

RAMAN SPECTROMETRIC STUDY OF INFLUENCE OF ACETAMINOPHEN ON FEMALE MICE'S REPRODUCTIVE HORMONE LEVELS IN BLOOD

ORIGINALITY REPORT	Seen and approved		26/11/2020
14%	Dr. Birech Zephania 		
	9%	10%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

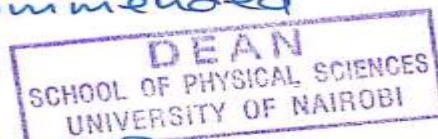
PRIMARY SOURCES

- | | | |
|----------|--|-----|
| 1 | Submitted to University of Nairobi
Student Paper | 1% |
| 2 | Zephania Birech, Annah M. Ondieki, Robert I.I. Opati, Peter W. Mwangi. "Low cost Raman sample substrates from conductive silver paint smear for Raman spectroscopic screening of metabolic diseases in whole blood", Vibrational Spectroscopy, 2020
Publication | 1% |
| 3 | Jonatan R. Ruiz, Jorge Ramirez-Lechuga, Francisco B. Ortega, José Castro-Piñero et al. "Artificial neural network-based equation for estimating VO2max from the 20m shuttle run test in adolescents", Artificial Intelligence in Medicine, 2008
Publication | 1% |
| 4 | Giuseppe Pezzotti, Wenliang Zhu, Tetsuya Adachi, Satoshi Horiguchi, Elia Marin, Francesco Boschetto, Eriko Ogitani, Osam | <1% |

SI index is within the recommended limit



30/11/2020



UNIVERSITY OF NAIROBI

Declaration of Originality Form

This form must be completed and signed for all works submitted to the University for examination.

Name of Student: Ondieki Moraa Annah

Registration Number: I56/14382/2018

College: College of Physical and Biological Sciences

Faculty/School/Institute: Physical Sciences

Department: Physics

Course Name: MSc Thesis

Title of the work: Raman Spectrometric Study of Influence of Acetaminophen on Female Mice's
Reproductive Hormone Levels in Blood

DECLARATION

1. I understand what Plagiarism is and I am aware of the University's policy in this regard.
2. I declare that this Thesis (Thesis, project, essay, assignment, paper, report, etc) is my original work and has not been submitted elsewhere for examination, award of degree or publication. Where other people's work, or my own work has been used, this has been properly been acknowledged and referenced in accordance with the University of Nairobi's requirements.
3. I have not sought or used the services of any professional agencies to produce this work.
4. I have not allowed, and shall not allow anyone to copy my work with an intention of passing it off as his/her own work.
5. I understand that any false claim in respect of this work shall result in disciplinary action, in accordance to the University Plagiarism Policy.

Signature:



Date: 26th November 2020



UNIVERSITY OF NAIROBI

**RAMAN SPECTROMETRIC STUDY OF INFLUENCE OF
ACETAMINOPHEN ON FEMALE MICE'S REPRODUCTIVE
HORMONE LEVELS IN BLOOD**

BY

ONDIEKI MORAA ANNAH

I56/14382/2018

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science
in Physics of the University of Nairobi

November 2020

DECLARATION

I declare that this thesis is my original research idea and has not been submitted anywhere towards approval for research. Where other peoples' work has been used, this has been properly acknowledged and referenced in accordance with the University of Nairobi requirements.

Signature:



Date: 26th November 2020

Ondieki Moraa Annah

I56/14382/2018

Department of Physics, University of Nairobi

This thesis is submitted with our approval as University supervisors

Signature

Date

Dr. Zephania Birech
Department of Physics
University of Nairobi
P.O. Box 30197-00100
Nairobi-Kenya
birech@uonbi.ac.ke



27th November 2020

Prof. Kenneth A. Kaduki
Department of Physics
University of Nairobi
P.O. Box 30197-00100
Nairobi-Kenya
kaduki@uonbi.ac.ke



26th November 2020

Dr. Catherine Kaluwa
Department of Veterinary
Anatomy and Physiology
University of Nairobi
P.O. Box 30197-00100
Nairobi-Kenya
kaingucatherine@gmail.com



27th November 2020

DEDICATION

This thesis is dedicated to my beloved parents (Dr. Peterson Ondieki Osero and Mrs. Hellen Jane Ondieki), my siblings (Sarah, Daniel, and David), my loving longtime friend (Hannington Anekha) and our lovely daughter (Kaylah Ashley).

ACKNOWLEDGEMENTS

I would like to recognize the presence of the Almighty God in my life and the strength he has given me throughout the time of my studies. In a special way, I would like to express my many thanks to my supervisors; Dr. Zephaniah Birech, Professor Kenneth Kaduki, and Dr. Catherine Kaluwa, for allowing me to perform research in a very interesting field and for the constant support they provided throughout my research period.

Many thanks to Miss Anne Ndeke and Mr. Daniel Kwoba, the Laboratory Technologist of the Department of Veterinary Anatomy and Physiology at the University of Nairobi for helping the research work during the sample preparation stage. Also, special appreciation to Mr. Dickson Omucheni, the Chief Technologist of the Department of Physics at the University of Nairobi for guiding me through the Raman spectroscopy measurements.

I sincerely express my gratitude to the International Science Programme (ISP) – Sweden for sponsoring my degree in Master of Science (Physics) through the payment of fees and monthly stipend without which I would not have been able to undertake this research. I can't forget to acknowledge Dr. Alix Dehayem Massop of the Department of Physics at the University of Nairobi for her support, encouragement, and advice as a female physicist and a mentor. Also wish to acknowledge the members of the laser spectroscopy research group for providing an encouraging research environment.

Many thanks to Ondieki family led by my lovely supportive father Dr. Peterson Ondieki, and my caring mother Mrs. Jane Ondieki, for always pushing me forward even when the research was challenging. Finally, special thanks to the love of my life and the father of my daughter Hannington, and our lovely daughter, Ashley for understanding and supporting me throughout the research period.

ABSTRACT

This study presents work on the effects of acetaminophen on female mice's reproductive hormone levels in blood using Raman spectroscopy together with chemometrics. Here weak Raman signals are enhanced by a factor of 3.94 when conductive silver paste smeared glass slides are used as Raman substrates. These Raman substrates were characterized using Raman spectroscopy upon 785 nm excitation and for the first seven days after preparation, they were found to possess chemical stability. The substrates were applied in four female hormones (estradiol, follicle-stimulating hormone (FSH), luteinizing hormone (LH) and progesterone) characterization and level determination in blood using Raman spectroscopy. The spectral profiles of respective standard hormones (with no blood) displayed common (480, 1244 and 1454 cm^{-1}) and other bands (540 cm^{-1} (in LH and progesterone), 1005 cm^{-1} (estradiol and FSH), 880 cm^{-1} (FSH), 837 cm^{-1} (FSH and LH) and 1360 cm^{-1} (progesterone). The spectral data set analyses to reveal these bands was done using Principal Component Analysis (PCA) and analysis of variance (ANOVA). In order to identify biomarker bands of the respective hormones in the blood, Raman spectroscopy of simulate samples (prepared by mixing each separately with male mouse's blood at different concentrations) were done. The Raman experimental parameters were: Excitation wavelength, 785 nm; spectra accumulation, 10; exposure time, 10 s; center wavelength, 1000 cm^{-1} ; microscope objective, $\times 10$, 0.3 numerical aperture; 600 lines grating; and $\approx 68.5 \mu\text{m}$ beam spot size. The biomarker bands were identified to be centered around wavenumbers 668, 902, 1011 cm^{-1} for estradiol; 1219 and 1296 cm^{-1} for FSH; 1440 cm^{-1} for LH and 1569 cm^{-1} for progesterone. These bands exhibited intensity variation with the concentration of the respective hormones in the blood and were used in determining the levels in the blood of treated (acetaminophen administered) and normal (untreated/control) mice. The bands (biomarker bands for each hormone) were used to build Artificial Neural Network (ANN) models to achieve quantitative analysis. A determination coefficient (R^2) and root mean square error (RMSE) values of greater than 98.05% and less than 0.0602 respectively were estimated. The high (low) value of R^2 (RMSE) indicated a good model for concentration prediction. In this work, upon applying the calibrated ANN model in hormone concentration level determination in blood, it was found that mice treated with acetaminophen had on average a level increase from normal of 29%, 138%, 1.1% and 44% in estradiol, FSH, LH and progesterone hormones respectively. Such large hormone level modification away from normal has adverse effects on

fertility. These results served to prove that application of Raman spectroscopy together with chemometrics has a great potential in use for hormone level determination in blood. Besides, the limits of detection were estimated to be 38 pg/ml, 0.45 mIU/ml, 0.69 mIU/ml and 7.14 ng/ml for estradiol, follicle-stimulating hormone, luteinizing hormone and progesterone respectively. These values were way lower than reported values of HPLC and ELISA methods thus indicating the high sensitivity of the Raman spectroscopic method.

TABLE OF CONTENTS

DECLARATION.....	i
DEDICATION.....	ii
ACKNOWLEDGEMENTS.....	iii
ABSTRACT	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ACRONYMS AND ABBREVIATIONS	xiii
CHAPTER ONE: INTRODUCTION.....	1
1.1: Background to the study	1
1.2: Statement of the problem	2
1.3: Objectives of the study.....	3
1.3.1 The general objective of the study	3
1.3.2 Specific objectives.....	3
1.4: Justification of the study	3
CHAPTER TWO: LITERATURE REVIEW.....	5
2.1: Introduction	5
2.2: Acetaminophen’s mechanism of action	5
2.3: Acetaminophen and female reproduction	5
2.4: Hormone detection techniques	6
2.5: Raman spectroscopic applicability in hormone detection	7
2.6: Application of chemometrics in Raman spectroscopy	7
THREE: THEORETICAL BACKGROUND	8
3.1: Introduction	8

3.2: Surface Enhancement Raman spectroscopy.....	8
3.2.1: Principles of Raman Spectroscopic technique.....	8
3.2.2: Surface Enhancement Raman Scattering phenomenon.....	9
.....	10
3.3: Raman instrumentation and interpretation of Raman peaks	10
3.4: Principles of chemometric techniques	11
3.4.1: Principle Component Analysis (PCA)	11
3.4.2: Artificial Neural Networks (ANNs).....	12
CHAPTER FOUR: MATERIALS AND METHODS	15
4.1: Introduction	15
4.2: Experimental Set-Up.....	15
4.2.1: Instrumentation of Laser Raman Spectroscopy	15
.....	16
4.2.2: Instrumental Calibration for laser Raman spectrometer.....	16
4.3: Substrate Preparation for Laser Raman Spectroscopy.....	17
.....	18
4.4: Sample Preparation for Laser Raman Spectroscopy	18
4.4.1: Preparation of pure hormones	18
4.4.2: Preparation of blood from female mice	18
4.4.3 Preparation of calibration standards (Male mice’s blood mixed with standard hormones)	
.....	19
4.5: Raman Spectral Data Preprocessing	20
4.6: Chemometric Analysis of Raman Spectra	20
4.6.1: Principal Component Analysis (PCA)	20
4.6.2: Artificial Neural Networks (ANN)	21

CHAPTER FIVE: RESULTS AND DISCUSSIONS.....	23
5.1: Introduction	23
5.2: Conductive silver smeared microscope glass Raman sample substrates	23
5.3: Characteristic Raman Spectra of female reproductive hormones	27
5.4: Raman Spectra of blood from male Swiss albino mice mixed with the female reproductive hormones.....	31
5.5: Qualitative analysis of blood from treated and normal mice	42
5.6: Quantitative analysis of hormones	45
5.6.1: ANN model on the standard hormones mixed with blood	45
5.6.2: Quantification of hormones in mice blood samples using ANN	49
CHAPTER SIX: CONCLUSION AND RECOMMENDATION	56
6.1: Conclusion.....	56
6.3: Recommendation.....	56
REFERENCES	57
APPENDICES	68
Appendix 1: Neural Network script in R	68
Appendix 2: Principal Component Analysis (PCA) script in R.....	75
Appendix 3: Raman Spectrometer set-up at the Physics department (University of Nairobi) ..	78
Appendix 4: (a) Swiss Albino Mouse and (b)Swiss Albino Mice in a cage	78

LIST OF TABLES

<i>Table 4.1: Table showing different concentrations of hormones in the blood (simulate samples) in units of pg/ml (picogram per milliliter), mIU/ml (milli-international units per milliliter, and ng/ml (nanogram per milliliter) respectively.</i>	19
<i>Table 5.1: Table showing the spectral bands for female reproductive standard hormone with the tentative vibrational assignment.</i>	39
<i>Table 5.2: Table showing the variation of Raman peak intensity observed in acetaminophen treated mouse from a normal mouse, with the tentative vibrational assignment for day one, six, eleven, sixteen and twenty of the study period (letter D, S and I represent decrease, same (no change) and increase respectively).</i>	44
<i>Table 5.3: Table showing the validation metrics of the ANN models constructed, trained and validated for training and validating the models.</i>	46
<i>Table 5.4: Table showing the results of predicted and actual hormone concentration obtained from the developed ANN Models</i>	48
<i>Table 5.5: Table showing the limit of detection for the ANN models developed in this study and the ones reported in the literature.</i>	49
<i>Table 5.6: Table showing the average of female reproductive hormone levels in the blood for different stages of the Estrous cycle from acetaminophen treated and normal mice for a period of 20 days. These values were obtained when Raman spectral data were applied onto the build ANN models.</i>	50
<i>Table 5.8: Table showing z_test results for Raman dataset of blood of acetaminophen treated and untreated mice</i>	55

LIST OF FIGURES

<i>Figure 3.1: Figure showing a simplified energy level diagram of Raman Energy state.</i>	9
<i>Figure 3.2: Figure showing a simplified illustration of SERS phenomenon.</i>	10
<i>Figure 3.3: Figure showing a simplified structure of an ANN model.</i>	12
<i>Figure 3.4: Figure showing a simplified structure of a neuron</i>	13
<i>Figure 4.1: Figure showing the schematic of the Raman spectrometer used in this study.</i>	16
<i>Figure 4.2: Figure showing the Raman Spectrum of a standard silicon wafer upon 785 nm laser excitation.</i>	17
<i>Figure 4.3: Figure showing an illustration of the substrate preparation process.</i>	18
<i>Figure 4.4: Figure showing how the sample was prepared.</i>	19
<i>Figure 4.5: Figure showing a schematic diagram of a simplified structure of ANN procedure used in this study.</i>	22
<i>Figure 5.1: Figure showing Raman spectra obtained from (a) thin and a thick smear of blood on a clean microscope glass slide and a smear of blood on a conductive silver-coated clean microscope glass slide, (b) of silver smear on a clean microscope glass slide, (c) of a clean microscope glass slide.</i>	24
<i>Figure 5.2: Figure showing Raman spectra obtained from conductive silver-coated clean microscope glass slide taken on (a) Day zero, (b) Day two, and (c) Day seven after preparation.</i>	26
<i>Figure 5.3: Figure showing the average of Raman Spectra (mean of 15 spectra each) of (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH), and (d) Progesterone standard hormones.</i>	27
<i>Figure 5.4: Figure showing the two and three-dimensional score plot PCA results for standard hormones. The explained variances are indicated in percentages and were 31, 25, and 11 % for PC1, PC2, and PC3 respectively.</i>	29
<i>Figure 5.5: Figure showing the (i) PC1, (ii) PC2, and (iii) PC3 loadings plot, and (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH) and (d) Progesterone as reference Raman spectra for PCA results of standard hormones.</i>	30
<i>Figure 5.6: Figure showing Raman spectra of different concentrations of (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing hormone (LH), and (d) Progesterone standard hormones mixed with male swiss albino mouse's blood.</i>	31

Figure 5.7: Figure showing Analysis of Variance results for Raman spectra of different concentrations of a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing hormone (LH), and (d) Progesterone standard hormones mixed with male swiss albino mouse's blood.....**32**

Figure 5.8: Figure showing the 3-Dimensional PCA score plots for (a) Estradiol, (b) FSH, (a) LH, and (d) Progesterone standard hormones mixed with blood to form different concentrations. The explained variances are indicated in percentages and were 25, 25, and 15 % for Estradiol, 88, 3.5, and 1.9% for FSH, 42, 23 and 8.5% for LH, and 41, 27, and 8.3% for Progesterone for PC1, PC2, and PC3 respectively.**34**

Figure 5.9: Figure showing the PC1, PC2, and PC3 loadings plot, with low, medium and high concentrations for (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH) and (d) Progesterone as reference Raman spectra for PCA results of standard hormones mixed with male mice's blood.....**35**

Figure 5.10: Figure showing (a) the 3-Dimensional PCA score plot with (b) PC1 versus PC2, (c) PC1 versus PC3, and (d) PC2 versus PC3 2-D PCA score plots for standard hormones mixed with blood. The explained variances are indicated in percentages and were 79, 8.1, and 4.8 % for PC1, PC2, and PC3 respectively.**37**

Figure 5.11: Figure showing the (i) PC1, (ii) PC2, and (iii) PC3 loadings plot, and (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH) and (d) Progesterone as reference Raman spectra for PCA results of standard hormones mixed with blood.....**38**

Figure 5.12: Figure showing (a) analysis of variance results and the average Raman spectra of blood from normal (black curve labeled C for "Control") and acetaminophen treated (red curve labeled T for "Treated") mice taken at (b) day one, (c) day six, (d) day eleven, (e) day sixteen, and (f) day twenty of the treatment period.....**43**

Figure 5.13: Figure showing the regression plots showing predicted against actual concentration obtained from biomarker-Raman band for (a) Estradiol (in picogram/milliliter(pg/ml)), (b) FSH (in milli-International Units/milliliter (mIU/ml)), (c) LH (in milli-International Units/milliliter (mIU/ml))and (d) Progesterone (in nanogram/milliliter (ng/ml))......**47**

Figure 5.14: Figure showing the average of female reproductive hormone levels in the blood for different stages of the Estrous cycle from acetaminophen treated and normal mice for a period of 20 days. These values were obtained when Raman spectral data were applied onto the build ANN models.....51

Figure 5.15: Figure showing a comparison between the predicted concentration of hormones in blood from treated and normal mice in day one, six, eleven, sixteen and twenty days of the treatment period.54

LIST OF ACRONYMS AND ABBREVIATIONS

ANN.....	Artificial Neural Network
ECS.....	Endocannabinoid system
ELISA.....	Enzyme-linked immunosorbent assay
FSH.....	Follicle-stimulating hormones
GnRH.....	Gonadotropin-releasing hormone
HPLC-RIA.....	High-pressure liquid chromatography-radioimmunoassay
LC-MS.....	Liquid chromatography-Mass spectroscopy
LC-MS/MS.....	Liquid chromatography-tandem mass spectroscopy
LH.....	Luteinizing hormone
LSPR.....	Localized Surface Plasmon Resonance
PCA.....	Principal component analysis
SERS.....	Surface enhancement Raman Spectroscopy

CHAPTER ONE: INTRODUCTION

1.1: Background to the study

Techniques for optical spectroscopy such as fluorescence, reflectance, infrared and Raman scattering have attracted interest in biomedical and pharmaceutical applications (Lakhwani *et al.*, 2013). This is because knowledge of the molecular composition of substances such as tissues (Li *et al.*, 2016), whole blood and drugs (Auner *et al.*, 2018) is given. Of these techniques, Raman spectroscopy is preferred because it is highly sensitive with the capability to detect a single molecule (Han *et al.*, 2011). This property has enabled it to be used in the determination of evasive changes in a variety of biological and chemical samples (Downes and Elfick, 2010). Raman spectroscopy works on the principle of light scattering, where the molecules in the sample that is irradiated with strong monochromatic light (laser) are excited and scattered radiation with frequency shift is collected in terms of Raman spectra. The spectra displayed give the fingerprint spectra that are unique to each specific compound making it easy to identify substances (O'Brien *et al.*, 2017).

Raman spectroscopy has shown the potential not only in detecting adulterants in foods such as milk (Alves da Rocha *et al.*, 2015), detecting diseases such as cancer (Vargis *et al.*, 2011) and identifying effects of drugs, but also in detecting hormones such as thyrotropin-releasing hormone (Zhu *et al.*, 2019), estrogen (Liu *et al.*, 2018a) and their variations (Duraipandian *et al.*, 2013). This is because unlike conventional methods, it is highly sensitive, has no sample destruction, requires small amounts of samples and requires virtually no sample preparation. Though Raman spectroscopy has been identified as the best analytical technique, vibrations involving strong dipole moments such as blood are very weak (Wahadoszamen *et al.*, 2015). Therefore, Surface Enhanced Raman Scattering (SERS) is used whereby the conductive silver paste is used as an appropriate substrate to improve Raman signals (Han *et al.*, 2011).

Recent research has shown that variation of female reproductive hormones can result from frequent use of analgesic drugs such as acetaminophen (paracetamol). Acetaminophen, an over the counter drug that is widely used by many people for pain relief, with or without a doctor's prescription (Botting, 2000), has been noted to disrupt the endocrine system (Guiloski *et al.*, 2017). This is because it has been seen to increase the effect of the Endocannabinoid System

(ECS) (Ghanem *et al.*, 2016), which is not only important in neuro-development and moderation of pain but also has an impact on the female reproductive system (Schultz *et al.*, 2012). According to Cohen *et al.*, (2018), acetaminophen use leads to an imbalance of female reproductive hormones thus affecting the maturation of ovarian follicle and ova (Arendrup *et al.*, 2018).

