixture20	0	0	2.04E-02	0	1.61E-02	0	0
Mixture21	6.57E-03	0	1.02E-02	0	8.03E-03	0	0
Mixture22	0	9.85E-03	1.02E-02	0	8.03E-03	0	0
Mixture23	6.57E-03	9.85E-03	1.02E-02	0	8.03E-03	0	0
Mixture24	0	0	0	2.14E-02	1.61E-02	0	0
Mixture25	6.57E-03	0	0	1.07E-02	8.03E-03	0	0
Mixture26	0	9.85E-03	0	1.07E-02	8.03E-03	0	0
Mixture27	6.57E-03	9.85E-03	0	1.07E-02	8.03E-03	0	0
Mixture28	0	0	1.02E-02	1.07E-02	8.03E-03	0	0
Mixture29	6.57E-03	0	1.02E-02	1.07E-02	8.03E-03	0	0
Mixture30	0	9.85E-03	1.02E-02	1.07E-02	8.03E-03	0	0
Mixture31	5.89E-03	8.82E-03	9.11E-03	9.59E-03	7.19E-03	0	0
Mixture32	0	0	0	0	0	8.89E-03	0
Mixture33	6.81E-03	0	0	0	0	8.89E-03	0
Mixture34	0	1.87E-02	0	0	0	1.78E-02	0
Mixture35	6.81E-03	9.35E-03	0	0	0	8.89E-03	0
Mixture36	0	0	9.37E-03	0	0	8.89E-03	0
Mixture37	6.81E-03	0	9.37E-03	0	0	8.89E-03	0
Mixture38	0	9.35E-03	9.37E-03	0	0	8.89E-03	0
Mixture39	6.81E-03	9.35E-03	9.37E-03	0	0	8.89E-03	0

Mixture40	0	0	0	1.85E-02	0	1.78E-02	0
Mixture41	6.81E-03	0	0	9.27E-03	0	8.89E-03	0
Mixture42	0	9.35E-03	0	9.27E-03	0	8.89E-03	0
Mixture43	6.81E-03	9.35E-03	0	9.27E-03	0	8.89E-03	0
Mixture44	0	0	9.37E-03	9.27E-03	0	8.89E-03	0
Mixture45	6.81E-03	0	9.37E-03	9.27E-03	0	8.89E-03	0
Mixture46	0	9.35E-03	9.37E-03	9.27E-03	0	8.89E-03	0
Mixture47	6.10E-03	8.37E-03	8.39E-03	8.30E-03	0	7.96E-03	0
Mixture48	0	0	0	0	1.61E-02	9.10E-03	0
Mixture49	6.57E-03	0	0	0	8.03E-03	8.89E-03	0
Mixture50	0	9.85E-03	0	0	8.03E-03	8.89E-03	0
Mixture51	6.57E-03	9.85E-03	0	0	8.03E-03	8.89E-03	0
Mixture52	0	0	1.02E-02	0	8.03E-03	8.89E-03	0
Mixture53	6.57E-03	0	1.02E-02	0	8.03E-03	1.03E-02	0
Mixture54	0	9.85E-03	1.02E-02	0	8.03E-03	1.03E-02	0
Mixture55	5.89E-03	8.82E-03	9.11E-03	0	0	8.15E-03	0
Mixture56	0	0	0	9.27E-03	0	9.10E-03	0
Mixture57	6.57E-03	0	0	9.27E-03	0	9.10E-03	0
Mixture58	0	9.85E-03	0	9.27E-03	0	9.10E-03	0
Mixture59	5.89E-03	8.82E-03	0	8.30E-03	0	1.59E-02	0
Mixture60	0	0	1.02E-02	9.27E-03	0	1.78E-02	0
Mixture61	5.89E-03	0	9.11E-03	9.59E-03	0	1.59E-02	0
Mixture62	0	4.93E-03	5.09E-03	5.35E-03	4.01E-03	5.17E-03	0
Mixture63	5.63E-03	4.93E-03	5.09E-03	5.35E-03	4.01E-03	5.17E-03	0
Mixture64	0	0	0	0	0	0	8.00E-03
Mixture65	6.81E-03	0	0	0	0	0	6.00E-03
Mixture66	0	9.35E-03	0	0	0	0	6.00E-03
Mixture67	6.81E-03	9.35E-03	0	0	0	0	6.00E-03
Mixture68	0	0	9.37E-03	0	0	0	6.00E-03
Mixture69	6.81E-03	0	9.37E-03	0	0	0	6.00E-03
Mixture70	0	9.35E-03	9.37E-03	0	0	0	6.00E-03
Mixture71	5.84E-03	4.67E-03	4.68E-03	0	0	0	4.00E-03
Mixture72	0	0	0	9.58E-03	0	0	6.00E-03
Mixture73	6.81E-03	0	0	9.58E-03	0	0	6.00E-03
Mixture74	0	9.35E-03	0	9.58E-03	0	0	6.00E-03
Mixture75	5.84E-03	4.67E-03	0	4.79E-03	0	0	4.00E-03
Mixture76	0	0	9.37E-03	9.58E-03	0	0	6.00E-03
Mixture77	5.84E-03	0	4.68E-03	4.79E-03	0	0	4.00E-03
Mixture78	0	4.67E-03	4.68E-03	4.79E-03	0	0	4.00E-03
Mixture79	5.84E-03	4.67E-03	4.68E-03	4.79E-03	0	0	4.00E-03
Mixture80	0	0	0	0	0	1.03E-02	6.21E-03

Mixture81	6.57E-03	0	0	0	0	1.03E-02	6.21E-03
Mixture82	0	4.93E-03	0	0	0	5.17E-03	4.14E-03
Mixture83	5.63E-03	4.93E-03	0	0	0	5.17E-03	4.14E-03
Mixture84	0	0	5.09E-03	0	0	5.17E-03	4.14E-03
Mixture85	5.63E-03	0	5.09E-03	0	0	5.17E-03	4.14E-03
Mixture86	0	4.93E-03	5.09E-03	0	0	5.17E-03	4.14E-03
Mixture87	5.63E-03	4.93E-03	5.09E-03	0	0	5.17E-03	4.14E-03
Mixture88	0	0	0	5.35E-03	0	5.17E-03	4.14E-03
Mixture89	5.63E-03	0	0	5.35E-03	0	5.17E-03	4.14E-03
Mixture90	0	4.93E-03	0	5.35E-03	0	5.17E-03	4.14E-03
Mixture91	5.63E-03	4.93E-03	0	5.35E-03	0	5.17E-03	4.14E-03
Mixture92	0	0	5.09E-03	5.35E-03	0	5.17E-03	4.14E-03
Mixture93	5.63E-03	0	5.09E-03	5.35E-03	0	5.17E-03	4.14E-03
Mixture94	0	4.93E-03	5.09E-03	5.35E-03	0	5.17E-03	4.14E-03
Mixture95	6.12E-03	5.35E-03	5.53E-03	2.33E-03	4.19E-03	0	4.50E-03
Mixture96	0	0	0	0	0	8.89E-03	6.00E-03
Mixture97	6.81E-03	0	0	0	0	8.89E-03	6.00E-03
Mixture98	0	9.35E-03	0	0	0	8.89E-03	6.00E-03
Mixture99	0	4.67E-03	0	0	0	4.44E-03	4.00E-03
Mixture100	5.84E-03	4.67E-03	0	0	0	4.44E-03	4.00E-03
Mixture101	5.84E-03	0	4.68E-03	0	0	4.44E-03	4.00E-03
Mixture102	0	4.67E-03	4.68E-03	0	0	4.44E-03	4.00E-03
Mixture103	5.84E-03	4.67E-03	4.68E-03	0	0	4.44E-03	4.00E-03
Mixture104	0	0	0	9.27E-03	0	8.89E-03	6.00E-03
Mixture105	5.84E-03	0	0	4.64E-03	0	4.44E-03	4.00E-03
Mixture106	0	4.67E-03	0	4.64E-03	0	4.44E-03	4.00E-03
Mixture107	5.84E-03	4.67E-03	0	4.64E-03	0	4.44E-03	4.00E-03
Mixture108	0	0	4.68E-03	4.64E-03	0	4.44E-03	4.00E-03
Mixture109	5.84E-03	0	4.68E-03	4.64E-03	0	4.44E-03	4.00E-03
Mixture110	0	4.67E-03	4.68E-03	4.64E-03	0	4.44E-03	4.00E-03
Mixture111	5.84E-03	4.67E-03	4.68E-03	4.64E-03	0	4.44E-03	4.00E-03
Mixture112	0	0	0	0	4.01E-03	5.17E-03	4.14E-03
Mixture113	5.63E-03	0	0	0	4.01E-03	5.17E-03	4.14E-03
Mixture114	0	4.93E-03	0	0	4.01E-03	5.17E-03	4.14E-03
Mixture115	5.63E-03	4.93E-03	0	0	4.01E-03	5.17E-03	4.14E-03
Mixture116	0	0	5.09E-03	0	4.01E-03	5.17E-03	4.14E-03
Mixture117	5.63E-03	0	5.09E-03	0	4.01E-03	5.17E-03	4.14E-03
Mixture118	0	4.93E-03	5.09E-03	0	4.01E-03	5.17E-03	4.14E-03
Mixture119	5.63E-03	4.93E-03	5.09E-03	0	4.01E-03	5.17E-03	4.14E-03
Mixture120	0	0	0	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture121	5.63E-03	0	0	5.35E-03	4.01E-03	5.17E-03	4.14E-03

Mixture122	0	4.93E-03	0	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture123	5.63E-03	4.93E-03	0	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture124	0	0	5.09E-03	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture125	5.63E-03	0	5.09E-03	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture126	0	4.93E-03	5.09E-03	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture127	5.63E-03	4.93E-03	5.09E-03	5.35E-03	4.01E-03	5.17E-03	4.14E-03
Mixture128	2.82E-03	2.46E-03	2.54E-03	2.68E-03	2.01E-03	2.59E-03	2.07E-03

Table 5. D-optimal design prediction set for synthetic human serum with seven components.

	Cholesterol	Linoleic	Linolenic	Arachidonic	EPA	DHA	Conjugated
Mixture A	2.52E-03	9.51E-03	0	0	0	0	0
Mixture B	0	4.76E-03	4.84E-03	4.75E-03	0	0	0
Mixture C	3.29E-03	0	0	5.35E-03	4.01E-03	0	0
Mixture D	0	0	5.09E-03	5.35E-03	4.01E-03	0	0
Mixture E	0	4.76E-03	4.84E-03	0	0	5.04E-03	0
Mixture F	2.21E-03	4.76E-03	0	4.75E-03	0	5.04E-03	0
Mixture G	0	0	4.84E-03	4.75E-03	0	5.04E-03	0
Mixture H	0	0	5.09E-03	0	4.01E-03	4.44E-03	0
Mixture I	2.21E-03	0	0	0	0	0	2.83E-03
Mixture J	0	0	4.84E-03	0	0	0	2.83E-03
Mixture K	1.89E-03	2.38E-03	2.42E-03	2.38E-03	0	0	1.89E-03
Mixture L	0	0	0	0	0	5.04E-03	2.83E-03
Mixture M	1.89E-03	2.38E-03	2.42E-03	0	0	2.52E-03	1.89E-03
Mixture N	1.89E-03	0	2.42E-03	2.38E-03	0	2.52E-03	1.89E-03
Mixture O	2.82E-03	0	2.54E-03	0	2.01E-03	2.59E-03	2.07E-03
Mixture P	2.82E-03	2.46E-03	0	2.68E-03	2.01E-03	2.59E-03	2.07E-03

4.1.3 CHEMOMETRIC ANALYSES OF TRAINING, PREDICTION, AND UNKNOWN HUMAN SERUM SAMPLES

Mean centering was initially not opted by the authors in this study as a common preprocessing step for the spectroscopic data due to the issue that calibrations produced with mean-centered data can respond to small instrumentation drifts by generating large errors in predicted concentrations.[198] K-matrix OLS, NNLS, RR, PLS, and PCR calculations were done in MATLAB using Chemometrics Toolbox.[199]

Determining the optimum number of factors (rank) to be used in the

calibration is a key step in both PCR and PLS. To select the number of factors for PLS and PCR methods, the cross-validation, leaving out one sample at a time, was used. This process was repeated 127 times until each sample had been left out once. The Predicted Residual Error Sum of Squares (PRESS) was used to determine the optimum number of factors in both algorithms. To calculate the PRESS we computed the errors between the expected and predicted concentrations for all of the samples, squared them, and summed them together as given by the equation (1) below:[200]

$$PRESS = \sum_{i=1}^{N} (y_i - y_i') \qquad (1)$$

where y and y' are the predicted and actual concentrations and N is the number of samples. The logarithmic plot of the PRESS values as a function of the number of factors indicates the rank to be used in the calibration. The root mean square error (RMSE) is also calculated for each algorithm. The general equation is:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - y_i')^2}{N}} \qquad (2)$$

The model with the minimum values for the root mean square error can indicate the appropriate model. In the initial work, all PLS calculations refer to PLS2 after an initial comparison showed that PLS2 yielded lower root mean square error of prediction (RMSEP) than PLS1. However, an attempt was also made to compare various chemometric algorithms (PLS vs PCR) by mean centering the training absorbance sets. Mean centering was also attempted in KM, RR, and PM but generated root mean square errors of predictions (RMSEPs) which are much larger than the non-mean centered

training data sets. Accordingly, the training data sets were not mean-centered in KM, RR, and PM regression models.

The various chemometric models can be referred to some references[201-207] and will be discussed further on the next sections.

4.1.4 CHEMOMETRIC ALGORITHMS USED FOR DATA ANALYSES

This section will describe the various chemometric techniques used for the data analyses. Chemometric techniques can be divided into two major approaches: the multiple linear regression (MLR) approach which includes the K-matrix (KM) and the P-matrix (PM) approaches and the soft modeling methods consisting of principal component regression (PCR) and partial least squares (PLS). For the MLR approach, several variants under such category consisting of the non-negative least squares (NNLS) and the ridge regression (RR) techniques applied to KM will also be discussed.

4.1.4.1 MULTIPLE LINEAR REGRESSION (MLR) APPROACH

KM and PM approaches were utilized for the MLR models. The KM was expressed in the matrix notation:

$$A = CK \qquad (3)$$

where A is the $n \times p$ matrix of absorbances, C is the $n \times m$ matrix of concentrations of constituents, K is the $m \times p$ matrix of absorptivities, n is the number of samples, p is the number of wavelengths, and m is the number of components.

Calibration is based on a set of *n* samples of known concentrations for which the spectra are measured. By means of the calibration sample set, estimation of absorptivities is possible by solving for the matrix *K* according to the general least squares solution:

$$K = (C^T C)^{-1} C^T A \quad (4)$$

The analysis was based on the spectrum a_0 (1 x *p*) of the unknown sample by use of:

$$c_0 = a_0 K^T (K K^T)^{-1} \quad (5)$$

where c_0 is the (1 x *m*) vector of sought-for concentrations.[202]

The P-matrix notation, on the other hand, was represented as:

$$C = AP \quad (6)$$

The calibration coefficients are now the elements of the P-matrix that are estimated by the generalized least squares solution according to:

$$P = (A^T A)^{-1} A^T C \quad (7)$$

Analyses were carried out by direct multiplication of the measured sample spectrum a_0 by the P-Matrix:[202]

$$c_0 = a_0 P \quad (8)$$

Equations (5) and (8) were solved following a quadratic optimization problem with *m* linear constraints according to:

$$\sum_{j=1}^{p} w_j [a_{oj} - \sum_{i=1}^{m} K_{ij} c_{oi}]^2 \rightarrow Minimum \qquad (9)$$

subject to $c_0 \geq 0$, i = 1,…,m and j = 1, …,p, where the weights w_j are usually chosen according to the estimated (proportional to the reciprocal) size of the errors $r(\lambda_j)$ based on

$$a_o(\lambda) = \sum_{i=1}^{m} K_i(\lambda) c_{oi} + r(\lambda) \qquad (10)$$

This is known as NNLS algorithm. The non-negative MLR approach (KM and PM) was used to quantitate the lipid analytes. The advantage of this method over the ordinary least squares estimation is the reliability of the computed amounts of constituents at low concentrations.[203]

Ridge regression (RR), on the other hand, is a variant of ordinary MLR whose goal is to circumvent the problem of independent variables collinearity.[204] The regression coefficients in this algorithm procedure were obtained from:

$$c_0 = a_0 K^T (KK^T + kI)^{-1} \qquad (11)$$

where k is a positive number (usually $0 < k < 1$) and I is the $m \times m$ identity matrix. Comparison of this expression with equation (5) reveals that a constant is added to the diagonal elements of the KK^T matrix of the normal equations. With k = 0, the least squares solution is obtained as in equation (5).[205] Hoerl and Kennard[206] suggested selecting a value of k by an examination of a ridge trace, which is a plot of the regression coefficients for different values of the ridge parameter. The value of k was chosen at a point where the regression coefficients should start to stabilize and that the root mean square error of prediction (RMSEP) should decrease.

4.1.4.2 SOFT MODELING METHODS: PRINCIPAL COMPONENT REGRESSION (PCR) AND PARTIAL LEAST SQUARES (PLS)

The methods of soft modeling are based on the inverse calibration model (i.e. PM approach) where concentrations are regressed on spectral data:

$$C = AB \tag{12}$$

C and *A* are again the *n x m* concentration and *n x p* absorbance matrices, respectively, and *B* is the *p x m* matrix of regression or *B*-coefficients.

PCR is best performed by means of SVD (singular value decomposition). This method involves the decomposition of the absorbance matrix *A* into two orthogonal matrices *U* and *V* joined by a diagonal matrix *W* of singular values:

$$A = UWV^T \tag{13}$$

Estimation of the matrix of regression coefficients *B* is performed column-wise by use of:

$$b = A^+c \tag{14}$$

with A^+ being the pseudo-inverse of the absorbance matrix *A*.[202]

Details of the PLS method can be referred to Otto.[202] It involves the decomposition of *A* and *C* according to:

$$A = TP^T + E \tag{15}$$

$$C = UQ^T + F \tag{16}$$

where *T* and *U* are the *n x d* scores matrices containing orthogonal rows; *P* are the *p* x *d* loadings of the *A* matrix; *E* is the *n x p* error (residual) matrix of *A* matrix; *Q* is the *m x d* loading matrix of the *C* matrix; and *F* is the *n x m* error (residual) matrix for the *C* matrix.

Computation of the *B*-coefficients for the general model gives:

$$B = W (P^T W)^{-1} Q^T \qquad (17)$$

with *W* as *d x p* matrix of PLS-weights.

All chemometric calculations were done in MATLAB using *Chemometrics Toolbox.*[199]

4.1.4.3 NEURAL NETWORKS

Neural network is divided into three layers comprised of input, hidden layers, and output. The input parameters are the absorbance at specified wavelengths. The parameters are connected to neurons in the hidden layer. The number of hidden layers and the number of neurons in each layer is flexible and is determined by the examination of errors in the results in the output layer in terms of concentrations. A basic network design is show in Fig. 10. More detailed information about neural network can be found in Hagan et al.[207] Neural network was performed using the *JMP Software Package*.[208]

Fig. 10. Example of a basic neural network design.

4.1.4.4 GENETIC ALGORITHM PARTIAL LEAST SQUARES (GAPLS)

PLS is known to treat very large data matrices, extracting the relevant part of the information and producing reliable but very complex models.[209] Previously, PLS was considered to be almost insensitive to noise, and therefore it was commonly stated that no feature selection at all was required.[210] However, over two decades ago, it has been widely recognized that a feature selection can be highly beneficial since a double goal can be achieved: improve the predictive ability of the model and highly simplify it.[211]

Genetic algorithms (GA) have been shown to be successful in selecting the most important features.[212-223] The first applications were reported around 1960 when Holland introduced the method.[224] It is, however, only since the 1980s that the number of publications increased exponentially, mainly due to advances in the computer technology.[225]

GA's ultimate goal is the optimization of a given response function. These algorithms are inspired by the theory of evolution: in a living environment, the "best" individuals have a greater chance to survive and a greater probability to spread their genomes by reproduction. The mating of two "good" individuals causes the mixing of their genomes, which may result in a "better" offspring. The terms "good", "best", and "better" are related to the fitness of the individuals to their environment.[215]

The GA consists of five steps which include creation of the initial population, fitness evaluation, the stop condition check, crossover, and mutation. These steps are explained in short details below. In GA, a chromosome is defined as a bit vector where every gene (bit) equals inclusion (1) or exclusion (0) of the spectral data for a certain

number of wavelengths defined by the *window size* in the model. The fitness evaluation to mutation steps are repeated until the stop condition is achieved.[226]

In the initial step, a population of random chromosomes is created in accordance to given probabilities for each gene to be active as defined by the initial probability-variable. Thereafter, fitness is measured as the inverse of the root mean square error of cross validation (RMSECV) value for the model based on the chromosome in question.[226]

The GA will stop when the amount of chromosomes defined by the convergence stop criterion is equal or when the maximum allowed number of generations, defined by the maximum number of generations-variable, is reached. Following the stop condition check, crossover follows wherein first as many of the fittest chromosomes, as defined by the elitism value, will be copied to the child generation. Consequently, the rest of the children will be generated by repeated crossover of two parent chromosomes selected randomly so that a chromosome with a high fitness value have a higher probability to be selected than one with a low fitness value. Single-point or uniform crossover can be used. The crossover probability defines the probability that each gene will be switched between the two parent chromosomes in the case of uniform crossover.[226]

After crossover, mutation follows. Two forms of the mutation operator called "gene" and "chromosome" were implemented. In the first one, mutation probability defines the probability for each gene to switch state in all of the child chromosomes, except those selected by elitism. The mutation probability variable gives in the second one the probability to switch, from a randomly chosen point, the state of the next 25% of all genes in the chromosome. Consequently, the switching stops if the end of the

chromosome is reached.[226]

The performance of the model is gauged by measuring the RMSEP and RMSECV using equation (2).[226] PLS1 regression method was used in GA method. GAPLS was done in MATLAB using the partial least squares-genetic algorithm (PLS-GA) Toolbox by Leardi.[227, 228] Appendix I shows the default parameters used for such algorithm. GAPLS was run three times for each trial and the average results were determined for the three trials.

4.1.5 GAS-CHROMATOGRAPHIC (GC-MS) QUANTITATION OF SERUM SAMPLES

Validation was done by quantitating the same serum samples using GC-MS detection. Blood serum was esterified using the method given by Guy Lepage and C. Roy.[229] A 1 µl of the upper benzene phase of the esterified serum was chromatographed as methyl esters on a 30-m fused silica column with an internal diameter of 0.320 mm. The column was wall-coated with 0.25 mm DB-23. Analysis was performed on a Shimadzu (GCMS-QP2010) gas chromatograph. Helium was used as the carrier gas. The injection temperature was held at 250°C, and the column oven temperature of 50°C. Splitless injection mode was used, and the oven temperature program was held for 2.0 minutes at 50°C and then raised 180°C at 10°C/min. After a 5.0 minute hold, the temperature was raised to 240°C at a rate of 5.0°C/min and held for 13 minutes. Peaks were identified by the use of pure reference compounds. Six PUFAs from 18 to 22-carbons were identified.

4.1.6 SPIKING OF HUMAN SERUM SAMPLES AND SYNTHETIC SETS

An attempt was also made of spiking the human serum samples with standard solutions of the individual fatty acids. Table 6 shows the molar concentrations of the fatty acids used to build the calibration model for the spiking of human serum samples. The concentration matrix was designed so as to obtain a spectral absorbance increment response between 0.1 and 1.2 absorbance units. The incremental absorbance responses for the additions are plotted with concentrations at different wavelengths, and the resulting slope is used as the molar extinction coefficient to determine the unknown serum molar concentrations. NNLS was used to determine the concentrations in the human serum samples. Spiking was also performed in a 7-component mixture (Mix 128, Table 4), and the RMSEP was determined according to equation (2). The concentrations of the samples for the spiking of the synthetic set were determined using NNLS, OLS, and RR as described in 4.1.4.1.

Table 6. Molar concentration matrix used to build a calibration model for the spiking of human serum model.

	Cholesterol	LA	LNA	AA	EPA	DHA	CLA
addition1	1.89E-05	1.89E-05	1.94E-05	1.82E-05	1.87E-05	1.94E-05	1.90E-05
addition2	3.74E-05	3.75E-05	3.83E-05	3.61E-05	3.71E-05	3.85E-05	3.76E-05
addition3	5.56E-05	5.57E-05	5.7E-05	5.36E-05	5.52E-05	5.73E-05	5.59E-05
addition4	7.34E-05	7.36E-05	7.53E-05	7.08E-05	7.29E-05	7.56E-05	7.39E-05
addition5	9.09E-05	9.12E-05	9.32E-05	8.77E-05	9.03E-05	9.37E-05	9.15E-05

4.2 VEGETABLE OILS

4.2.1 VEGETABLE OIL TRAINING, PREDICTION, AND UNKNOWN SETS

Oleic, linoleic, and linoleic fatty acid methyl esters should be as close as possible in concentration ranges of 0.0025 to 0.02 M in chloroform solutions in order to maintain the absorbance units from 0.1 to 1.2. It should be noted that the linolenic acid discussed

throughout in this paper refers to the alpha form. Tables 7 and 8 show the actual molar concentrations of the training and prediction set standards prepared. These sets were also used as training and prediction sets in human serum analysis adding oleic as the eighth component. Olive and sunflower oil samples were used as unknown sets. All samples were obtained from Sigma-Aldrich.

Table 7. Oleic, linoleic, and linolenic fatty acid methyl esters (FAME) training matrix by central composite and simplex lattice designs in chloroform solutions.

	Oleic	Linoleic	Linolenic
Mixture 1	8.31E-03	2.42E-03	1.92E-02
Mixture 2	1.91E-02	1.90E-02	2.33E-03
Mixture 3	8.31E-03	1.90E-02	2.33E-03
Mixture 4	8.31E-03	2.42E-03	2.33E-03
Mixture 5	1.91E-02	2.42E-03	2.33E-03
Mixture 6	1.72E-02	1.71E-02	1.74E-02
Mixture 7	8.31E-03	2.42E-03	1.92E-02
Mixture 8	8.31E-03	1.90E-02	1.92E-02
Mixture 9	1.37E-02	1.07E-02	2.33E-03
Mixture 10	1.37E-02	1.07E-02	1.92E-02
Mixture 11	1.37E-02	1.90E-02	1.08E-02
Mixture 12	1.37E-02	2.42E-03	1.08E-02
Mixture 13	1.91E-02	1.07E-02	1.08E-02
Mixture 14	8.31E-03	1.07E-02	1.08E-02
Mixture 15	1.37E-02	1.07E-02	1.08E-02
Mixture 16	1.37E-02	1.07E-02	1.08E-02
Mixture 17	1.37E-02	1.07E-02	1.08E-02
Mixture 18	1.37E-02	1.07E-02	1.08E-02
Mixture 19	1.37E-02	1.07E-02	1.08E-02
Mixture 20	1.37E-02	1.07E-02	1.08E-02
Mixture 21	1.43E-02	2.42E-03	2.33E-03
Mixture 22	9.55E-03	1.90E-02	2.33E-03
Mixture 23	9.55E-03	2.42E-03	1.92E-02
Mixture 24	1.24E-02	6.67E-03	2.33E-03
Mixture 25	8.31E-03	1.24E-02	2.33E-03
Mixture 26	1.24E-02	2.42E-03	6.65E-03
Mixture 27	8.31E-03	2.42E-03	1.24E-02
Mixture 28	8.31E-03	1.24E-02	6.65E-03
Mixture 29	8.31E-03	6.67E-03	1.24E-02
Mixture 30	8.31E-03	6.67E-03	2.33E-03

Table 8. Oleic, linoleic, and linolenic fatty acid methyl esters prediction matrix by an independent simplex lattice design in chloroform solutions.

	Oleic	Linoleic	Linolenic
Mixture 1	1.24E-02	3.45E-03	3.50E-03
Mixture 2	8.42E-03	1.73E-02	3.50E-03
Mixture 3	8.42E-03	3.45E-03	1.75E-02
Mixture 4	1.40E-02	7.48E-03	3.50E-03
Mixture 5	1.01E-02	1.44E-02	3.50E-03
Mixture 6	1.40E-02	3.45E-03	8.16E-03
Mixture 7	1.01E-02	3.45E-03	1.34E-02
Mixture 8	1.01E-02	1.44E-02	8.16E-03
Mixture 9	1.01E-02	7.48E-03	1.34E-02
Mixture 10	1.01E-02	7.48E-03	3.50E-03

4.2.2 ANALYSIS OF FAME STANDARDS AND VEGETABLE OILS USING THE ASSAY AND VALIDATION

The procedure for the acetyl chloride/perchloric acid (AC/PA) color assay reaction is conceptually simple. It entails placing a 10 microL aliquot of the fatty acid methyl esters (FAME) standards or vegetable oils into a 13 x 100 mm borosilicate disposable test tube, followed immediately by the careful addition of 1.0 mL AC then 40 microL of PA. The test tube is sealed tightly with parafilm and gently shaken for 20 seconds. The supernate is then transferred by pipette to a 10 mm pathlength optical glass cuvette and placed in the sample holder of a diode-array spectrophotometer (HP8452A). Analysis is done after 15 minutes from 350-550 nm at every 2 nm and 5 s integration time. Using the developed calibration matrix, the FAME standard concentrations in prediction sets and vegetable oils were determined using the chemometric techniques. The obtained chemometric molar concentrations were then validated with the existing USDA database concentrations.[37]

4.2.3 CHEMOMETRIC ANALYSES OF VEGETABLE OIL TRAINING, PREDICTION, AND UNKNOWN SETS

The training, prediction, and unknown set spectra were deconvoluted using various chemometric algorithms. The KM, NN, PCR, and PLS algorithms were utilized in this study. Mean centering was performed prior to the chemometric analyses. Chemometric analyses were performed in MATLAB using *Chemometric Toolbox*.[199] Neural network was performed using the *JMP Software Package*.[208]

Determining the number of factors (rank) to be used in the calibration is a key step in both PCR and PLS. To select the number of factors for PLS and PCR methods, the cross validation, leaving out one sample at a time, was used. This process was repeated 29 times, until each sample had been left out once. The PRESS was used to determine the optimum number of factors in both algorithms as given by equation (1) on page 44.[199]

The plot of the PRESS values as a function of the number of factors indicates the rank to be used in the calibration. The RMSE is also calculated for each algorithm. The general equation is given by equation (2) on page 44. The model with the minimum values for the RMSE indicated the appropriate model.

4.3 FOOD AND BIOLOGICAL SAMPLES

Various food and biological samples were treated with the Purdie assay with the same procedures as mentioned in 4.2.2 and pattern recognition was used to determine whether the assay can discriminate samples according to their types. These said samples include fat free milk, chicken liver, chicken, Braum's chocolate whole milk, salmon,

veal, ham, Heart Healthy Butter, egg yolk regular and free range, beef, ω-3 enriched oil, salmon in can, chicken dried, salmon dried for longer period of time, and turkey. All samples were purchased from Walmart. The solid samples were measured in uniform masses of 0.025 grams so as not to exceed 1.2 absorbance units. Liquid samples were also scaled so as not to exceed the sample absorbance of 1.2 absorbance units.

The pattern recognition technique used in this study, principal component analysis (PCA), is an unsupervised multivariate statistical method useful for reducing multidimensional data down to 2 or 3 dimensions that can readily be comprehended. The graphical representations presented utilize the first 2 or 3 principal components as the axes. Using PCA, the resulting principal components were plotted versus each other to produce 2- and 3-dimensional representations of the data to determine if any clustering patterns were separable. If a pattern was seen, then hierarchical cluster analysis was used to group together the data points using the *JMP* software package.[208] In cluster analysis, the process will start with one piece of data and combines groups based on distances from one another in the principal component space.[202] The cluster analysis in this study was agglomerative hierarchical with Ward's method being used for the distances.

CHAPTER 5

RESULTS AND DISCUSSIONS

Part of this chapter has been published in the *Journal of Biotech Research*, the *Lecture Notes in Engineering and Computer Science*, and the *Lipid Technology Journal* and appears in this dissertation with the journals' permission.[1-4, 230]

5.1 HUMAN SERUM SAMPLES

5.1.1 CHEMOMETRIC ALGORITHMS APPLIED TO HUMAN SERUM

Figure 11 shows the comparison of the root mean square error of prediction (RMSEP) for the seven different algorithms in each lipid analyte as calculated according to equation (2). It is very clear that the non-negative least squares (NNLS) algorithm when applied to K-matrix yielded lower RMSEPs than their ordinary regression approaches. It is well known that the introduction of physically induced constraints reduces the error amplification factor of so-called incorrectly posed problems (highly sensitive to measurement errors), sometimes by an order of magnitude.[231, 232] Gayle and Bennet[233] showed examples demonstrating the advantages of NNLS. Jochum and Schrott[234] also showed in their study the striking advantage of NNLS and its reliability of the computed amounts of constituents at low concentrations. The large RMSEPs contributed by the K-matrix approach is very evident in Figure 11. Though this approach

offers the advantage of representing the genuine absorptivities with reference to the spectra of the individual constituents as shown in Figure 12, it does, however, have the disadvantage in that the calibration and analysis are connected to the inversion of a matrix.[201] Although this is not a problem from the point of view of computational time, it might become a problem if ill-conditioned (less selective) systems are applied, where the spectra of the constituents are very similar.[202] In Figure 12, it is evident that several constituents have similar spectra save for cholesterol and conjugated linoleic. The RR technique was attempted in order to improve the prediction errors in such cases. The ridge parameter obtained by using the value of k = 5.00E-7, which is the value taken from the plot of the standardized ridge coefficients vs. ridge parameter (Figure 13) resulted in improved results in the RMSEPs over the ordinary K-matrix least squares solution as shown in Figure 11. In a comparative simulation study by Frank and Friedman,[235] it was shown that often the RR performs as well as PCR or PLS, all of them outperforming multiple linear regression (MLR) with forward variable selection.

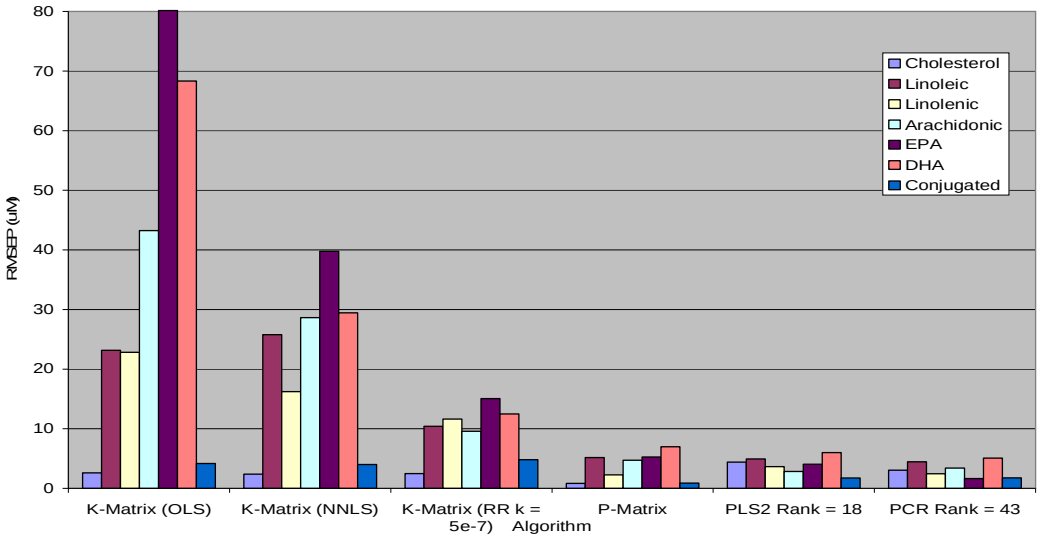


Figure 11. RMSEP comparison for each algorithm comparing the different lipid analytes.

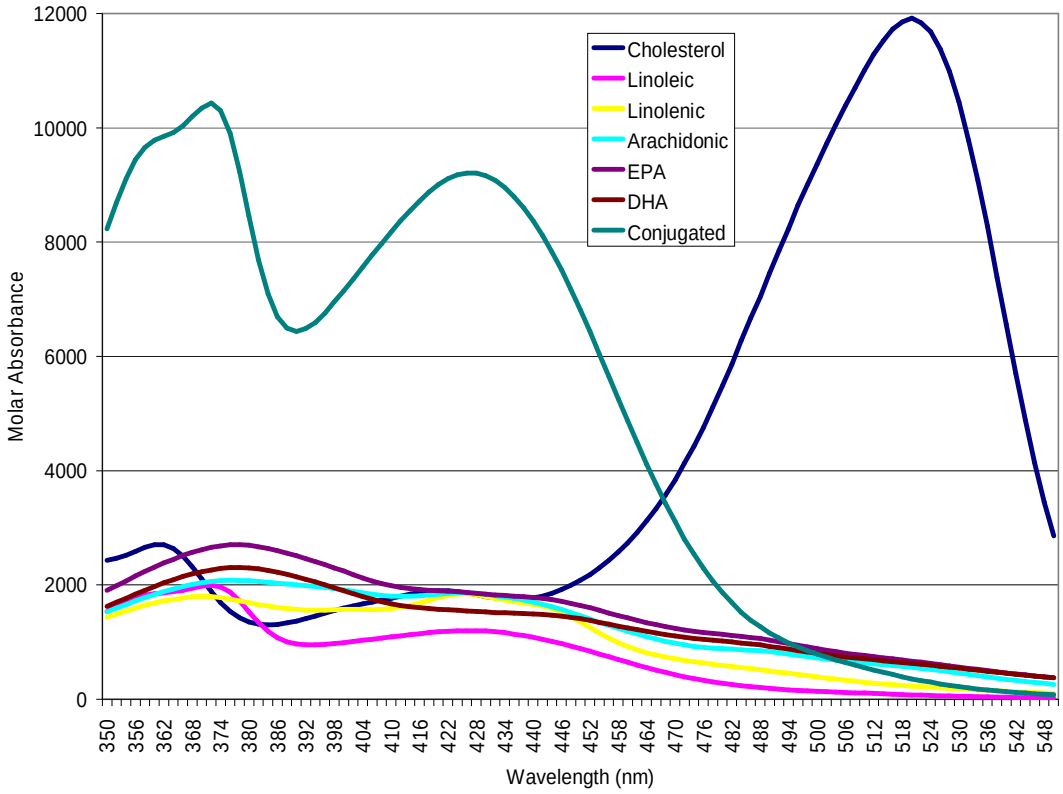


Figure 12. Molar absorptivities of the seven lipid analytes determined by the K-matrix model.

An alternative to the K-matrix approach is to calibrate the concentrations directly on the spectra. This is known as the P-matrix approach (or inverse model). A disadvantage of this calibration method is that the calibration coefficients (elements of the P-matrix) have no physical meaning, since they reflect the spectra of the individual components. Figure 14 shows the P-matrix regression coefficients obtained from the seven lipid analytes. It is evident in Figure 11 that the P-matrix technique also yielded comparable results with that of PCR and PLS. The P-matrix approach offers a slight advantage over the classical K-matrix approach because a second matrix inversion is avoided.[202] One of the assumptions made in MLR is that the independent variables are truly independent. To the degree that this assumption is invalid, the resulting model parameters will be more affected by noise, eventually leading to loss of full rank.[236]

Attempts to eliminate this collinearity problem have led to such developments as PCR and RR. Among all algorithms attempted, P-matrix, PCR, and PLS performed quite equally well, exhibiting low RMSEP values. The number of factors in PCR in Figure 11 might be high enough, but this number of factors was determined to be the optimum number after cross validation calculations.

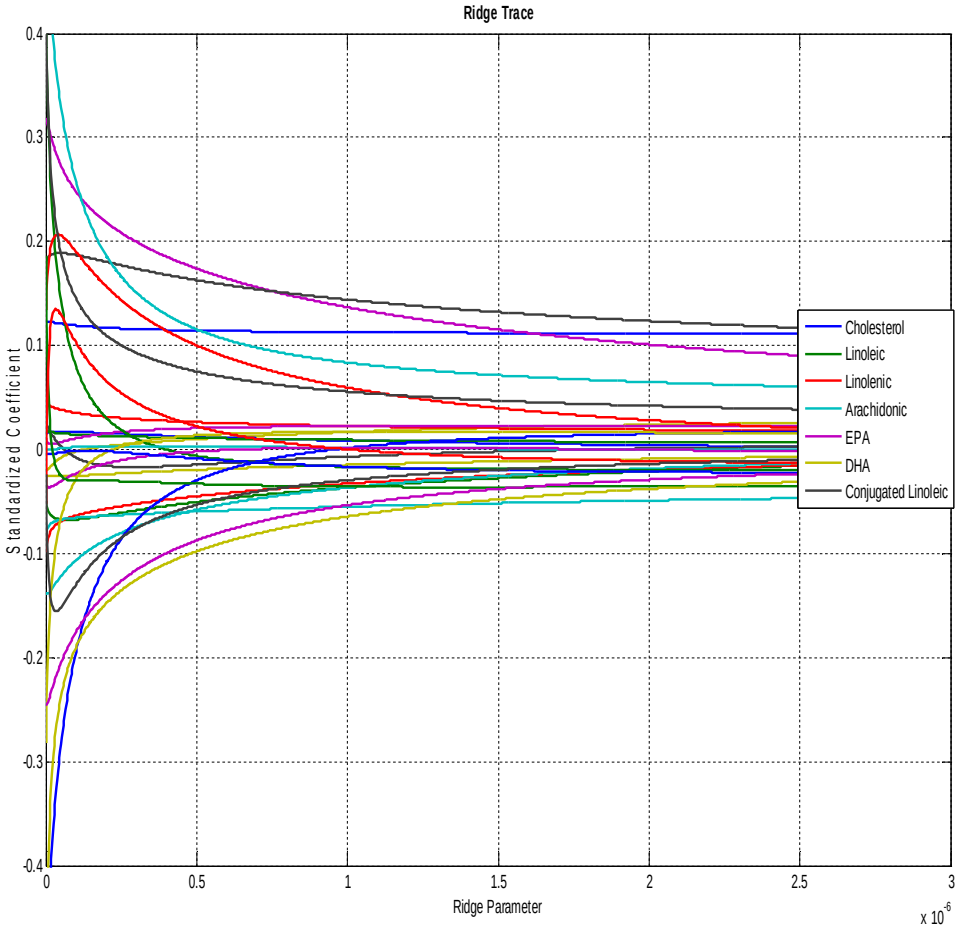


Figure 13. Plot of standardized coefficient vs ridge parameter for the RR approach. The point at which the ridge parameter, k = 5.00E-7, represents the optimum parameter value leading to lowest RMSEPs.

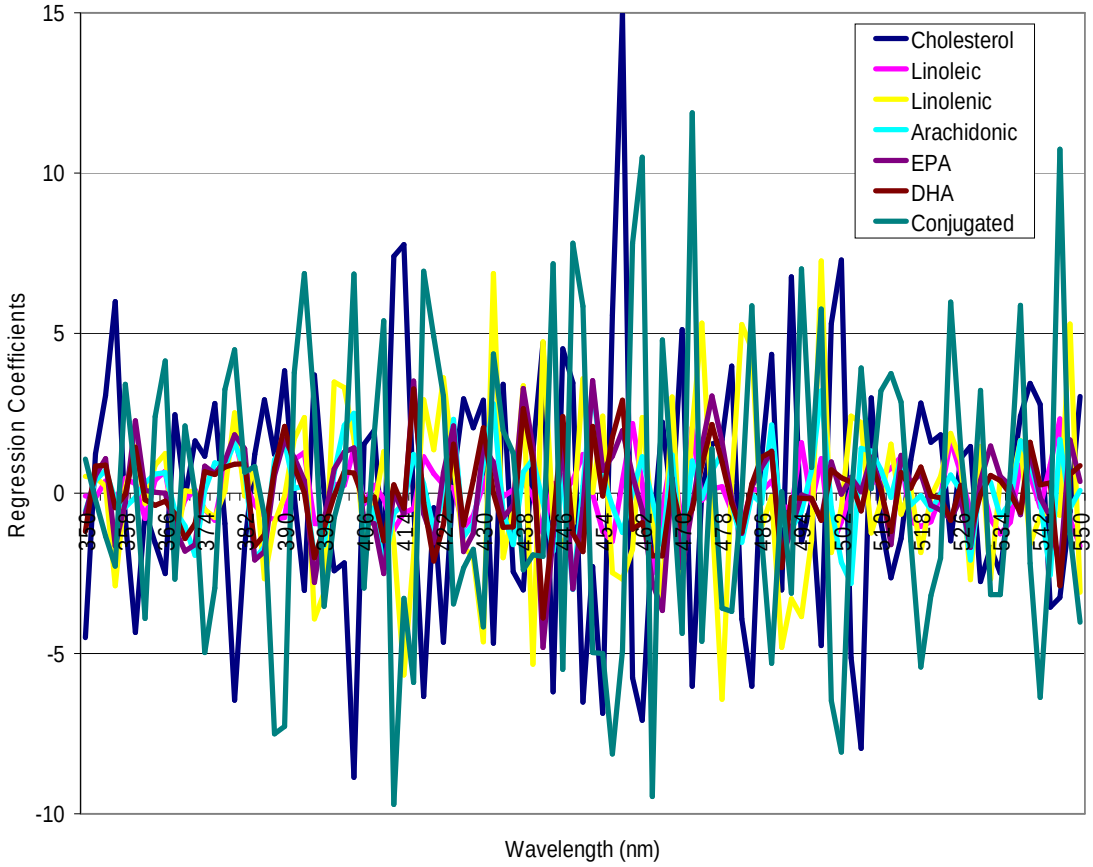


Figure 14. P-matrix regression coefficients obtained by using the P-matrix approach.

The PLS is a powerful multivariate statistical tool that has been successfully applied to the quantitative analysis with ultraviolet,[237, 238] near-infrared,[239-242] chromatographic,[169, 242, 243] and electrochemical data.[244] The PLS offers the signal-to-noise advantage gained by making use of all the measurements. Furthermore, by using only the significant number of latent variables in the procedure, a noise filtering effect is obtained which results in its improved predictive stability.[245] In a study conducted by Cassel , et al[246] wherein PLS was tested in the presence of three inadequacies, namely (i) skew instead of symmetric distributions for manifest variables, (ii) multi-collinearity within blocks of manifest and between latent variables, and (iii) misspecification of the

structural model (omission of regressors), the algorithm showed quite a robustness in these three inadequacies.

An alternative diagram showing how the different constituents behave with respect to the different algorithms is shown by their RMSEPs in Figure 15. Clearly from Figure 15, cholesterol and conjugated linoleic yielded the lowest RMSEPs. As mentioned earlier, these analytes have distinctive characteristic spectra that could be clearly distinguished from the other lipid analytes (Figure 11). EPA, DHA, linolenic, and arachidonic exhibit similar spectra and, thus, yielded high RMSEPs.

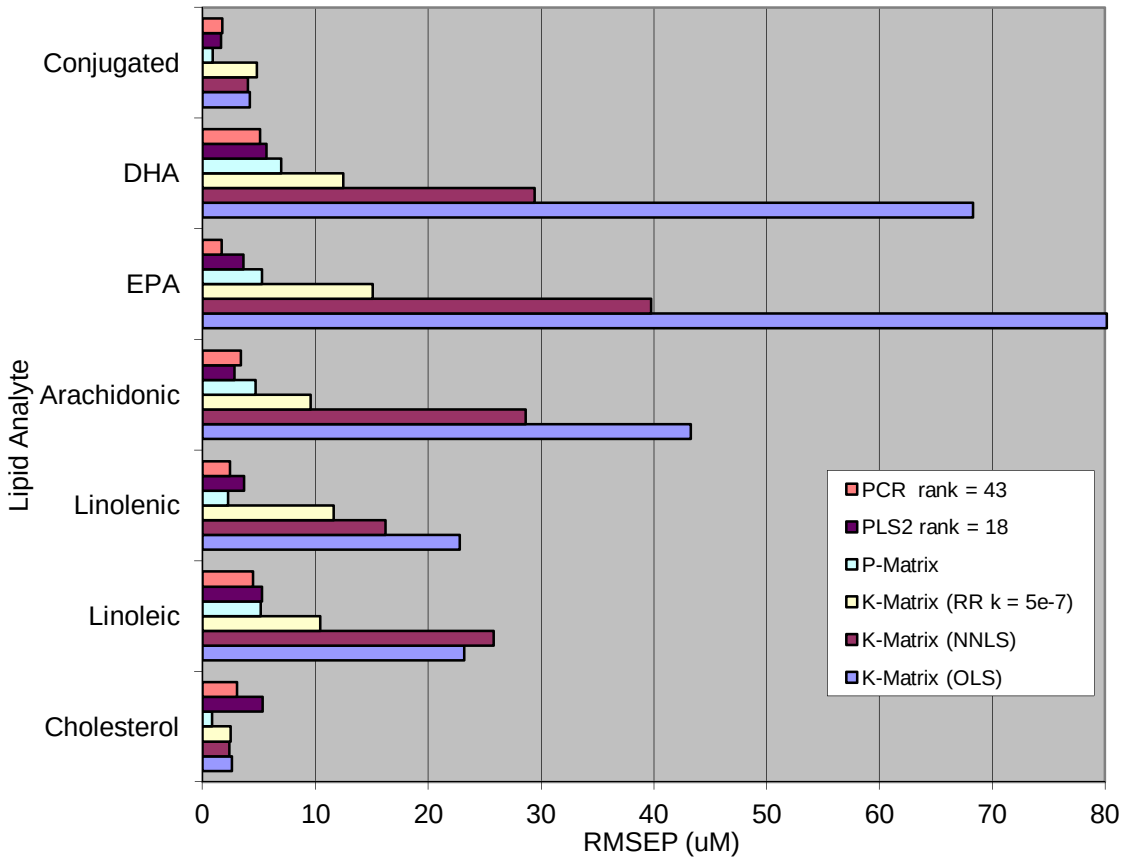


Figure 15. RMSEPs for the different algorithms as clustered in each lipid analyte.

The PLS2 calibration coefficients were tested on some serum samples. The same serum samples were validated using GC-MS. Table 9 shows the comparison for the total ω-6 and ω-3 PUFA concentrations expressed as percentages and their ratios. It can be noted that the ω-6 and ω-3 total % concentration and the ratio of ω-6 to ω-3 were substantially identical between the two methods. Table 10 shows the comparison between the PLS and the standard enzymatic test for cholesterol.

Table 9. Comparing ω-6 total conc % and ω-3 total conc % between PLS and GC-MS of five serum samples.

Category	ω-6 total conc %	ω-6 total conc %	ω-3 total conc %	ω-3 total conc %	ω-6/ω-3 total conc %	
Patient's code	PLS	GC-MS	PLS	GC-MS	PLS	GC-MS
P1	47.86	47.17	52.14	52.83	0.92	0.89
P2	50.12	54.05	49.88	45.95	1.00	1.18
P3	49.59	46.59	50.41	53.41	0.98	0.87
P4	48.83	47.93	51.17	52.07	0.95	0.92
P5	49.05	48.22	50.95	51.78	0.96	0.93

Table 10. Comparing cholesterol PLS and enzymatic test of five serum samples.

Sample	Cholesterol, PLS pred. (mg/dL)	Cholesterol, Enzymatic (mg/dL)	Percent Error
P1	203.61	187	-8.88
P2	207.27	188	-10.25
P3	187.98	189	0.54
P4	193.31	189	-2.28
P5	174.05	163	-6.78

5.1.2 PARTIAL LEAST SQUARES (PLS1) ALGORITHM APPLIED TO HUMAN SERUM

Neural network (NN) was first attempted in this study. Using three hidden nodes, four number of tours, and with a 0.01 overfit penalty, the RMSEP in the training model was still considerably higher than any other algorithms. Though neural networks (NNs) can implicitly detect complex non-linear relationships between independent and

dependent variables, they suffer from disadvantages of being prone to "overfitting," and are "black box" and have limited ability to identify possible causal relationships.[247]

As with the previous results wherein PLS2 outperformed all other algorithms in the training model,[230] partial least squares in the form of PLS1 yielded lesser RMSEP than PLS2 in the same training model in this paper after mean centering of the training data set (Figure 16). In PLS1, the highest RMSEP is obtained for DHA. The possibility of similarity in the molar absorbance spectra for EPA and DHA would be the reason why the RMSEP is higher for DHA (Figure 17). Nevertheless, this results show that despite similarities in the molar absorbance of the lipid components, the RMSEP of all components using PLS1 is still low as compared to other algorithms.