Several techniques that have been used to measure levels of female reproductive hormone include Liquid chromatography-Mass spectroscopy (LC-MS) (Díaz-Cruz *et al.*, 2003), Liquid chromatography-tandem mass spectroscopy (LC-MS/MS), High-pressure liquid chromatography-radioimmunoassay (HPLC-RIA), Enzyme-linked immunosorbent assay (ELISA) (Gavrilova and Lindau, 2009) and plasma biosensor (Li *et al.*, 2016). Previous studies showed that these techniques require complex sample pretreatment, long testing time, and results from them are not widely applied since they are hormone specific (Díaz-Cruz *et al.*, 2003). There is a need to use a highly sensitive and selective technique that is rapid, requires very little or no sample pretreatment, and can be used to measure all hormone levels. SERS has been identified to be a vibrational technique that has the capability of overcoming the limitations of the above-mentioned hormone detection techniques (Liu *et al.*, 2018a).

This research focuses on the use of the SERS technique and application of chemometrics such as principal component analysis (PCA) and artificial neural network (ANN) in identifying the effect of acetaminophen on the female reproductive hormone levels in the blood.

1.2: Statement of the problem

Several methods that have been used in the investigation of the effect of acetaminophen in the menstrual cycle, implantation, and fertility include chromatographic methods, mass spectroscopy (Hombal *et al.*, 2016) and ELISA. These techniques are relatively insensitive, utilize a lot of samples, hormone specific and results obtained from them can't be widely applied (Duraipandian *et al.*, 2013). Therefore there is a need to investigate whether the Raman Spectroscopy can be used to determine the effect of acetaminophen on female reproductive hormone levels in the blood since it is very sensitive, requires very little sample preparation and its results can be used in many applications.

In this study, we succeeded in detecting levels of female reproductive hormones in mice's blood using Raman spectroscopy. This has been achieved by first characterizing respective female

reproductive standard hormones (with no blood and mixed with blood) using Raman spectroscopy, and then applying chemometrics (PCA and ANN) to the Raman data obtained. The results here revealed that estradiol, FSH, LH and progesterone standard hormones have very different spectral patterns allowing each hormone to possess its biomarker band. The biomarker bands obtained were then used to make predictive models. The developed ANN models were then used in estimating concentration levels of each hormone in blood from acetaminophen treated and normal mice. It was seen that in treated mice, the levels of the four hormones were on average elevated in the blood of the treated animals compared with those of the control group (untreated).

1.3: Objectives of the study

1.3.1 The general objective of the study

To demonstrate the ability of Raman spectroscopy coupled with chemometrics in identifying the influence of acetaminophen on female reproductive hormone levels in Swiss Albino mice's blood.

1.3.2 Specific objectives

- i. To determine the Raman spectra of blood samples from acetaminophen treated and untreated mice.
- ii. To determine the Raman spectra of estradiol, follicle-stimulating hormone (FSH), luteinizing hormone (LH), and progesterone standard hormones.
- iii. To analyze the Raman spectra/ Raman spectroscopic data using chemometrics such as PCA for segregation between spectral dataset, and ANN for prediction of hormone levels.

1.4: Justification of the study

Infertility or difficulty in conceiving and maintaining a pregnancy is on the increase for it has been shown to affect 15% of females of reproductive age (Seli *et al.*, 2007). It has been shown that acetaminophen can be one of the driving forces to these problems (Hurtado-Gonzalez and Mitchell, 2017). It would be imperative to investigate, the effect of acetaminophen on female reproductive hormone levels in the blood. Mice have a short estrus cycle, are easy to handle and manipulate (Rudolph *et al.*, 2012) hence ideal as an animal model. Reproduction is driven by the hypothalamic-gonadal-pituitary axis (Paccola *et al.*, 2013) leading to follicle-stimulating

hormone (FSH) and luteinizing hormone (LH) secretion, and ovaries releasing estrogen and progesterone (Ramírez-González *et al.*, 2016).

The several techniques that have been used to study the effect of acetaminophen on female reproductive hormones include sampling, ELISA (Matyas *et al.*, 2015) and some spectroscopic techniques such as mass spectroscopy and vibrational spectroscopy (i.e. infrared spectroscopy). Most of these techniques are insensitive, have a tedious sample pretreatment regime, and require a lot of samples, therefore there is a need to use a technique that overcomes these limitations. Raman spectroscopic technique has been identified as a non-invasive vibrational technique that is non-destructive and can be applied in qualitative and quantitative research (Lakhwani *et al.*, 2013). The method virtually requires no sample preparation, minimal blood samples, and has no sample destruction (Liu *et al.*, 2014). It is also very fast, robust, and highly sensitive (Seli *et al.*, 2007). Despite Raman spectroscopy having all these advantages, it is expensive, cannot eliminate fluorescence. Vibrations involving strong dipole moments such as blood are very weak (Wahadoszamen *et al.*, 2015). To overcome these problems, the conductive silver paste is smeared on glass slides to act as an active substrate enhancing Raman signals. This is because silver nanoparticles are not only used as optical sensors and catalysts, but are also cost-effective, robust to provide the highest electrical field enhancement due to its favorable dielectric function, and can be used in small quantities as compared to gold metallic nanoparticles (Wahadoszamen *et al.*, 2015). Fluorescence is reduced by using a 785nm wavelength diode laser which is capable of reducing the noise level and has a large ratio of signal to noise as compared to a 532nm diode laser (Choquette *et al.*, 2007).

CHAPTER TWO: LITERATURE REVIEW

2.1: Introduction

Studies have shown that many users of analgesic drugs such as acetaminophen and ibuprofen have little or no knowledge of the mechanism of their action (Cohen *et al.*, 2018). Over the last decade, most of the countries in the world have made acetaminophen available without a prescription, because it is thought to be safe at recommended doses (Ngeranwa, 2015). However increasing evidence from recent studies has shown that reproductive age (15-49 years) females on constant acetaminophen use, have disrupted endocrine function (Guiloski *et al.*, 2017), leading to negative reproductive effects (Arendrup *et al.*, 2018).

2.2: Acetaminophen's mechanism of action

Endocannabinoids, their receptors, and enzymes that break them down constitute the Endocannabinoid system (ECS) that exists within the body (Shenglong and Ujendra, 2018). As a postsynaptic neuron is stimulated, endocannabinoids shape, unlock, and migrate back to the presynaptic neuron, where cannabinoid receptors are triggered, thus slowing neuronal function (Mouslech and Valla, 2009). Endocannabinoid receptors include cannabinoid receptor-1 (CB₁) that are mostly located in the brain region and nerve endings where they act to reduce pain sensations, and cannabinoid receptor-2 (CB₂) is located in the peripheral nervous system where they act to reduce inflammation (Fine and Rosenfeld, 2013). Klinger-Gratz *et al.* (2018), reports that acetaminophen alleviates pain through inhibiting CB₁ on the presynaptic membrane, hindering the reuptake of cannabinoids (anandamide) into the neurons thus affecting the function of the brain (Schultz *et al.*, 2012).

2.3: Acetaminophen and female reproduction

Recent studies have shown that the inhibition of endocannabinoids results in female fertility reduction (Lauretta *et al.*, 2019), due to an imbalance of ovarian hormones (estrogen and progesterone) thus affecting ovulation, implantation, and fertility (Gonzaga *et al.*, 2012). According to Arendrup *et al.* (2018), exposure to acetaminophen leads to the irregular menstrual cycle, due to the variation of female reproductive hormone levels bringing about reduced fertility or premature menopause. Matyas *et al.* (2015), reports that acetaminophen use increases progesterone levels during the follicular phase leading to anovulation. According to Hurtado-

Gonzalez and Mitchell (2017), acetaminophen use reduces the fertility of the litter as seen in the study on pregnant rats. Acetaminophen reduced the number of follicles in babies whose mothers were on acetaminophen during gestation (Toda, 2018). Hurtado-Gonzalez *et al.* (2018), reported a drop of gonocyte levels by 17% and 30% on day one and day seven respectively for mice treated with acetaminophen. Cramer *et al.* (1998) suggested that women who use acetaminophen frequently during the menstrual cycle had lower levels of LH than women who have never used the drug.

2.4: Hormone detection techniques

Trace levels of female reproductive hormones such as FSH, estrogen, progesterone, and LH have been detected using different techniques. The first techniques used to quantitate these hormones were chromatographic techniques such as gas chromatography (GC) and liquid chromatography (LC) where LC was reported to have better detection limits than GC. These techniques were reported to have low sensitivity, complex sample preparation, and destroyed the samples (Borio *et al.*, 2012). Spectroscopic methods such as mass spectroscopy were developed as an alternative to chromatographic methods and were later combined with chromatographic techniques to have liquid chromatography-mass spectroscopy (LC-MS) and High-performance liquid chromatography-tandem quadruple mass spectroscopy (HPLC-MS/MS) (Sui *et al.*, 2016). These techniques were grouped as direct methods with high sensitivity but showed limitations. They required relatively complex sample preparation, were time-consuming and the generated results could not be widely applied (Vedad *et al.*, 2018). Indirect methods were later developed which included high-pressure liquid chromatography-radioimmunoassay (HPLC-RIA), enzyme-linked immunosorbent assay (ELISA) and plasma biosensor, which were reported to have high sensitivity but still required complex sample pretreatment and could only probe a specific hormone at a time (Matyas *et al.*, 2015).

Vibrational methods such as Infrared (IR) and Raman spectroscopy have recently been developed and used for structural determination (Downes and Elfick, 2010). They are also used to probe molecular interactions of the biochemical samples (Auner *et al.*, 2018) such as hormones. Raman spectroscopy in the study of estrogen was noted to produce inherently weak signals and fluorescence interference which resulted from the fluorescence of contaminants in the sample (Ullah *et al.*, 2019). Therefore Surface-Enhanced Raman Scattering (SERS) has been

identified to overcome these limitations (Sui *et al.*, 2016). It has advantages such as require minimal sample preparation, has non-destructive nature, is highly sensitive and has rich information content (Gonzaga *et al.*, 2012), enabling quantification and identification of molecules (Seli *et al.*, 2007).

2.5: Raman spectroscopic applicability in hormone detection

Recently, Raman spectroscopy has been known to be among the best vibrational spectroscopy techniques for biological sample analysis (Bergholt *et al.*, 2011). It has been used to identification of bacteria, early detection of cancer (Kanter *et al.*, 2009), detection of adulterants in food, identification of components in a sample (Atkins *et al.*, 2017a), and detection of certain effects of drugs in humans. In hormone detection, SERS spectroscopic technique has been utilized in increasing the sensitivity of detecting the trace amounts of phenolic estrogen (Liu *et al.*, 2018b), and identifying hormonal variations (Duraipandian *et al.*, 2013) where it has been reported to provide the most detailed chemical composition of the substances. Their study showed that high wavenumber confocal Raman spectroscopy along with bio-molecular simulation can be an effective hormonal variation detection diagnostic tool.

2.6: Application of chemometrics in Raman spectroscopy

Researches have shown that to interpret data from spectroscopic techniques such as Raman spectroscopy, chemometrics such as PCA, ANN or both are required (Rousseau *et al.*, 2008; and Ioele *et al.*, 2011). Ioele *et al.* (2011) succeeded in quantifying caffeine in pharmaceutical mixtures using PCA combined with ANN. According to Bonnier and Byrne (2012), PCA can enhance a better understanding of the molecular content of biological samples such as cells. Determination of concentration levels of glucose in whole blood has been achieved using ANN based on data from Raman spectroscopy (Wang *et al.*, 2015). Based on Raman data, Alves da Rocha *et al.* (2015) succeeded in quantifying whey in liquid milk. In the pharmaceutical industry classification of tablets has been achieved using ANN based on spectroscopic data (Galata *et al.*, 2019).

THREE: THEORETICAL BACKGROUND

3.1: Introduction

In this chapter, we discuss principles of Raman Spectroscopic technique, Surface Enhancement Raman Scattering, interpretation of Raman peaks and principles of chemometric tools focusing on PCA and ANN.

3.2: Surface Enhancement Raman spectroscopy

Surface-enhanced Raman spectroscopy (SERS) is a method of enhancing the weak signal intensity of a sample placed close to a metallic nano-surface.

3.2.1: Principles of Raman Spectroscopic technique

Raman spectroscopic technique is a non-destructive technique for detecting the system's rotational, vibrational, and other low-frequency modes (Ferrari, 2001). It uses the concept of light-scattering whereby a sample is illuminated with a powerful monochromatic (laser) source that interacts with the sample's molecular vibrations, phonons, and other excitations (Kalantri *et al.*, 2010). The dispersed radiation ends up being either Rayleigh or Raman scattered. In Rayleigh dispersion, the scattered radiation is of the same wavelength as the incoming laser wavelength and therefore has no information on the sample's vibrational energy levels (Butler *et al.*, 2016). Raman scattering or inelastic scattering is where the wavelength of scattered radiation varies from the wavelength of the incoming laser. This change in wavelength gives information about system vibration modes and leads to qualitative measurement of molecules' biochemical composition (Liu *et al.*, 2018a). Raman scattering is of two forms: Stokes Raman scattering (frequency of scattered radiation is lower than the frequency of the excitation laser) and Anti-stokes Raman scattering (scattered radiation's frequency is higher than the excitation laser frequency). When a molecule interacts with a photon's electrical field as in figure 3:1 below, the energy transfer must satisfy Bohr's frequency condition in which the energy difference, ΔE of two quantized states (Auner *et al.*, 2018) is given as:

$$\Delta E = h\nu_k = \frac{hc}{\lambda_k} = hc\bar{\nu} \quad (3.1)$$

where c is the speed of light, h is Planck's constant, $\bar{\nu}$ is the wavenumber, and

$$\bar{\nu} = \frac{1}{hc} \Delta E \quad (3.2)$$

Equation 3.2 shows that the wavenumber $\bar{\nu}$ is directly proportional to the energy difference ΔE . This relation is used in vibrational spectroscopy where the detected vibrational lines observed are represented as a function of change in the wavenumber rather than as a function of change in frequency (Auner *et al.*, 2018). The relationship between the vibrational frequencies and the molecular structure leads to distinguishable Raman spectra for each molecule (Han *et al.*, 2011). The energy level diagram of Figure 3.1 illustrates Raman energy states (Kalantri *et al.*, 2010).

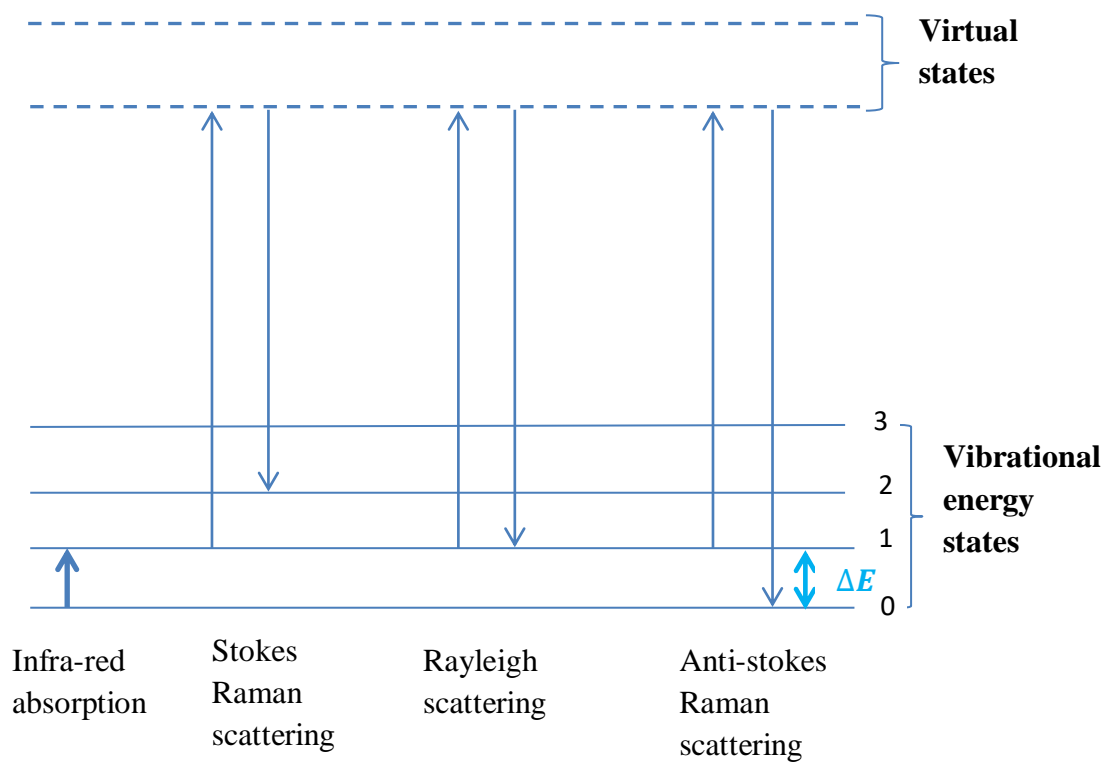


Figure 3.1: Figure showing a simplified energy level diagram of Raman Energy state.

3.2.2: Surface Enhancement Raman Scattering phenomenon

SERS is a phenomenon that greatly enhances a molecule's weak Raman signal by increasing the electron cloud density around metallic nanostructures (Liu *et al.*, 2018b). This enhancement can be of two types, namely electromagnetic enhancement, and chemical enhancement. For both processes to operate at the same time, the sample must be placed very close to a metallic nano-surface (Behnammorshedi and Nazem, 2015). Chemical enhancement occurs when a molecule is directly bound to the metal surface such that there is a generation of charge-transfer

(Wahadoszamen *et al.*, 2015), for example, the energy transfer from the metal to the sample's molecular bonds can be mediated by an electron-hole pair. Electromagnetic enhancement results from the action of localized surface plasmon resonance (LSPR) which occurs when incident radiation (Raman excitation source) of some wavelength interacts with a metallic nanostructured surface generating a strong electromagnetic field around the nanostructured surface (Auner *et al.*, 2018). The large electric field produced and then acts directly on the target molecule like the molecule excited by an intensified light. Of these two mechanisms, researches focus more on electromagnetic enhancement because it is the main cause of SERS (Han *et al.*, 2011). Figure 3.2 shows a simplified illustration of the SERS phenomenon.

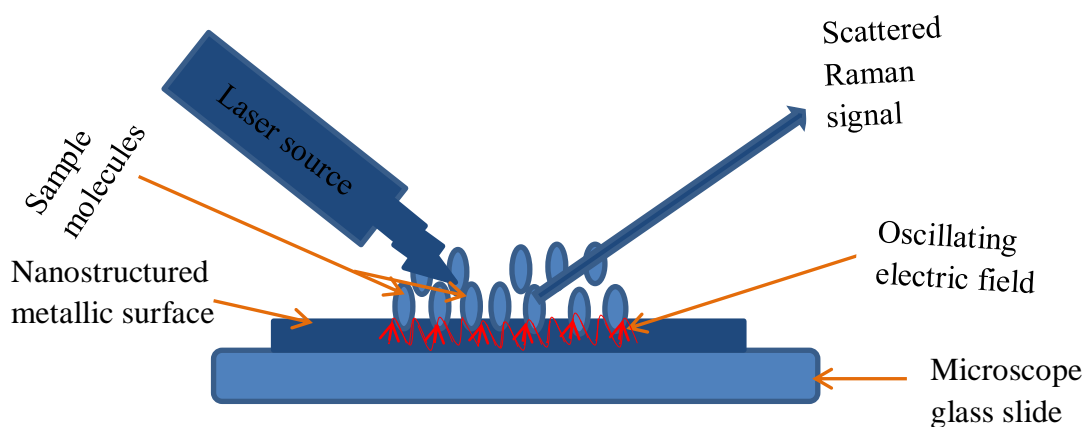


Figure 3.2: Figure showing a simplified illustration of SERS phenomenon.

3.3: Raman instrumentation and interpretation of Raman peaks

In Raman spectroscopy, a laser source is required to excite the sample (Kalantri *et al.*, 2010), a microscope with lenses that focus the laser beam on the sample and capture the scattered radiation, filters that purify the reflected and dispersed light so that only Raman light is collected, a spectrometer (diffraction grating) that splits light into spectral components, a very sensitive detector for detecting weak signals, and a computer to control the whole system, display the spectrum and enable an analysis of the information (Lakhwani *et al.*, 2013).

The spectra displayed give the fingerprint spectra that are unique to each specific compound making it easy to identify substances (O'Brien *et al.*, 2017). To interpret the spectra, therefore, one needs to know that in the Raman spectra each peak corresponds to a specific vibration of the molecule (Winnard *et al.*, 2017).

3.4: Principles of chemometric techniques

Although Raman spectra can be obtained quite easily, to understand and interpret it is quite cumbersome, therefore, chemometric tools are used (Borio *et al.*, 2012). This is because the spectral data obtained are complex and involve large data sets (many variables in dataset). Chemometric tools employ mathematical and statistical methods to extract relevant information from complex data to achieve quality control, quality assurance, and quantitative information (Auner *et al.*, 2018). Chemometrics have ability to analyze the huge spectral distribution and differentiate between different sample's spectra. Thorough data preprocessing helps in implementation of these chemometrics. Here we focused on the use of PCA and ANN as chemometrics for multivariate data analysis.