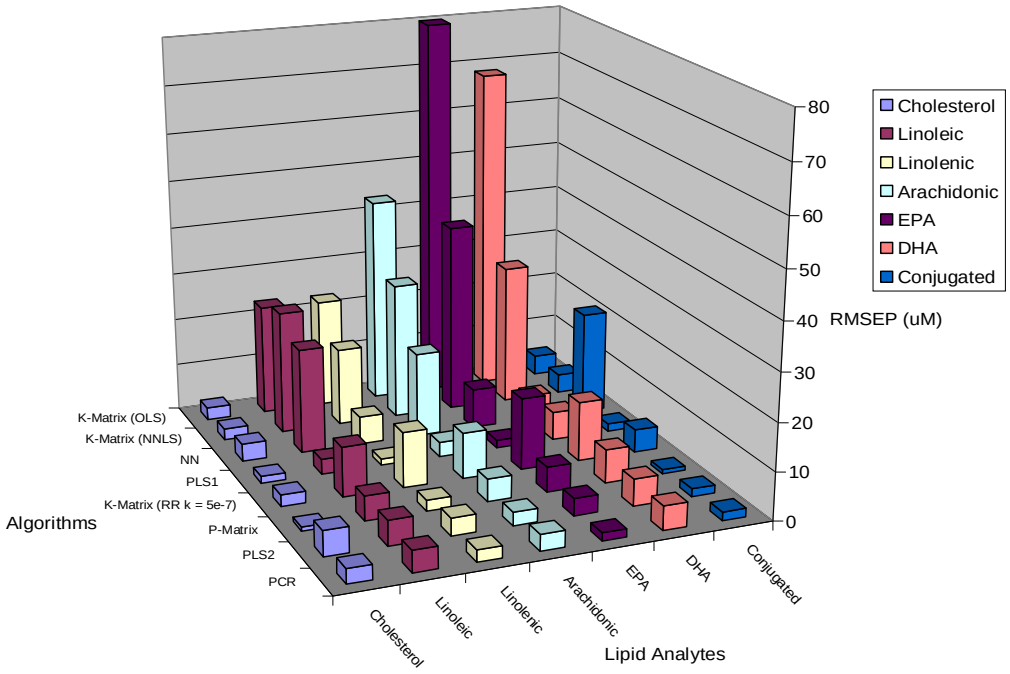


Figure 16. 3-dimensional (3D) diagram of the RMSEP of the lipid analytes in each algorithm. PLS1 yielded the least RMSEP for all analytes. Other algorithms are identical as in previous result in 5.1.1 and are shown here for comparison purposes only.[230]

Table 11 below shows the RMSEP values comparing all algorithms as seen in Figure 16. It is apparent that PLS1 algorithm showed the lowest RMSEP for almost of all of the components as compared to other algorithms.

Table 11. RMSEP (μM) values comparing all algorithms as seen in Fig. 16. (Linoleic acid (LA); linolenic acid (LNA); arachidonic acid (AA); eicosapentaenoic acid (EPA); docosahexaenoic acid (DHA); conjugated linoleic acid (CLA); ordinary least squares (OLS); non-negative least squares (NNLS); neural network (NN); principal component regression (PCR); partial least squares (PLS)).

	Cholesterol	LA	LNA	AA	EPA	DHA	CLA
K-matrix (OLS)	2.60	23.17	22.82	43.25	80.15	68.32	4.16
K-matrix (NNLS)	2.39	25.78	16.24	28.63	39.75	29.45	4.01
NN	3.63	21.97	5.84	17.74	8.34	5.25	21.50
PLS1	1.46	3.14	1.35	3.06	1.67	5.94	1.60
K-matrix (RR)	2.47	10.41	11.62	9.58	15.07	12.48	4.81
P-matrix	0.83	5.16	2.27	4.71	5.25	6.97	0.88
PLS2	5.29	5.27	3.66	2.80	3.59	5.64	1.65
PCR	3.05	4.47	2.44	3.40	1.65	5.08	1.76

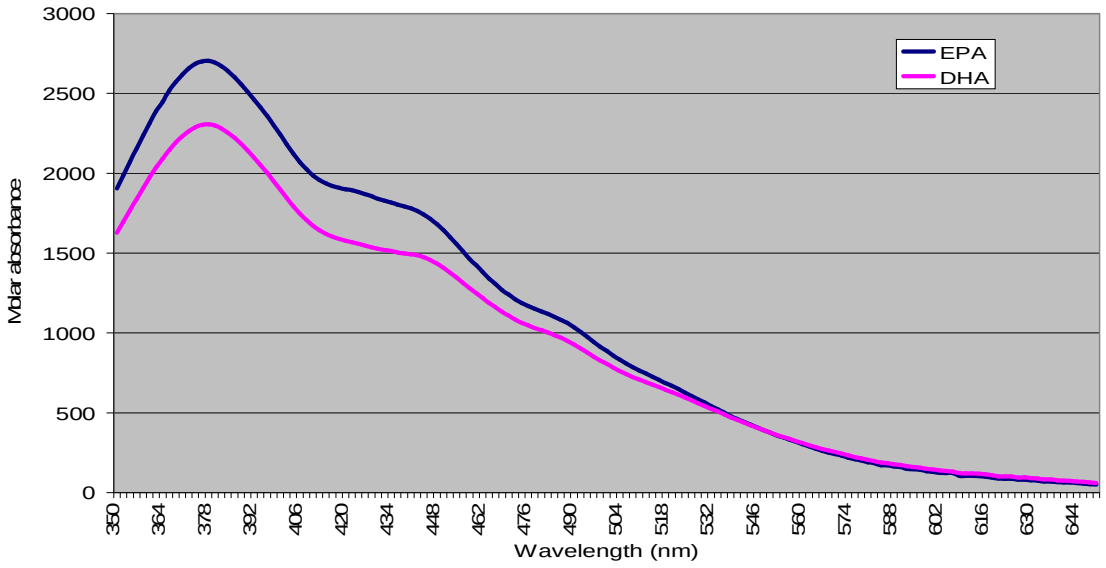


Figure 17. Molar absorptivities of EPA and DHA determined by the K-matrix model as in 5.1.1.[230] These are shown for comparison purposes only.

The PLS2 differs from PLS1 in the approach used to perform the signal decomposition and the regression analysis. Thus, PLS2 calculates the number of factors on all the components simultaneously and one weighed number of factors is optimized. PLS1 performs the optimization of the number of factors for only one component at a time. The application of PLS in spectroscopic data can be referred to some bibliographic references.[209, 248-251]

Choosing the optimum number of factors in PLS1 is the key to obtain a good calibration model. The trick is to keep only those factors that contain analytical information. The discarded factors should contain only noise. If too many factors are kept, there is danger of overfitting the data and adding noise to the calibration. If there are not enough factors, a proper calibration model cannot be generated.[199]

From Figures 18 and 19, it is readily apparent that prediction errors are minimized when calibrations are developed using the indicated number of factors as stated in the analyte's respective figure captions. The obtained PLS2 and PLS1 calibration models were applied to five serum samples obtained from HMC. Save for conjugated linoleic, all lipid components yielded positive molar concentrations in PLS1. As compared to the previous result wherein PLS2 used 18 factors,[230] the factors were reduced to an optimum number of 6 in PLS2 in this paper. The possibility of including a wide range of cholesterol and PUFAs concentration ranges calibration matrix is still collected, and when done, this assay will serve as a direct, time and cost saving method for simultaneously quantitating cholesterol and PUFAs in human serum.

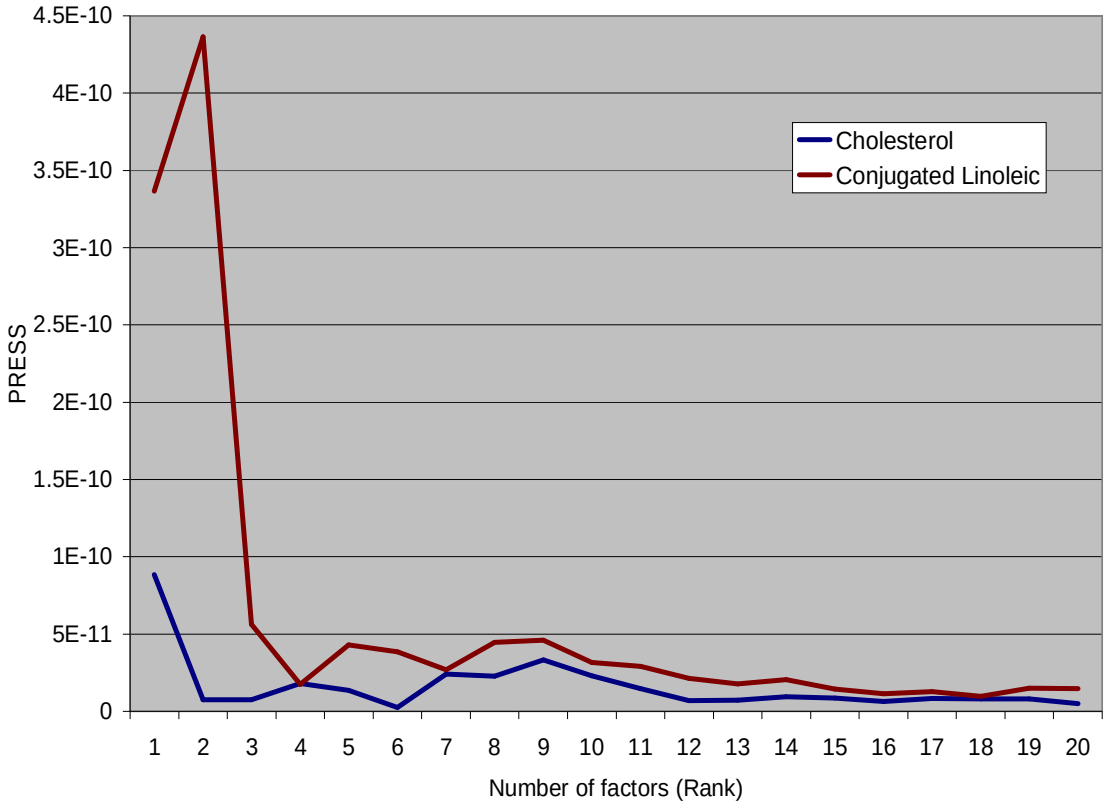


Figure 18. Plot of PRESS vs ranks for cholesterol and conjugated linoleic. 6 and 4 factors were chosen for cholesterol and conjugated linoleic, respectively.

In Table 12, molar concentrations of cholesterol, linoleic, linolenic, arachidonic, EPA, and DHA compared quite equally well with the GC-MS method. In Table 13, molar concentrations of conjugated linoleic compared quite equally well with the first two samples but not with the remaining three. CLA in normal physiological human serum exists in low concentrations (10-70 µM) as compared to other fatty acids, linoleic (2270-3850 µM), α-linolenic (50-130 µM), arachidonic (520-1490 µM), EPA (14-100 µM), and DHA (30-250 µM).[168, 252] This would be the most probable reason of the negative molar concentrations for the PLS1 in CLA.

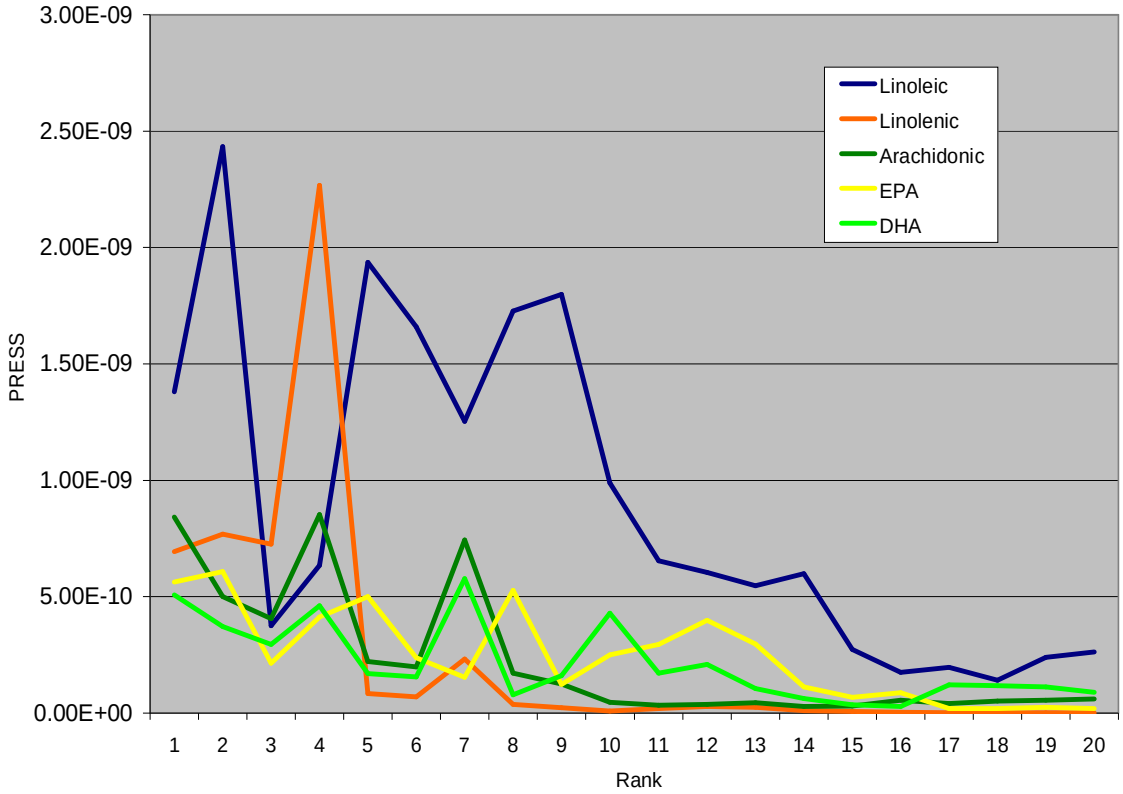


Figure 19. Plot of PRESS vs ranks for linoleic, linolenic, arachidonic, EPA, and DHA with ranks 3, 8, 17, 7, and 3, respectively.

Although successful, especially, in comparing relative percentage change in fatty acids for clinical studies, GC's disadvantages include the derivatization steps which can alter the structure of the fatty acid or create side-products that can overlap with the analytes needed.[253] Short chain fatty acid methyl esters can be eluted quickly and missed.[254] Also, the procedures are quite labor intensive. It is also difficult to obtain meaningful concentrations when using only a limited number of standards. These limitations strengthen the case for the development of this simple and direct method assay that does not require separation and reacts directly with the PUFAs.

Table 12. PLS1 molar concentrations of cholesterol, linoleic, linolenic, arachidonic, EPA, and DHA in human serum samples compared to GC-MS.

	PLS1	GC-MS	% Difference	PLS1	GC-MS	% Difference
	Cholesterol			Linoleic		
P1	4.21E-03	4.09E-03	2.89	2.08E-03	1.98E-03	4.89
P2	3.25E-03	3.01E-03	8.13	3.27E-03	2.87E-03	14.0
P3	2.25E-03	2.39E-03	6.05	2.87E-05	3.19E-05	9.88
P4	3.13E-03	3.39E-03	7.55	2.62E-03	2.44E-03	7.18
P5	3.11E-03	3.19E-03	2.41	2.06E-03	3.01E-03	31.4
	Linolenic			Arachidonic		
P1	3.13E-05	2.67E-05	17.2	1.46E-03	1.50E-03	2.42
P2	3.83E-05	3.48E-05	10.0	2.86E-04	2.52E-04	13.7
P3	5.79E-04	5.31E-04	9.05	2.40E-03	2.31E-03	3.84
P4	5.76E-05	6.10E-05	5.58	2.86E-04	2.52E-04	13.7
P5	3.75E-05	6.16E-05	39.1	9.70E-05	1.38E-04	29.7
	EPA			DHA		
P1	2.83E-03	3.52E-03	19.5	2.28E-03	2.73E-03	16.5
P2	4.05E-03	4.30E-03	5.86	9.77E-04	1.05E-03	6.91
P3	2.39E-03	2.97E-03	19.6	1.80E-03	1.83E-03	1.81
P4	2.10E-03	1.85E-03	13.3	1.74E-03	2.77E-03	37.1
P5	2.17E-03	2.20E-03	1.39	1.80E-03	1.60E-03	12.3

Table 13. PLS1 molar concentrations of conjugated linoleic in human serum samples compared to GC-MS.

	PLS1	GC-MS	% Difference
P1	1.22E-04	1.21E-04	0.67
P2	5.10E-04	5.17E-04	1.39
P3	-8.26E-05	5.70E-04	-
P4	-4.12E-04	1.03E-04	-
P5	-5.63E-04	7.05E-05	-

5.2 VEGETABLE OIL SAMPLES

The molar absorbance spectra for oleic, linoleic, and linolenic fatty acid methyl esters obtained using the training set are shown in Fig. 20. Central composite design and simplex lattice design training sets were used because they have demonstrated to be a useful method in formulations of experiments, fits nicely into the sequential experimentation that is involved with the experimental design, requires fewer experiments, and provides convenience and high accuracy.[255-257] It is readily apparent

that linoleic and linolenic molar absorbance spectra are six times greater than that of the oleic. Oleic is characterized by two smooth valley type peaks found at 368 and 442 nm. Linoleic, on the other hand has maxima that occur at 376 and 426 nm. Linolenic has two maximum peaks also occurring at 376 and 426 nm, with the latter peak about 1000 molar absorbance more than the 426 nm peak of the linoleic. It is also apparent from Fig. 20, a small shoulder is found at 444 nm for linolenic.

Oleic fatty acid is a monounsaturated fatty acid with a double bond occurring at carbon 9 relative to the -COOH terminal. Linoleic, on the other hand has two double bonds occurring at carbons 9 and 12; while linolenic has three double bonds found at carbons 9, 12, and 15 all relative to the -COOH terminal (Fig. 21). The most probable reason why the molar absorbance of oleic is buried under that of linoleic and linolenic fatty acids is due to its monounsaturated property.

The first attempt of deconvoluting the spectra is by the KM model. However, the KM approach yielded RMSEP high enough that the regression coefficients may yield high errors in the actual sample sets. The RMSEP is especially useful in comparing the prediction errors of the different regression models.[258] A high RMSEP (greater than 10 μM) as in this study simply means the regression model will give high errors in the unknown set samples. It is observed that there are less samples (n=30) in the training sets than the number of variables (p=101). In such case, the KM calibration model has limited applicability, yielding high RMSEP.[259] Although KM approach offers the advantage of estimating the true constituent spectra of the components in the training sets, it has, however, the disadvantage of requiring the knowledge of the concentrations of all interfering chemical constituents with a spectral profile in the training set and that the

calibration and analysis are connected to the inversion of the matrix.[235, 260]

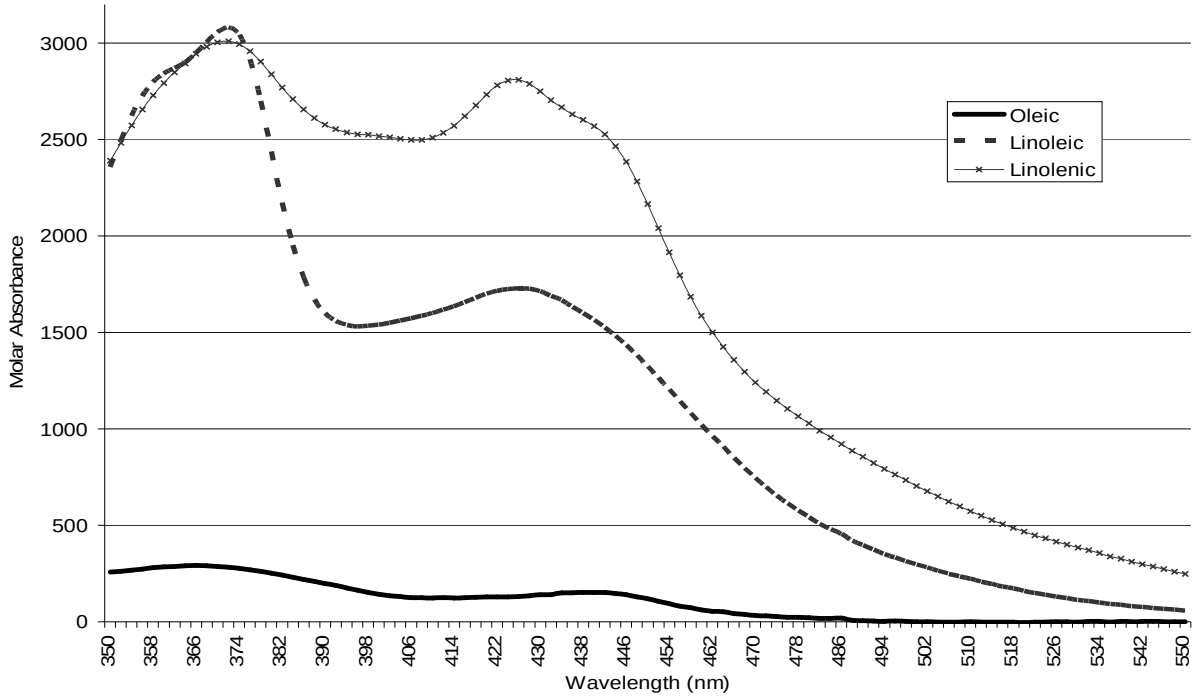


Fig. 20. Molar absorbance spectra of oleic, linoleic, and linolenic fatty acids obtained by the K-matrix chemometric approach.

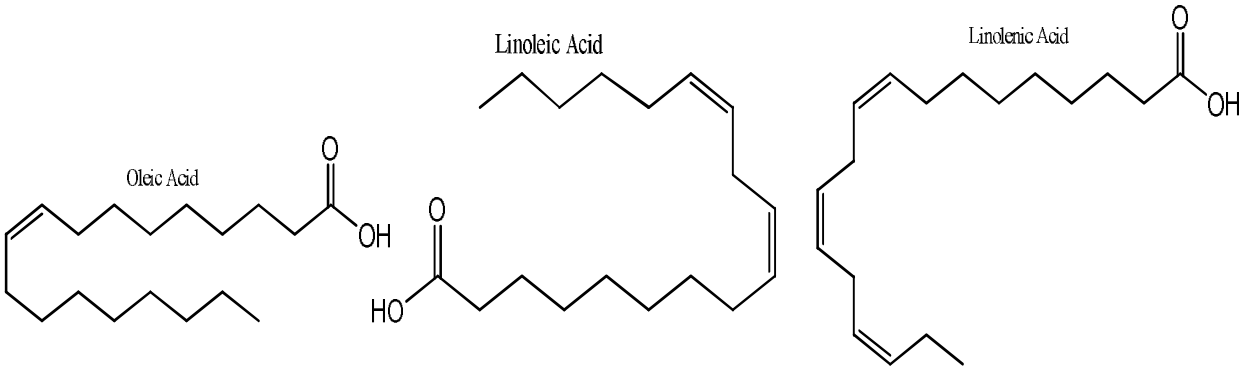


Fig. 21. Structures of oleic, linoleic, and linolenic fatty acids.

NN was next attempted in the study using 3 hidden nodes and 200 maximum iterations. NN improved the RMSEP over the KM approach yet not low enough to be used for the unknown sets (Fig. 22). NN offers the disadvantage of requiring a large amount of data to ensure that the results are statistically accurate and the networks adapt their analysis of data in response to the training which is connected to the network.[261]

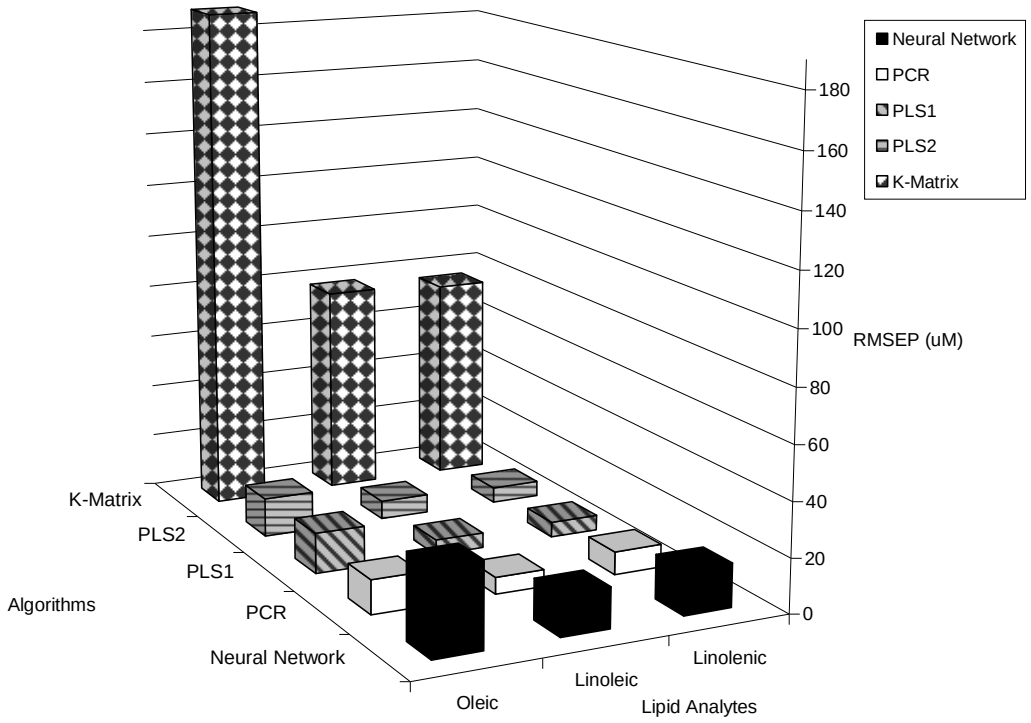


Fig. 22. RMSEP of oleic, linoleic, and linolenic fatty acid methyl esters compared against the different algorithms.

PCR and PLS algorithms were then attempted to obtain their RMSEPs, and the results are satisfying over the KM and NN models. A quite equal performance for PCR, PLS2, and PLS1 algorithms were obtained for their respective RMSEPs (Fig. 22). PLS differs from PCR in that it uses the concentration data from the training set and the spectral data in modeling, whereas PCR only uses the spectral data.[262] However, the solutions and hence the performance of PLS and PCR tend to be quite similar in most situations, largely because they are applied to problems involving high collinearity.[235]

A decision about the number of suitable PLS/PCR factors is necessary. Too few factors lead to underfitting leading to inadequate predictions, since the information extracted by the model is not enough to explain the data. On the other hand, too many

factors leads to overfitting, that is, the model cannot be generalized to new data that did not contribute to the model construction.[263]

For PCR, 6 factors were used for the model while 18 factors were used for PLS2. For PLS1, 5, 6, and 12 factors were used for oleic, linoleic, and linolenic, respectively (Fig.23). These factors were chosen based on the plot of PRESS vs number of factors/rank chart as shown in Fig. 23.

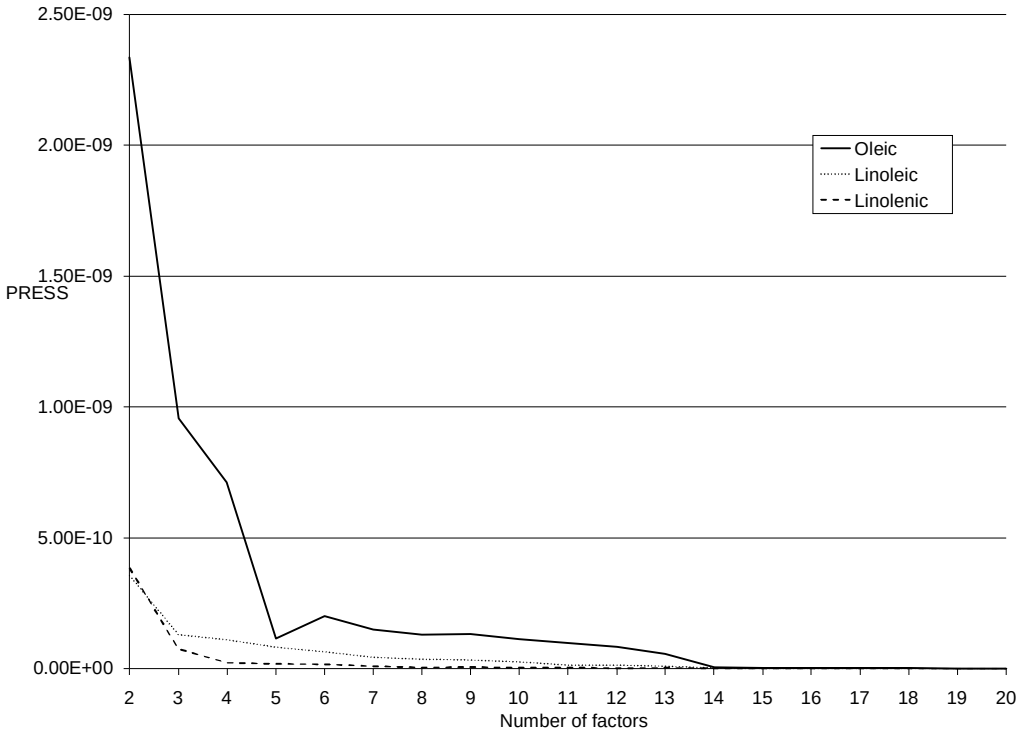


Fig. 23. PRESS chart for oleic, linoleic, and linolenic fatty acid methyl esters. 5, 6, and 12 factors were chosen for oleic, linoleic, and linolenic, respectively, in PLS1.

The plot with the lowest PRESS indicated the number of factors to be used for PLS1. After choosing the number of factors for each algorithm, the RMSEP was calculated to indicate the appropriate model. The model(s) with the lowest RMSEP indicated to be the appropriate model to be tested for the unknown samples.

After choosing the number of factors for PCR, PLS2, and PLS1 algorithms (Fig. 23), the obtained regression coefficients were applied to determine the molar concentrations of olive and sunflower oil samples. In Table 14, PLS2 yielded the lowest % error for oleic but not so much difference with PLS1 and PCR. For linoleic, PLS2 also yielded the lowest % error compared to PLS1 and PCR algorithms. In Table 15, PLS2 also yielded the lowest % error for oleic while PCR yielded the lowest % error for linoleic. Based on the obtained results, PLS2 mostly yielded a better performance than PLS1 and PCR algorithms.

Table 14. Molar concentrations of oleic linoleic, and linolenic fatty acid methyl esters in olive oil samples compared using the three most robust algorithms.

Component	PLS2	Database	% Error	PLS1	Database	% Error	PCR	Database	% Error
Oleic	1.50E-02	2.37E-02	36.7	1.49E-02	2.37E-02	37.2	1.46E-02	2.37E-02	38.3
Linoleic	1.06E-03	1.05E-03	-1.3	9.17E-04	1.05E-03	12.7	9.99E-04	1.05E-03	4.9
Linolenic	1.50E-03	0	-	1.47E-03	0	-	1.50E-03	0	-

Table 15. Molar concentrations of oleic linoleic, and linolenic fatty acid methyl esters in sunflower oil samples compared using the three most robust algorithms.

Component	PLS2	Database	% Error	PLS1	Database	% Error	PCR	Database	% Error
Oleic	2.08E-03	2.30E-03	9.6	6.88E-03	6.10E-3	-12.7	1.09E-03	2.30E-03	52.5
Linoleic	4.41E-03	3.84E-03	-14.9	1.27E-02	1.06E-2	-19.9	4.20E-03	3.84E-03	-9.4
Linolenic	8.59E-04	0	-	1.11E-03	0	-	5.85E-04	0	-

Normally we expect PLS1 to give a better model than PLS2. However, PLS2 gives better results than PLS1 especially if the analyte concentrations are strongly correlated.[259] Noticeable zero concentrations were obtained for linolenic fatty acid using the USDA database for the primary reason that it exists in low quantities in vegetable oil samples relative to both oleic and linoleic.[176-178, 260]

Expanding the training and prediction sets and testing the PLS and PCR algorithms to other types of vegetable oil samples would probably improve the differentiation as to which algorithm would be the most appropriate one to be employed in this study.

As part of an unpublished work and will be discussed herewith, an attempt was made of using PCA and cluster analysis to determinate the capability of the Purdie Assay to discriminate the prepared training set of vegetable oils made by central composite design (Table 6, mixtures 1 to 20). PCA was first performed in spectral data (20 samples *x* 101 variables matrix) of synthetic vegetable oil mixtures in chloroform solutions consisting of oleic, linoleic, and linolenic samples. Central composite design consists of two-level full factorial, star, and replicate designs and was used in this study because it is considered to be economical[205] and has several advantages over the three-level design in that the total number of runs in the former is frequently less than that required for a three-level full factorial factorial design. For example, with $p = 5$ variables, 243 runs would be required for the three-level full factorial design, whereas with single replicates for the cube and star portions and four center points, the total number of runs required for a central composite design would 16 + 10 + 4 = 30. A second advantage of the mentioned

design is that it lends to a sequential approach to experimentation, since it can be built in sections.[264]

The spectral data matrix for the central composite design was decomposed in PCA by singular value decomposition algorithms according to the equation below:[205]

$$\underset{nxp}{X} = \underset{nxnc}{C} \; \underset{ncxp}{S^T} \qquad (18)$$

where X is an n (=20) spectra at p (=101) wavelengths; C, 20 x 3 concentration matrix; S^T, 3 x 101 matrix of the pure spectra (n=20 is the number of mixture spectra, nc(=3) is the number of components, and p(=101) is the number of wavelengths). Eq. (18) shows the decomposition of the spectral matrix in real factors, a product of S^T of the spectra with a matrix C of concentration profiles.

By decomposing matrix X with a PCA as many significant principal components (PCs) should be found as there are chemical species in the mixtures.[205]

The decomposition in the wavelength space, for a system with three components is given by:

$$\underset{nxp}{X} = \underset{nx3}{T^*} \; \underset{3xp}{V^{*T}} + \underset{nxp}{E} \qquad (19)$$

Eq. (19) shows the decomposition of the spectral matrix in abstract factors T^* and V^{*T} (E is the error). The score matrix T^* gives the location of the spectra defined by the three principal components.

In this study, all spectral wavelengths (350-550 nm) were used and consequent variable reduction was performed as discussed earlier. For 101-dimensional space, feature reduction was performed to a 2-dimensional data and the resulting principal components, PC1 and PC2 were plotted against each other. The first two PCs were used in this study for they retain the largest information and the most variations.[205] After plotting the first two PCs, agglomerative hierarchical clustering using Ward's algorithm was performed. The method starts with each point as its own cluster. At each step the clustering process calculates the Ward's distance between each cluster and combines the two clusters that are closest together. This combining continues until all the points are in one final cluster.[196] Specifically, for clusters S_{w1} and S_{w2} whose cardinalities are N_{w1} and N_{w2} and centroids c_{w1} and $c_{w1,}$ respectively, Ward's distance is defined as:

$$dw(S_{w1}, S_{w2}) = \frac{N_{w1} N_{w2}}{N_{w1} + N_{w2}} d(c_{w1}, c_{w2}) \qquad (20)$$

where $d(c_{w1}, c_{w2})$ is the squared Euclidean distance between c_{w1} and c_{w2}.[265] Fig. 24 shows the dendogram resulting from hierarchical clustering. Each number corresponds to the clusters containing the samples of similar observations. Ward's method was used in this study for it leads to well-structured dendograms. There are no completely satisfactory methods for determining the number of population clusters for any type of cluster analysis. In this study, the number of clusters was determined by using a Scree Plot found below the dendogram in Fig. 24. The place where the Scree Plot changes from a sharp downward slope to a more level slope is an indication of the number of clusters. Five clusters were identified corresponding to different levels of the prepared lipids according

to the central composite design of synthetic sets. The resulting plot of PC1 versus PC2 and the resulting clusters with 0.90 confidence density ellipses are shown in Fig. 25.

Cluster 1 consisting of mixtures 1, 7, and 10 is characterized by low oleic and linoleic, and high linolenic concentrations. Cluster 2 comprising of mixtures 4, 5, 9, and 12 consists of low concentrations in all three components. Cluster 3, on the other hand, consisting of mixtures 2 and 3, is high in oleic but low in linoleic and linolenic concentrations. Cluster 4 has mixtures 6, 8, and 11 consists of average oleic and high linoleic and linolenic concentrations. Lastly cluster 5 includes mixtures 13-20 and has average concentrations for the three components. The results indicate that the assay has the capability of discriminating patterns generated from different mixtures.

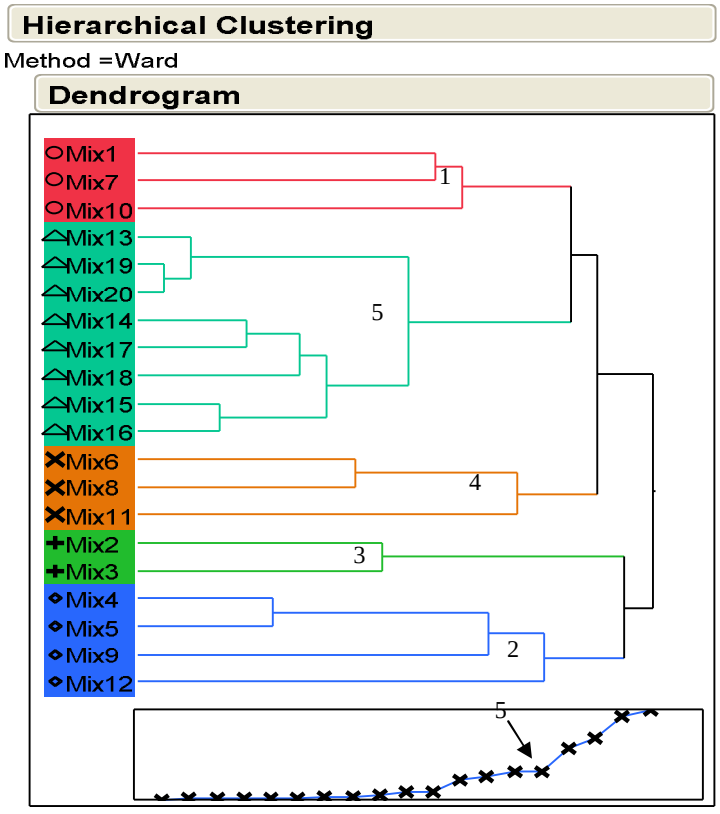


Figure 24. Dendogram for vegetable oil central composite design training set. Below the diagram is the Scree Plot with five clusters.

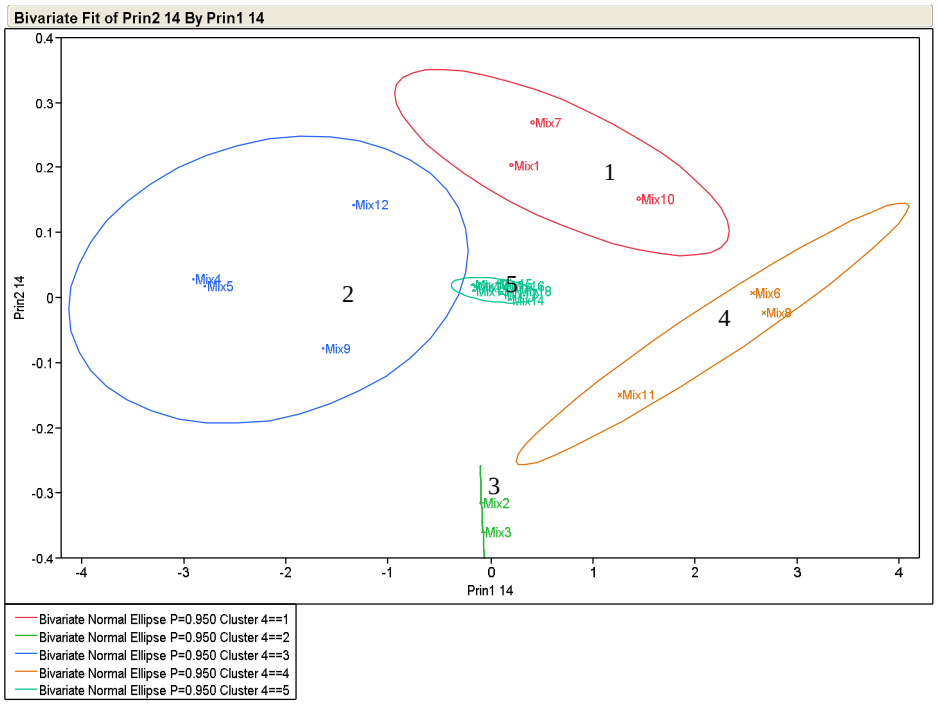


Figure 25. Clusters corresponding to the dendogram in Fig. 24. Five clusters were identified corresponding to different concentrations of lipids prepared.

The next work for this project will be to test a wide variety of actual vegetable oil samples to find out whether the assay can discriminate such samples according to the PUFA levels present in such. Once completed, the assay may provide a direct test to discriminate vegetable oil samples according to their PUFA contents.

5.3 FOOD AND BIOLOGICAL SAMPLES

Pattern recognition of several food and biological samples led to the discrimination of eleven clusters with each corresponding to the particular sample types. After hierarchical clustering, the Scree Plot was able to identify eleven clusters as shown in Figure 26. All of the information in this section is subject to change as more biological/food samples are analyzed. The general trend that we see to this point is that the samples cluster together according to their types (Fig. 27).

Cluster 1 consists of Walmart Fat Free Milk. Cluster 2 consists of chicken liver, chicken, and Braum's Chocolate Whole Milk while cluster 3 has salmon and veal. Cluster 4 consists of ham and Heart Healthy Butter, cluster 5 has egg yolk regular and free range, while cluster 6 has beef. Omega-3 enriched oil, on the other hand belongs to cluster 7, salmon in can for cluster 8, chicken dried and salmon dried belong to cluster 9, while cluster 10 is a salmon dried for longer period of time.

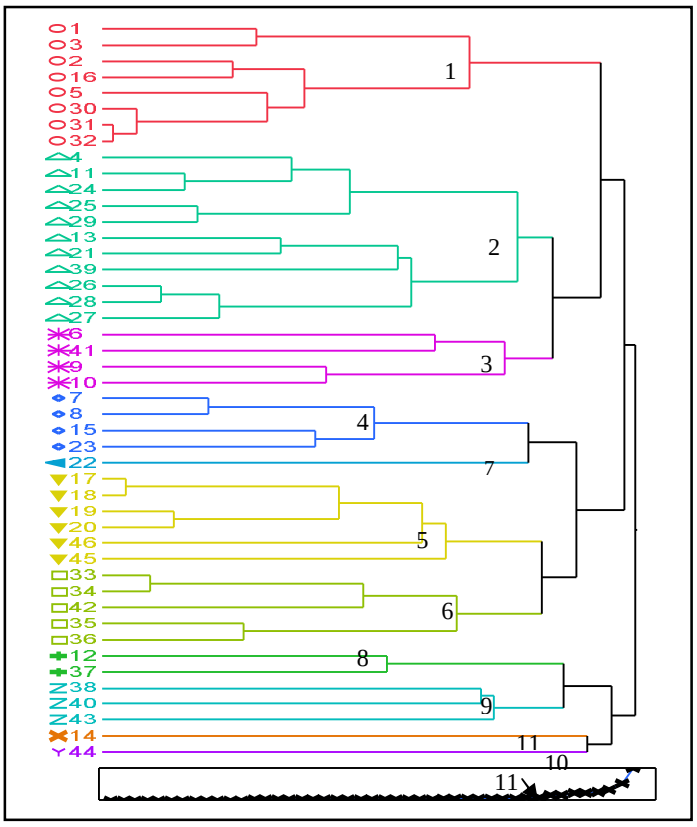


Figure 26. Dendogram for food and biological samples.
Below the diagram is the Scree Plot with eleven clusters.

Lastly, cluster 11 is a turkey sample. The separation suggests that the pattern can be used to examine the patterns generated from different mixtures. Although an initial analysis of the clusters was completed, all of the following groups are subject to change as the fatty acid profile is increased and as more information about the background of

various food and biological samples is obtained. Further collection of the samples' information including amounts of the different PUFAs and food processing conditions will have to be deduced to determine if this pattern recognition can be used as a screening tool to discriminate various biological and food samples according to their types and/or PUFAs contents.

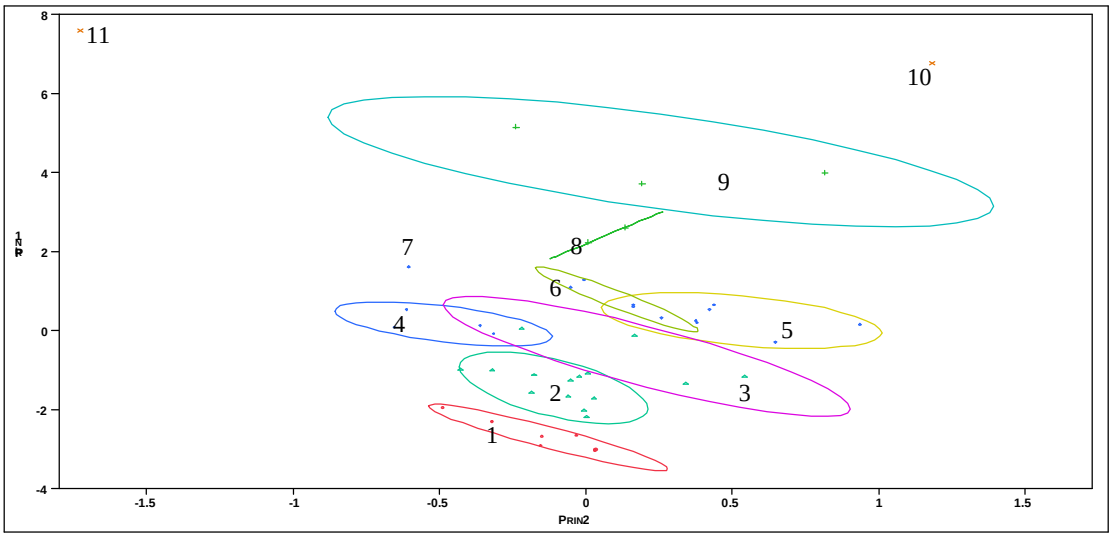


Figure 27. Clusters of different food and biological samples after treatment with the Purdie Assay.

5.4 INCLUSION OF OLEIC ACID IN SYNTHETIC HUMAN SERUM SAMPLES

Oleic acid (OA) was included as the eighth component in the training and prediction sets of the synthetic human serum samples. Using the three most robust soft modeling methods consisting of PCR, PLS2, and PLS1, PLS1 is known to reduce the RMSEP for all the components (Fig. 28). However, the inclusion of a monounsaturated fatty acid (MUFA) led to a higher RMSEP than the ones without OA in all components (Table 16). It is readily apparent from Fig. 20 that OA has small molar absorbance spectrum relative to that of the cholesterol and PUFAs. This would be the most probably

reason that despite attempts of trying out the most robust chemometric algorithms consisting of PLS1, PLS2, and PCR, the RMSEP for OA is relatively higher than most of the other components.

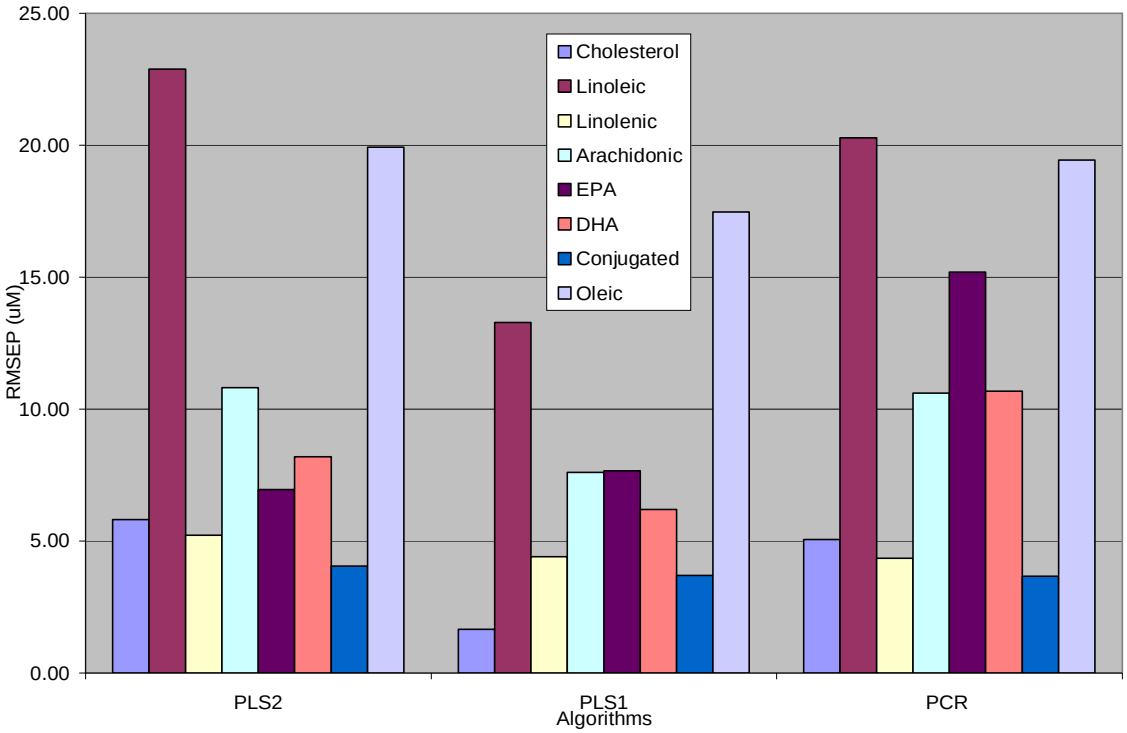


Figure 28. RMSEP for the different components of human serum with the inclusion of oleic acid.

GAPLS was attempted as a solution to reduce the RMSEP for the eight-component solution. We have given 101 variables for the training set. The GA was run for 101 variables (in the range 350-550 nm) using a PLS1 regression where the maximum number of factors allowed is the optimal number of components determined by cross-validation on the model containing all the variables, and the selected variables were used for the running of PLS1. The GA used in this study was adapted from Leardi *et al*,[227] whose parameters are described in Appendix I. In order to obtain the optimum set of

wavelengths for the determination of the eight lipids, the GA procedure was repeated three times for each lipid component. Multiple GA runs were used because it was found out that multiple isolated runs can be advantageous than a single GA run by reaching the global solution using fewer function evaluations.[266] Wavelength selection has clearly taken place, as indicated by the differences in selection frequency of the wavelength variables. Figure 29 shows the sample GA histogram of frequency of selection of each variable for cholesterol. The green horizontal line shows the cutoff for the model with the minimum RMSECV and is of interest in the study because of a larger number of variables selected.[227, 228] Appendices II to IX show the complete trial GA histogram of frequency of selection of each variable for each component. Table 17 shows the selected common wavelengths by the GA approach for each component after three GA trial runs.

RMSEP considerably reduced for all the components over the non-GA PLS1 approach except for EPA and DHA (Table 16). The similarity in the spectra of EPA and DHA as seen in Fig. 17 might probably be the reason of the increase in RMSEP for EPA and DHA in GAPLS. It is also well established that high spectral overlap causes a large prediction error as evident in this case.[267] Moreover, inspection of the selected wavelength variables revealed short intervals of selected wavelengths in Table 17 (12 and 5 for EPA and DHA, respectively), which might indicate over-fitting.[268] The average number of factors used for cholesterol, LA, LNA, AA, EPA, DHA, CLA, and OA for the GAPLS approach were 4, 14, 9, 11, 5, 11, 7, and 19, respectively.

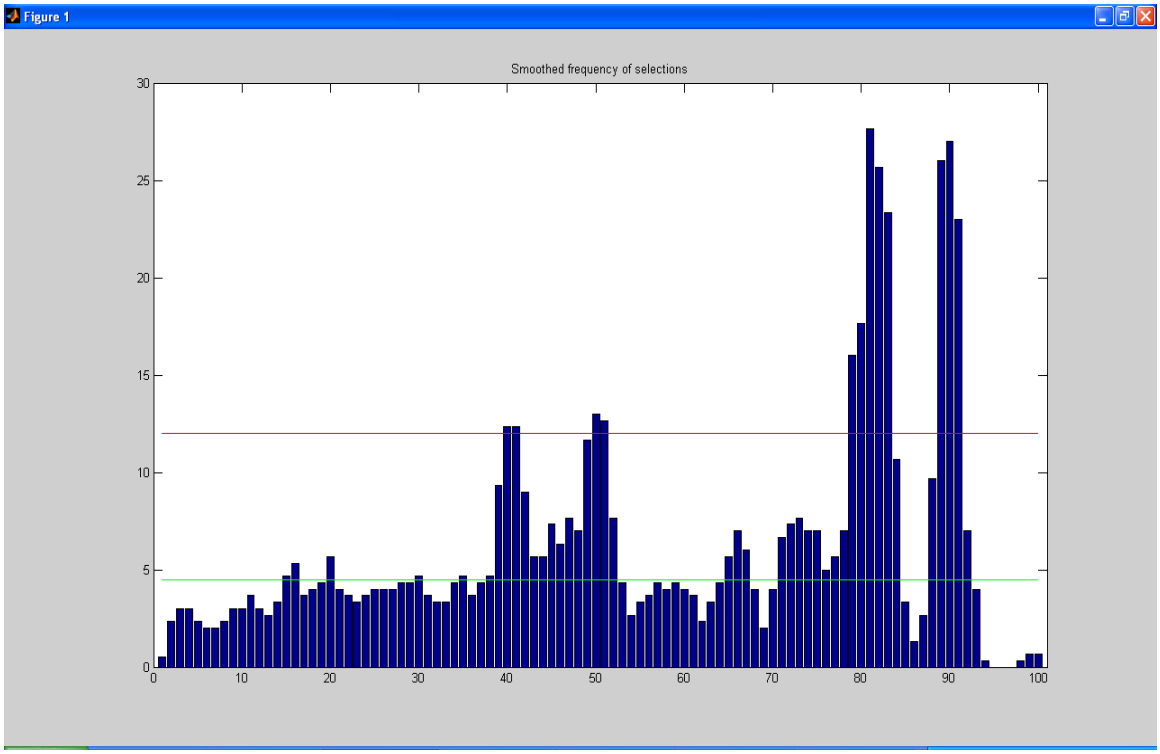


Figure 29. Histogram of frequency of selection of each variable for cholesterol.

Table 16. RMSEP values for PLS1 with** and without* oleic. GAPLS was applied for the eight-component mixture and reduced the RMSEP for almost all of the components.

	Cholesterol	LA	LNA	AA	EPA	DHA	CLA	Oleic
PLS1*	1.46	3.14	1.35	3.06	1.67	5.94	1.60	
PLS1**	1.66	13.28	4.41	7.61	7.67	6.20	3.70	17.47
GAPLS	1.37	11.45	4.42	1.67	13.93	11.66	2.53	15.89

Table 17. Selected common wavelengths by the GA approach for each component as an average of three GA trial runs.