3.4.1: Principle Component Analysis (PCA)

Principal Component Analysis (PCA) is a data dimensionality reduction technique often used to compress and visualize large amounts of data (Rousseau *et al.*, 2008). It uses an unsupervised method for exploratory data analysis and making predictive models bring out the similarities or differences in data (Ullah *et al.*, 2019). PCA works by clustering data according to their spectral patterns. This is achieved by projecting the original data on the principal components (Li *et al.*, 2011). With PCA the original matrix X which is $m \times n$, is decomposed into:

$$X = YZ^T \tag{3.3}$$

where:

Y is $m \times r_{max}$ representing the first principal component (PC1)

Z^T is the transpose of the matrix Z that is a $n \times r_{max}$ the matrix representing the second principal component (PC2), and

r_{max} is the number of components that are equivalent to the rank of X .

The PC1 describes most of the variance in X while PC2 captures the second largest variance and must be orthogonal to PC1 (Potcoava *et al.*, 2014). The PCA graph contains loadings, scores, and residues. With the help of the score plot, PCA does not only reduce the dimensionality of data but also detects outliers and shows covariance between samples (Davidson *et al.*, 2013). This enables one to find expected or unexpected trends in data. The loadings reveal qualitative

information that leads to molecular band identification in the spectra. The relevant features obtained are then used to make the predictive model, thus reducing the overfitting of the model, enhancing the accuracy of the model, and model training time reduction.

3.4.2: Artificial Neural Networks (ANNs)

ANNs are simplified computational models derived from a branch of computing science known as artificial intelligence that aims to understand and model human brain activity (Haykin, 1999). It is a mathematical technique that uses a collection of parallel and symbolic processing algorithms for data grouping, classification, clustering, and prediction and has been used, for instance, to extract qualitative and quantitative information from Raman spectra (Galata *et al.*, 2019). This is because the quantification process of the raw Raman spectra is quite complicated as its intensity depends not only on the concentration of the sample but also on the power of the laser and the instrumental effects. ANNs have demonstrated the ability to effectively identify patterns, extract features, and extract nonlinear relationships from data (Alves da Rocha *et al.*, 2015). The structure of ANN consists of weights, input, hidden, and output layers as shown in Figure 3.3:

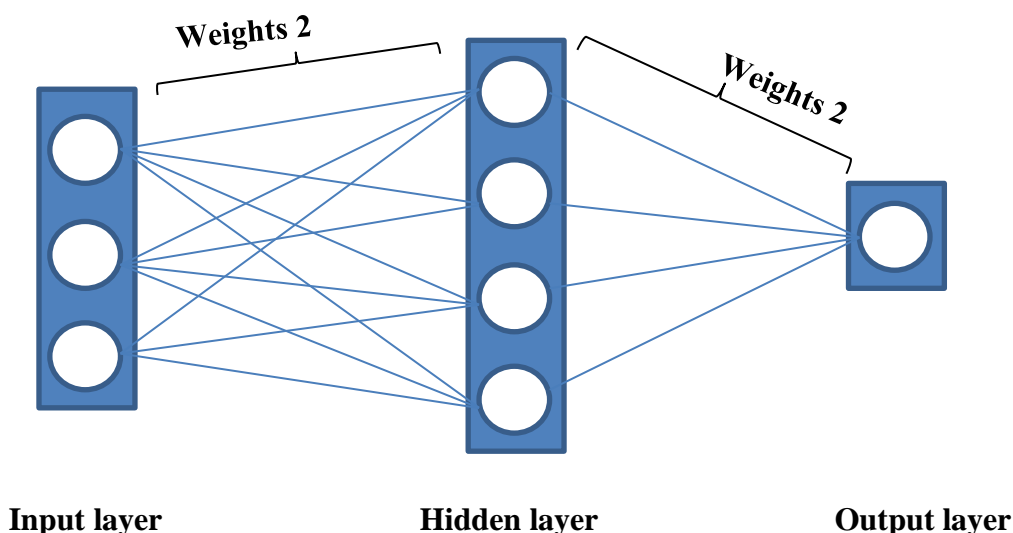


Figure 3.3: Figure showing a simplified structure of an ANN model.

The input layer consists of the independent variables (Raman spectra), while the output layer has the dependent variable (concentration). The hidden layer consists of all possible relations between the input layer and the output layer and allows for the put together the effect on the

output layer of a variety of independent variables. This is like testing all possible interactions in a regression model, without the addition of extra degrees of freedom. The activation function for the hidden and output nodes in the logistic function is given as:

$$f(y) = \frac{1}{1+e^{-y}} \quad (3.4)$$

Where y is the input variables

Hence the function computed by the network is ANN

$$y_1, \dots, y_N = f(w_1 h_1 + w_2 h_2 + \dots + w_N h_N) \quad (3.5)$$

Where w is the weights connected from hidden units to the output unit; and h_i is the function computed by the hidden unit i ,

$$h_i(y_1, \dots, y_N = f(w_{1i}y_1 + w_{2i}y_2 + \dots + w_{Ni}y_N) \quad (3.6)$$

where w_{ji} is the weight of the link from input j to hidden unit i .

In order to obtain the synaptic weights of the ANN, the resilient backpropagation algorithm (Rprop+) was used which is the result of backpropagation. This algorithm changes the weights and bias network with direct adaptation process of weighting based on local gradient information from an iteration of learning, so that the number of iterations needed to reach the target.

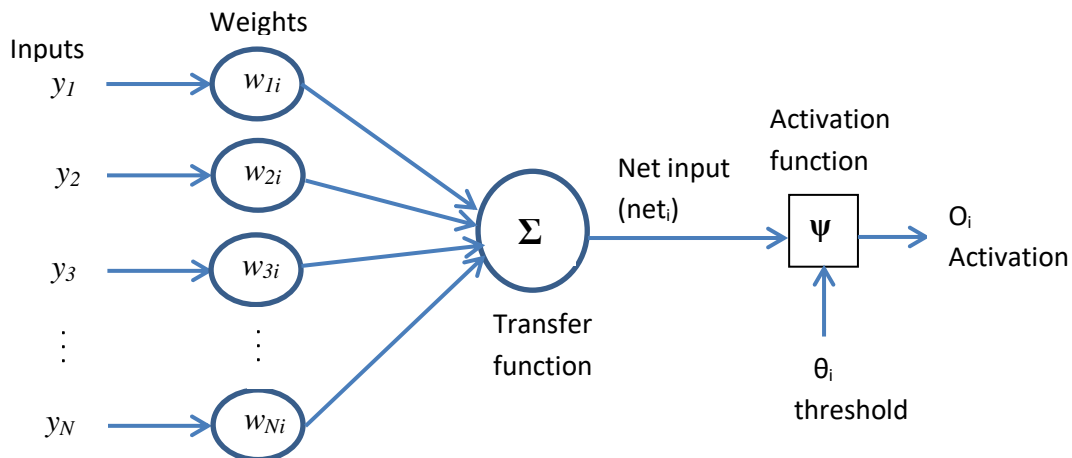


Figure 3.4: Figure showing a simplified structure of a neuron

As in the biological neuron, the integration of the input signals is performed by the cell body. The neuron is activated when the result exceeds a certain minimum value, thus producing a potential action that is sent to the axon (Gautam *et al.*, 2015). The received signals are then processed by each node sending the result to succeeding nodes. Likewise in the artificial neural network, the input layer consists of the independent variables information that is interpreted by

the weights and biases between layers, and then gives dependent variables information at the output layer, thus revealing the relationships in the dataset (Cheng and Sun, 2015). An artificial neuron is formed from three basic elements that include; a set of synapses or connections, each of which is characterized by a weight, an adder that sums the weighted inputs from the respective signals in the connection, producing in the output a linear combination of the inputs, and an activation function to limit the output neuron amplitude (Gautam *et al.*, 2015).

The procedure for developing a neural network includes selecting data, constructing and training a network to obtain the desired values at the output (Wang *et al.*, 2015). The training algorithm modifies the weights on the connections through which the signal is transmitted to minimize the gap between the network outputs and the desired one. Finally, the network's output is evaluated using root mean square error and regression analysis (Galata *et al.*, 2019).

CHAPTER FOUR: MATERIALS AND METHODS

4.1: Introduction

This chapter introduces the materials used in this study and how the experiment was set. Substrate preparation, sample pretreatment, and application of Raman spectroscopy for quantification of female reproductive hormones in simulate samples (i.e. female reproductive standard hormones mixed with blood from male mouse) using ANN models have been discussed.

4.2: Experimental Set-Up

4.2.1: Instrumentation of Laser Raman Spectroscopy

The instrumentation for this study consisted of a laser Raman spectrometer (STR, Airix Corporation) with a spectrometer having diffraction grating of 600, 1200 and 1800 lines/mm, 785nm and 532nm excitation lasers (Princeton Instruments), microscope with x10 to x100 magnification lenses (Max Plan), a CCD camera, and a computer. The microscope (Olympus) focused light on the sample and collected the sample's backscattered radiation using the same objective lens. Stokes and Rayleigh's scattering were filtered out using the notch filter which allowed only the anti-stokes scattering to get into the spectrometer. The spectrometer's diffraction grating split the anti-stokes scattering radiation, dispersed each spectral element by wavenumber, and projected it onto the computer via a charge-coupled device (CCD) camera. Figure 4.1 shows a schematic of the Raman spectrometer.

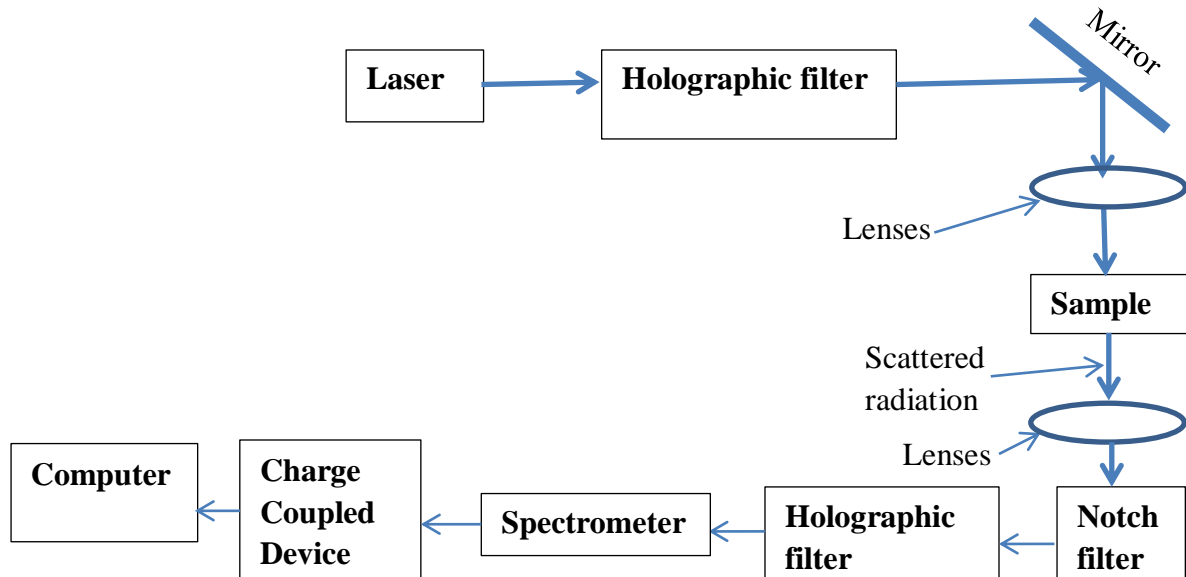


Figure 4.1: Figure showing the schematic of the Raman spectrometer used in this study.

4.2.2: Instrumental Calibration for laser Raman spectrometer

Before starting the experiment, optimization parameters were set as follows to obtain the appropriate Raman spectra; 600lines grating with 1000 cm^{-1} center wavelength to cover a wider spectral range, 10 seconds exposure time, the number of accumulations at 10, the $\times 10$ magnification objective lens with 0.3 numerical aperture, $\approx 68.5\ \mu\text{m}$ beam spot size, 785 nm excitation laser to minimize fluorescence, 50% of $\approx 18.20\text{ mW}$ excitation power got to the sample. On ensuring that the optimization parameters set were maintained, and just before starting the experiment, calibration of the instrument was done using a standard silicon wafer. This was done by replacing the sample under the microscope with a standard silicon wafer and excited using 785nm diode laser to obtain the spectra, and Raman scatter bands maintained at approximately 520.5 cm^{-1} as in figure 4.2.

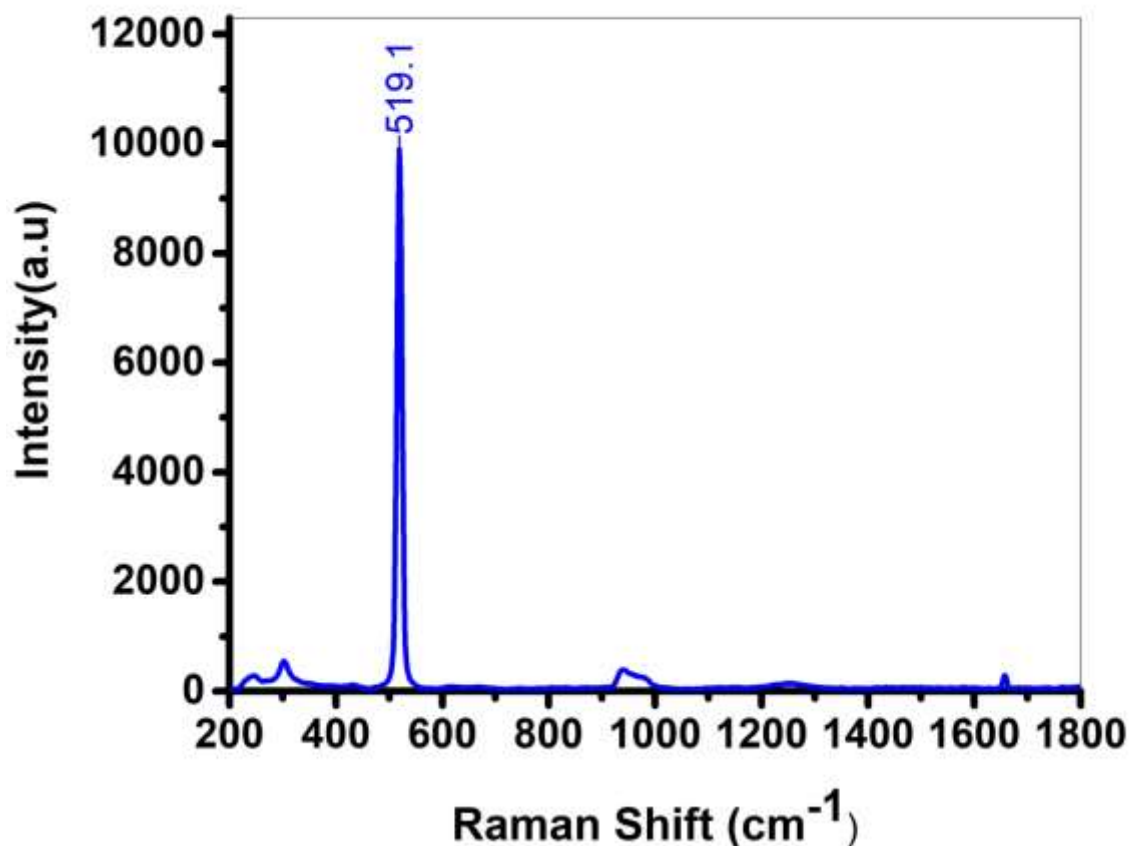


Figure 4.2: Figure showing the Raman Spectrum of a standard silicon wafer upon 785 nm laser excitation.

On ensuring that the system was correctly calibrated, the sample was put under the microscope, irradiated by 785 nm laser to obtain Raman spectral data. Raman Measurements were taken between 7 am and 7 pm. This was done in a dark room to avoid background interference. Five spectra were obtained from different locations of each sample. The procedure above was repeated each time the spectra were collected.

4.3: Substrate Preparation for Laser Raman Spectroscopy

Since SERS substrates used in bioanalytical applications help in obtaining detailed chemical information of biomolecules attached to them, they, therefore, need to be highly sensitive, uniform, and stable (Behnammorshedi and Nazem, 2015). Conductive silver smear substrates have been used in metabolic disease screening, where they have been described to be cheap, easy to prepare, and chemically stable (Birech *et al.*, 2020). In this study, we used the conductive

silver paste obtained from SPI-CHEM suppliers. The silver paste was smeared using a small brush on the glass slides and left in the air for about 20 minutes to dry. Figure 4.3 is an illustration of how the substrates were prepared.

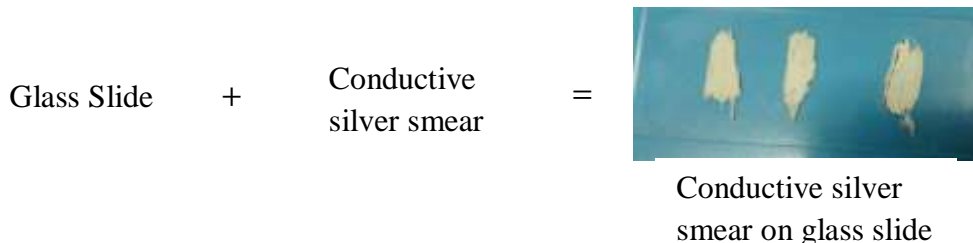


Figure 4.3: Figure showing an illustration of the substrate preparation process.

4.4: Sample Preparation for Laser Raman Spectroscopy

The three sets of the sample used in this study included;

- a. Pure female reproductive hormones obtained from the Department of Veterinary Anatomy and Physiology, College of Agriculture and Veterinary Sciences at the University of Nairobi.
- b. Blood from ‘normal’ and acetaminophen treated female Swiss albino mice obtained from the Department of Veterinary Anatomy and Physiology, College of Agriculture and Veterinary Sciences at the University of Nairobi.
- c. Calibration standards (Pure hormones mixed with blood from male Swiss albino mice).

4.4.1: Preparation of pure hormones

Here, $\approx 5\mu\text{l}$ of each female reproductive standard hormone were measured using a micro-pipette and smeared on the already prepared substrate. The standard hormone concentrations used were 20 pg/ml (for estradiol), 5 mIU/ml (for FSH and LH) and 0.2 ng/ml (for progesterone). The prepared pure hormone samples were immediately taken for Raman measurements.

4.4.2: Preparation of blood from female mice

In this study, ten female Swiss albino mice were grouped into two groups with each group consisting of five mice. One group was treated with acetaminophen syrup according to their weight, and the other group was left controlled for twenty days. Blood was drawn from each

mouse's tail tip daily and smeared on the already prepared substrate. This was done for twenty days.

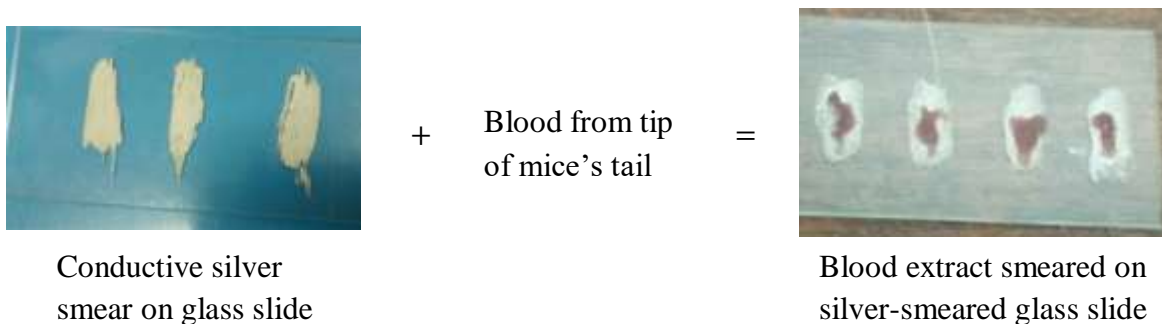


Figure 4.4: Figure showing how the sample was prepared.

4.4.3 Preparation of calibration standards (Male mice's blood mixed with standard hormones)

The Raman spectral data from calibration standards i.e. female reproductive standard hormones mixed with normal male Swiss Albino mouse's blood was used. For each hormone, a set of spectra measured at fifteen different concentrations were used to develop the ANN models in R software for predicting hormones in the blood of treated and untreated mice. This was done to mimic the real samples as much as possible. Table 4.1 displays the prepared different concentrations of female reproductive hormones in the blood.

Table 4.1: Table showing different concentrations of hormones in the blood (simulate samples) in units of pg/ml (picogram per milliliter), mIU/ml (milli-international units per milliliter, and ng/ml (nanogram per milliliter) respectively.