Component	Number of wavelengths selected	Selected common wavelengths (nm) in three trials
Cholesterol	21	430-434, 438-440, 446-452, 492, 498, 502-514, 526-530
LA	20	370-394, 458-468, 478
LNA	33	366, 396, 416, 430-482, 488-492
AA	52	356, 370, 382-406, 410-420, 430-486, 494-496
EPA	12	414-436
DHA	5	442-450
CLA	16	374-386, 408-410, 422, 448-458
OA	25	354, 358-362, 382-406, 412, 430-438, 480, 492

An attempt of combining EPA and DHA into one and determining the RMSEP will be the next approach in completing this project. Moreover, validation of GAPLS with GC-MS and expanding the training and prediction sets will also be taken into consideration. Another interesting part of extending the project will be to use GA for other chemometric algorithms like PLS2, PCR, and MLR and compare the predictive ability of the models.

5.5 SPIKING OF HUMAN SERUM SAMPLES AND SYNTHETIC SETS

Attempts were also done by spiking a human serum sample matrix with standard solutions of fatty acids. Molar concentrations of the individual fatty acids were added to fresh human serum samples independently at different increments as shown in Table 6. For each lipid analyte, the spectra resulting from the addition of the spikes were then subtracted with the spectra of the original human serum samples used to determine the increase in spectral response. A plot is then derived between absorbance values vs concentrations at each wavelength and the slope obtained then describes the molar absorbance value for such analyte at a specific wavelength. Figure 30 shows the absorbance spectra for cholesterol after subtraction of the original human serum sample from which it was spiked. As an example, the spiking procedure exhibited linearity especially at the wavelength of maximum absorption as shown in Figure 31 with R^2 values of 0.983, 0.9734, and 0.9902 at 362 nm, 420 nm, 520 nm peaks, respectively for cholesterol. Figure 32 shows the molar absorbance values obtained for the different components after the spiking procedure. The obtained molar absorbance values were then used to determine the concentrations of an unknown human serum sample using the NNLS method. Table 19 shows the molar concentrations of some human serum samples obtained using the spiking NNLS method. Due to the similarity in the spectra of EPA and

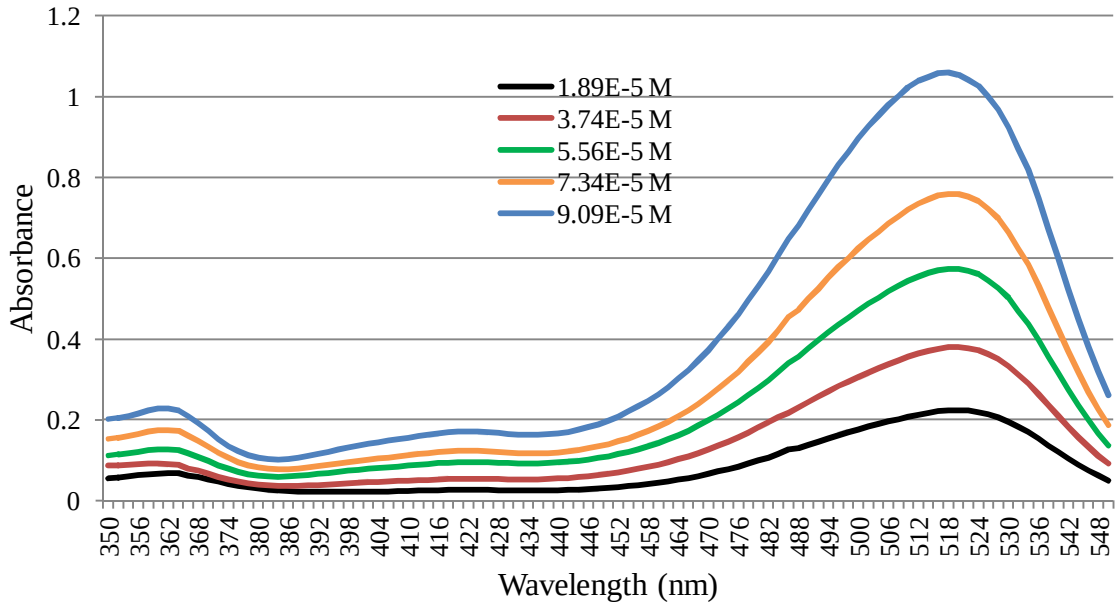


Figure 30. Absorbance spectra for cholesterol after subtraction of the original human serum sample from which it was spiked.

DHA, the average molar absorbance spectrum was obtained from the two components, and was used to determine the concentrations of the unknown. The experimentally obtained cholesterol serum concentrations were then compared with the enzymatic concentrations determined independently by the OSU Seretean Wellness Center. Table 18 shows this comparison of concentrations for the cholesterol component as an initial work.

While we were able to obtain a comparison for the cholesterol analyte, comparing the concentrations for the other components with the GC-MS would be the next step in completing this project. However, due to the tedious task of spiking a human serum sample, this technique is not ideal in a typical clinical setting, especially if batches of samples will be analyzed.

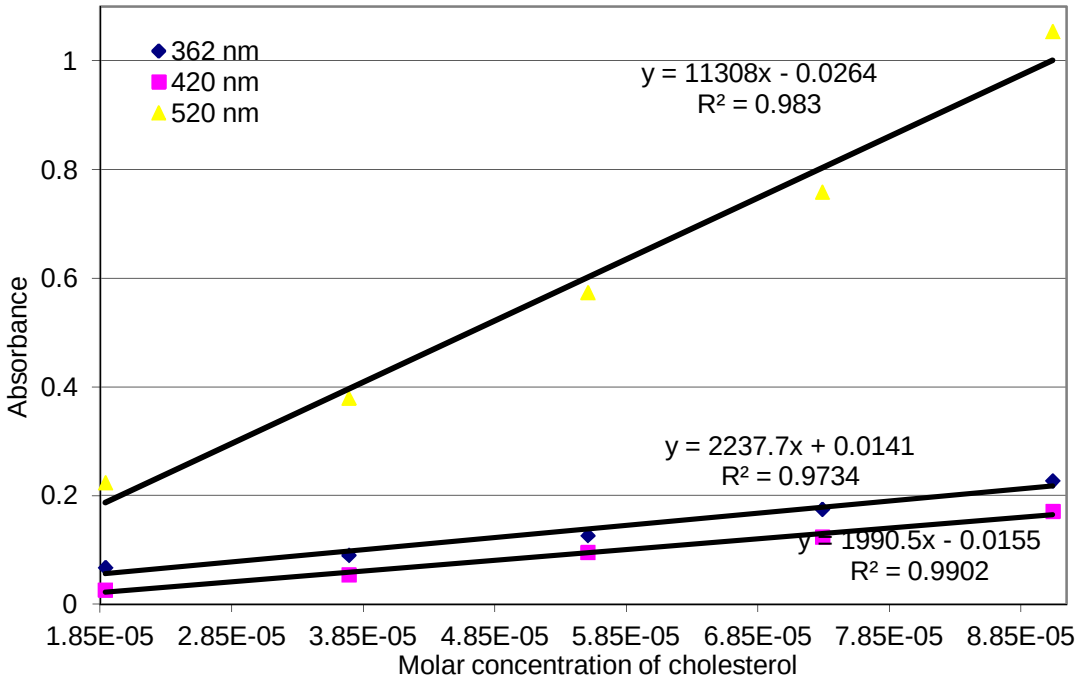


Figure 31. Linearity of response at the three peaks of cholesterol after spiking of human serum samples. R^2 values are 0.983, 0.9734, and 0.9902 for 362 nm, 420 nm, 520 nm, respectively.

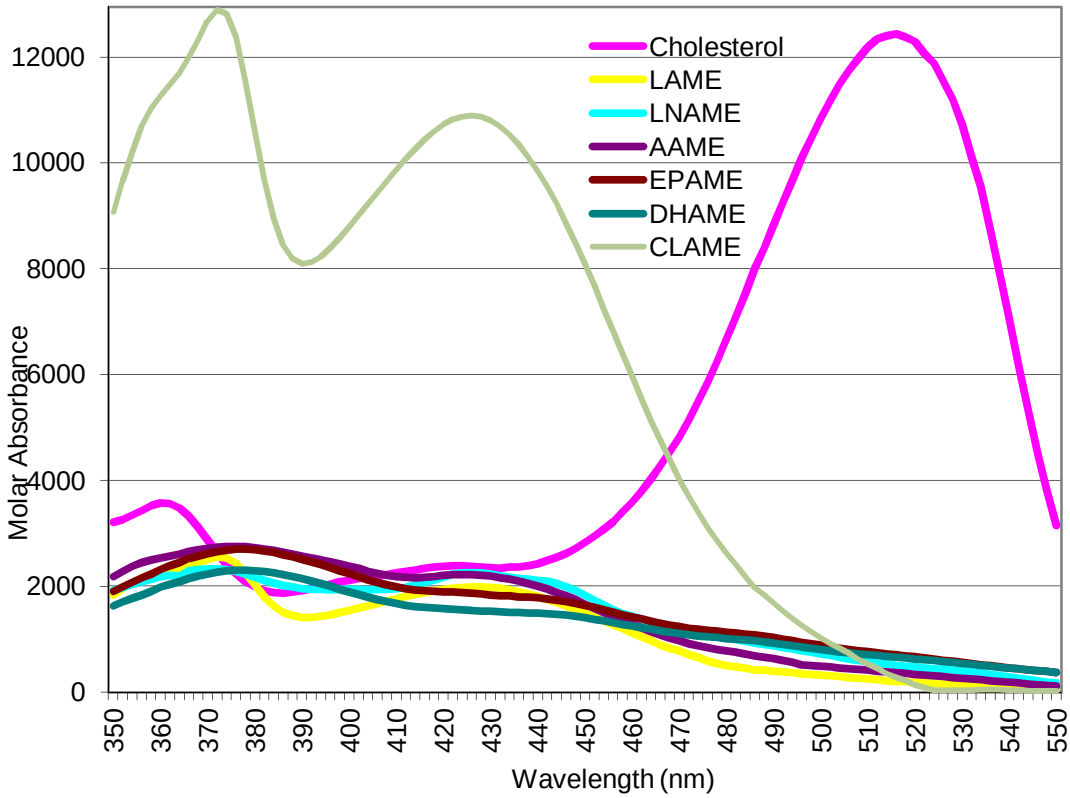


Figure 32. Molar absorbance from spiking of human serum samples using the spiking matrix from Table 6.

Table 18. Molar concentrations of actual human serum samples using the molar absorbance matrix obtained from Figure 32. EPA and DHA were averaged due to the similarity in the molar spectra.

Components	P1	P2	P3	P4
Cholesterol	4.47E-03	4.24E-03	5.58E-03	3.71E-03
LA	3.87E-03	3.84E-03	5.44E-03	1.70E-03
LNA	0	0	0	0
AA	3.06E-03	1.31E-03	7.91E-04	1.50E-03
EPA/DHA average	3.83E-04	1.50E-03	2.78E-03	0
CLA	0	0	0	0

Table 19. Cholesterol concentrations (mg/dL) comparison obtained using the spiking matrix and the enzymatic tests.

Patient Code	Spiking	Enzymatic Test	% Error
P1	172.99	180.00	-3.89
P2	164.09	184.00	-10.82
P3	215.95	226.00	-4.45
P4	194.49	227.00	-14.32

Spiking was also performed in a 7-component synthetic mixture solution (Mixture 128 of full factorial design) using similar spiking matrix as in Table 6. Figure 33 shows the obtained molar absorbance values from such experiment. These obtained molar absorbance values were then used to determine the RMSEP for an independent prediction set of Table 5 using the OLS, NNLS, and RR algorithms. Table 20 shows the RMSEP values comparing the three algorithms as seen in Figure 34 for the spiking procedure. Despite attempts of lowering the RMSEP, the prediction errors are considerably high enough to be used as training model to determine the unknown concentrations. Examining the condition number for Table 6 as a measure of error sensitivity, and as a criterion for data evaluation and experimental design, the value obtained was 1.98E4, which was high enough. When the condition number is a minimum (i.e. close to 1.00), small errors in the experimental measurements have the least effect on the unknown variables.[269] Ideally, the condition number for the matrix experimental design of additions should be 1 to allow no amplification of error.[203, 270]

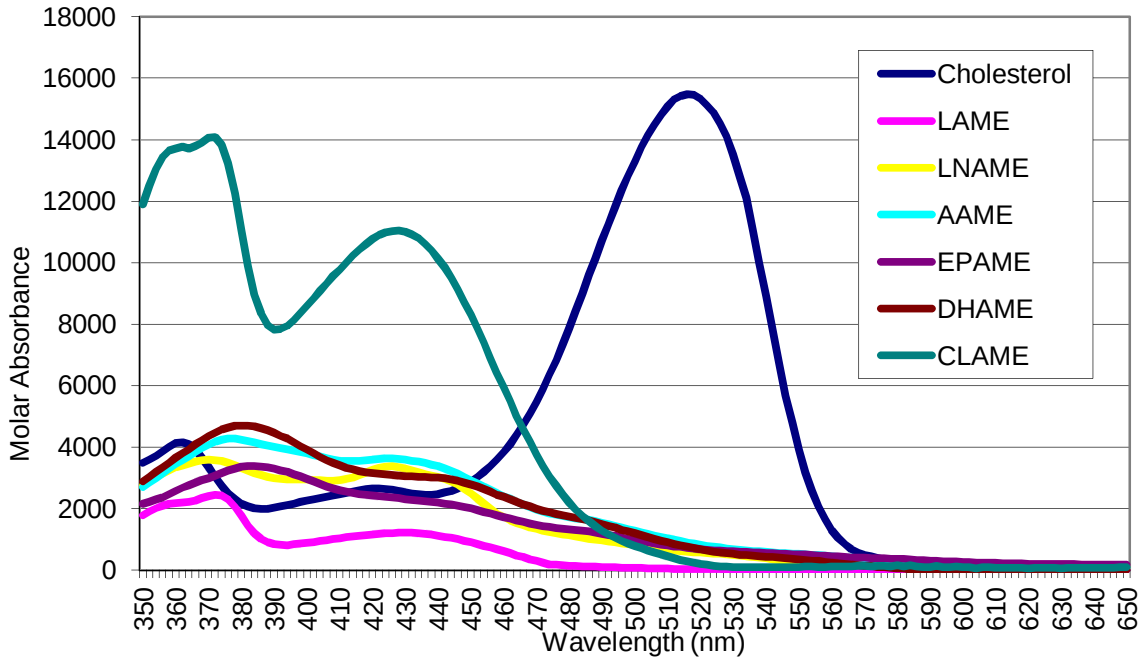


Figure 33. Molar absorbance from the spiking of a 7-component mixture using the spiking matrix from Table 6.

Table 20. RMSEP (µM) values comparing the three algorithms for the spiking procedure.

Spiking (OLS)	24.89	46.75	33.86	365.87	37.69	212.84	36.66
Spiking (NNLS)	24.53	47.69	33.95	199.11	18.04	120.51	38.02
Spiking (RR k = 5e-7)	22.45	67.59	30.43	101.45	14.63	76.02	26.19

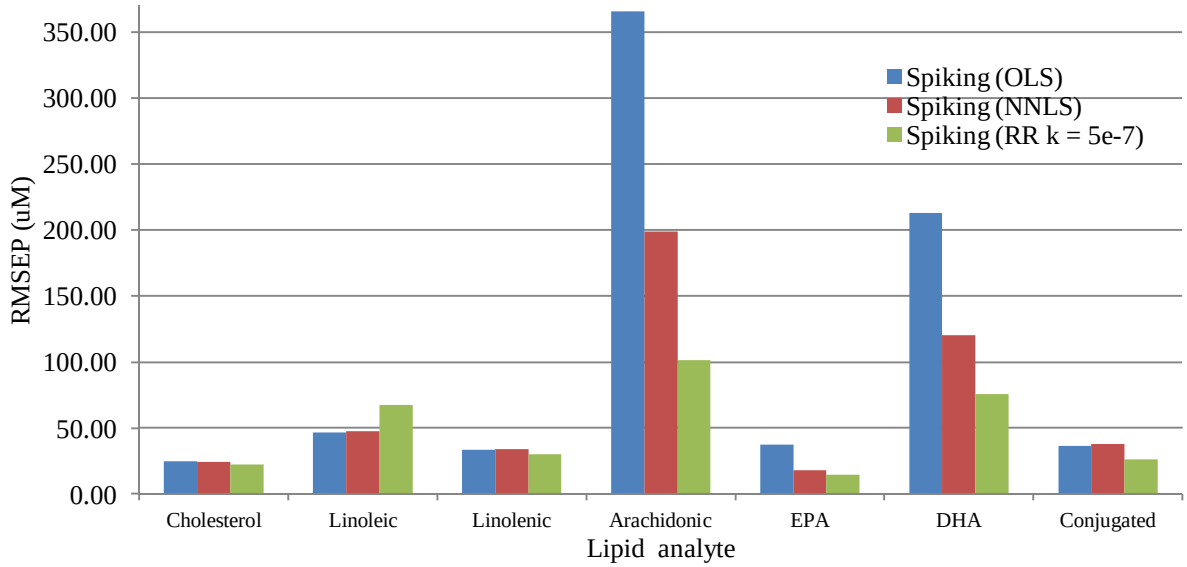


Figure 34. RMSEP (µM) for the lipids by spiking of a 7-component synthetic mixture.

CHAPTER 6

CONCLUSION

Part of this chapter has been published in the *Journal of Biotech Research*, the *Lecture Notes in Engineering and Computer Science*, and the *Lipid Technology Journal* and appears in this dissertation with the journals' permission.[1-4, 230]

A number of independent chemometric algorithms were tested that included MLR-NNLS, RR, PCR, and PLS. The principal outcome was that the RR, P-matrix, PCR, and PLS algorithms performed equally well enough as compared to the K-matrix approach when applied to the study of prepared mixtures (synthetic sera) in chloroform solutions. The PLS in the form of PLS2 model was tested for intact human serum specimens, and yielded results for ω-3 and ω-6 PUFA data that are comparable when using the GC-MS gold standard method. Similar results were also derived for the between-methods ω-6/ω-3 ratios. The first part of the study, therefore, showed the dominance of PLS2 over the other chemometric models. The study has also shown how the Purdie assay, coupled with chemometric algorithms, might provide alternatives to separations methods for the direct determination of lipids in human serum and its synthetic models. The advantages of this simple technology are the reduction in time and costs.[230]

For the second part of the study, as with the previous research results wherein PLS2 outperformed KM, PM, RR, and PCR in prepared mixtures in chloroform solutions (synthetic sera), PLS1 yielded the least RMSEP for all the lipid components as compared to all other algorithms in this study. This study has also attempted to determine the molar concentrations of cholesterol and PUFAs in human serum by the PLS1 algorithm. PLS1 yielded molar concentrations quite comparable with the GC-MS method in the actual human serum samples. The consistencies in the validation are evidence that the assay can be used as an alternative to the GC-MS procedures. While the GC-MS procedures gives only percentage values of the PUFAs, and obtaining a calibration curve in terms of peak areas and heights is a very tedious task, this new spectroscopic technology offers the advantages of being direct, simple, rapid, and cost efficient. The assay has a potential market for a wide array of clinical settings wherein GC-MS is impossible.[1]

The assay was extended to include OA as the eighth component for the simultaneous determination of unsaturated fatty acids in synthetic human serum models. However, despite attempts of utilizing the most robust algorithms like PLS1, PLS2, and PCR, the RMSEP was still high for the OA. GAPLS was able to successfully reduce the RMSEP for all the components over the non-GA PLS1 approach except for EPA and DHA. An attempt of combining EPA and DHA into one and determining the RMSEP will be the next approach in completing this project. Moreover, validation of GAPLS with GC-MS and expanding the training and prediction sets will also be taken into consideration.

Spiking of the human serum sample and synthetic sets of lipids was also performed but the procedure is considered tedious to perform and is not ideal for a typical

clinical setting especially when batches of samples are to be analyzed. An extension of the assay was performed for the pattern recognition of biological and food samples. The assay was able to discriminate eleven clusters corresponding to different food and biological samples.

The last part of the study attempted the simultaneous spectrophotometric and chemometric determination of the most abundant mono- and PUFAs in vegetable oils. The most important aspect of this work is the possibility of simultaneous determination of oleic, linoleic, and linolenic fatty acids in vegetable oil samples using the patented assay developed. No extraction step is required, and hence the use of organic solvents for separation, which are generally toxic pollutants, is avoided. It has been shown in this study that PCR, PLS2, and PLS1 algorithms compared equally well in the prediction sets, and that PLS2 mostly yielded a better performance than PLS1 and PCR algorithms in the unknown samples. Moreover, the assay was also able to discriminate the training set samples according to levels of lipids prepared. Compared to most other existing methods, the proposed methods are very simple, cheap, rapid and especially selective.[4]

REFERENCES

1. Dumancas, G. G., Muriuki, M., Purdie, N., Reilly, L., Partial least squares (PLS1) algorithm for quantitating cholesterol and polyunsaturated fatty acids in human serum. *Journal of Biotech Research* **2010,** *2*, 121-130.
2. Dumancas, G. G., Muriuki, M., Purdie, N., Reilly, L., Simultaneous spectrophotometric and chemometric determination of lipids in synthetic mixtures and human serum. *Lipid Technology* **2009,** *21* (5/6), 127-130.
3. Dumancas, G. G., Muriuki, M., Purdie, N., Reilly, L., Pattern recognition for discrimination of dyslipidemic states. *Lecture Notes in Engineering and Computer Science* **2011**.
4. Dumancas, G. G., Muriuki, M., Purdie, N., Reilly, L., Simultaneous spectrophotometric and chemometric determination of oleic, linoleic, and linolenic fatty acids in vegetable oils. *Lecture Notes in Engineering and Computer Science* **2011**.
5. Xu, J., Kochanek, K. D., Murphy, S. L., Tejada-Vera, B., Deaths: Final data for 2007. *National Vital Statistics Reports Centers for Disease Control and Prevention* **2010,** *58* (19), 1-135.
6. Anitschkow, N., Chalatow, S., Über experimentelle cholesterinsteatose und ihre bedeutung für die einiger pathologischer prozesse. *Zentralbl. Allg. Pathol* **1913,** *24*, 1-9.
7. Verschuren, W. M. M., Jacobs, D. R., Bloemberg, B. P. M., Kromhout, D., Menotti, A., Aravanis, C., Blackburn, H., Buzina, R., Dontas, A. S., Fidanza, F., Karvonen, M. J., Nedelijkovic, S., Nissinen, A., Toshima, H., Serum total cholesterol and long-term coronary heart disease mortality in different cultures. *Journal of the American Medical Association* **1995,** *274* (2), 131-136.
8. Goodnight, S. H., Jr.; Harris, W. S.; Connor, W. E.; Illingworth, D. R., Polyunsaturated fatty acids, hyperlipidemia, and thrombosis. *Arteriosclerosis* **1982,** *2* (2), 87-113.
9. Sturdevant, R. A.; Pearce, M. L.; Dayton, S., Increased prevalence of cholelithiasis in men ingesting a serum-cholesterol-lowering diet. *N Engl J Med* **1973,** *288* (1), 24-7.
10. Ederer, F., Leren, P., Turpelnen, O., Frantz, I.D., Cancer among men on cholesterol-lowering diets: experience from five clinical trials. *Lancet* **1971,** *2*, 203-206.
11. Rambjor, G. S.; Walen, A. I.; Windsor, S. L.; Harris, W. S., Eicosapentaenoic acid is primarily responsible for hypotriglyceridemic effect of fish oil in humans. *Lipids* **1996,** *31*, S45-S49.
12. Harris, W. S., n-3 fatty acids and serum lipoproteins: human studies. *Am J Clin Nutr* **1997,** *65* (5 Suppl), 1645S-1654S.
13. Harris, W. S.; Hustvedt, B. E.; Hagen, E.; Green, M. H.; Lu, G.; Drevon, C. A., N-3 fatty acids and chylomicron metabolism in the rat. *J Lipid Res* **1997,** *38* (3), 503-15.

14. Mori, T. A.; Burke, V.; Puddey, I. B.; Watts, G. F.; O 'Neal, D. N.; Best, J. D.; Beilin, L. J., Purified eicosapentaenoic and docosahexaenoic acids have differential effects on serum lipids and lipoproteins, LDL particle size, glucose, and insulin in mildly hyperlipidemic men. *American Journal of Clinical Nutrition* **2000,** *71* (5), 1085-1094.
15. Nordoy, A., Dietary fatty acids and coronary heart diseases. *Lipids* **1999**, 34:S19–S22.
16. Sellmayer, A.; Hrboticky, N.; Weber, P. C., Lipids in vascular function. *Lipids* **1999,** *34 Suppl*, S13-8.
17. Leaf, A., Diet and sudden cardiac death. *J Nutr Health Aging* **2001,** *5* (3), 173-8.
18. Hwang, D., Fatty acids and immune responses - A new perspective in searching for clues to mechanism. *Annual Review of Nutrition* **2000,** *20*, 431-456.
19. Storlien, L., Hulbert, A. J., Else, P. L., Polyunsaturated fatty acids, membrane function and metabolic diseases such as diabetes and obesity. *Curr. Opin. Clin. Nutr. Metab. Care.* **1998,** *1*, 559-563.
20. Storlien, L. H.; Kriketos, A. D.; Calvert, G. D.; Baur, L. A.; Jenkins, A. B., Fatty acids, triglycerides and syndromes of insulin resistance. *Prostag Leukotr Ess* **1997,** *57* (4-5), 379-385.
21. Salem, N.; Litman, B.; Kim, H. Y.; Gawrisch, K., Mechanisms of action of docosahexaenoic acid in the nervous system. *Lipids* **2001,** *36* (9), 945-959.
22. Cave, W. T., Dietary N-3 (Omega-3) Polyunsaturated Fatty-Acid Effects on Animal Tumorigenesis. *Faseb J* **1991,** *5* (8), 2160-2166.
23. World review of nutrition and dietetics, Volume 92, Omega-6/Omega-3 essential fatty acid ratio: the scientific evidence. Simopoulus, A. P., Cleland, L. G. , Ed. Karger: 2003.
24. Simopoulos, A. P., Omega-3-Fatty-Acids in Health and Disease and in Growth and Development. *American Journal of Clinical Nutrition* **1991,** *54* (3), 438-463.
25. Eaton, S. B.; Konner, M., Paleolithic Nutrition - a Consideration of Its Nature and Current Implications. *New Engl J Med* **1985,** *312* (5), 283-289.
26. Simopoulos, A. P., Genetic variation and evolutionary aspects of diet. *Antioxidant status, diet, nutrition, and health*, Papas, A. M., Ed. CRC Press: Washington, D. C., 1999a.
27. Simopoulos, A. P., Evolutionary aspects of omega-3 fatty acids in the food supply. *Prostag Leukotr Ess* **1999,** *60* (5-6), 421-429.
28. Simopoulos, A. P., New products from the agri-food industry: The return of n-3 fatty acids into the food supply. *Lipids* **1999,** *34*, S297-S301.
29. Simopoulos, A. P., The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp Biol Med* **2008,** *233* (6), 674-688.
30. Rischlau, P., Bernt, E., Gruber, W, Enzymatische bestimmung des gesamt-cholesterins im serum. *J Clin Chem Clin Biochem.* **1974,** *12*, 403-407.
31. Allain, C. C.; Poon, L. S.; Chan, C. S. G.; Richmond, W.; Fu, P. C., Enzymatic Determination of Total Serum-Cholesterol. *Clin Chem* **1974,** *20* (4), 470-475.
32. Liu, J.-W., Huang, Y.S., Separation and quantitation of polyunsaturated fatty acids and eicosanoids by HPLC *HPLC of Acyl Lipids*, Lin, J.-T., McKeon, T.A., Ed. HNB: 2005; pp 117-157.
33. Studer, J.; Purdie, N.; Krouse, J. A., Friedel-Crafts acylation as a quality control assay for steroids. *Appl Spectrosc* **2003,** *57* (7), 791-796.

34. Savage, G. P.; Dutta, P. C.; McNeil, D. L., Fatty acid and tocopherol contents and oxidative stability of walnut oils. *J Am Oil Chem Soc* **1999,** *76* (9), 1059-1063.
35. Ayorinde, F. O.; Clifton, J.; Afolabi, O. A.; Shepard, R. L., Rapid Trans-Esterification and Mass-Spectrometric Approach to Seed Oil Analysis. *J Am Oil Chem Soc* **1988,** *65* (6), 942-947.
36. Jennings, B. H.; Akoh, C. C., Enzymatic modification of triacylglycerols of high eicosapentaenoic and docosahexaenoic acids content to produce structured lipids. *J Am Oil Chem Soc* **1999,** *76* (10), 1133-1137.
37. USDA National Nutrient Database for Standard Reference. Nutrient Data Laboratory. www.nal.usda.gov/fnic/foodcomp/search
(accessed November 2011).
38. Focus on cholesterol research. Kramer, M. A., Ed. Nova Science Publishers: Hauppauge, 2006.
39. Li, J. J., *Triumph of the heart: The story of statins*. Oxford University Press: 2009.
40. Cabot, S., Jasinka, M., *Cholesterol: The real truth*. Ten Speed Pr: 2006.
41. Garrett, R., Grisham, C. M., *Biochemistry*. Brooke/Cole: Boston, 2010.
42. Brody, T., *Nutritional biochemistry*. Academic Press: San Diego, 1999.
43. Sardesai, V. M., *Introduction to clinical nutrition*. 2nd edition revised and expanded ed.; Marcel Dekker, Inc.: New York, 2003.
44. Campbell, M. K., Farrell, S. O., *Biochemistry*. 6th Edition ed.; Thomson Brooks/Cole: Belmont, 2009.
45. Dancygier, H., *Clinical hepatology: Principles and practice of hepatobiliary diseases*. Berlin: Berlin, 2010; Vol. 1.
46. Institute of Medicine, F. N. B., Dietary Fats: Total Fat and Fatty Acids. *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein and Amino Acids*, National Academy Press: Washington DC, 2005; pp 422-541.
47. Bruhl, L., Determination of trans fatty acids in cold pressed oil. *Eur J Med Res* **1995,** *1*, 89-93.
48. *Handbook of food science, technology, and engineering*, Hui, Y. H., Ed. CRC Press: Boca Raton, 2006; Vol. 1.
49. Spector, A. A., Essentiality of fatty acids. *Lipids* **1999,** *34*, S1-S3.
50. Riemersma, R. D., Armstrong, R., Kelly, R. W., Wilson, R., *Essentiality of fatty acids and eicosanoids*. AOCS Press: Champaign, 1999.
51. James, M. J.; Gibson, R. A.; Cleland, L. G., Dietary polyunsaturated fatty acids and inflammatory mediator production. *American Journal of Clinical Nutrition* **2000,** *71* (1), 343s-348s.
52. Lands, W. E. M.; Morris, A.; Libelt, B., Quantitative Effects of Dietary Polyunsaturated Fats on the Composition of Fatty-Acids in Rat-Tissues. *Lipids* **1990,** *25* (9), 505-516.
53. Simopoulos, A. P., The importance of the ratio of omega-6/omega-3 essential fatty acids. *Biomed Pharmacother* **2002,** *56* (8), 365-379.
54. Harris, W. S.; Mozaffarian, D.; Rimm, E.; Kris-Etherton, P.; Rudel, L. L.; Appel, L. J.; Engler, M. M.; Engler, M. B.; Sacks, F., Omega-6 Fatty Acids and Risk for Cardiovascular Disease A Science Advisory From the American Heart Association Nutrition Subcommittee of the Council on Nutrition, Physical Activity, and Metabolism;

Council on Cardiovascular Nursing; and Council on Epidemiology and Prevention. *Circulation* **2009,** *119* (6), 902-907.
55. Calder, P. C., n-3 polyunsaturated fatty acids, inflammation, and inflammatory diseases. *American Journal of Clinical Nutrition* **2006,** *83* (6), 1505s-1519s.
56. Sonestedt, E.; Ericson, U.; Gullberg, B.; Skog, K.; Olsson, H.; Wirfalt, E., Do both heterocyclic amines and omega-6 polyunsaturated fatty acids contribute to the incidence of breast cancer in postmenopausal women of the Malmo diet and cancer cohort? *Int J Cancer* **2008,** *123* (7), 1637-1643.
57. Pala, V.; Krogh, V.; Muti, P.; Chajes, V.; Riboli, E.; Micheli, A.; Saadatian, M.; Sieri, S.; Berrino, F., Erythrocyte membrane fatty acids and subsequent breast cancer: A prospective Italian study. *J Natl Cancer I* **2001,** *93* (14), 1088-1095.
58. Holub, D. J.; Holub, B. J., Omega-3 fatty acids from fish oils and cardiovascular disease. *Mol Cell Biochem* **2004,** *263* (1), 217-225.
59. Purnamadjaja, A. H.; Russell, R. A., Pheromone communication in a robot swarm: necrophoric bee behaviour and its replication. *Robotica* **2005,** *23*, 731-742.
60. Thomas, A., Fats and fatty oils. In *Ullmann's encyclopedia of industrial chemistry*, Wiley-VCH: Weinheim, 2002.
61. Villarreal-Lozoya, J. E.; Lombardini, L.; Cisneros-Zevallos, L., Phytochemical constituents and antioxidant capacity of different pecan [Carya illinoinensis (Wangenh.) K. Koch] cultivars. *Food Chem* **2007,** *102* (4), 1241-1249.
62. Moore, K. M.; Knauft, D. A., The Inheritance of High Oleic-Acid in Peanut. *J Hered* **1989,** *80* (3), 252-253.
63. Untoro, J.; Schultink, W.; West, C. E.; Gross, R.; Hautvast, J. G., Efficacy of oral iodized peanut oil is greater than that of iodized poppy seed oil among Indonesian schoolchildren. *American Journal of Clinical Nutrition* **2006,** *84* (5), 1208-1214.
64. Nutter, M. K., Lockhart, E. E., Harris, R. S. , The chemical composition of depot fats in chicken and turkeys. *Journal of the American Oil Chemists' Society* **1943,** *20* (11), 231-234.
65. Kokatnur, M. G.; Oalmann, M. C.; Johnson, W. D.; Malcom, G. T.; Strong, J. P., Fatty-Acid Composition of Human Adipose-Tissue from 2 Anatomical Sites in a Biracial Community. *American Journal of Clinical Nutrition* **1979,** *32* (11), 2198-2205.
66. Rizzo, W. B.; Watkins, P. A.; Phillips, M. W.; Cranin, D.; Campbell, B.; Avigan, J., Adrenoleukodystrophy - Oleic-Acid Lowers Fibroblast Saturated C22-26 Fatty-Acids. *Neurology* **1986,** *36* (3), 357-361.
67. Teres, S.; Barcelo-Coblijn, G.; Benet, M.; Alvarez, R.; Bressani, R.; Halver, J. E.; Escriba, P. V., Oleic acid content is responsible for the reduction in blood pressure induced by olive oil. *P Natl Acad Sci USA* **2008,** *105* (37), 13811-13816.
68. Obici, S.; Feng, Z. H.; Morgan, Y.; Stein, D.; Karkanias, G.; Rossetti, L., Central administration of oleic acid inhibits glucose production and food intake. *Diabetes* **2002,** *51* (2), 271-275.
69. Zak, B., Cholesterol Methodologies - Review. *Clin Chem* **1977,** *23* (7), 1201-1214.
70. Pearson, S.; Stern, S.; Mcgavack, T. H., A Rapid Procedure for the Determination of Serum Cholesterol. *J Clin Endocr Metab* **1952,** *12* (9), 1245-1246.
71. Zlatkis, A.; Zak, B.; Boyle, A. J., A New Method for the Direct Determination of Serum Cholesterol. *J Lab Clin Med* **1953,** *41* (3), 486-492.

72. Wybenga, D. R.; Pileggi, V. J.; Dirstine, P. H.; Digiorgi.J, Direct Manual Determination of Serum Total Cholesterol with a Single Stable Reagent. *Clin Chem* **1970,** *16* (12), 980-&.
73. Huang, T. C.; Chen, C. P.; Raftery, A.; Wefler, V., A Stable Reagent for Liebermann-Burchard Reaction - Application to Rapid Serum Cholesterol Determination. *Anal Chem* **1961,** *33* (10), 1405-&.
74. Etienne, G., Etienne, J., Cottlet, J., Crockett, R. , Dosage du cholesterol total serique par la methods directe de Huang. *Ann. Biol.Clin.* **1964,** *22*, 923.
75. Ness, A. T.; Pastewka, J. V.; Peacock, A. C., Evaluation of Recently Reported Stable Liebermann-Burchard Reagent + Its Use for Direct Determination of Serum Total Cholesterol. *Clin Chim Acta* **1964,** *10* (3), 229-&.
76. Leppanen, V., Evaluation of the Para-Toluene-Sulfonic Acid Method for Quantitative Determination of Total Cholesterol in Serum. *Scand J Clin Lab Inv* **1956,** *8* (3), 201-206.
77. Figueroa, I. S.; Pike, R. L., An Ultramicro Method for the Determination of Total Serum Cholesterol. *Anal Biochem* **1960,** *1* (2), 103-106.
78. Richards, R. W.; Setchell, K. D.; Woodman, D. D., Improved Procedure for Estimation of Serum Cholesterol. *Clin Chim Acta* **1971,** *31* (2), 403-407.
79. Kim, E., Serum Cholesterol Assay Using a Stable Liebermann-Burchard Reagent. *Clin Chem* **1969,** *15* (12), 1171-&.
80. Perlstein, M. T.; Thibert, R. J.; Zak, B., Spectrophotometric Study of Influences on Direct Ferric Perchlorate Method for Determination of Serum-Cholesterol. *Microchem J* **1975,** *20* (4), 428-439.
81. Rappaport, F.; Eichhorn, F., Sulfosalicylic Acid as a Substitute for Paratoluene Sulfonic Acid .A. In the Estimation of Cholesterol in the Diagnostic Test for Systemic Lupus Erythematosus. *Clin Chim Acta* **1960,** *5* (2), 161-163.
82. Vanboetzelaer, G. L.; Zondag, H. A., A Rapid Modification of the Pearson Reaction for Total Serum Cholesterol. *Clin Chim Acta* **1960,** *5* (6), 943-944.
83. Jamieson, A., Method for Rapid Determination of Serum Total Cholesterol Using Modification of Pearson Reaction. *Clin Chim Acta* **1964,** *10* (6), 530-&.
84. Watson, P., A simple method for the direct determination of serum cholesterol. *Clin. Chim. Acta.* **1960,** *5*, 637.
85. Zak, B.; Weiner, L. M.; Welsh, B., Spectrophotometric Study of Bilirubin Interference in Huang Reaction for Cholesterol. *Clin Chim Acta* **1970,** *30* (3), 697-&.
86. Postma, T.; Stroes, J. A. P., Lipid Screening in Clinical Chemistry. *Clin Chim Acta* **1968,** *22* (4), 569-&.
87. Drezga, Z.; Mikacdev.D, Evaluation of Some Methods for Serum Total Cholesterol. *Clin Chim Acta* **1969,** *26* (2), 317-&.
88. Hunteler, J. L.; Vandersl.W, Modification of Serum-Cholesterol Determination by Continuous-Flow Analysis. *Clin Chim Acta* **1972,** *42* (2), 449-451.
89. Annan, W., Isherwood, D. M. , An automated method for the direct determination of total serum cholesterol. *J. Med. Lab. Technol.* **1969,** *26*, 202.
90. Assous, E. F., Girard, M. L. , Micromethode nouvelle de dosage du cholesterol total directement sur le serum sanguin. *Ann.Bioi. Clin.* **1962,** *20*, 973.

91. Sommers, P. B.; Jatlow, P. I.; Seligson, D., Direct Determination of Total Serum-Cholesterol by Use of Double-Wavelength Spectrophotometry. *Clin Chem* **1975,** *21* (6), 703-707.
92. Jurand, J.; Albertrecht, F., Estimation of Serum Cholesterol. *Clin Chim Acta* **1962,** *7* (4), 522-&.
93. Papastathopoulos, D. S.; Rechnitz, G. A., Enzymatic Cholesterol Determination Using Ion-Selective Membrane Electrodes. *Anal Chem* **1975,** *47* (11), 1792-1796.
94. Noma, A.; Nakayama, K., Polarographic Method for Rapid Microdetermination of Cholesterol with Cholesterol Esterase and Cholesterol Oxidase. *Clin Chem* **1976,** *22* (3), 336-340.
95. Babson, A. L.; Shapiro, P. O.; Phillips, G. E., A New Assay for Cholesterol and Cholesterol Esters in Serum Which Is Not Affected by Bilirubin. *Clin Chim Acta* **1962,** *7* (6), 800-&.
96. Klungsoyr, L.; Haukenes, E.; Closs, K., A Method for the Determination of Cholesterol in Blood Serum. *Clin Chim Acta* **1958,** *3* (6), 514-518.
97. Saifer, A., Photometric Determination of Total and Free Cholesterol and the Cholesterol Ester Ratio of Serum by a Modified Liebermann-Burchard Reaction. *Am J Clin Pathol* **1951,** *21* (1), 24-32.
98. Saifer, A.; Kammerer, O. F., Photometric Determination of Total Cholesterol in Plasma or Serum by a Modified Liebermann-Burchard Reaction. *J Biol Chem* **1946,** *164* (2), 657-677.
99. Bloor, W. R., The determination of cholesterol in blood. *J Biol Chem* **1916,** *24* (3), 227-231.
100. Bloor, W. R., The determination of small amounts of lipid in blood plasma. *J Biol Chem* **1928,** *77* (1), 53-73.
101. Kingsley, G. R.; Schaffert, R. R., Determination of Free and Total Cholesterol by Direct Chloroform Extraction. *J Biol Chem* **1949,** *180* (1), 315-328.
102. Zak, B.; Dickenman, R. C.; White, E. G.; Burnett, H.; Cherney, P. J., Rapid Estimation of Free and Total Cholesterol. *Am J Clin Pathol* **1954,** *24* (11), 1307-1315.
103. Kabara, J. J.; Mclaughlin, J. T.; Riegel, C. A., Quantitative Microdetermination of Cholesterol Using Tomatine as Precipitating Agent. *Anal Chem* **1961,** *33* (2), 305-&.
104. Reinhold, J. G., The determination of blood cholesterol. II. Factors influencing the accuracy of various methods. *Am. J. Clin.Pathol.* **1936,** *6*, 31.
105. Kenny, A. P., The Determination of Cholesterol by the Liebermann-Burchard Reaction. *Biochem J* **1952,** *52* (5), 611-619.
106. Sperry, W. M.; Brand, F. C., The colorimetric determination of cholesterol. *J Biol Chem* **1943,** *150* (2), 315-324.
107. Sperry, W. M.; Webb, M., A Revision of the Schoenheimer-Sperry Method for Cholesterol Determination. *J Biol Chem* **1950,** *187* (1), 97-106.
108. Luden, G., Studies on cholesterol. III. The influence of bile derivatives in Bloor's cholesterol determination. Preliminary report. . *J Biol Chem* **1917,** *29* (3), 463-476.
109. Martinek, R. G., Review of methods for determining cholesterol and cholesterol esters in serum. *J. Am. Med. Technol.* **1970,** *32*, 64.
110. Rosenthal, H. L.; Pfluke, M. L.; Buscaglia, S., A Stable Iron Reagent for Determination of Cholesterol. *J Lab Clin Med* **1957,** *50* (2), 318-322.

111. Mann, G. V., A Method for Measurement of Cholesterol in Blood Serum. *Clin Chem* **1961,** *7* (3), 275-&.
112. Myers, V. C.; Wardell, E. L., The colorimetric estimation of cholesterol in blood, with a note on the estimation of coprosterol in feces. *J Biol Chem* **1918,** *36* (1), 147-156.
113. Reinhold, J. G., Quantitative determination of free cholesterol and cholesterol as esters without digitonin. *P Soc Exp Biol Med* **1935,** *32* (4), 614-616.
114. Reinhold, J. G., The determination of blood cholesterol. I. A comparison of standard colorimetric methods and a modified method with gravimetric determination of digitonin precipitates. *Am. J. Clin.Pathol.* **1936,** *6*, 23.
115. Abell, L. L.; Levy, B. B.; Brodie, B. B.; Kendall, F. E., A Simplified Method for the Estimation of Total Cholesterol in Serum and Demonstration of Its Specificity. *J Biol Chem* **1952,** *195* (1), 357-366.
116. Richmond, W., Use of Cholesterol Oxidase for Assay of Total and Free Cholesterol in Serum by Continuous-Flow Analysis. *Clin Chem* **1976,** *22* (10), 1579-1588.
117. Windaus, A., Uber die Entgiftung der Saponine durch Cholesterin. *Bar. Dtsch. Chem. Ges.* **1909,** *42*, 238.
118. Brown, H. H.; Zlatkis, A.; Zak, B.; Boyle, A. J., Rapid Procedure for Determination of Free Serum Cholesterol. *Anal Chem* **1954,** *26* (2), 397-399.
119. Obermer, E.; Milton, R., The estimation of cholestrol in blood - Supplementary notes on a method utilizing the bernoulli reaction. *J Lab Clin Med* **1937,** *22*, 943-949.
120. Delsal, J. L., Microdosage Colorimetrique Du Cholesterol Libre. *B Soc Chim Biol* **1946,** *28* (4-6), 441-445.
121. Sobel, A. E.; Drekter, I. J.; Natelson, S., Estimation of small amounts of cholesterol as the pyridine cholesteryl sulfate. *J Biol Chem* **1936,** *115* (2), 381-390.
122. Sobel, A. E.; Goodman, J.; Blau, M., Cholesterol in Blood Serum - Studies of Microestimation as the Pyridinium Cholesteryl Sulfate. *Anal Chem* **1951,** *23* (3), 516-519.
123. Eskelson, C. D.; Dunn, A. L.; Cazee, C. R., Effects of Tomatine on Colorimetric Determination of Cholesterol by Zak Procedure. *Clin Chem* **1967,** *13* (6), 468-&.
124. Driscoll, J. L.; Aubuchon, D.; Descotea.M; Martin, H. F., Semiautomated, Specific Routine Serum Cholesterol Determination by Gas-Liquid Chromatography. *Anal Chem* **1971,** *43* (10), 1196-&.
125. Schoenheimer, R.; Sperry, W. M., A micromethod for the determination of free and combined cholesterol. *J Biol Chem* **1934,** *106* (2), 745-760.
126. Fontell, K., Holman, R. T., Lambertsen, G. , Some new methods for separation and analysis of fatty acids and other lipids. *J. Lipid Res.* **1960,** *1*, 391.
127. Vanden Heuvel, W. J. A., Sweeley, C. C., Horning, E. C., Separation of steroids by gas chromatography. *J. Am. Chem. Soc.* **1960,** *82*, 3481.
128. Watts, R.; Carter, T.; Taylor, S., Results for Serum-Cholesterol and Triglycerides by Gas-Liquid-Chromatography, as Compared with a Continuous-Flow Technique. *Clin Chem* **1976,** *22* (10), 1692-1696.
129. Munster, D. J.; Lever, M.; Carrell, R. W., Contributions of Other Sterols to Estimation of Cholesterol. *Clin Chim Acta* **1976,** *68* (2), 167-175.
130. Bjorkhem, I.; Blomstra.R; Svensson, L., Serum-Cholesterol Determination by Mass Fragmentography. *Clin Chim Acta* **1974,** *54* (2), 185-193.

131. Lillienberg, L.; Svanborg, A., Determination of Plasma Cholesterol - Comparison of Gas-Liquid-Chromatographic, Colorimetric and Enzymatic Analyses. *Clin Chim Acta* **1976,** *68* (3), 223-233.
132. Macgee, J.; Ishikawa, T.; Miller, W.; Evans, G.; Steiner, P.; Glueck, C. J., Micromethod for Analysis of Total Plasma Cholesterol Using Gas-Liquid Chromatography. *J Lab Clin Med* **1973,** *82* (4), 656-662.
133. Ishikawa, T. T.; Brazier, J. B.; Stewart, L. E.; Fallat, R. W.; Glueck, C. J., Direct Quantitation of Cholestanol in Plasma by Gas-Liquid-Chromatography. *J Lab Clin Med* **1976,** *87* (2), 345-353.
134. Wycoff, H. D.; Parsons, J., Chromatographic Microassay for Cholesterol and Cholesterol Esters. *Science* **1957,** *125* (3243), 347-348.
135. Shin, Y. S., Silicic Acid Column Chromatography for Microdetermination of Cholesterol, Cholesterol Ester, and Phospholipid from Human Cerebrospinal Fluid. *Anal Biochem* **1963,** *5* (5), 369-&.
136. Egge, H., Murawski, U., Muller, J., Zilliken, F., Microlipidanalysen aus Serum mit dem Eppendorfsystem 3000. *Z. Kim. Chem KIm. Biochem. 8,* 488.
137. Hansen, P. W.; Dam, H., Paper Chromatography and Colorimetric Determination of Free and Esterified Cholesterol in Very Small Amounts of Blood. *Acta Chem Scand* **1957,** *11* (10), 1658-1662.
138. Kim, H. S.; Suzuki, M.; Oneal, R. M., Leukocyte Lipids of Human Blood. *Am J Clin Pathol* **1967,** *48* (3), 314-&.
139. Kritchev.D; Davidson, L. M.; Kim, H. K.; Malhotra, S., Quantitation of Serum-Lipids by a Simple Tlc-Charring Method. *Clin Chim Acta* **1973,** *46* (1), 63-68.
140. Tochston.Jc; Wortmann, W.; Murawec, T.; Kasparow, M., Quantitative Spectrodensitometry of Silica-Gel Thin-Layers Impregnated with Sulfuric-Acid. *J Chromatogr Sci* **1972,** *10* (8), 490-&.
141. Beukers, H.; Veltkamp, W. A.; Hooghwin.Gj, A Method for Determination of Molecular Distribution of Free and Esterified Cholesterol in Serum by Thin-Layer Chromatography. *Clin Chim Acta* **1969,** *25* (3), 403-&.
142. Baets, J. D.; Lezy, W., Improved Method for Lipoprotein Electrophoresis on Cellogel. *Clin Chim Acta* **1971,** *32* (1), 142-&.
143. Anderson, J. T.; Keys, A., Cholesterol in Serum and Lipoprotein Fractions - Its Measurement and Stability. *Clin Chem* **1956,** *2* (3), 145-159.
144. Cohen, L.; Jones, R. J.; Batra, K. V., Determination of Total, Ester and Free Cholesterol in Serum and Serum Lipoproteins. *Clin Chim Acta* **1961,** *6* (5), 613-&.
145. Jencks, W. P.; Hyatt, M. R.; Jetton, M. R.; Mattingly, T. W.; Durrum, E. L., Study of Serum Lipoproteins in Normal and Atherosclerotic Patients by Paper Electrophoretic Techniques. *J Clin Invest* **1956,** *35* (9), 980-990.
146. Jordan, W. J., A Modified Total Serum Cholesterol Method to Eliminate Effect of High Bilirubin Levels. *Clin Chem* **1968,** *14* (1), 31-&.
147. Jordan, W. J.; Knoblock, K. C., Micromethod for Total Serum Cholesterol That Eliminates Interference by High Bilirubin Concentrations. *Clin Chem* **1970,** *16* (1), 18-&.
148. Chin, H. P.; Blankenh.Dh; Chin, T. J., Altered Partition of Serum Cholesterol and Cholesteryl Ester in a Petroleum Ether-Ethanol-Water System after Incubation. *Lipids* **1966,** *1* (4), 285-&.

149. Windaus, A., Uber die quantitative Bestimmung des Cholesterins und der Cholesterinester in einigen normalen und pathologischen Nieren. *Z. Physiol. Chem.* **1910,** *65*, 110.
150. Levine, J., Morganstern, S., Vlastelica, D. , A direct Liebermann-Burchard method for serum cholesterol. In *Technicon Symposiua, Automation in Analytical Chemistry*, N, B. Scova.: 1967; p 1.
151. Zak, B., Epstein, E., Baginski, E. S. , Review and critique of cholesterol methodology. *Ann. Clin. Lab. Sci.* **1972,** *2*, 101.
152. Macgee, J., Enzymatic Determination of Polyunsaturated Fatty Acids. *Anal Chem* **1959,** *31* (2), 298-302.
153. Herb, S. F.; Riemenschneider, R. W., Spectrophotometric Micromethod for Determining Polyunsaturated Fatty Acids. *Anal Chem* **1953,** *25* (6), 953-955.
154. Miles, J.; Glasscock, R.; Aikens, J.; Gerich, J.; Haymond, M., A Micro-Fluorometric Method for the Determination of Free Fatty-Acids in Plasma. *Journal of Lipid Research* **1983,** *24* (1), 96-99.
155. Brice, B. A.; Swain, M. L.; Herb, S. F.; Nichols, P. L.; Riemenschneider, R. W., Standardization of Spectrophotometric Methods for Determination of Polyunsaturated Fatty Acids Using Pure Natural Acids. *J Am Oil Chem Soc* **1952,** *29* (7), 279-287.
156. Herb, S. F.; Riemenschneider, R. W., Influence of Alkali Concentration and Other Factors on the Conjugation of Natural Polyunsaturated Acids as Determined by Ultraviolet Absorption Measurements. *J Am Oil Chem Soc* **1952,** *29* (11), 456-461.
157. *HPLC of acyl lipids*, Lin, J.-T., McKeon, T. A. , Ed. HNB Publishing: New York, 2005.
158. *High-performance liquid chromatography and lipids: A practical guide*, Christie, W. W., Ed. Pergamon Press: Oxford, 1987; pp 144-156.
159. Nikolova-Damyanova, B., Reversed-phase high performance liquid chromatography: General principles and applications to the analysis of fatty acids and triacylglycerols. *Advances in Lipid Methodology—Four*, Christie, W. W., Ed. The Oily Press: Dundee, 1997; pp 193-251.
160. Chen, S. H.; Chuang, Y. J., Analysis of fatty acids by column liquid chromatography. *Anal Chim Acta* **2002,** *465* (1-2), 145-155.
161. Scholfield, C. R., High performance liquid chromatography of fatty acid methyl esters: Analytical separations. *Journal of the American Oil Chemists' Society* **1975,** *52*, 85-91.
162. Nakashima, K., Wada, M., , HPLC determination of fatty acids using derivatization methods. *HPLC of acyl lipids*, Lin, J.-T., McKeon, T. A, Ed. HNB Publishing: New York, 2005; pp 17-41.
163. Gerard, H. C.; Moreau, R. A.; Fett, W. F.; Osman, S. F., Separation and Quantitation of Hydroxy and Epoxy Fatty-Acids by High-Performance Liquid-Chromatography with an Evaporative Light-Scattering Detector. *J Am Oil Chem Soc* **1992,** *69* (4), 301-304.
164. Moreau, R. A., The evaporative light-scattering detector as a tool for the analysis of lipids by HPLC. *HPLC of acyl lipids*, Lin, J.-T., McKeon, T. A, Ed. HNB Publishing: New York, 2005; pp 93-116.
165. Folch, J.; Lees, M.; Stanley, G. H. S., A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues. *J Biol Chem* **1957,** *226* (1), 497-509.