Hormone	Concentration
Estradiol (pg/ml)	15, 16, 20, 24, 30, 40, 50, 60, 75, 80, 90, 96, 100, 120,150, 200, 225, 240, 250, 300, 333.3 and 400
FSH and LH (mIU/ml)	0.5, 0.83, 1, 1.25, 1.6 7, 2, 2.5, 3.3, 3.75, 4, 4.2, 5, 6.25, 6.7, 7.5, 8, 8.33, 10, 12.5, 16, 16.7, 18.7, 20, 21, 25, 33.3 and 40
Progesterone (pg/ml)	0.1, 0.2, 0.25, 0.33, 0.5, 0.67, 0.75, 2, 2.67, 4, 5.3, 6, 6.7, 8, 10, 13.3, 20 and 26.6

4.5: Raman Spectral Data Preprocessing

While Raman spectroscopy has attracted attention in biological sample analysis, the raw Raman spectral data obtained is vulnerable to background contributions and fluorescence that should be extracted through data preprocessing (Gautam *et al.*, 2015). In this study, raw Raman spectral data obtained were first passed through Vancouver Raman software (*version 1.0.0, VB.Net programming language*) for fluorescence subtraction. Using origin software (*open source, OriginPro 9.1.0 version*), baseline correction and smoothing using a 20-point average algorithm to minimize baseline variation and background noise at the region between 350 and 1850 cm^{-1} was done on them. All spectra finally underwent normalization (to maximum intensity) procedure using R software (*open source, version 1.2.5033 2009-2019 RStudio, Inc.*). The normalized spectra were determined using the following equation from the original spectral data:

$$X_{Normalized} = \frac{X_{actual} - X_{min}}{X_{max} - X_{min}} \quad (4.1)$$

in which $X_{Normalized}$ was the normalized Raman intensity of the spectra at each wavenumber, X_{actual} was the Raman intensity of each spectrum at the same wavenumber, X_{min} was the minimum Raman intensity of the spectrum, and X_{max} was the maximum Raman intensity of the spectrum.

4.6: Chemometric Analysis of Raman Spectra

4.6.1: Principal Component Analysis (PCA)

Feature selection was done on the Raman spectral data already pre-processed through the use of an unsupervised technique called Principal Component Analysis (PCA). This technique not only defined the spectral patterns in the dataset and represented the data in such a way that associations between variables and correlations between observations were observed, but also reduced data dimensionality without much information loss, thereby allowing data to be visualized. PCA analysis was done on Raman spectral data for; pure hormones (each consisting of fifteen spectra), and simulate samples that were grouped into low, medium and high concentration (each containing seven spectra). This was achieved through the use of R software

using chemo ChemoSpecUtils, ChemoSpec, knitr and R.utils libraries (see Appendix 2 for code listing).

4.6.2: Artificial Neural Networks (ANN)

The identified spectral bands from feature selection were used to make the ANN predictive models. The Raman spectral data from calibration standards i.e. estradiol, FSH, LH, and progesterone standard hormones, each consisting of fifteen different concentrations were used to develop the ANN models in R software. The development of these models was based on three steps that comprised of construction, training, and prediction. The dataset was first imported and read in R software, then partitioned such that seventy percent of it formed the training data and thirty percent tested the model. Resilient backpropagation (Rprop) algorithm was used to train the model since it is faster than other commonly used backpropagation algorithms, and it doesn't require any specified learning rate values. Finally, the model was validated using the leave one out cross-validation that involved making $n - 1$ sub-models, where n was the sample number, and the left-out sample to test the model. Figure 4.5 is an illustration of how the ANN predictive model was structured.

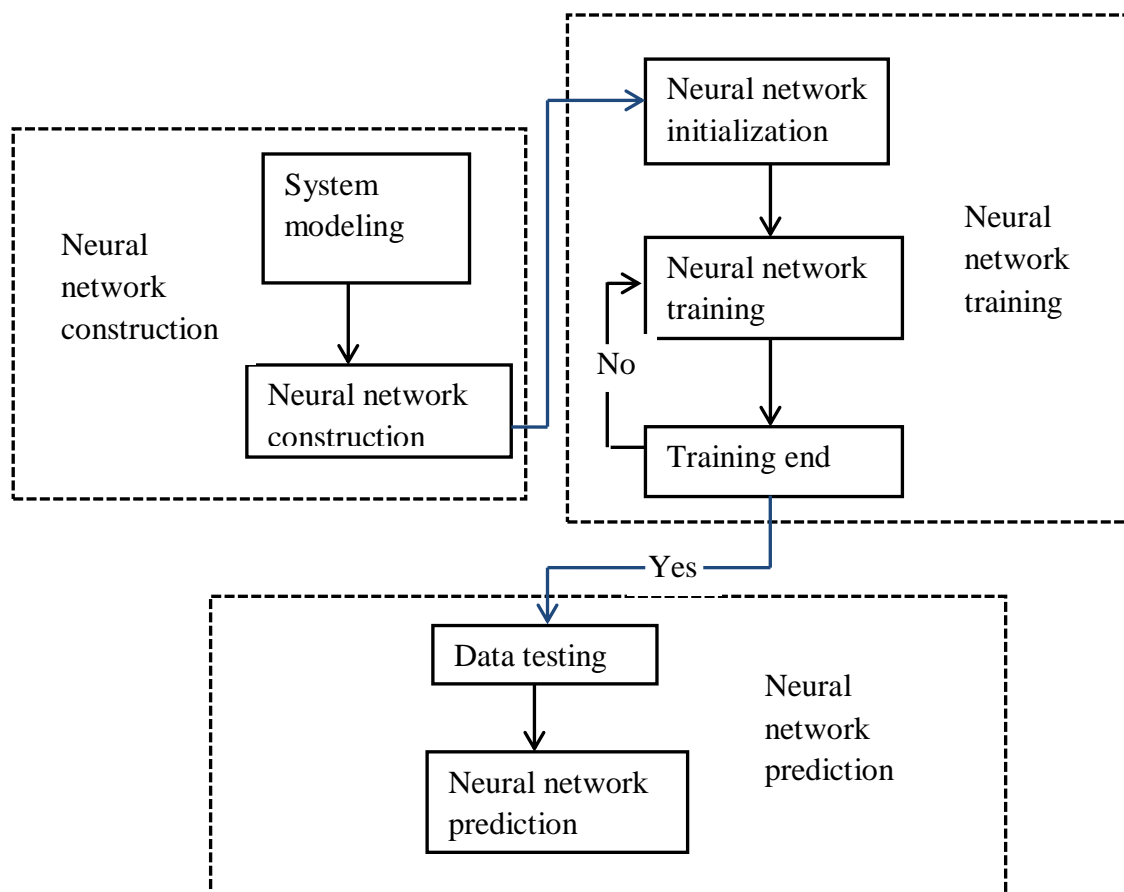


Figure 4.5: Figure showing a schematic diagram of a simplified structure of ANN procedure used in this study.

Before using the built ANN model to predict unknown hormone concentrations, its performance was assessed by computing validation metrics which included Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and determination coefficient (R^2). All of these metrics were intended to calculate the distance between the predicted value and the real value. MAE is defined as the sum of individual absolute errors standardized by the number of samples in which the difference between ground value and predicted value for a sample was defined as the individual error and is given by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - r_i| \quad (4.2)$$

where p_i represents the predicted concentration value, r_i the real concentration value and n the total sample number. RMSE is a modification of the mean absolute value, with the absolute value replaced by a square of an individual error term and is defined as:

$$RMSE = \sqrt{\left(\frac{1}{n} \sum_{i=1}^n (p_i - r_i)^2\right)} \quad (4.3)$$

The mean difference between the predicted and real concentrations was calculated by both MAE and RMSE. Unlike MAE, however, RMSE was the most favored objective metrics that evaluated the ANN model's accuracy because, due to its square definition, it gives more attention to the large errors.

The determination coefficient R^2 was used to approximate the complete association between spectral data and concentration and is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (r_i - p_i)^2}{\sum_{i=1}^n (r_i - \bar{r}_i)^2} \quad (4.4)$$

where \bar{r}_i represents the average actual concentration value.

CHAPTER FIVE: RESULTS AND DISCUSSIONS

5.1: Introduction

In this chapter, we have presented results of Raman spectroscopy of conductive silver-painted Raman sample substrates, female reproductive standard hormones, simulate samples of four-female reproductive standard hormones in the blood of male Swiss albino mouse at different concentrations, and spectra obtained from blood samples of subject animals (i.e. treated and normal mice). Included are the results of the Raman spectral data analyzed using PCA and ANN.

5.2: Conductive silver smeared microscope glass Raman sample substrates

Raman signals are known to be enhanced when the sample's molecules are in very close proximity with metallic nanostructures that support localized surface plasmon resonance (LSPR) (Zhang *et al.*, 2005). Raman substrates, therefore, are very essential in the collected signals, especially when working with biological samples (Sui *et al.*, 2016). In this study, conductive silver paste coated microscope glass slides were used as low-cost Raman substrates. Similar substrates have shown excellent results in Raman spectroscopy detection and differentiation between blood from HIV infected and non-infected (Otange *et al.*, 2017), diabetic and non-diabetic subjects (Birech *et al.*, 2017), obese and non-obese subjects (Birech *et al.*, 2019). Figure 5.1 shows the Raman spectra obtained from a clean microscope glass slide, conductive silver paste smeared glass slide, blood on conductive silver coated glass slide, and a thin and thick smear of blood on clean microscopic glass slide upon 785 nm excitation.

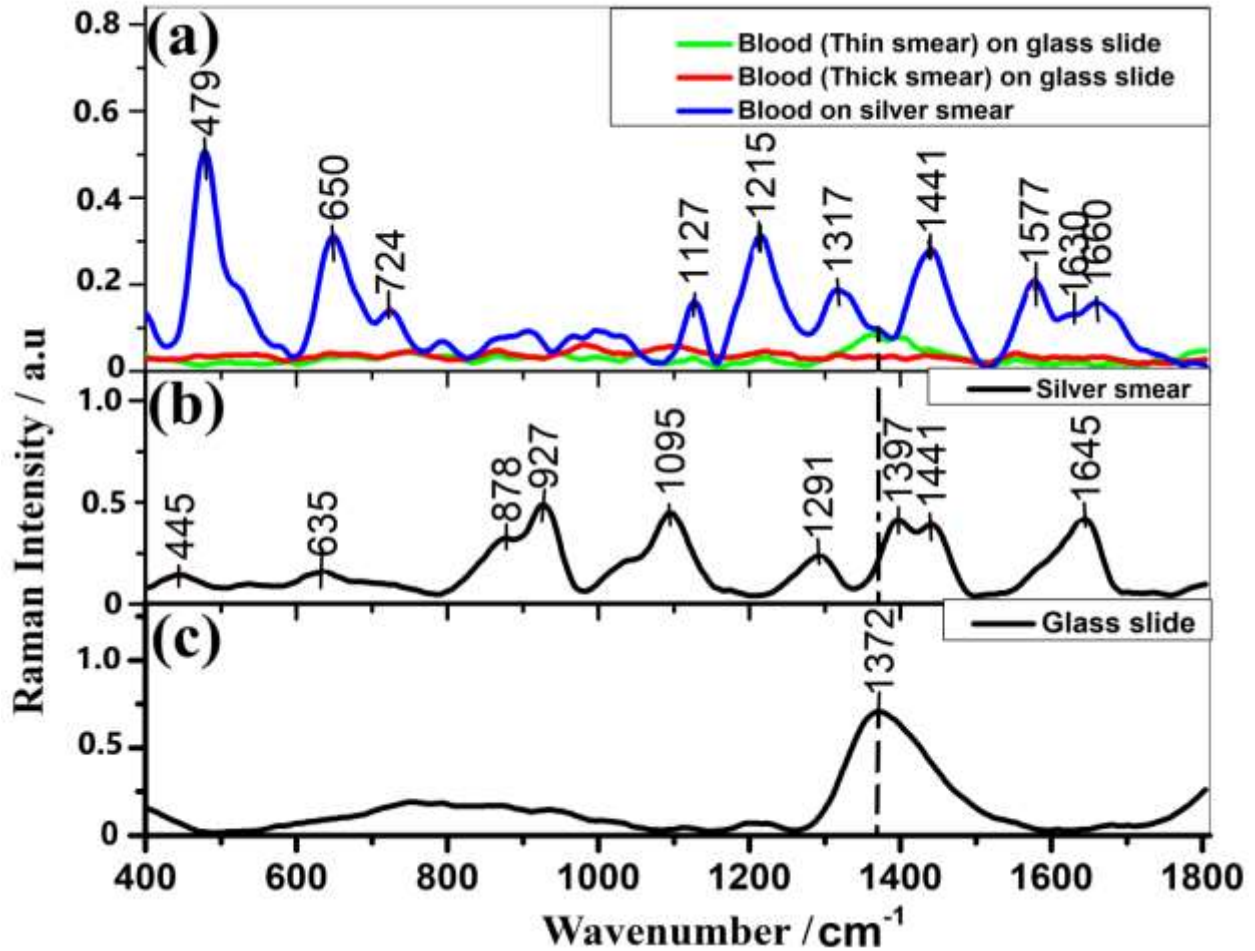


Figure 5.1: Figure showing Raman spectra obtained from (a) thin and a thick smear of blood on a clean microscope glass slide and a smear of blood on a conductive silver-coated clean microscope glass slide, (b) of silver smear on a clean microscope glass slide, (c) of a clean microscope glass slide.

The influence of the conductive silver coated glass substrates on the Raman signals from the blood samples was evident as displayed in Fig. 5.1(a). The Raman spectra obtained from a thin spread of blood of $\approx 15 \mu\text{m}$ in thickness on the silver coated glass slide was more intense and displayed pronounced spectral features as compared to those obtained from a thin ($\approx 15 \mu\text{m}$ in thickness) and a thick ($\approx 30 \mu\text{m}$ in thickness) blood smears on a clean glass slide. The Raman signal was found to be enhanced by a factor of 3.94. This factor was estimated as was done by Birech *et al.* (2020) where the ratio of the area under curve (AUC) values of the SERs signal to the normal Raman signal was computed as:

$$\text{Enhancement factor } (E.F) = \frac{AUC_{\text{Blood on Ag smear}}}{AUC_{\text{Blood on clean glass slide}}}$$

Birech and the co-authors reported on Raman spectroscopic characterization of similar substrates, in which an enhancement factor of 1.7 was reported (Birech *et al.*, 2020). This implied that the factor obtained in this study was higher. The Raman signals obtained from a thin blood spread on a clean glass slide (green line in Fig. 5.1 a) were suppressed by the intense photoluminescence background signals centered at wavenumber 1372 cm^{-1} emanating from the glass slide (Birech *et al.*, 2020; Downes and Elfick, 2010). Comparing Raman spectrum of thin with a thick smear of blood on a glass slide figure 5.1(a), it was evident that the Raman spectral band centered around 1372 cm^{-1} was only seen on the spectra of the thin smear of blood on a glass slide, which implied that signals emanating from glass substrates were a significant noise problem upon 785 nm laser excitation of the thin blood smear on a glass slide. The glass's photoluminescence signal bands are displayed in figure 5.1(c).

SERS substrates should be chemically stable (i.e. they should neither be easily oxidized nor be affected by the weather). Silver substrates used as SERS substrates have shown excellent results in the study of examining their stability at high temperatures (Suzuki and Yoshimura, 2017). Figure 5.2 shows Raman spectra obtained from one of the conductive silver smeared glass Raman substrates at different days after preparation and storage at room temperature.

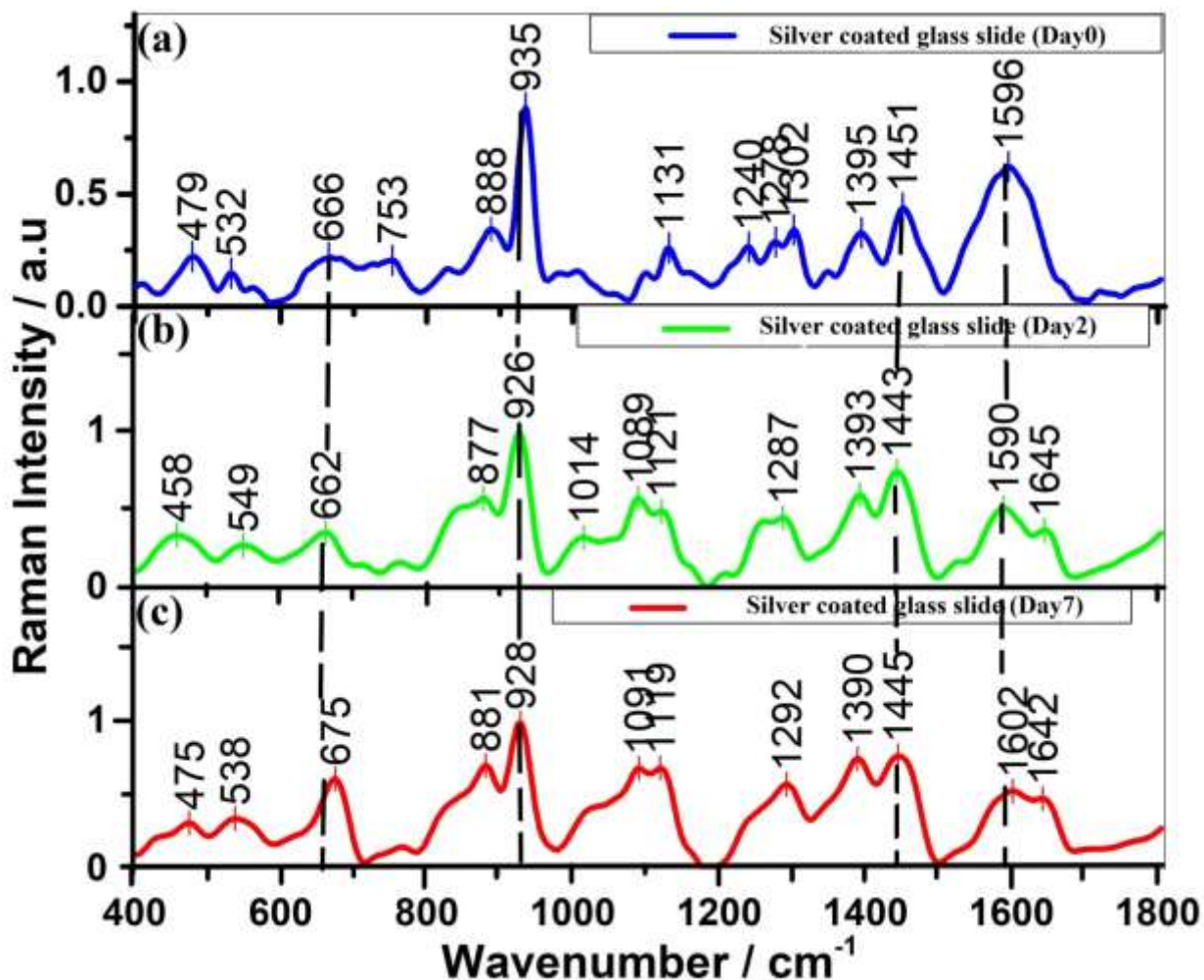


Figure 5.2: Figure showing Raman spectra obtained from conductive silver-coated clean microscope glass slide taken on (a) Day zero, (b) Day two, and (c) Day seven after preparation.

The spectral profiles obtained for the first seven days were found to be identical thus indicating chemical stability. The prominent invariant bands observed were centered around wavenumbers 666 cm^{-1} attributed to CS stretching (Dingari *et al.*, 2012), 935 cm^{-1} attributed to C-C stretching (Zheng *et al.*, 2007), 1451 cm^{-1} attributed to CH_2 bending (Potcoava *et al.*, 2014), and 1596 cm^{-1} attributed to C-C stretching (Han *et al.*, 2011). The fact that the Raman spectra obtained for the different days on the same conductive silver painted glass slide were identical implied that these non-conventional Raman substrates were not chemically modified and hence stable within the first seven days. It was shown in Birech *et al.* (2020) that when the substrates are left over a long time (99 days) the influence of the environment on these substrates becomes significant (Birech *et al.*, 2020).

5.3: Characteristic Raman Spectra of female reproductive hormones

The prepared conductive silver coated glass slides were used as Raman substrates for Raman spectroscopic characterization of the standard female reproductive hormones (Estradiol, FSH, LH and Progesterone) samples upon 785 nm laser excitation. From the obtained Raman spectra (Figure 5.3), it is evident that the spectral profiles of each of the hormones were different indicating the sensitivity of the Raman technique in distinguishing between the various standard hormones when placed onto the conductive silver smeared glass substrates.

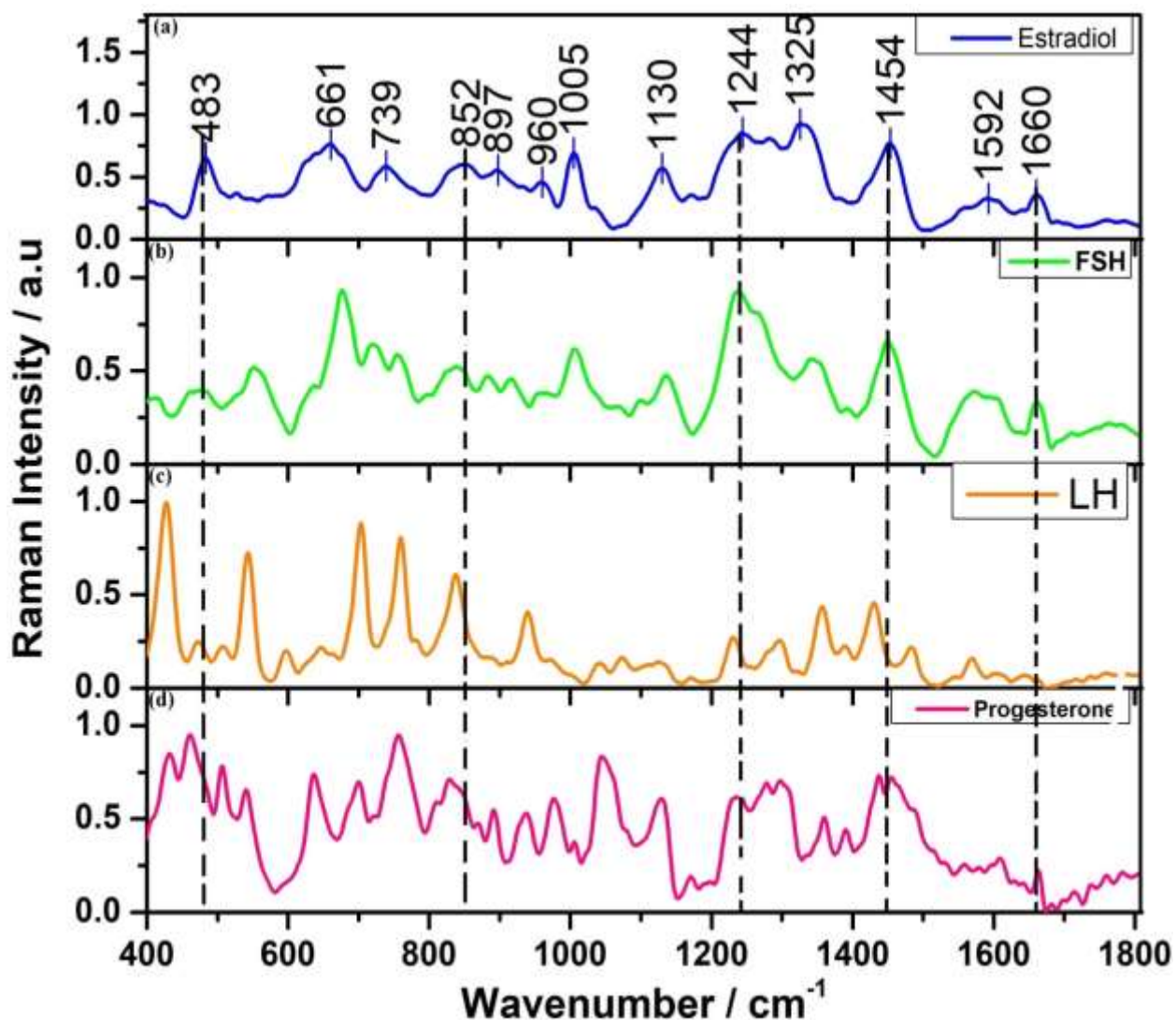


Figure 5.3: Figure showing the average of Raman Spectra (mean of 15 spectra each) of (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH), and (d) Progesterone standard hormones.

Prominent bands associated with amide I (around 1660 cm^{-1} in Estradiol, FSH and progesterone) and amide III ($1230\text{-}1300\text{ cm}^{-1}$ in Estradiol, FSH, LH, and progesterone) (Herrero, 2008) proteins are evident in the spectra for the hormones. Also seen are bands ascribed to tryptophan at around 760 cm^{-1} (LH and progesterone), 880 cm^{-1} (FSH), and 1360 cm^{-1} (progesterone). Bands associated with tyrosine, phenylalanine, and tryptophan are displayed in Table 5.1. The isoenergetic bands in all the hormones are those centered around wavenumbers 480 cm^{-1} ascribed to C-C stretching of glucose; 1244 cm^{-1} ascribed to amide III bands (Ullah *et al.*, 2019), 1454 cm^{-1} ascribed to CH_2/CH_3 scissoring of proteins (Huang *et al.*, 2004). Others are those centered at around 540 cm^{-1} (in LH and progesterone) attributed to C-C wagging, 760 cm^{-1} (in FSH, LH, and progesterone) attributed to tryptophan, 837 cm^{-1} (in FSH and LH) attributed to tyrosine ring stretching, 1005 cm^{-1} (in estradiol and FSH) attributed to phenylalanine ring stretching, and 1130 cm^{-1} (in estradiol, FSH and progesterone) attributed to C-N/C-C stretch of proteins (Herrero, 2008). More details on component and vibrational assignments of the bands are given (Table 5.1).

To segregate Raman spectral data from each of the reproductive hormones studied here, PCA was used. We used fifteen spectra for each of the estradiol, FSH, LH and progesterone. Figure 5.4 displays the score plot which shows clear segregation of the spectral data sets from the hormones thus supporting the idea that the Raman spectroscopic technique is a potentially sensitive alternative hormonal assaying method. The three PCs explained 67% of the variability of the data, where PC1, PC2 and PC3 accounted for 31%, 25% and 11% respectively. The bands responsible for the segregation are displayed in the loadings plots of Figure 5.5. These bands are those centered at wavenumbers 1005, 1244 and 1452 cm^{-1} (in estradiol), 1662 cm^{-1} (for FSH), 428, 544, 703, 761, 837 cm^{-1} (for LH), and 1044 and 1604 cm^{-1} (in progesterone). The overlap is seen in estradiol, FSH and progesterone hormones in the PCA scores plot was due to the common bands seen around 483 and 1660 cm^{-1} which was attributed to C-C stretching of glucose and C=O stretching of Amide I. This was attributed to the other components in the blood such as hormones, glucose and proteins (remaining 33% of data not explained by the three PCs).

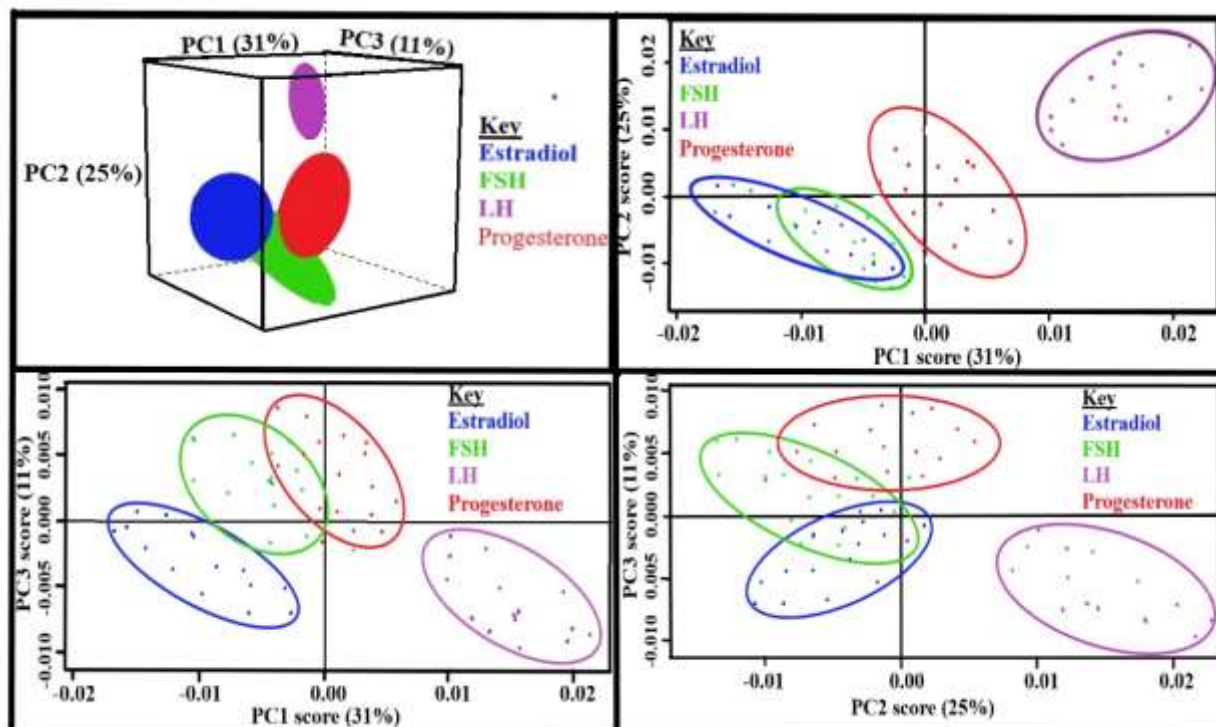


Figure 5.4: Figure showing the two and three-dimensional score plot PCA results for standard hormones. The explained variances are indicated in percentages and were 31, 25, and 11 % for PC1, PC2, and PC3 respectively.

These bands were assigned to C-C stretching in Amide II proteins and phenylalanine (1005 cm^{-1}) (Bergholt *et al.*, 2011; Liu *et al.*, 2014); CH_2/CH_3 scissoring vibrations in proteins (1244 cm^{-1} and 1452 cm^{-1}) (Huang *et al.*, 2004), 1662 cm^{-1} (C-C/C=O stretching of Amide I) (Herrero, 2008), 761 cm^{-1} (C-C bending vibrations), 837 cm^{-1} (CH_3 scissoring for tyrosine) (Ullah *et al.*, 2019), 544 cm^{-1} (C-C wagging), 1064 cm^{-1} and 1604 cm^{-1} (C-C stretching) (Huang *et al.*, 2004; Liu *et al.*, 2014; Otage *et al.*, 2017; and Zhu *et al.*, 2019).

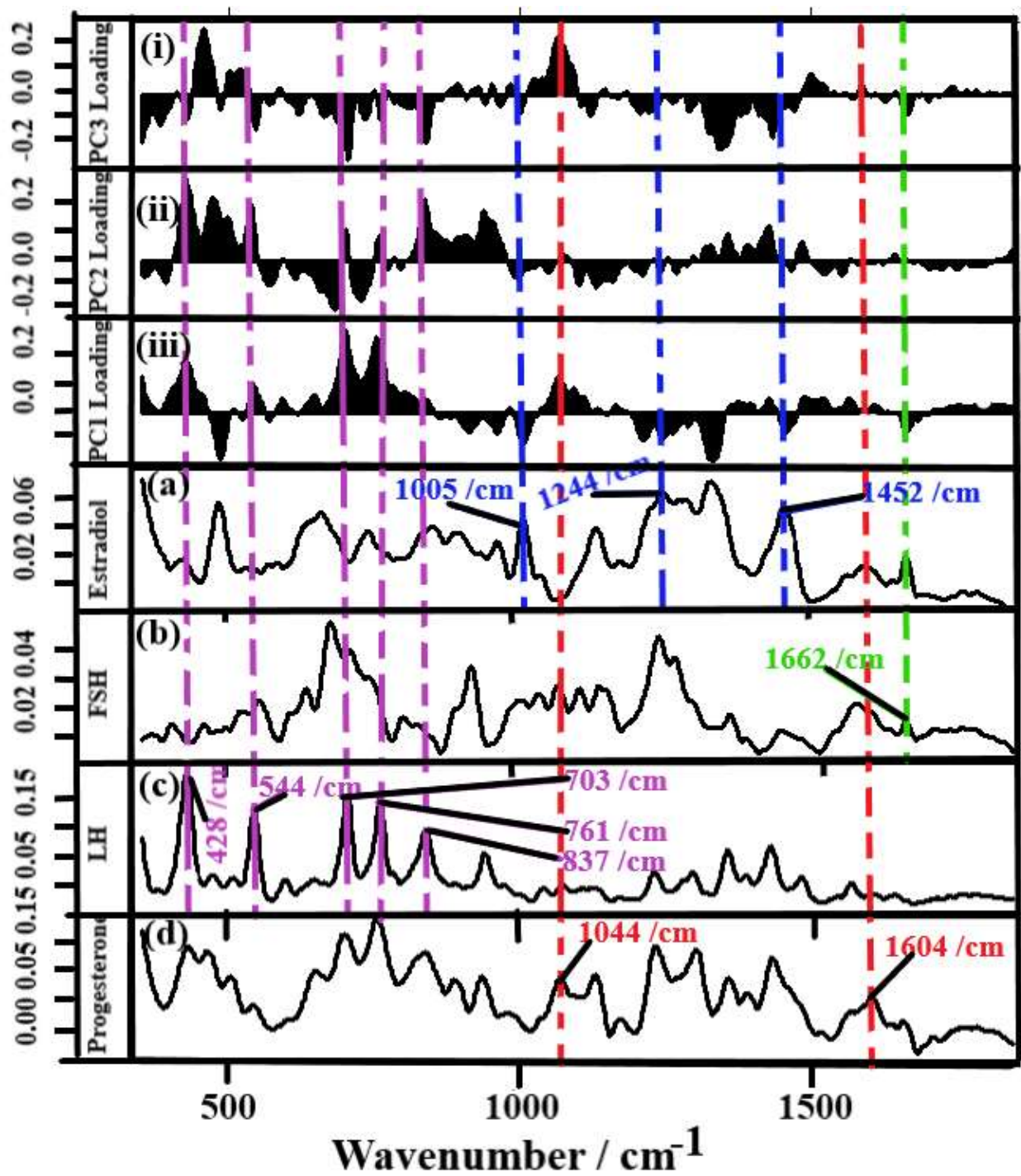


Figure 5.5: Figure showing the (i) PC1, (ii) PC2, and (iii) PC3 loadings plot, and (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH) and (d) Progesterone as reference Raman spectra for PCA results of standard hormones.

5.4: Raman Spectra of blood from male Swiss albino mice mixed with the female reproductive hormones

To explore further if the Raman spectroscopy technique can be used to detect the variation of the female reproductive hormones in the blood, blood obtained from a male Swiss albino mouse was mixed with the respective standard hormones. Blood from a male mouse was chosen since the levels of these hormones (LH, FSH, estradiol, progesterone) are low and fairly stable (Konforte *et al.*, 2013). This was done to make simulate samples that mimic real/field samples as much as possible. This would also facilitate the identification of biomarker Raman bands in blood for each of the reproductive hormones under study here. Figure 5.6 shows some spectra obtained from blood mixed with estradiol, FSH, LH, and progesterone hormones at different concentrations.

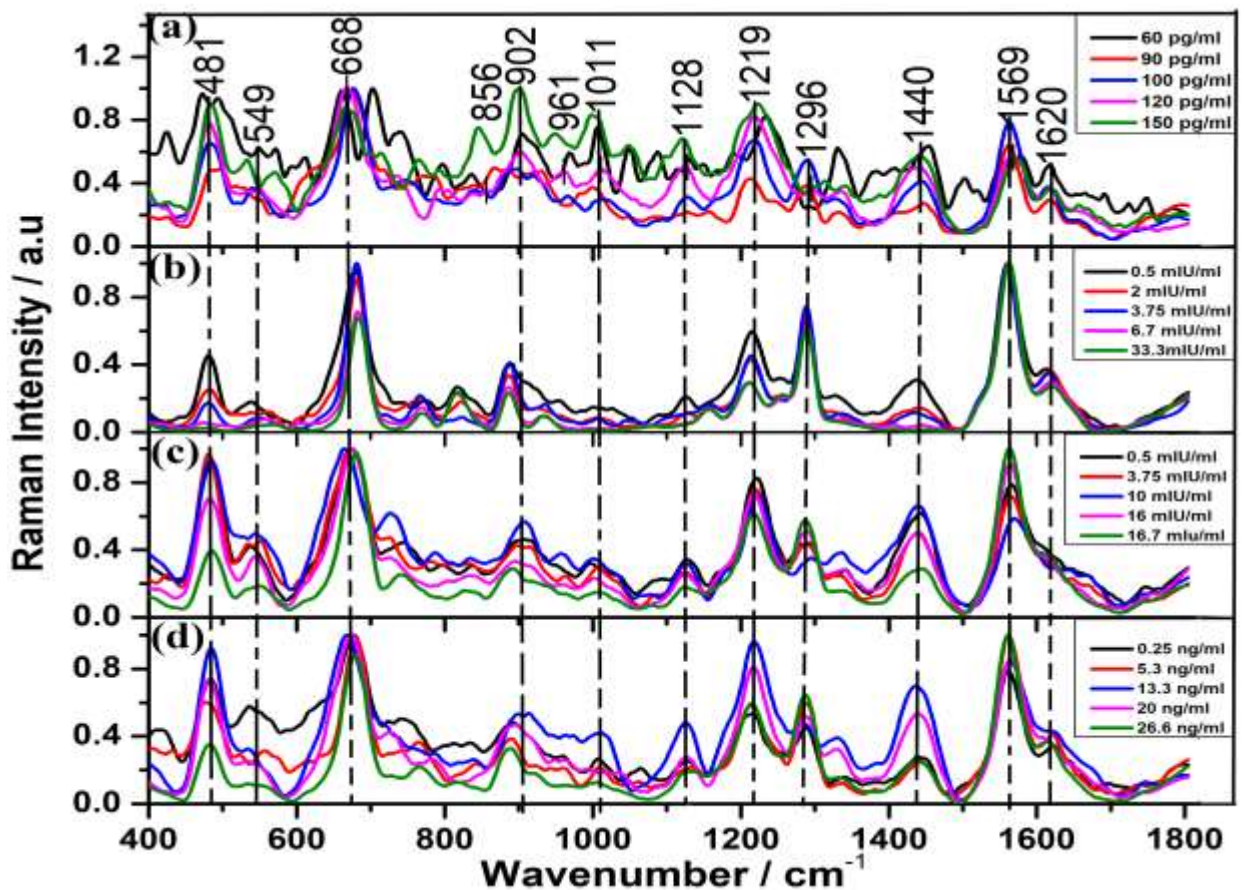


Figure 5.6: Figure showing Raman spectra of different concentrations of (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing hormone (LH), and (d) Progesterone standard hormones mixed with male swiss albino mouse's blood.

As can be seen from figure 5.6, the spectral profiles are fairly identical with variations seen in the intensity of some bands such as those centered at wavenumbers; 481, 668, 902, 1219 and 1440 cm^{-1} (for all hormones); 1620 cm^{-1} (for estradiol, FSH and progesterone); 1011, 1128 and 1569 cm^{-1} (for estradiol, LH and progesterone); and 549 cm^{-1} (for LH). The profiles were identical since all the hormones investigated here are all naturally present in blood at various concentration levels as mentioned earlier. To verify further the significantly varying bands, analysis of variance (ANOVA) was done on each of the hormones in figure 5.6, and the results are displayed in figure 5.7.