166. Sattler, W.; Puhl, H.; Hayn, M.; Kostner, G. M.; Esterbauer, H., Determination of Fatty-Acids in the Main Lipoprotein Classes by Capillary Gas-Chromatography - Bf3 Methanol Transesterification of Lyophilized Samples Instead of Folch Extraction Gives Higher Yields. *Anal Biochem* **1991,** *198* (1), 184-190.
167. Brondz, I., Development of fatty acid analysis by high-performance liquid chromatography, gas chromatography, and related techniques. *Anal Chim Acta* **2002,** *465* (1-2), 1-37.
168. Lagerstedt, S. A.; Hinrichs, D. R.; Batt, S. M.; Magera, M. J.; Rinaldo, P.; McConnell, J. P., Quantitative determination of plasma C8-C26 total fatty acids for the biochemical diagnosis of nutritional and metabolic disorders. *Mol Genet Metab* **2001,** *73* (1), 38-45.
169. Lindberg, W.; Ohman, J.; Wold, S.; Martens, H., Simultaneous Determination of 5 Different Food Proteins by High-Performance Liquid-Chromatography and Partial Least-Squares Multivariate Calibration. *Anal Chim Acta* **1985,** *174* (Aug), 41-51.
170. Otvos, J. D.; Jeyarajah, E. J.; Bennett, D. W., Quantification of Plasma-Lipoproteins by Proton Nuclear-Magnetic-Resonance Spectroscopy. *Clin Chem* **1991,** *37* (3), 377-386.
171. Jeyarajah, E. J.; Cromwell, W. C.; Otvos, J. D., Lipoprotein particle analysis by nuclear magnetic resonance spectroscopy. *Clin Lab Med* **2006,** *26* (4), 847-+.
172. Serrai, H.; Nadal, L.; Leray, G.; Leroy, B.; Delplanque, B.; de Certaines, J. D., Quantification of plasma lipoprotein fractions by wavelet transform time-domain data processing of the proton nuclear magnetic resonance methylene spectral region. *Nmr Biomed* **1998,** *11* (6), 273-280.
173. Ala-Korpela, M.; Lankinen, N.; Salminen, A.; Suna, T.; Soininen, P.; Laatikainen, R.; Ingman, P.; Jauhiainen, M.; Taskinen, M. R.; Heberger, K.; Kaski, K., The inherent accuracy of H-1 NMR spectroscopy to quantify plasma lipoproteins is subclass dependent. *Atherosclerosis* **2007,** *190* (2), 352-358.
174. Wishart, D. S., Quantitative metabolomics using NMR. *Trends in Analytical Chemistry* **2008,** *27* (3), 228-237.
175. Fuchs, B.; Suss, R.; Schiller, J., An update of MALDI-TOF mass spectrometry in lipid research. *Prog Lipid Res* **2010,** *49* (4), 450-475.
176. *Oleochemical manufacture and applications*, Gunstone, F., Hamilton, R., Ed. Sheffield Academic Press: Sheffield, 2001.
177. Gerpen, J. V., Shanks, B., Pruszko, R., Clements, D., Knothe, G., *Biodiesel production technology*; 2004; p 1.
178. Vereshch, A. G.; Novitska, G. V., Triglyceride Composition of Linseed Oil. *J Am Oil Chem Soc* **1965,** *42* (11), 970-4.
179. Craske, J. D., Separation of Instrumental and Chemical Errors in the Analysis of Oils by Gas-Chromatography - a Collaborative Evaluation. *J Am Oil Chem Soc* **1993,** *70* (4), 325-334.
180. Geladi, P.; Kowalski, B.; Givens, J., The Use of Video in Chemometrics Education - Introduction to Multivariate Data-Analysis. *Chemometr Intell Lab* **1986,** *1* (1), 7-7.
181. Miyake, Y.; Yokomizo, K.; Matsuzaki, N., Determination of unsaturated fatty acid composition by high-resolution nuclear magnetic resonance spectroscopy. *J Am Oil Chem Soc* **1998,** *75* (9), 1091-1094.

182. Shoolery, J. N., Applications of high resolution nuclear magnetic resonance to study of lipids. *Dietary Fats and Health* Perkins, E. G., Ed. American Oil Chemists' Society: Champaign, 1983; pp 220–240.
183. Sacchi, R.; Medina, I.; Aubourg, S. P.; Addeo, F.; Paolillo, L., Proton Nuclear-Magnetic-Resonance Rapid and Structure-Specific Determination of Omega-3 Polyunsaturated Fatty-Acids in Fish Lipids. *J Am Oil Chem Soc* **1993,** *70* (3), 225-228.
184. Aursand, M.; Jorgensen, L.; Grasdalen, H., Positional Distribution of Omega-3-Fatty-Acids in Marine Lipid Triacylglycerols by High-Resolution C-13 Nuclear-Magnetic-Resonance Spectroscopy. *J Am Oil Chem Soc* **1995,** *72* (3), 293-297.
185. Diehlund, B. W. K., Ockels, W. , Ermittelung der fettsäureverteilung durch 13C-NMR spectroskopie. *Fat Sci. Technol.* **1995,** *97*, 115-118.
186. Matsui, T., Iwasaki, H., Matsumoto, F., Osajima, J., Quality evaluation of edible oils by proton nuclear magnetic resonance measurement. *Quality evaluation of edible oils by proton nuclear magnetic resonance measurement.* **1995,** *1*, 94-97.
187. Gunstone, F. D., The Composition of Hydrogenated Fats by High-Resolution C-13 Nuclear-Magnetic-Resonance Spectroscopy. *J Am Oil Chem Soc* **1993,** *70* (10), 965-970.
188. Wanasundara, U. N., Shahidi, F. , Application of NMR spectroscopy to assess oxidative stability of canola and soybean oils. *J. Food Lipids* **1993,** *1*, 15-24.
189. Gunstone, F. D., High-Resolution C-13 Nmr - a Technique for the Study of Lipid Structure and Composition. *Prog Lipid Res* **1994,** *33* (1-2), 19-28.
190. Craig, B. M.; Murty, N. L., Quantitative Fatty Acid Analysis of Vegetable Oils by Gas-Liquid Chromatography. *J Am Oil Chem Soc* **1959,** *36* (11), 549-552.
191. Ayorinde, F. O.; Garvin, K.; Saeed, K., Determination of the fatty acid composition of saponified vegetable oils using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Sp* **2000,** *14* (7), 608-615.
192. Phillips, F. C.; Erdahl, W. L.; Schmit, J. A.; Privett, O. S., Quantitative-Analysis of Triglyceride Species of Vegetable-Oils by High-Performance Liquid-Chromatography Via a Flame Ionization Detector. *Lipids* **1984,** *19* (11), 880-887.
193. Adlof, R. O.; Copes, L. C.; Emken, E. A., Analysis of the Monoenoic Fatty-Acid Distribution in Hydrogenated Vegetable-Oils by Silver-Ion High-Performance Liquid-Chromatography. *J Am Oil Chem Soc* **1995,** *72* (5), 571-574.
194. Tshugaev, L., Gastev, A., Application of the xanthogenic reaction. *Ber. Dtsch. Chem. Ges.* **1910,** *42*, 4631–4634.
195. Hanel, H. K.; Dam, H., Determination of Small Amounts of Total Cholesterol by the Tschugaeff Reaction with a Note on the Determination of Lathosterol. *Acta Chem Scand* **1955,** *9* (4), 677-682.
196. SAS Institute, I. SAS Institute, Inc.: Cary, 2008.
197. Zak, B.; Artiss, J. D., Decreased water: increased solids: distorted serum concentrations. *Microchem J* **2002,** *72* (3), 235-240.
198. *Factor Analysis in Chemistry*, 2nd ed.; Malinowski, E. R., Ed. John Wiley and Sons: New York, 1991.
199. Kramer, R. The Mathworks, Inc.: Natick, 1994.
200. 1998. *Chemometric Techniques for Quantitative Analysis*, Kramer, R., Ed. Marcel Dekker, Inc.: New York.

201. Jurs, P. C.; Kowalski, B. R.; Isenhour, T. L.; Reilley, C. N., Computerized Learning Machines Applied to Chemical Problems - Investigation of Convergence Rate and Predictive Ability of Adaptive Binary Pattern Classifiers. *Anal Chem* **1969,** *41* (6), 690-&.
202. *Chemometrics: Statistics and Computer Application in Analytical Chemistry*, Otto, M., Ed. Wiley-VCH: 1999.
203. Jochum, C.; Jochum, P.; Kowalski, B. R., Error Propagation and Optimal Performance in Multicomponent Analysis. *Anal Chem* **1981,** *53* (1), 85-92.
204. Newell, G. J.; Lee, B., Ridge-Regression - an Alternative to Multiple Linear-Regression for Highly Correlated Data. *J Food Sci* **1981,** *46* (3), 968-969.
205. *Data Handling in Science and Technology Handbook of Chemometrics and Qualimetrics: Part A*, Massart, D. L., Vandeginste, B. G. M., Buydens, L.M.C., De Jong, S., Lewi, P. J., Smeyers-Verbeke, J., Ed. Elsevier Science: 1997; Vol. 20.
206. Hoerl, A. E.; Kennard, R. W., Ridge Regression - Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970,** *12* (1), 55-&.
207. Geladi, P.; Kowalski, B. R., An Example of 2-Block Predictive Partial Least-Squares Regression with Simulated Data. *Anal Chim Acta* **1986,** *185*, 19-32.
208. SAS Institute, I. *SAS JMP Statistical Discovery Software 8.0.2*, Cary, 2009.
209. Geladi, P.; Kowalski, B. R., Partial Least-Squares Regression - a Tutorial. *Anal Chim Acta* **1986,** *185*, 1-17.
210. Thomas, E. V.; Haaland, D. M., Comparison of Multivariate Calibration Methods for Quantitative Spectral-Analysis. *Anal Chem* **1990,** *62* (10), 1091-1099.
211. Thomas, E. V., A Primer on Multivariate Calibration. *Anal Chem* **1994,** *66* (15), A795-A804.
212. Hibbert, D. B., Genetic algorithms in chemistry. *Chemometr Intell Lab* **1993,** *19*, 277-293.
213. Leardi, R.; Boggia, R.; Terrile, M., Genetic Algorithms as a Strategy for Feature-Selection. *J Chemometr* **1992,** *6* (5), 267-281.
214. Leardi, R., Application of a Genetic Algorithm to Feature-Selection under Full Validation Conditions and to Outlier Detection. *J Chemometr* **1994,** *8* (1), 65-79.
215. Jouanrimbaud, D.; Massart, D. L.; Leardi, R.; Denoord, O. E., Genetic Algorithms as a Tool for Wavelength Selection in Multivariate Calibration. *Anal Chem* **1995,** *67* (23), 4295-4301.
216. Wegner, J. K.; Zell, A., Prediction of aqueous solubility and partition coefficient optimized by a genetic algorithm based descriptor selection method. *J Chem Inf Comp Sci* **2003,** *43* (3), 1077-1084.
217. Hemmateenejad, B.; Akhond, M.; Miri, R.; Shamsipur, M., Genetic algorithm applied to the selection of factors in principal component-artificial neural networks: Application to QSAR study of calcium channel antagonist activity of 1,4-dihydropyridines (nifedipine analogous). *J Chem Inf Comp Sci* **2003,** *43* (4), 1328-1334.
218. Cho, H. W.; Kim, S. B.; Jeong, M. K.; Park, Y.; Ziegler, T. R.; Jones, D. P., Genetic algorithm-based feature selection in high-resolution NMR spectra. *Expert Syst Appl* **2008,** *35* (3), 967-975.
219. Cho, S. J.; Hermsmeier, M. A., Genetic algorithm guided selection: Variable selection and subset selection. *J Chem Inf Comp Sci* **2002,** *42* (4), 927-936.

220. Ding, Q.; Small, G. W.; Arnold, M. A., Genetic algorithm-based wavelength selection for the near-infrared determination of glucose in biological matrixes: Initialization strategies and effects of spectral resolution. *Anal Chem* **1998,** *70* (21), 4472-4479.
221. Lucasius, C. B. K., G., Understanding and using genetic algorithms: Part 1. Concepts, properties and context. *Chemometr. Intell. Lab. Syst.* **1993,** *19*, 1-33.
222. Leardi, R., Genetic algorithms in feature selection. *Genetic algorithms in molecular modeling*, Devillers, J., Ed. Academic Press: 1996; p 97.
223. Leardi, R., Genetic algorithms in chemometrics and chemistry: a review. *J Chemometr* **2001,** *15* (7), 559-569.
224. Fraser, A. S., Simulation of Genetic Systems. *J Theor Biol* **1962,** *2* (3), 329-&.
225. Massart, D. L. V., B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J., Genetic algorithms and other global search strategies. *Data handling in science and technology: Handbook of chemometrics and qualimetrics: Part A*, Elsevier: Amsterdam, 1997; Vol. 20A, p 840.
226. Nordling, T. E. M. K., J.; Alander, J. T.; Geladi, P. In *Genetic algorithms as a tool for wavelength selection*, The Proceedings of the 11th Finnish Artificial Intelligence Conference, Vantaa (Finland), 1-3 Sep. 2004; Alander, J. T. A., P.; Hyotyniemi, H., Ed. Finnish Artificial Intelligence Society (FAIS): Vantaa (Finland), 2004; pp 99-113.
227. Leardi, R.; Gonzalez, A. L., Genetic algorithms applied to feature selection in PLS regression: how and when to use them. *Chemometr Intell Lab* **1998,** *41* (2), 195-207.
228. Leardi, R., Application of genetic algorithm-PLS for feature selection in spectral data sets. *J Chemometr* **2000,** *14* (5-6), 643-655.
229. Lepage, G.; Roy, C. C., Specific Methylation of Plasma Nonesterified Fatty-Acids in a One-Step Reaction. *Journal of Lipid Research* **1988,** *29* (2), 227-235.
230. Dumancas, G., Muriuki, M., Purdie, N., and Reilly, L, Chemometric algorithms for the direct determination of lipids in synthetic mixtures and human serum. *Journal of Biotech Research* **2010,** *2* (34-43).
231. Eckhardt, U., Numerical Treatment of Improperly Posed Problems. *Computing* **1976,** *17* (3), 193-206.
232. Elden, L. *A program for interactive regularization*; Dept. of Mathematics, LinkÄoping University, Sweden: 1979.
233. Gayle, J. B.; Bennett, H. D., Consequences of Model Departures on Resolution of Multicomponent Spectra by Multiple-Regression and Linear-Programming. *Anal Chem* **1978,** *50* (14), 2085-2089.
234. Jochum, P.; Schrott, E. L., Deconvolution of Multicomponent Ultraviolet Visible Spectra. *Anal Chim Acta* **1984,** *157* (2), 211-226.
235. Frank, I. E.; Friedman, J. H., A Statistical View of Some Chemometrics Regression Tools. *Technometrics* **1993,** *35* (2), 109-135.
236. *Applied Regression Analysis*, 3rd ed.; Draper, N., Smith, H., Ed. Wiley: New York, 1998.
237. Lindberg, W.; Persson, J. A.; Wold, S., Partial Least-Squares Method for Spectrofluorimetric Analysis of Mixtures of Humic-Acid and Ligninsulfonate. *Anal Chem* **1983,** *55* (4), 643-648.
238. Otto, M.; Wegscheider, W., Spectrophotometric Multicomponent Analysis Applied to Trace-Metal Determinations. *Anal Chem* **1985,** *57* (1), 63-69.

239. Martens, H., Jensen, S.A., Partial least squares regression: A new two-stage NIR calibration method. *Progress in Cereal Chemistry and Technology 5a* Holas, J., Kratochvit, J., Ed. Elsevier Publ.: Amsterdam, 1983; pp 607–647.
240. Geladi, P.; Macdougall, D.; Martens, H., Linearization and Scatter-Correction for near-Infrared Reflectance Spectra of Meat. *Appl Spectrosc* **1985,** *39* (3), 491-500.
241. Martens, M.; Martens, H., Near-Infrared Reflectance Determination of Sensory Quality of Peas. *Appl Spectrosc* **1986,** *40* (3), 303-310.
242. Dunn, W. J.; Stalling, D. L.; Schwartz, T. R.; Hogan, J. W.; Petty, J. D.; Johansson, E.; Wold, S., Pattern-Recognition for Classification and Determination of Polychlorinated-Biphenyls in Environmental-Samples. *Anal Chem* **1984,** *56* (8), 1308-1313.
243. Lindberg, W.; Ohman, J.; Wold, S.; Martens, H., Determination of the Proteins in Mixtures of Meat, Soymeal and Rind from Their Chromatographic Amino-Acid Pattern by the Partial Least-Squares Method. *Anal Chim Acta* **1985,** *171* (May), 1-11.
244. Otto, M.; Thomas, J. D. R., Model Studies on Multiple Channel Analysis of Free Magnesium, Calcium, Sodium, and Potassium at Physiological Concentration Levels with Ion-Selective Electrodes. *Anal Chem* **1985,** *57* (13), 2647-2651.
245. *Chemometrics*, Sharaf, M. A., Illman, D.L., Kowalski, B.R., Ed. John Wiley & Sons, Inc.: New York, 1986; Vol. 82.
246. Cassel, C.; Hackl, P.; Westlund, A. H., Robustness of partial least-squares method for estimating latent variable quality structures. *J Appl Stat* **1999,** *26* (4), 435-446.
247. Tu, J. V., Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemiol* **1996,** *49* (11), 1225-1231.
248. McLaurin, P., Worsfold, P. J., Crane, M., and Norman, P., Multicomponent Analysis by Ultraviolet/Visible Spectrophotometry with Multivariate Calibration. *Anal. Proc.* **1992,** *29*, 65-68.
249. Espinosamansilla, A.; Delapena, A. M.; Salinas, F.; Zamoro, A., Simultaneous Determination of Pesticides by Multivariate Spectral-Analysis and Derivative Spectrophotometry. *Anal Chim Acta* **1992,** *258* (1), 47-53.
250. Espinosa-Mansilla, A., Munoz de la Pena, A., Salinas, F., and Martinez Galera, M., Simultaneous determination of 2-furfuraldehyde, 5-hydroxymethyl-furfuraldehyde and malonaldehyde in mixtures by derivative spectrophotometry and partial least squares analysis. *Anal. Chim. Acta* **1993,** *276*, 141.
251. Espinosa-Mansilla, A., Salinas, F., and de Orbe, I., Simultaneous determination of sulphadiazine, doxycycline, faraltadone and trimethoprim by partial least square multivariate calibration. *Anal. Chim. Acta* **1995,** *313*, 103-112.
252. Wang, L. S.; Huang, Y. W.; Liu, S. L.; Yan, P.; Lin, Y. C., Conjugated linoleic acid induces apoptosis through estrogen receptor alpha in human breast tissue. *Bmc Cancer* **2008,** *8*, -.
253. Eder, K., Gas-Chromatographic Analysis of Fatty-Acid Methyl-Esters. *Journal of Chromatography B-Biomedical Applications* **1995,** *671* (1-2), 113-131.
254. Lima, E. S.; Abdalla, D. S. P., High-performance liquid chromatography of fatty acids in biological samples. *Anal Chim Acta* **2002,** *465* (1-2), 81-91.

255. Rotthauser, B.; Kraus, G.; Schmidt, P. C., Optimization of an effervescent tablet formulation containing spray dried L-leucine and polyethylene glycol 6000 as lubricants using a central composite design. *Eur J Pharm Biopharm* **1998,** *46* (1), 85-94.
256. *Design and analysis in chemical research*, Tranter, R. L., Ed. Sheffield Academic Press: Sheffield 2000.
257. Yan-Jie, L., Rong-Wu, X., Application of central composite design/response surface methodology in pharmacy experiment design. *Chinese Journal of Modern Applied Pharmacy* **2007**, 6-7.
258. Faber, N. K. M., Estimating the uncertainty in estimates of root mean square error of prediction: application to determining the size of an adequate test set in multivariate calibration. *Chemometr Intell Lab* **1999,** *49* (1), 79-89.
259. *Data Handling in Science and Technology: Handbook of Chemometrics and Qualimetrics: Part B.* , Massart, D. L., Vandeginste, B. G. M., Buydens, L.M.C., De Jong, S., Lewi, P. J., Smeyers-Verbeke, J., Ed. Elsevier Science: 1997; Vol. 20.
260. *Practical spectroscopy series: Handbook of near-infrared analysis* 3rd ed.; Burns, D. A., Ciurczak, E.W. , Ed. CRC Press: 2008; Vol. 35.
261. Cannady, J., Artificial neural networks for misuse detection. In *National Information Systems Security Conference*, 1998; pp 368-81.
262. Frenich, A. G.; Galera, M. M.; Vidal, J. L. M.; Garcia, M. D. G., Partial least-squares and principal-component regression of multi-analyte high-performance liquid chromatography with diode-array detection. *J Chromatogr A* **1996,** *727* (1), 27-38.
263. Zhang, M. H.; Xu, Q. S.; Massart, D. L., Boosting partial least squares. *Anal Chem* **2005,** *77* (5), 1423-1431.
264. Jones, S. P., An overview of response surface methodology: second order designs. *Data handling in science and technology: robustness of analytical chemical methods and pharmaceutical technological products*, Hendriks, M. M. W. B. D. B., J. H.; Smilde, A. K., Ed. Elsevier: Amsterdam, Vol. 19, pp 27-29.
265. *Clustering for data mining: a data recovery approach* Mirkin, B., Ed. Chapman & Hall/CRC: 2005; p 114.
266. Shonkwiler, R. In *Parallel genetic algorithms*, Proceedings of the Fifth International Conference on Genetic Algorithms, Forrest, S., Ed. 1993; pp 199-205.
267. Lorber, A. K., B. R., The effect of interferences and calbiration design on accuracy: Implications for sensor and sample selection. *J Chemometr* **1988,** *2* (1), 67-79.
268. Nordling, T. E. M. K., J.; Nyström, J., Bodén, I.; Lindholm-Sethson, B.; Geladi, P.; Alander, J. T. In *Wavelength selection by genetic algorithms in near infrared spectra for melanoma diagnosis* The 3rd European Medical and Biological Engineering Conference, Prague, Czech Republic, Prague, Czech Republic, 2005; pp 1727-1983.
269. Kaplan, C. R.; Gentry, J. W., Use of Condition Numbers for Shortcut Experimental-Design. *Aiche J* **1987,** *33* (4), 681-685.
270. Lorber, A., Error Propagation and Figures of Merit for Quantification by Solving Matrix Equations. *Anal Chem* **1986,** *58* (6), 1167-1172.

APPENDICES

APPENDIX I

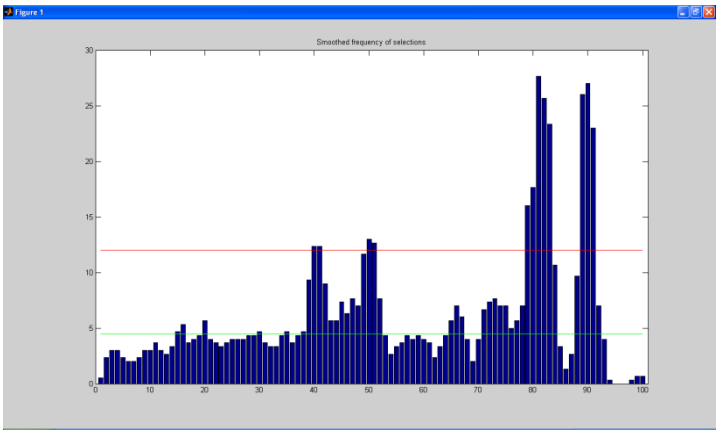
GAPLS default parameters used

Response to be maximized: cross-validated explained variance (%);
Regression method: PLS (the maximum number of components allowed is the optimal number of components determined by cross-validation on the model containing all the variables);
Backward elimination after every 100th evaluation and at the end (if the number of evaluations is not a multiple of 100)
Population size: 30 chromosomes;
Deletion groups: 5;
Maximum number of variables selected in the same chromosome: 30;
Probability of mutation: 1%;
Probability of cross-over: 50%;
Preprocessing: Autoscaling;
Maximum number of components: 15;
Number of runs: 100;
Window size for smoothing: 3.