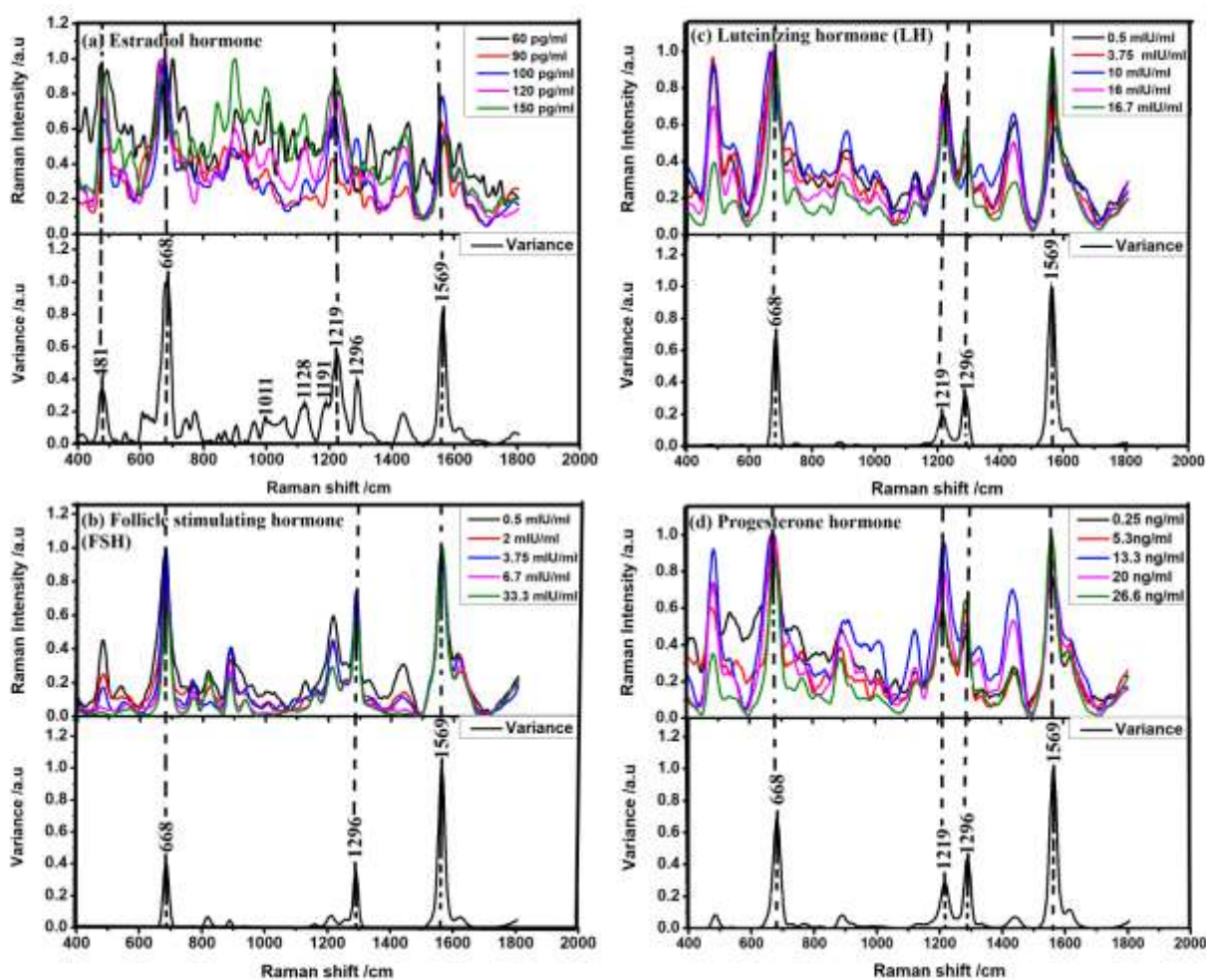
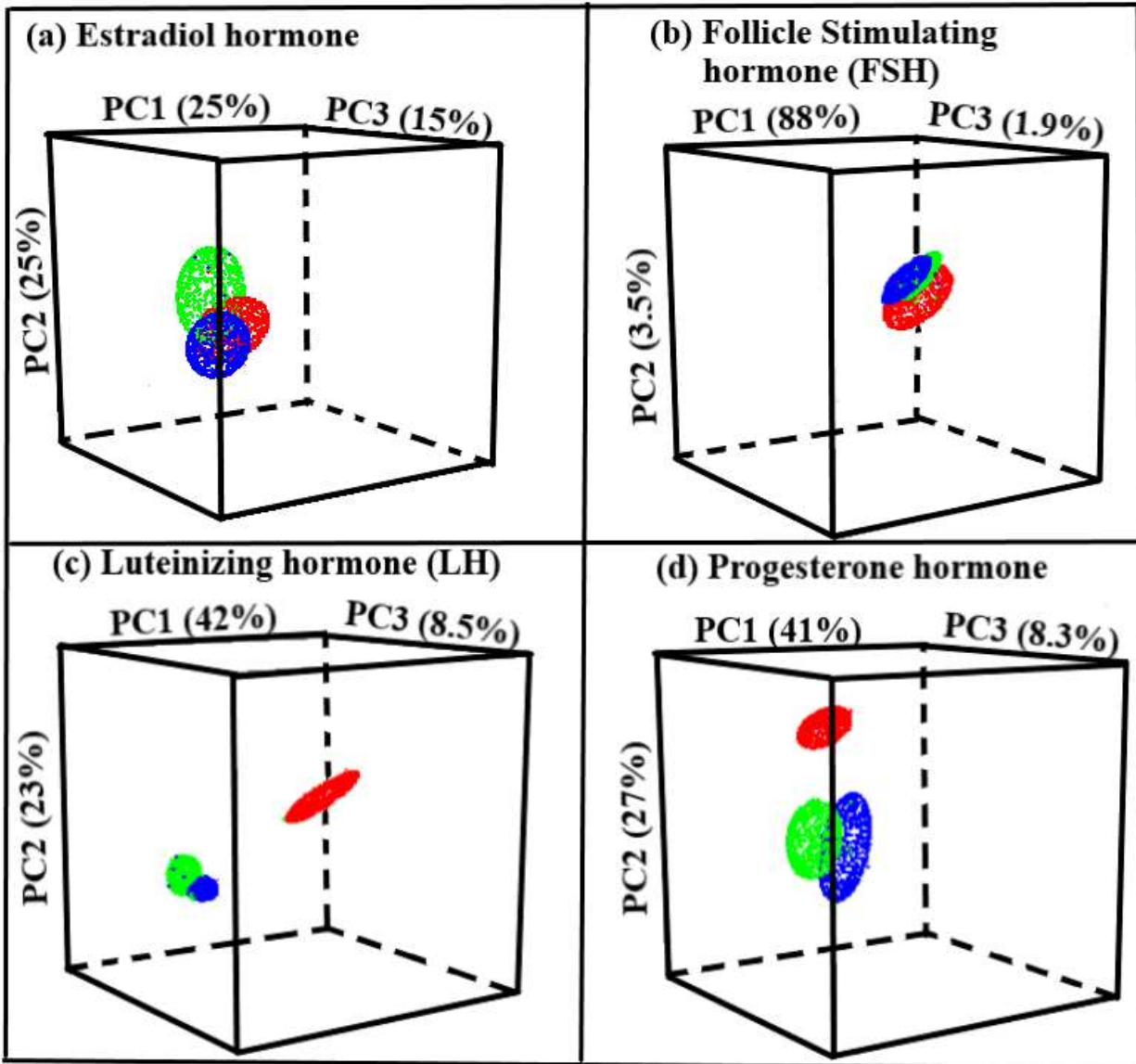


Figure 5.7: Figure showing Analysis of Variance results for Raman spectra of different concentrations of (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing hormone (LH), and (d) Progesterone standard hormones mixed with male swiss albino mouse's blood.

The bands with significant variance (large variance values in Fig. 5.7) were those whose intensities varied with concentration and were found to be those centered around wavenumbers (see figure 5.7); 668 and 1569 cm^{-1} (for all hormones); 1296 cm^{-1} (for FSH, LH and progesterone); 1219 cm^{-1} (for estradiol, FSH and progesterone); and 481, 1128 and 1191 cm^{-1} (for estradiol). PCA was also done on the combined spectral dataset with different concentrations for each hormone mixed with blood to see whether low, medium and high concentration of each hormones in blood were any different. These groups (low, medium and high concentrations) were as follows: Estradiol (low (15-60), medium (15-60) and high (200-400))pg/ml; FSH and LH (low(0.5-3.3), medium (3.75-8.33) and high (10-40)) mIU/ml; and progesterone (low (0.1-0.67), medium (0.75-6) and high (6.7-26.6)) ng/ml. This was to further show that Raman spectroscopy was sensitive to concentration as Raman spectral data sets will be segregated on a score plot based on concentration of each hormone in blood. Figure 5.8 shows 3-Dimensional score plots for standard hormones mixed with blood to form different concentrations (see Table 4.1). It was evident that spectra of the low, medium, and high hormone concentrations in blood are fairly different giving rise to spectral differentiation. The spectral bands exhibiting higher variance with respect to concentration, hence responsible for the observed segregation on score plots, were expected to have large loadings values as exhibited on PC loadings plots. These bands (with large loadings) were identified as follows: In estradiol (481, 668, 902, 1011, 1128, 1219, 1296 and 1569 cm^{-1}); in FSH (481, 668, 1219, 1296 and 1569 cm^{-1}); in LH (481, 668, 1296, 1440 and 1569 cm^{-1}); and in progesterone (481, 902, 1011, 1219, 1440 and 1569 cm^{-1}). In Figure 5.9, it is seen that the common bands with similar variations are centered around wavenumbers: 481, 688, 1219, 1296 and 1569 cm^{-1} . The bands which could be used as biomarker bands are centered around wavenumbers 668, 902, 1011 cm^{-1} for estradiol; 1219 and 1296 cm^{-1} for FSH; 1440 cm^{-1} for LH and 1569 cm^{-1} for progesterone. The similar bands seen in ANOVA and PCA included those centered around 668 and 1569 (for all hormones), 1296 (for FSH and LH); 1219 (for estradiol, FSH and progesterone); and 481 and 1128 (for estradiol). More bands were noticed to vary with concentration in PCA than in ANOVA which implied that PCA is more sensitive to data variance than ANOVA.



KEY

- Low hormone concentration
- Medium hormone concentration
- High hormone concentration

Figure 5.8: Figure showing the 3-Dimensional PCA score plots for (a) Estradiol (low (15-60), medium (75-150) and high (200-400)) pg/ml, (b) FSH (low (0.5-3.3), medium (3.75-8.33) and high (10-40)) mIU/ml, (c) LH (low (0.5-3.3), medium (3.75-8.33) and high (10-40)) mIU/ml, and (d) Progesterone (low (0.1-0.67), medium (0.75-6) and high (6.7-26.6)) ng/ml standard hormones in blood. The explained variances are indicated in percentages and were 25, 25, and 15 % for Estradiol, 88, 3.5, and 1.9% for FSH, 42, 23 and 8.5% for LH, and 41, 27, and 8.3% for Progesterone for PC1, PC2, and PC3 respectively.

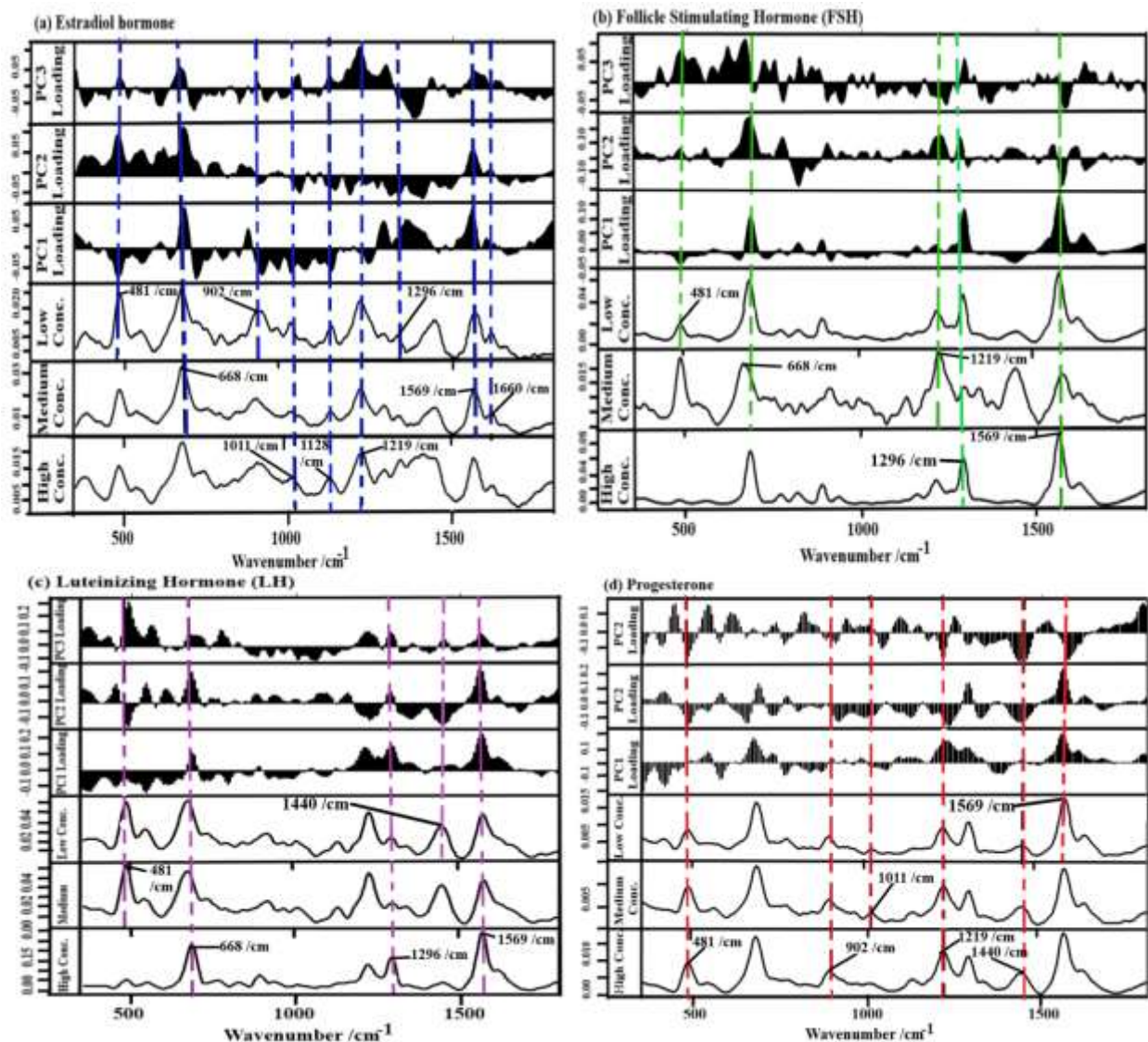


Figure 5.9: Figure showing the PC1, PC2, and PC3 loadings plot, with low, medium and high concentrations for (a) 32, 102 and 244 pg/ml Estradiol, (b) 1.6, 6 and 21.3 mIU/ml Follicle Stimulating Hormone (FSH), (c) 1.6, 6 and 21.3 mIU/ml Luteinizing Hormone (LH) and (d) 0.3, 3.5 and 14.1 ng/ml Progesterone respectively as reference Raman spectra for PCA results of standard hormones mixed with male mice's blood.

In order to reveal further the subtle differences in the combined Raman spectral data sets obtained from blood mixed with the hormones at varied concentrations, PCA was used. This was done so as to find out how different the respective hormones are from each other in terms of their Raman spectral pattern. The spectra for blood mixed with the different hormones (estradiol, FSH, LH, and progesterone) were shown by the PCA score plots (see Figure 5.10) to have

differing spectral patterns. This was made clear by the distinct segregation between the respective spectral data sets. The 2-Dimensional score plots (Figure 5.10) and loadings plots for these data sets (Figure 5.11) were used to identify the bands responsible for the segregation. This was achieved by obtaining the bands on the loadings plot (Figure 5.11) that merged with 2-D score plots (Figure 5.10) where estradiol is on the negative side of PC1, PC2 and PC3; FSH on the positive side of PC1, PC2 and PC3; LH on the negative side of PC1, either positive or negative side of PC2 and negative side of PC3; progesterone on the either negative or positive of PC1, negative of PC2 and positive of PC3. From the score plots displayed in Figure 5.10, it was observed that the data sets from the different hormones were segregated indicating that they were different hence showing the power of Raman spectroscopy in hormone detection and differentiation in blood. The bands responsible for the segregation were identified (from loadings plot of Figure 5.11) as those centered around wavenumber; 1011 cm^{-1} for estradiol; 668 and 1296 cm^{-1} for FSH; 481 and 1440 cm^{-1} for LH; and 1569 cm^{-1} for progesterone. The vibrational assignments of these prominent band (which were largely similar to those identified from ANOVA) were as follows : 481 cm^{-1} ascribed to C-C stretching of starch (Li *et al.*, 2011); 668 ascribed to C-C stretching of proteins (Downes and Elfick, 2010); 1011 ascribed to C-C stretching of phenylalanine (Bergholt *et al.*, 2011); 1219 ascribed to C-C stretching of proteins (Atkins *et al.*, 2017a; Han *et al.*, 2011); 1296 ascribed to C-C stretching of proteins (Vedad *et al.*, 2018); 1440 ascribed to CH_2/CH_3 bending of proteins (Han *et al.*, 2011; Ullah *et al.*, 2019); and 1569 cm^{-1} ascribed to C-C stretching of phenyl rings (Han *et al.*, 2011).

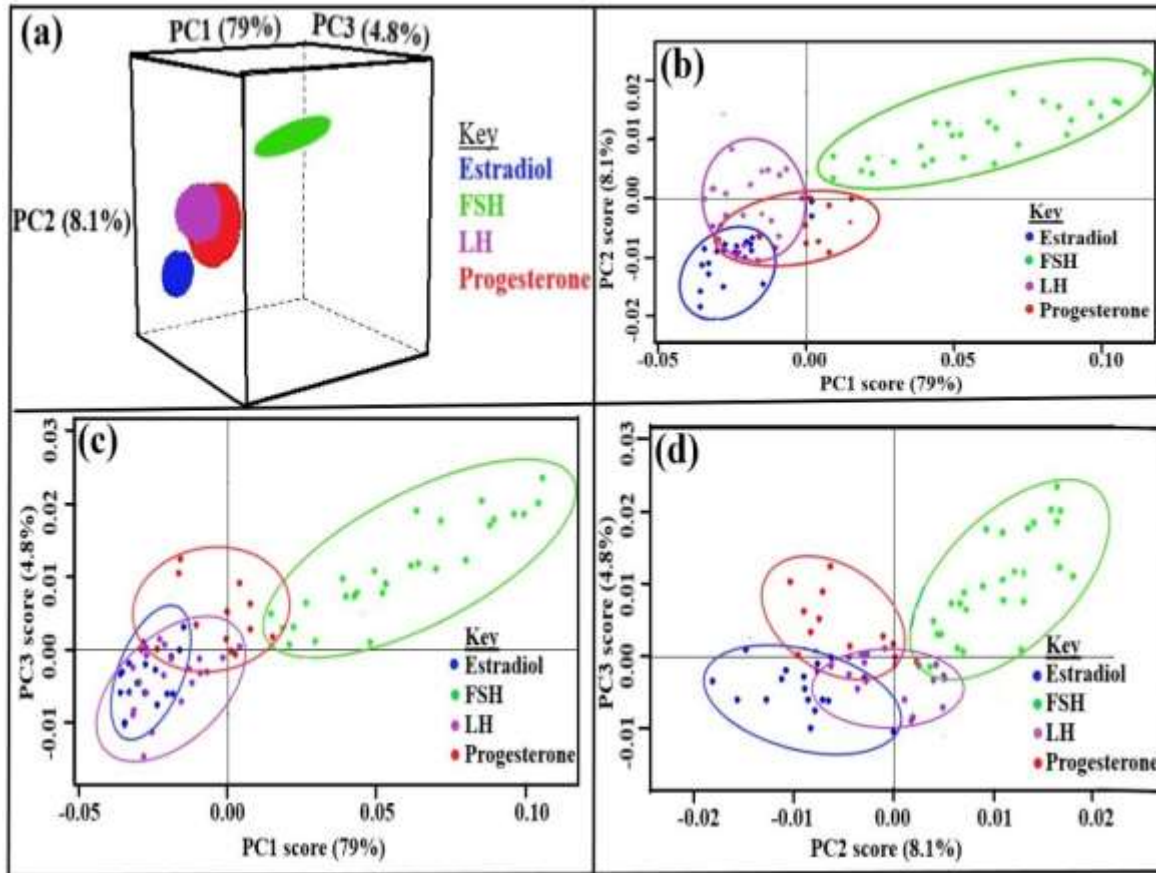


Figure 5.10: Figure showing (a) the 3-Dimensional PCA score plot with (b) PC1 versus PC2, (c) PC1 versus PC3, and (d) PC2 versus PC3 2-D PCA score plots for standard hormones mixed with blood. The explained variances are indicated in percentages and were 79, 8.1, and 4.8 % for PC1, PC2, and PC3 respectively.

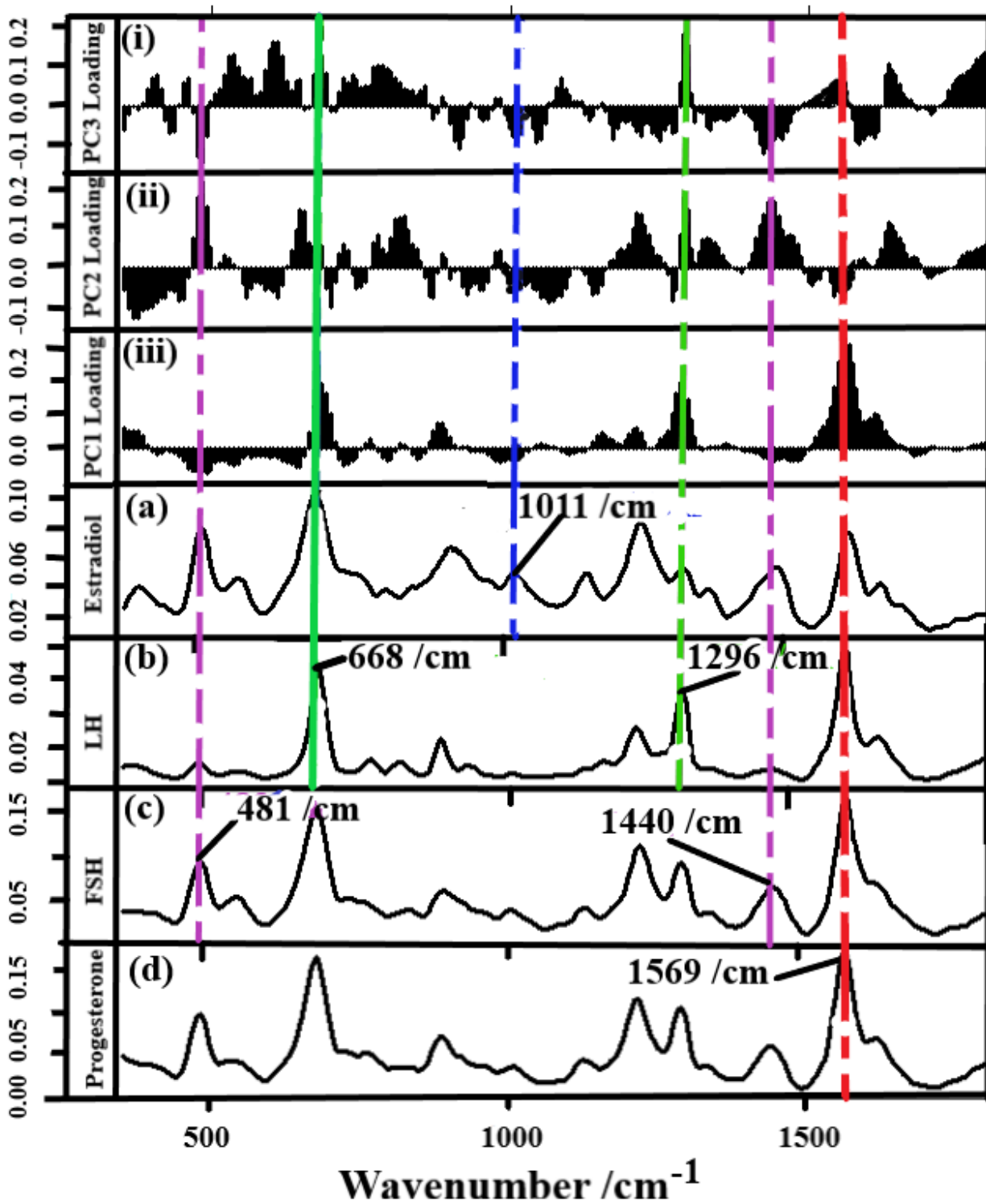


Figure 5.11: Figure showing the (i) PC1, (ii) PC2, and (iii) PC3 loadings plot, and (a) Estradiol, (b) Follicle Stimulating Hormone (FSH), (c) Luteinizing Hormone (LH) and (d) Progesterone as reference Raman spectra for PCA results of standard hormones mixed with blood.

Table 5.1: Table showing the spectral bands for female reproductive standard hormone with the tentative vibrational assignment.

Raman peak value (cm ⁻¹)	Standard hormones				Tentative vibrational assignment	References
	Estradiol (C ₁₈ H ₂₄ O ₂)	FSH (C ₇₉ H ₁₂₅ N ₁₉ O ₂₃ S)	LH (C ₆₀ H ₂₃ N ₁₅ O ₁₃)	Progesterone (C ₂₁ H ₂₀ O ₂)		
483	✓	✓	✓	✓	C-C (Stretching)	(Li <i>et al.</i> , 2011)
544		✓	✓	✓	C-C (deformation)	(Vedad <i>et al.</i> , 2018)
645	✓	✓			C-C (twisting of tyrosine)	(Zheng <i>et al.</i> , 2007)
937			✓	✓	C-C /C-S (stretching)	(Ullah <i>et al.</i> , 2019), (Herrero, 2008)
1005	✓	✓			C-C (phenylalanine Stretching)	(Bergholt <i>et al.</i> , 2011)
1044			✓	✓	C-C (stretching)/ C-H (bending)	(Potcoava <i>et al.</i> , 2014), (Bergholt <i>et al.</i> , 2011), (Wahadoszamen <i>et al.</i> , 2015)
1129	✓	✓		✓	C-C (stretching)	(Potcoava <i>et al.</i> , 2014)
1156		✓	✓		C-C (stretching)	(Zhu <i>et al.</i> , 2019)
1238	✓	✓	✓	✓	C=C (Stretching of amide III)	(Ullah <i>et al.</i> , 2019)

1296	✓		✓	✓	C-N (Stretching) \$ N-H (Bending) Amide III	(Huang <i>et al.</i> , 2004), (Herrero, 2008)
1303				✓	CH ₂ twist /Scissoring in lipids	Huang <i>et al.</i> , 2004, O'Brien <i>et al.</i> , 2017
1438	✓	✓	✓	✓	CH ₂ (Scissoring)	(Zhu <i>et al.</i> , 2019); (Ullah <i>et al.</i> , 2019)
1569	✓	✓	✓		-C=C- (Stretching in tryptophan)	(Zheng <i>et al.</i> , 2007)
1662	✓		✓	✓	C=C (stretching)/ C-H-C of CH ₃ (bending)	(Wahadoszame n <i>et al.</i> , 2015)
1763		✓		✓	C=O (stretching in Amide I)	(Vedad <i>et al.</i> , 2018), (Herrero, 2008)
Raman peak value (cm ⁻¹)	Standard hormones mixed with blood				Tentative vibrational assignment	References
	Estradiol	FSH	LH	Progesterone		
483	✓	✓	✓	✓	C-C (Stretching)	(Li <i>et al.</i> , 2011)
679	✓	✓	✓	✓	C-C (Stretching in proteins) or	(Downes and Elfick, 2010); (Dingari <i>et al.</i> ,

					C-S, Stretching in cysteine	2012)
769		✓			NH ₂ (Bending of lipids)	(Wahadoszame n <i>et al.</i> , 2015); (Downes and Elfick, 2010)
817		✓			C-C (stretching of glucose)	(Potcoava <i>et al.</i> , 2014)
886	✓	✓	✓	✓	C-O-C / CH ₂ (Stretching)	(Davidson <i>et al.</i> , 2013), (Auner <i>et al.</i> , 2018)
1006	✓				C-C (Stretching of phenylalanine)	(Auner <i>et al.</i> , 2018)
1128	✓		✓	✓	C-H/ C-CH ₃ (Stretching)	(Dingari <i>et al.</i> , 2012), (Atkins <i>et al.</i> , 2017a)
1158		✓			C-C (Stretching)	(Zheng <i>et al.</i> , 2007)
1219	✓	✓	✓	✓	C-H (Bending) + C-C (Stretching)	(Han <i>et al.</i> , 2011), (Atkins <i>et al.</i> , 2017b)
1287	✓	✓	✓	✓	C-H/C-C-H (Stretching) & CH ₂ (twisting)	(Potcoava <i>et al.</i> , 2014)

1440	✓		✓	✓	CH ₂ (Scissoring)	(Han <i>et al.</i> , 2011)
1569	✓	✓	✓	✓	amide II (C–N stretching, N– H bending) -C=C- (Stretching)	(Vedad <i>et al.</i> , 2018); (Borio <i>et al.</i> , 2012)
1619	✓	✓		✓	C=C (stretching)	(Wahadoszame n <i>et al.</i> , 2015), (Vedad <i>et al.</i> , 2018)

5.5: Qualitative analysis of blood from treated and normal mice

After finding the Raman spectral bands associated with each standard hormone (see Figures 5.9, 5.10, and 5.11), it was now possible to use them in determining their level variation in blood obtained from female mice. Figure 5.12 displays the analysis of variance results and Raman spectra for the blood of female mice both treated and normal (untreated) mice. The treated mice were administered acetaminophen which was expected to influence the concentration levels of their reproductive hormones in the blood hence their fertility. The ANOVA results with the Raman spectra of blood from normal and acetaminophen treated mice, taken on days one, six, eleven, sixteen, and twenty of the treatment period have been plotted using the same wavenumber axis for ease of comparison.

Figure 5.12 reveals the presence of female reproductive hormones in blood from treated and normal mice. This is because most of the identified biomarker bands for each hormone seen in previous results (see section 5.4) are observed. The significantly varying bands observed are those centered around wavenumber 481, 668, 1296, 1440 and 1569 cm⁻¹. These bands match with the identified biomarker bands seen i.e. 481, 668, 1011, 1219, 1296, 1440 and 1569 cm⁻¹. We, therefore, discuss their Raman intensity variation in treated mice when compared to normal (see Table 5.2). It is seen that the estradiol biomarker band centered at wavenumber; 1011 cm⁻¹ vibrated more on days one, six, eleven, sixteen and twenty in treated mice when compared to

normal. The FSH biomarker band at 668 cm^{-1} vibrated more on days one, six, eleven, sixteen and twenty, and 1296 cm^{-1} vibrated more on days one, six, eleven and sixteen, and vibrated less on day twenty in treated mice when compared to normal. The LH biomarker band centered at 481 cm^{-1} vibrated more on days six, sixteen and twenty, and less on days one and eleven; 1440 cm^{-1} vibrated more on days six, eleven and sixteen, and vibrated less on days one and twenty in treated mice when compared to normal. The progesterone biomarker band centered at 1569 cm^{-1} vibrated more on days one, six, sixteen and twenty, and less on day eleven in treated mice when compared to normal.

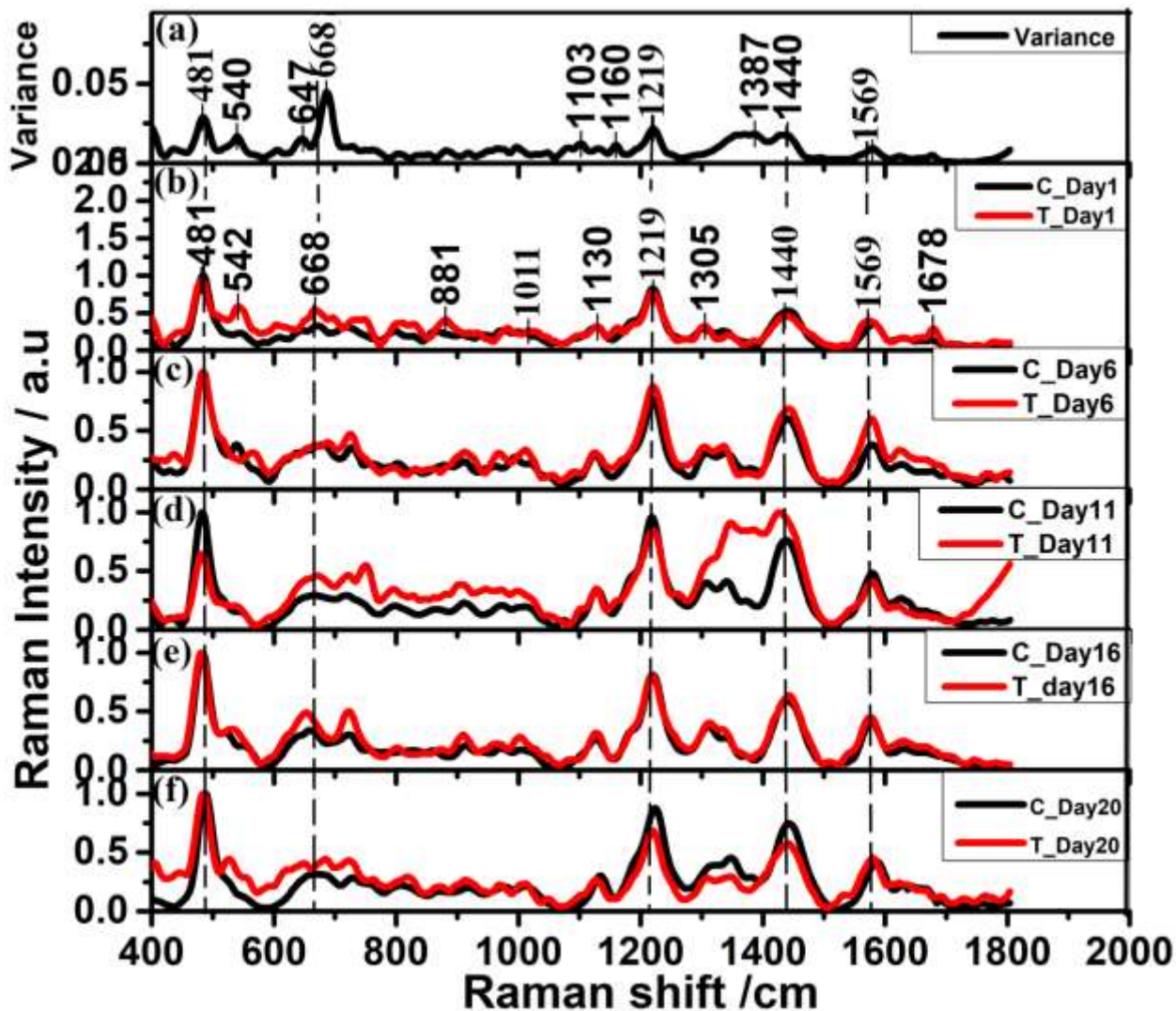


Figure 5.12: Figure showing (a) analysis of variance results and the average Raman spectra of blood from normal (black curve labeled C for “Control”) and acetaminophen treated (red curve

labeled T for “Treated”) mice taken at (b) day one, (c) day six, (d) day eleven, (e) day sixteen, and (f) day twenty of the treatment period.

Table 5.2 displays the Raman spectral bands present in figure 5.12, their vibrational assignment, and the variation of Raman spectra of acetaminophen treated mice from the normal mice in terms of Raman intensity of the peaks observed.

Table 5.2: Table showing the variation of Raman peak intensity observed in acetaminophen treated mouse from a normal mouse, with the tentative vibrational assignment for day one, six, eleven, sixteen and twenty of the study period (letter D, S and I represent decrease, same (no change) and increase respectively).

Raman peak value (cm ⁻¹)	Raman peak intensity value of treated compared to normal mouse (I, Increased / D, Decreased/S, Same)					Tentative vibrational assignment	References
	Day 1	Day 6	Day 11	Day 16	Day 20		
481	D	S	D	S	S	C-C, Stretching in starch	(Li <i>et al.</i> , 2011)
668	I	I	I	I	I	C-S, Stretching in cysteine	Dingari <i>et al.</i> , 2012, Auner <i>et al.</i> , 2018
881	I	I	I	I	I	C-C, Stretching of proteins	(Auner <i>et al.</i> , 2018)
1011	I	I	I	I	I	C-C, Stretching of proteins (phenylalanine)	Zheng <i>et al.</i> , 2007, Potcoava <i>et al.</i> , 2014
1128	S	I	S	I	D	C-CH ₃ , Stretching of protein/ C-C stretching within phenol group	Zheng <i>et al.</i> , 2007, Han <i>et al.</i> , 2011, Atkins <i>et al.</i> , 2017
1219	D	I	D	D	D	CH Rocking/O-H bending/C-C Stretching from phenol groups and C-N Bending/CNN bending in phenyl N	Han <i>et al.</i> , 2011,
1296	I	I	I	I	D	CH ₂ /CH ₃ , Deformation within the phenyl ring	Han <i>et al.</i> , 2011, Sui <i>et al.</i> , 2016, Ullah <i>et al.</i> , 2019
1440	D	I	I	I	D	CH ₂ , Scissoring of proteins/lipids/fatty acids	Zheng <i>et al.</i> , 2007, Dingari <i>et al.</i> , 2012, Potcoava <i>et al.</i> , 2014, Auner <i>et al.</i> , 2018

1569	I	I	D	I	I	C-C, Stretching within phenyl ring	Han <i>et al.</i> , 2011, Sui <i>et al.</i> , 2016, Ullah <i>et al.</i> , 2019
------	---	---	---	---	---	------------------------------------	--

Biologically, when hormones get elevated, the intensity of the biomarker band is expected to increase due to an increase in vibrating molecules adding to the signal. But this may not be the case because in this study since Raman intensity depends not just on the concentration but also factors like instrumental effects and laser power (Galata *et al.*, 2019). Therefore, it was not easy to quantify compounds/molecules/substances based on Raman spectral data unless combined with multivariate analytical techniques such as ANN.

5.6: Quantitative analysis of hormones

5.6.1: ANN model on the standard hormones mixed with blood

The ANN technique was used to quantify the hormone levels present in the blood samples. This technique uses the relationship between the input and output which is controlled by a transfer function. First, the input is multiplied by a weighting factor which determines to what degree the input affects the output. Then, the weighted inputs are added together to create a pre-neuron number, which is analyzed and adjusted to generate the output via a transfer function. Quantification using ANN has been achieved in quantifying whey in liquid milk (Alves da Rocha *et al.*, 2015), and the determination of glucose in whole blood (Wang *et al.*, 2015) based on the spectral dataset.

Based on the biomarker bands for each hormone obtained from the feature selection process (i.e. using figure 5.10 and figure 5.11), four ANN models were trained, tested, and validated. The input layer of these models was the spectra of biomarker bands. The hidden layer was one with 10 neurons, logistic activation function and resilient backpropagation (rprop+) algorithm. The output was the predicted hormone concentration. The samples used were about fifteen with seventy percent forming the training set and thirty percent making the test set. Table 5.3 displays the validation results obtained from these developed ANN models. The test of the model's accuracy was performed using the validation metrics that included Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and correlation coefficient (R^2) which were calculated using equation 4.2, equation 4.3, and equation 4.4 respectively. From table 5.3, it is evident that for each ANN model, the R^2 (coefficient of determination) value obtained was close to one, and

the RMSE and MAE values were low. We can, therefore, conclude that the models were accurate.

Table 5.3: Table showing the validation metrics of the ANN models constructed, trained and validated for training and validating the models.

ANN model	RMSE		MAE		R ²	
	Calibration (±0.0094)	Test (±0.012)	Calibration (±0.0018)	Test (±0.0050)	Calibration (±0.0048)	Test (±0.0042)
Estradiol pg/ml	0.0286	0.0169	0.0074	0.0069	0.9963	0.9986
FSH (mIU/ml)	0.0540	0.0602	0.0121	0.0213	0.9805	0.9828
LH (mIU/ml)	0.0544	0.0521	0.0122	0.0184	0.9815	0.9862
Progesterone (ng/ml)	0.0171	0.0113	0.0049	0.0050	0.9985	0.9988

The performance of the models was again evaluated using the regression plots of predicted concentration versus the known (see Figure 5.13).

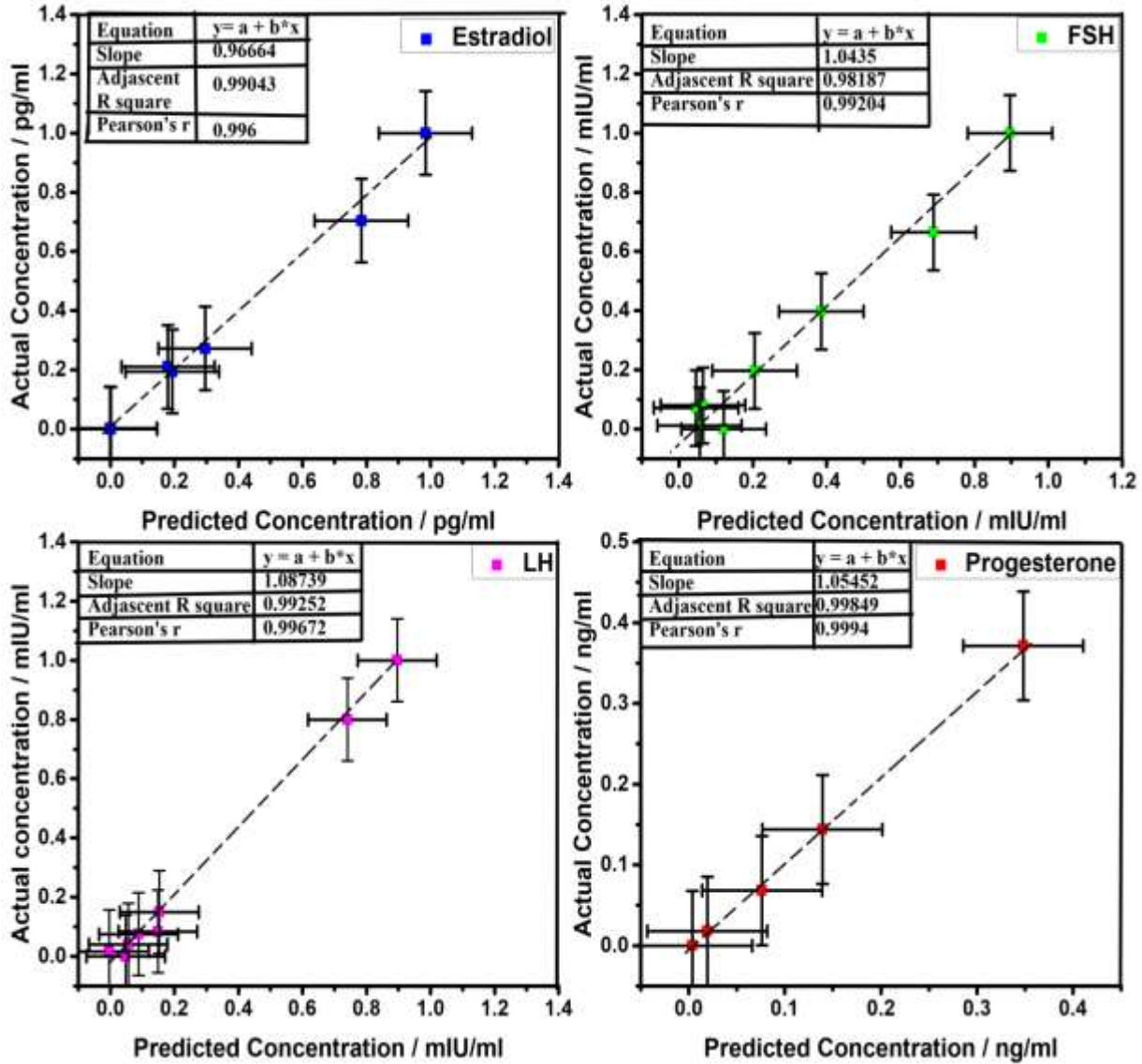


Figure 5.13: Figure showing the regression plots showing predicted against actual concentration obtained from biomarker-Raman band for (a) Estradiol (in picogram/milliliter(pg/ml)), (b) FSH (in milli-International Units/milliliter (mIU/ml)), (c) LH (in milli-International Units/milliliter (mIU/ml))and (d) Progesterone (in nanogram/milliliter (ng/ml)).

For each hormone (Figure 5.13), the Pearson correlation coefficient (r), and the correlation squared (R^2) values are further away from zero but close to one, indicating the strong linear relationship between the predicted and actual concentration. The data used to plot the regression plot in figure 5.13 was presented in table 5.4.

Table 5.4: Table showing the results of predicted and actual hormone concentration obtained from the developed ANN Models

Estradiol hormone (±0.15 pg/ml)		Follicle-stimulating hormone (±0.11 mIU/ml)		Luteinizing hormone (±0.12 mIU/ml)		Progesterone (±0.06 ng/ml)	
Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual
0.001	0	0.12107	0	0.04776	0	0.00371	0
0.00102	0.0026	0.05578	0.01265	0.00358	0.01666	0.01938	0.0178
0.19356	0.19481	0.04665	0.07129	0.05584	0.04	0.07635	0.06818
0.17995	0.21039	0.06501	0.08032	0.08803	0.075	0.13905	0.14394
0.29574	0.27273	0.20475	0.19679	0.14871	0.084	0.34826	0.37121
0.78419	0.70439	0.38495	0.39759	0.15283	0.15		
0.98419	1	0.68886	0.66466	0.74029	0.8		
		0.89594	1	0.89575	1		

In all cases, data were close to the regression lines revealing that the predicted concentration of the testing set was in acceptable good agreement with the experimental (actual) concentration. This confirms the excellent performance of the models.