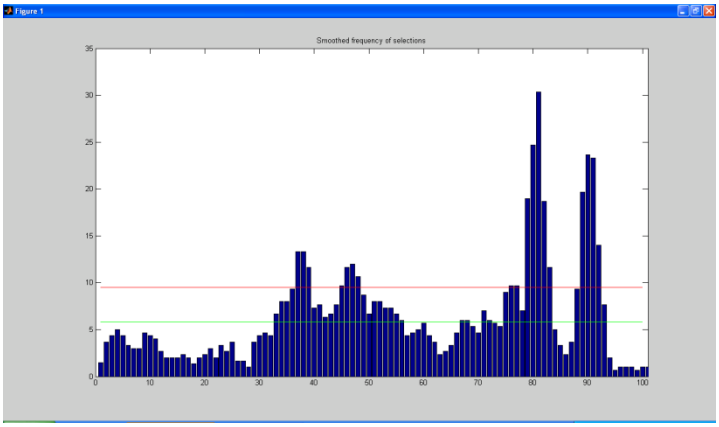
APPENDIX II

Histogram of frequency of selection of each variable for cholesterol trials using the model with minimum RMSECV (green line)

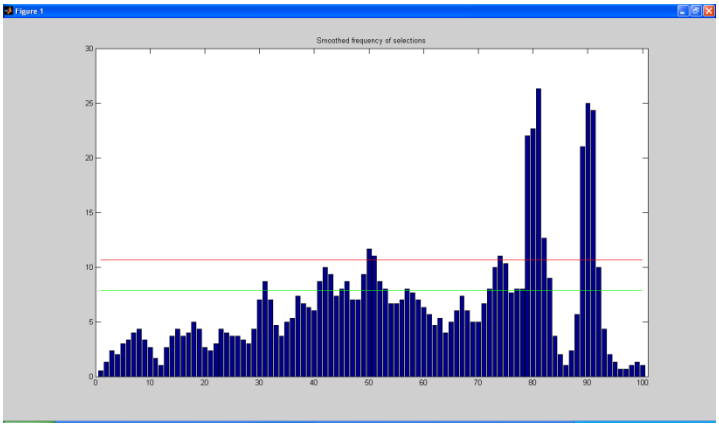
Trial 1



Trial 2



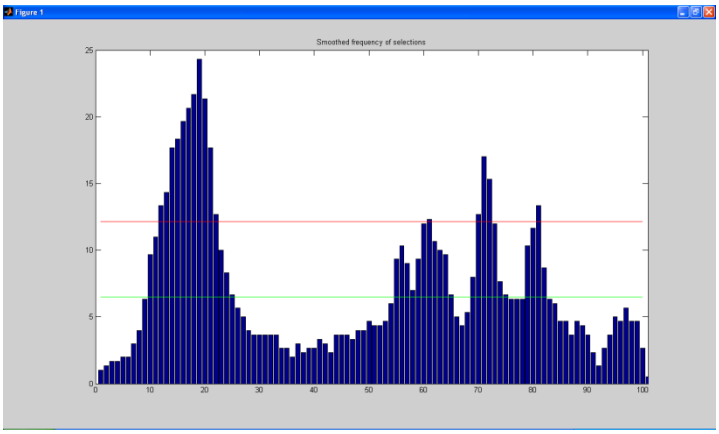
Trial 3



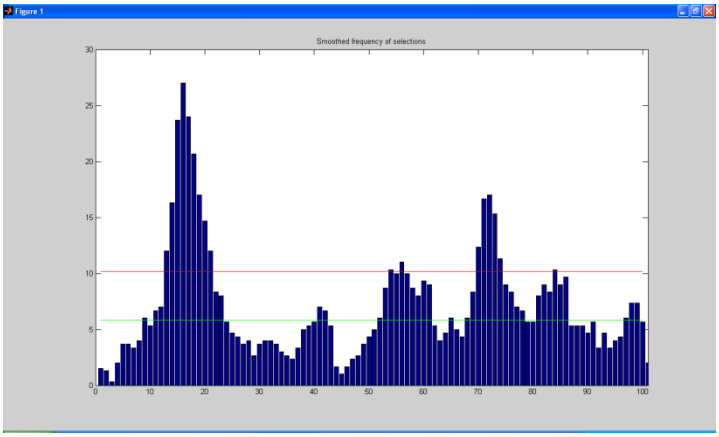
APPENDIX III

Histogram of frequency of selection of each variable for LA trials using the model with minimum RMSECV (green line)

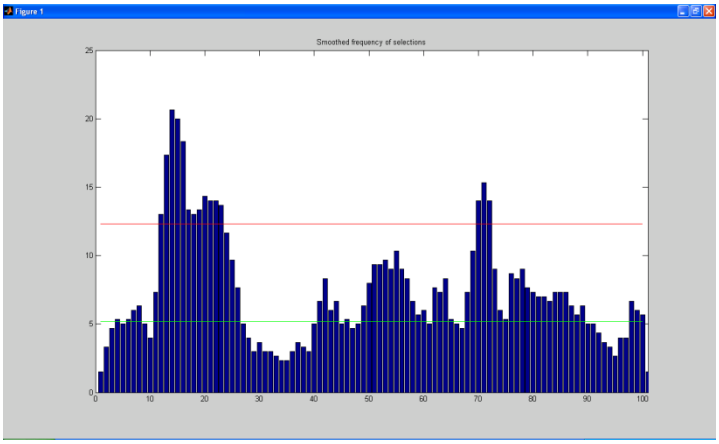
Trial 1



Trial 2



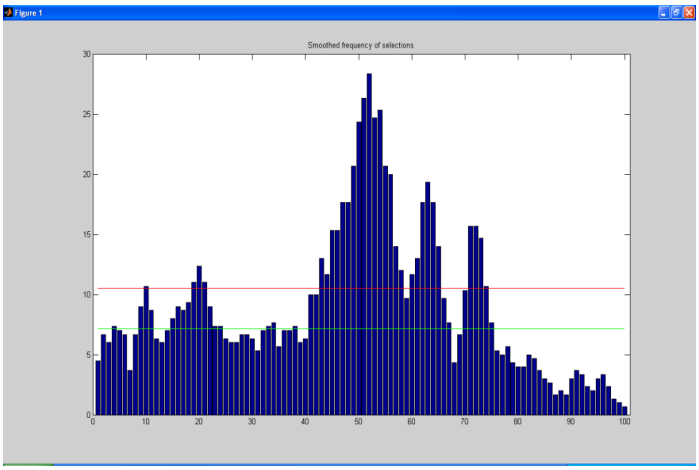
Trial 3



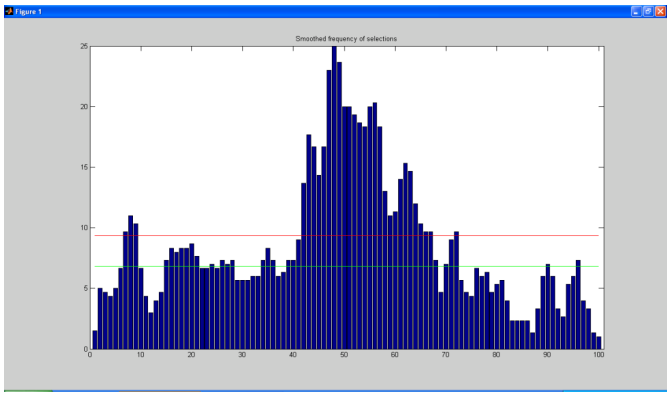
APPENDIX IV

Histogram of frequency of selection of each variable for LNA trials using the model with minimum RMSECV (green line)

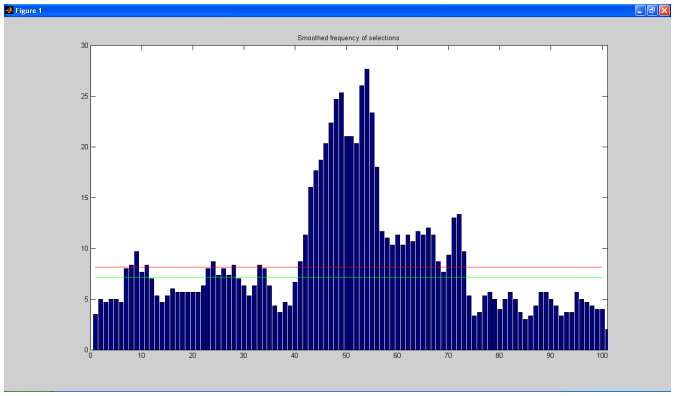
Trial 1



Trial 2



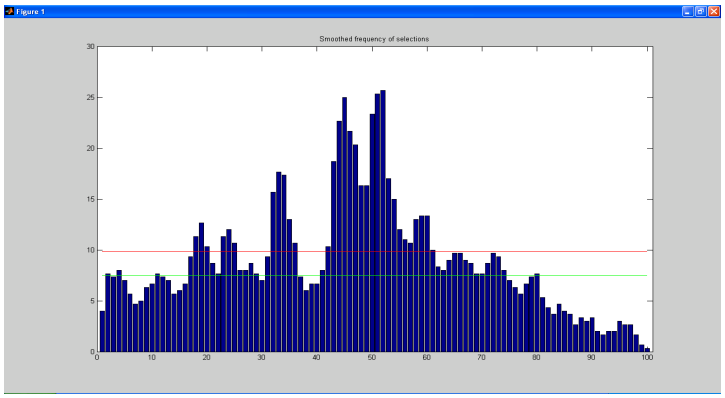
Trial 3



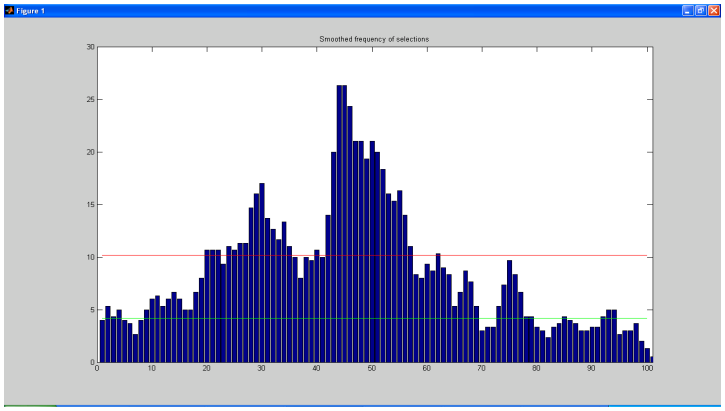
APPENDIX V

Histogram of frequency of selection of each variable for AA trials using the model with minimum RMSECV (green line)

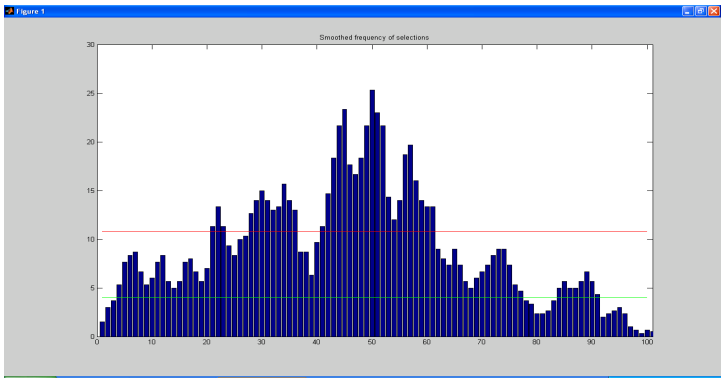
Trial 1



Trial 2



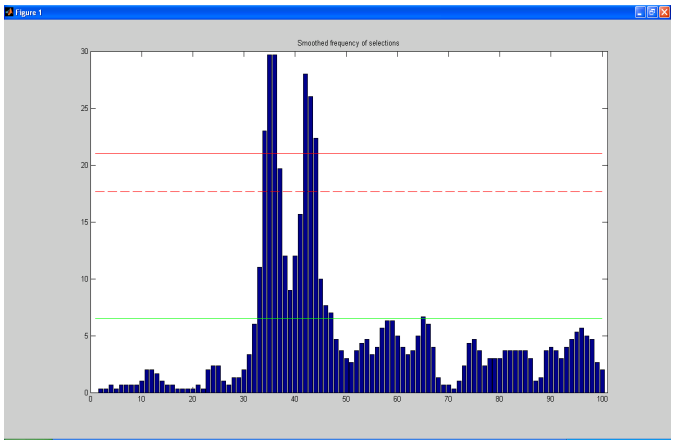
Trial 3



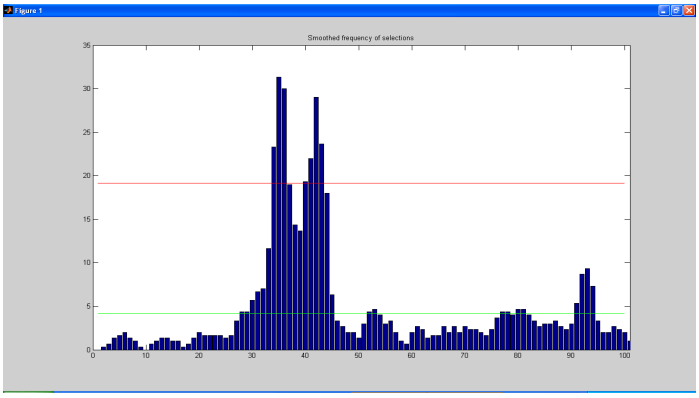
APPENDIX VI

Histogram of frequency of selection of each variable for EPA trials using the model with minimum RMSECV (green line)

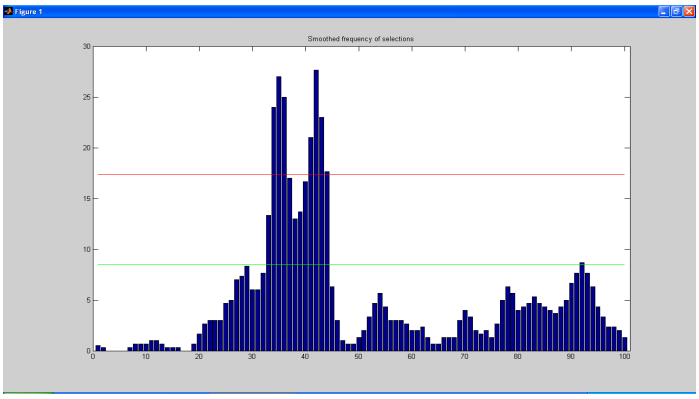
Trial 1



Trial 2



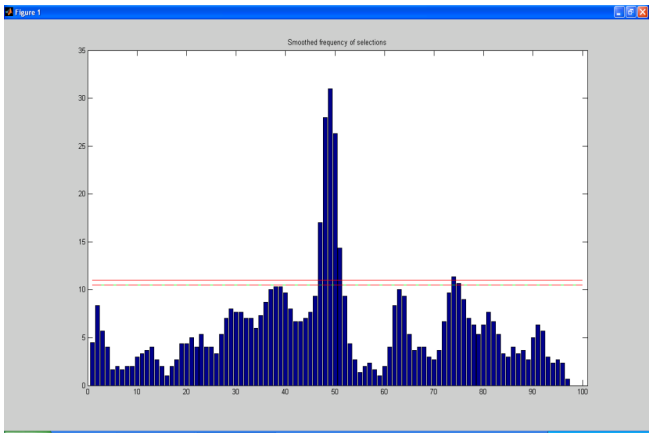
Trial 3



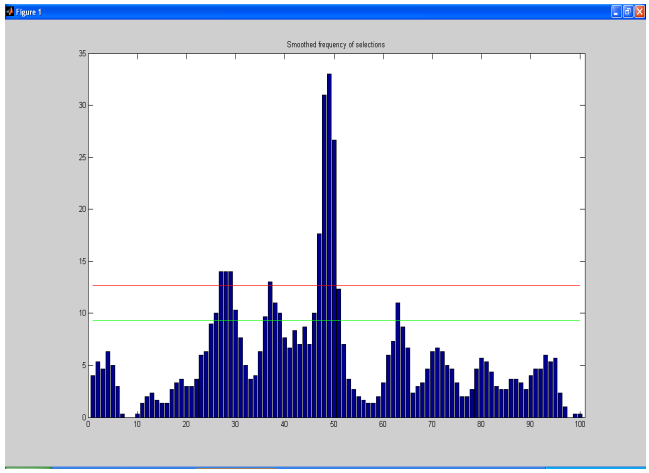
APPENDIX VII

Histogram of frequency of selection of each variable for DHA trials using the model with minimum RMSECV (green line)

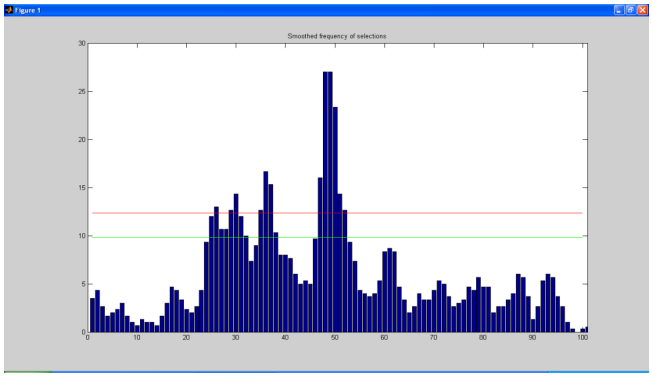
Trial 1



Trial 2



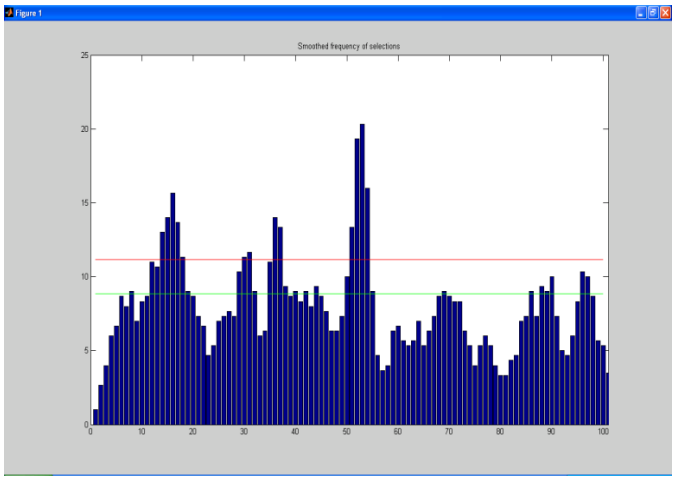
Trial 3



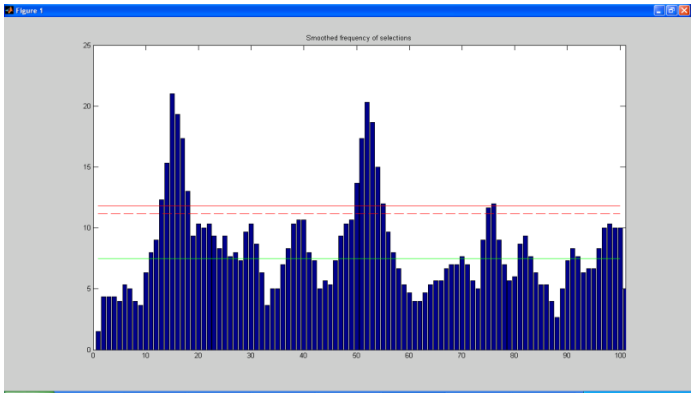
APPENDIX VIII

Histogram of frequency of selection of each variable for CLA trials using the model with minimum RMSECV (green line)

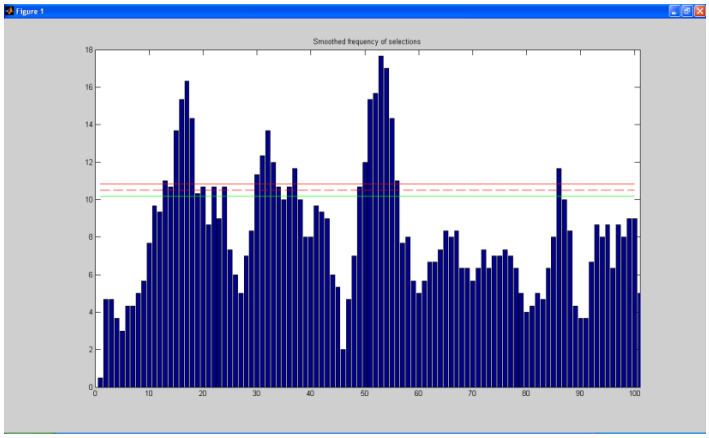
Trial 1



Trial 2



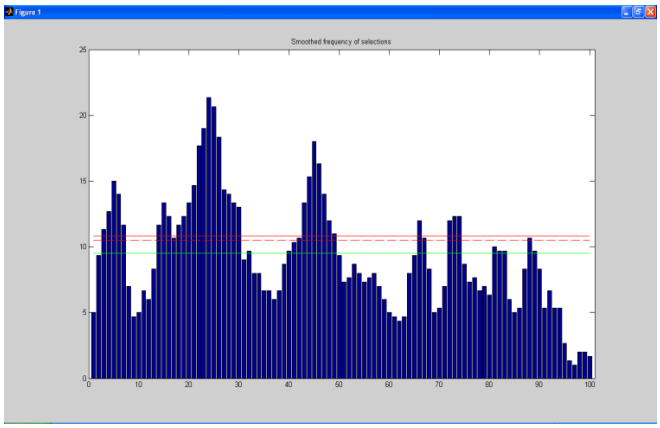
Trial 3



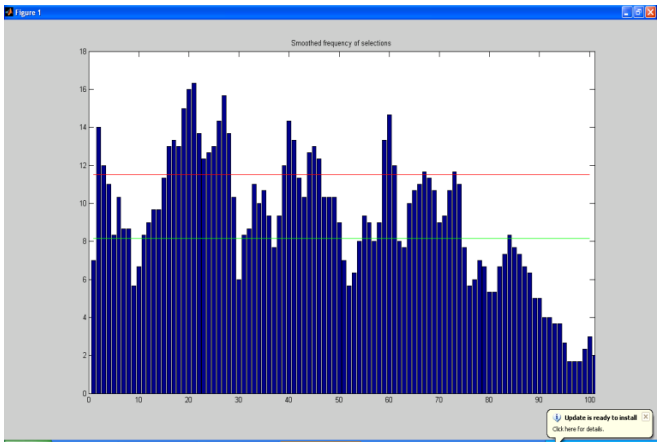
APPENDIX IX

Histogram of frequency of selection of each variable for OA trials using the model with minimum RMSECV (green line)

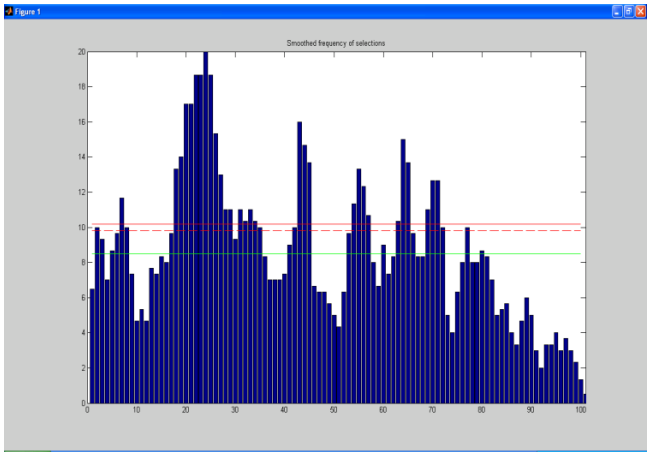
Trial 1



Trial 2



Trial 3



APPENDIX X

Copy of Oklahoma State University Wellness Center blood testing consent form.

OSU WELLNESS CENTER
INFORMED CONSENT FORM

BLOOD TESTING

Explanation of Test

The blood test you are about to undergo is part of the Oklahoma State University Wellness Program. The test includes selected blood variables analyzed from the fingerstick method or from a venous sample.

It will be determined, prior to testing, that this test is appropriate and safe for you. All testing will be conducted by trained personnel and procedures will be explained to your satisfaction at the outset.

Possible Risks

The potential risks associated with the venipuncture/fingerstick are: (1) Venipuncture/fingerstick may cause some pain or discomfort. The exact amount, if any will be dependent upon individual preconceptions and pain threshold levels. (2) Possible hematoma (bruising) at the venipuncture/fingerstick site following the procedure. The occurrence or non-occurrence will be dependent upon bleeding/coagulation times and adherence to instructions pertaining to holding a cotton ball against the venipuncture/fingerstick site, with pressure, for five minutes following extraction of the needle or following the fingerstick. (3) Slight risk of infection. Any break in the integrity of the skin is associated with a small degree of infection risk. However, if directions are followed the risk is very small.

Consent by Subject

Information, which is obtained in the health screening, will be treated as privileged and confidential. IF USED FOR STATISTICAL RESEARCH PURPOSES, INDIVIDUAL IDENTITIES WILL NOT BE RELEASED. IF RESULTS FROM THE RESEARCH ARE PUBLISHED, NO INDIVIDUALS WILL BE NAMED IN THE ANALYSIS OF THE DATA. DATA will be statistically analyzed in an aggregate manner. The Wellness Center professional/medical staff RESERVE THE RIGHT TO contact you if your results are outside normal REFERENCE limits for recommendations and educational opportunities. In addition, if indicated, a small amount of the blood drawn may be used for research in alternative cholesterol testing.

I have read the foregoing, I understand it, and any questions which may have occurred to me have been answered to my satisfaction.

Date_____

Subject Signature_____

VITA

Gerard G. Dumancas

Candidate for the Degree of

Doctor of Philosophy

Thesis: SIMULTANEOUS SPECTROPHOTOMETRIC AND CHEMOMETRIC DETERMINATION OF CHOLESTEROL AND MONO-/POLYUNSATURATED FATTY ACIDS

Major Field: Chemistry

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Chemistry at Oklahoma State University, Stillwater, Oklahoma in May 2012.

Completed the requirements for the Bachelor of Science in Chemistry at the University of the Philippines, Miag-ao, Iloilo, Philippines in 2005.

Experience: Teaching Assistant (General Chemistry and Organic Chemistry), Oklahoma State University, 2008-2012;
Research Assistant, Oklahoma State University, 2007-2008;
General Chemistry, Physical Chemistry, and Mathematics Instructor, Iloilo American Memorial School, 2006

Professional Memberships: American Chemical Society
Royal Society of Chemistry
American Oil Chemists Society
International Association of Engineers
Society for Applied Spectroscopy
Phi Kappa Phi
Phi Beta Delta
Golden Key International Honor Society

Name: Gerard G. Dumancas Date of Degree: May 2012

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: SIMULTANEOUS SPECTROPHOTOMETRIC AND CHEMOMETRIC DETERMINATION OF CHOLESTEROL AND MONO-/POLYUNSATURATED FATTY ACIDS

Pages in Study: 118 Candidate for the Degree of Doctor of Philosophy

Major Field: Chemistry

Scope and Method of Study: The ultimate goal of this research project was to complete the development of a simple, direct alternative method for the simultaneous quantitative determination of cholesterol and polyunsaturated fatty acids (PUFAs) in human serum by exploitation of various chemometric algorithms and consequent validation with the gas chromatography-mass spectrometry (GC-MS). In addition, oleic acid (OA) was also added as the eighth component and the performance of the various chemometric algorithms were compared. The study was also extended to various food and biological samples and chemometric algorithms were applied to obtain meaningful information of the data set.

Findings and Conclusions: For the first part of the study, ridge regression (RR), P-matrix (PM), principal component regression (PCR), and partial least squares (PLS2) algorithms performed quite equally well enough than the K-matrix (KM) approach when applied to the study of prepared mixtures (synthetic sera) in chloroform solutions. The PLS in the form of PLS2 model was tested for intact human serum specimens, and yielded results for ω-3 and ω-6 PUFA data that are comparable when using the GC-MS. Similar results were also derived for the between-methods ω-6/ω-3 ratios. The first part of the study, therefore, showed the dominance of PLS2 over the other chemometric models.

The second part of the study showed that PLS1 algorithm yielded the least root mean square error of prediction (RMSEP) for all the lipid components as compared to all other algorithms. PLS1 yielded molar concentrations quite comparable with the GC-MS in the actual human serum samples. Inclusion of OA yielded high RMSEP despite attempts of utilizing the most robust algorithms like PLS1, PLS2, and PCR. GAPLS was able to successfully reduce the RMSEP for all the components over the non-GA PLS1 approach except for EPA and DHA. The spiking of human serum samples was also done in the study but the task is considered tedious for a typical clinical setting.

In a study involving OA, LA, and LNA in vegetable oils, it has been shown that PCR, PLS2, and PLS1 algorithms compared quite equally well in the prediction sets and that PLS2 mostly yielded a better performance than PLS1 and PCR algorithms in the unknown samples.

An extension of the assay was performed for the pattern recognition of biological and food samples. The assay was able to discriminate eleven clusters corresponding to different food and biological samples.

The study has shown how the Purdie assay coupled with chemometric algorithms might provide alternatives to separations methods for the direct determination of lipids in human serum, vegetable oils, and their synthetic models. The advantages of this simple technology are the reduction in time and costs.

ADVISER'S APPROVAL: Dr. Neil Purdie