5.6.1.1: Determination of Limit of Detection (LOD)

The LOD was calculated using the following equation (Desimoni and Brunetti, 2015):

$$LOD = K \frac{\sigma_B}{S} \dots\dots\dots(5.1)$$

in which σ_B was the standard deviation of blank signals (Raman spectra of male mouse's blood), S was the slope of the graph of the area under the curve against concentration, and k was the expansion factor chosen according to the acceptable false positive error (α) and false-negative

error (β) values, and is approximately three ($k \approx 3$) (Desimoni and Brunetti, 2015). The LOD results for the four female reproductive hormones displayed in table 5.4 were calculated using equation 5.1.

Table 5.5: Table showing the limit of detection for the ANN models developed in this study and the ones reported in the literature.

Female Reproductive hormone	Limit of Detection (LOD)	Reported LOD in literature	References
Estradiol (pg/ml)	38	20- 42 (HPLC)	(Kumar <i>et al.</i> , 2014)
Follicle-stimulating hormone (mIU/ml)	0.45	1.92 (ELISA)	(Robertson <i>et al.</i> , 2001)
Luteinizing hormone (mIU/ml)	0.69	1.92 (ELISA)	(Robertson <i>et al.</i> , 2001)
Progesterone (ng/ml)	7.14	0.6 (HPLC); 0.2 (ELISA)	(Gautam and Purvis, 2017)

The reported LOD for these hormones include; For estradiol they range from 20-42 pg/ml (0.02-0.042 ppb) using HPLC (Kumar *et al.*, 2014), and 2.5 ng/l for ELISA (Farré *et al.*, 2007). For progesterone were 0.6 μ g/ml (0.6 ppm) using HPLC (Gautam and Purvis, 2017) and 0.2 ng/ml for ELISA (Khatun *et al.*, 2009). For FSH and LH were 0.010 IU/l (5 ppm) for ELISA (Robertson *et al.*, 2001) and 5 pg/ml for LC-MS/MS. The LOD in Table 5.4 is incredibly lower than those reported in the literature suggesting the accuracy or high sensitivity of our Raman spectroscopic technique.

5.6.2: Quantification of hormones in mice blood samples using ANN

Reduction in the reproductive life span is on the increase due to the imbalance of reproductive hormone levels in the blood (Kaingu and Odumaa, 2019). This is probably due to the use of analgesic drugs such as acetaminophen (Gadd *et al.*, 2002). Gates *et al.* (2010) reported on the effects of acetaminophen on reproductive hormone levels using the chromatographic technique. The study reported a reduction in estrogen levels in female mice treated with acetaminophen. Cohen and the co-authors reported a decline in testosterone levels in males treated with acetaminophen using high-performance liquid chromatography-tandem mass spectrometry

(Cohen *et al.*, 2018). In contrast to Cohen and Gates, Oyedeji *et al.* (2013) reported non-significant changes in ovarian and extra-ovarian hormone levels in rats treated with acetaminophen to normal rats.

Once the ANN models were ascertained to be accurate, they were used to determine levels of estradiol, Follicle Stimulating Hormone (FSH), Luteinizing Hormone (LH) and progesterone hormones in blood from treated and normal mice. The average values of the predicted concentration levels of the respective hormones were calculated per stage of the estrous cycle across the treatment period in each of the groups (treated and normal) and displayed in Table 5.6. Each group had five mice, where the treated group received acetaminophen for 20 days according to their weight.

Table 5.6: Table showing the average of female reproductive hormone levels in the blood for different stages of the Estrous cycle from acetaminophen treated and normal mice for a period of 20 days. These values were obtained when Raman spectral data were applied onto the build ANN models.

Estrous cycle	Estradiol hormone (pg/ml)		Follicle-Stimulating hormone (FSH) (mIU/ml)		Luteinizing hormone (LH) (mIU/ml)		Progesterone hormone (ng/ml)	
	Normal (± 2.77)	Treated (± 8.18)	Normal (± 0.30)	Treated (± 0.46)	Normal (± 0.73)	Treated (± 0.72)	Normal (± 0.87)	Treated (± 0.98)
Proestrus	114.59	101.59	3.033	3.413	8.939	9.591	12.08	9.252
Estrus	125.39	102.73	3.666	2.890	10.21	9.029	13.39	10.41
Metestrus	119.34	136.67	3.500	4.900	6.728	12.10	15.27	13.79
Diestrus	113.13	110.31	2.341	3.085	9.130	9.166	11.29	11.80

The displayed data of female reproductive hormone levels (on average) in blood taken from acetaminophen treated and normal Swiss albino mice throughout the estrous cycles in Table 5.6 was plotted in Figure 5.14 for clear visualization.

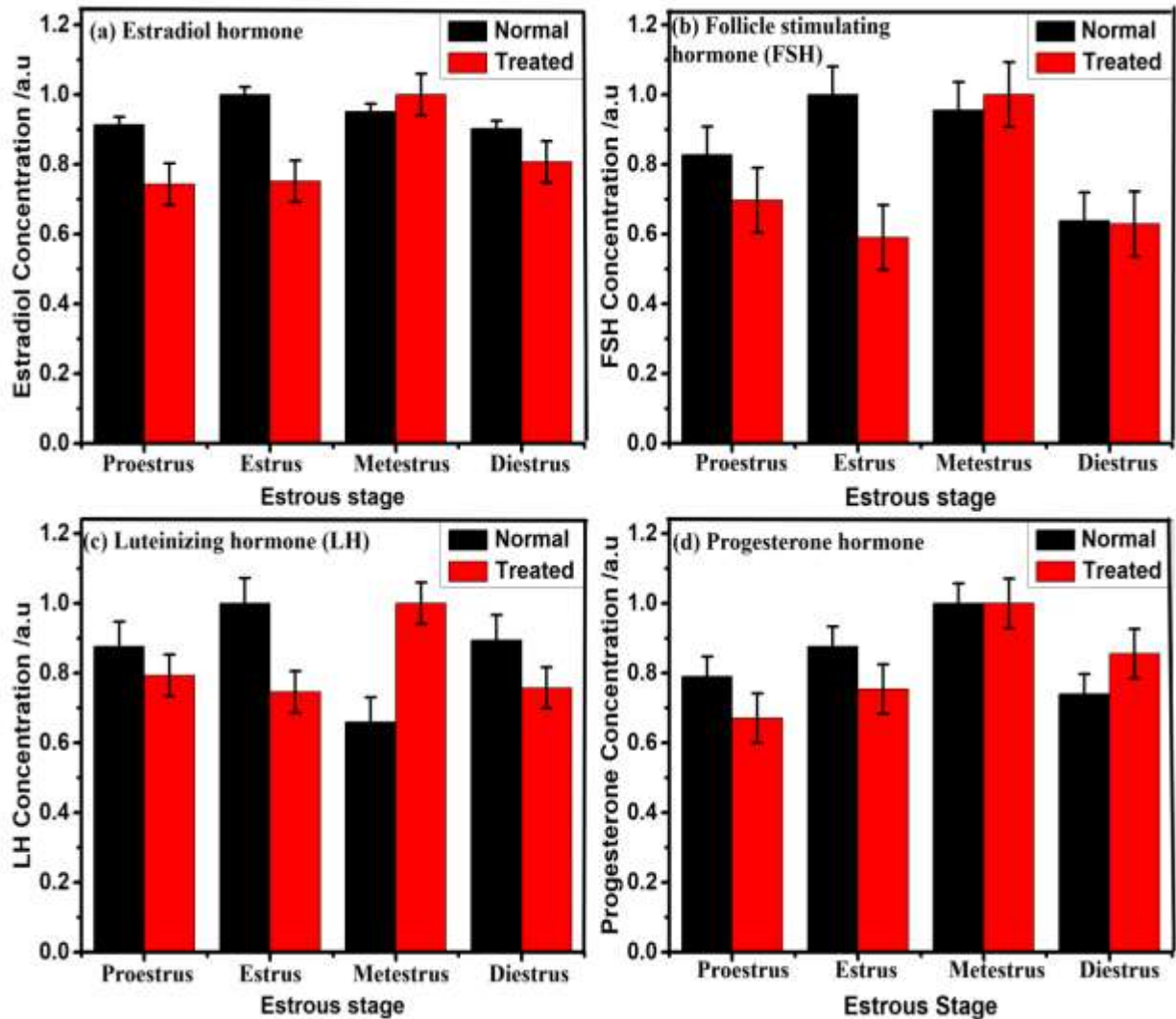


Figure 5.14: Figure showing the average of female reproductive hormone levels in the blood for different stages of the Estrous cycle from acetaminophen treated and normal mice for a period of 20 days. These values were obtained when Raman spectral data were applied onto the build ANN models.

$$\text{hormone level \% increase or decrease} = \frac{\text{Mean treated hormone level}}{\text{mean normal hormone level}} \times 100 - 100 \dots\dots\dots 5.1$$

Using the equation 5.1, the percentage increase or decrease in the hormone levels (on average) obtained from the four ANN model as displayed in Table 5.6 was calculated. It was seen that in treated mice estradiol, FSH, LH and progesterone increased by 29%, 138%, 1.1% and 44% from the normal respectively.

The levels of estradiol in treated mice declined in the proestrus, estrus, and diestrus phases compared to the normal. With a subsequent increase in estradiol levels during the metestrus phase compared to the normal (Table 5.6 and Figure 5.14). The levels of FSH increased in

treated mice in the proestrus, metestrus, and diestrus phases compared to the normal. With a subsequent non-significant decline during the estrus phase compared to the normal (Table 5.6 and Figure 5.14). The levels of LH increased in treated mice in the proestrus, metestrus, and diestrus phases compared to the normal. With a subsequent non-significant decline during the estrus phase compared to the normal. The levels of progesterone declined in treated mice in the proestrus, estrus, and metestrus phases compared to the normal. With a subsequent non-significant decline increase during the diestrus phase compared to the control (Table 5.6 and Figure 5.14).

According to Caligioni (2009) and Paccola *et al* (2013) under normal circumstances, the levels of estradiol increase during the proestrus phase of the cycle, continue to increase to the estrus phase of the cycle. During the metestrus phase of the cycle, the estradiol levels decline and start to increase during the diestrus phase of the cycle. But this is not the case in treated mice (Table 5.6 and Figure 5.14) where the levels of estradiol are the highest in metestrus. FSH levels increase from the proestrus phase to the estrus phase of the cycle where they peak towards the end of estrus, then decline in metestrus and the diestrus phase of the cycle (Byers *et al.*, 2012). On the contrary, in treated mice, the FSH levels decline in estrus instead of increasing (Table 5.6 and Figure 5.14). The LH levels in proestrus increase towards estrus, where they are supposed to peak at the beginning of estrus, decline during the metestrus and diestrus phase of the cycle (Caligioni, 2009). This is not what was observed in treated mice for the LH level is highest in the metestrus phase of the cycle (Table 5.6 and Figure 5.14). Progesterone levels decline in proestrus and estrus, then rise in the metestrus and diestrus phase of the cycle (Gava *et al.*, 2004). There are non-significant changes in progesterone levels in treated mice (Table 5.5).

To explore further, the predicted concentration of hormones in blood from treated and normal mice for some of the treatment days is shown in Table 5.7. Mice take 4-5days to complete the estrous cycle (Rudolph *et al.*, 2012), we thus assumed that the mice took five days. This gave us the reason to show data for days one, six, eleven, sixteen and twenty. For clear visualization, data in Table 5.6 was plotted to have predicted concentration of hormones in blood from treated and normal mice against the treatment days (Figure 5.15). As can be seen in figure 5.15 and Table 5.7, the estradiol levels in treated mice increased in day six, eleven and sixteen, and declined in day one and twenty when compared to the normal. The levels of FSH increased in treated mice on day eleven, and significantly declined in day one, six and sixteen and twenty when compared

to the normal. The LH levels increased in treated mice in days one, six and twenty and declined in day eleven and sixteen when compared to normal. The levels of progesterone in treated mice increased in day six and eleven, and declined in day one, sixteen and twenty when compared to normal.

Table 5.7: Table showing a comparison between the predicted concentration of hormones in blood from treated and normal mice in day one, six, eleven, sixteen and twenty days of the treatment period.

Treatment Days	Estradiol hormone (pg/ml)		Follicle-stimulating hormone (FSH) (mIU/ml)		Luteinizing hormone (LH) (mIU/ml)		Progesterone hormone (ng/ml)	
	Normal (± 20.4)	Treated (± 17.1)	Normal (± 0.36)	Treated (± 1.46)	Normal (± 1.08)	Treated (± 1.26)	Normal (± 2.24)	Treated (± 2.22)
One	122.64	108.18	0.593	1.098	6.309	10.83	13.52	10.98
Six	43.70	156.70	1.865	3.043	6.060	7.286	2.707	15.98
Eleven	129.71	166.98	1.632	9.260	10.94	10.22	8.678	20.98
Sixteen	59.76	73.61	0.985	1.512	11.06	4.373	7.086	8.242
Twenty	146.30	140.59	2.698	3.625	8.678	10.80	15.13	11.78
Mean	100.422	129.212	1.5546	3.7076	8.6094	8.7018	9.4242	13.5924

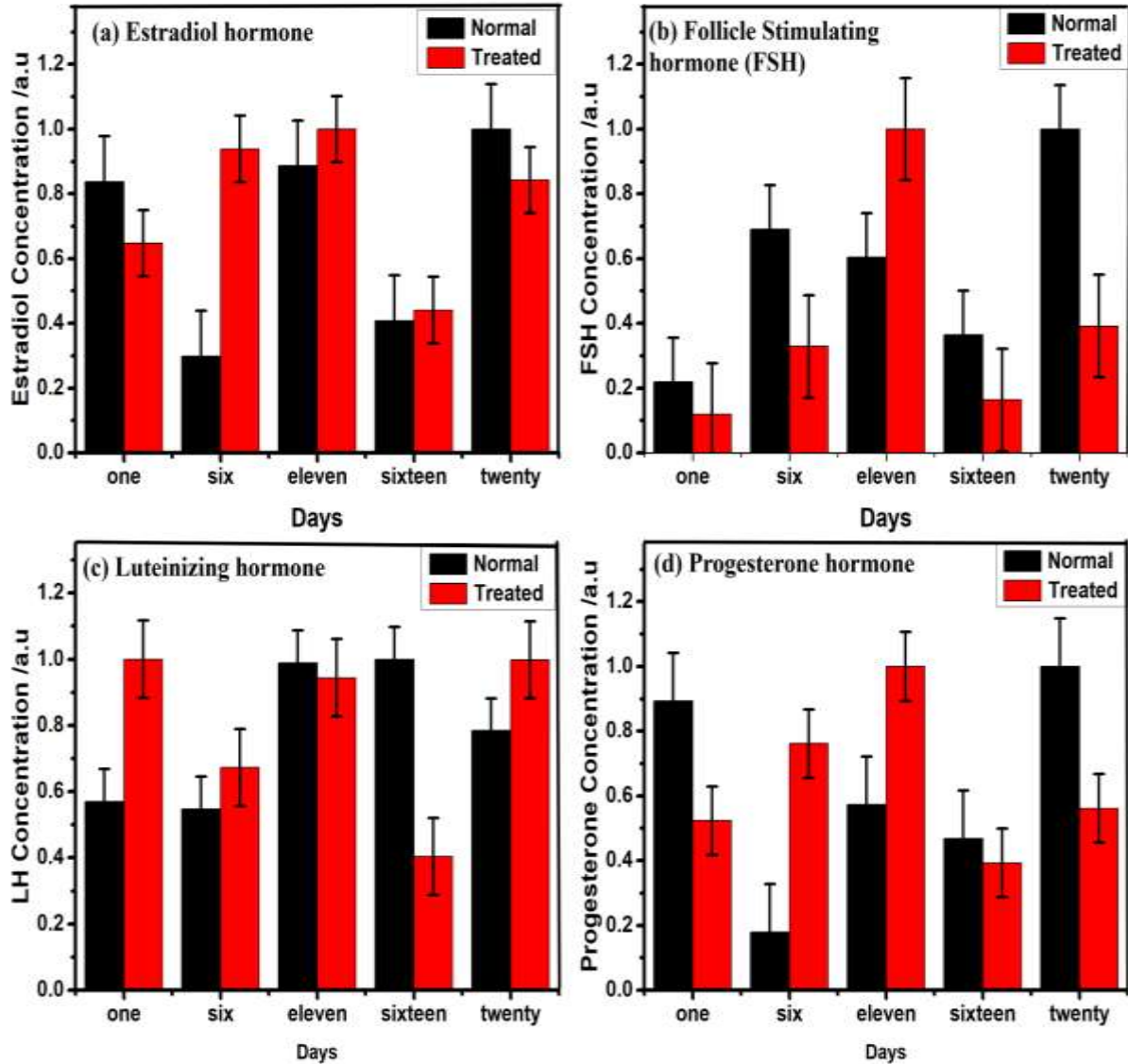


Figure 5.15: Figure showing a comparison between the predicted concentration of hormones in blood from treated and normal mice in day one, six, eleven, sixteen and twenty days of the treatment period.

The last day of the treatment period shows a significant decline in levels of estradiol, FSH and progesterone in treated mice when compared to the normal, and a significant increase in levels of LH (Figure 5.14). Kaingu *et al* (2017) reported that imbalances/alterations in the levels of female reproductive hormones might have led to ovarian function irregularity.

Table 5.8 displays the results of statistical analysis performed on the Raman dataset of actual samples (blood of treated and normal mice).

Table 5.8: Table showing z -test results for Raman dataset of blood of acetaminophen treated and untreated mice

Sample	Mean	Variance	Standard deviation	p_value	z-critical
Normal	0.238978644	0.028583427	0.16906634	3.25×10^{-12}	1.644854
Acetaminophen Treated	0.26044	0.025156711	0.158608673	6.51×10^{-12}	1.959964

Here the alpha level was left as default i.e. 0.05. Since p_values were less than the alpha level (see Table 5.8), the null hypothesis was rejected that there was no significant difference in the means of normal and acetaminophen treated mice's blood. We thus conclude that the results in Table 5.6, Table 5.7, Table 5.8, Figure 5.14 and Figure 5.15 suggest that acetaminophen causes hormonal variation thereby compromising fertility. This study, therefore, agrees with previous studies (Cohen *et al.* (2018); Gates *et al.* (2010)), which found that acetaminophen indeed affects hormone variations thus reducing fertility.

CHAPTER SIX: CONCLUSION AND RECOMMENDATION

6.1: Conclusion

This work has demonstrated the ability of SERS together with chemometrics in identifying the effects of acetaminophen on female mice's reproductive hormone levels in blood. Here, Raman substrates used (i.e. the conductive silver paste smear) were noted to be chemically stable since their Raman spectral profiles were the same for the first seven days after preparation. These Raman substrates were applied to female hormones (estradiol, FSH, LH and progesterone) and characterized using Raman spectroscopy upon 785 nm laser excitation. It was seen that Raman spectroscopy has the potential to distinguish between these hormones since their spectral profiles were different from each hormone. Using Principal Component Analysis (PCA) and analysis of variance (ANOVA), the spectral data set analyses to reveal these bands was done. Biomarker bands of the respective hormones in the blood were identified from the Raman spectroscopy of simulates (prepared by mixing each separately with male mouse's blood at different concentrations). The bands (biomarker bands for each hormone) were used to build Artificial Neural Network (ANN) models to achieve quantitative analysis. A test for model's accuracy was done using determination coefficient (R^2) and root mean square error (RMSE) values which were seen to be greater than 98.05% and less than 0.0602 respectively. Using the calibrated ANN models in determining the concentration hormone level in blood of mice, it was noted that mice treated with acetaminophen had on average a level increase of 29%, 138%, 1.1% and 44% in estradiol, FSH, LH and progesterone hormones respectively from normal. This implied that prolonged use of acetaminophen can indeed alter some female reproductive hormone levels in the blood thus compromising fertility.

6.3: Recommendation

We recommend the need for a more effective standard calibration procedure in which standard hormones aren't mixed with blood physically, but instead standards with blood in them.

REFERENCES

- Alves da Rocha, R., Paiva, I. M., Anjos, V., Furtado, M. A. M., and Bell, M. J. V. (2015). Quantification of whey in fluid milk using confocal Raman microscopy and artificial neural network. *J. Dairy Sci.* **98**, 3559–3567.
- Arendrup, F. S., Mazaud-Guittot, S., Jégou, B., and Kristensen, D. M. (2018). EDC IMPACT: Is exposure during pregnancy to acetaminophen/paracetamol disrupting female reproductive development? *Endocr. Connect.* **7**, 149–158.
- Atkins, C. G., Buckley, K., Blades, M. W., and Turner, R. F. B. (2017a). Raman Spectroscopy of Blood and Blood Components. *Appl. Spectrosc.* **71**, 767–793.
- Atkins, C. G., Buckley, K., Blades, M. W., and Turner, R. F. B. (2017b). Raman Spectroscopy of Blood and Blood Components. *Appl. Spectrosc.* **71**, 767–793.
- Auner, G. W., Koya, S. K., Huang, C., Broadbent, B., Trexler, M., Auner, Z., Elias, A., Mehne, K. C., and Brusatori, M. A. (2018). Applications of Raman spectroscopy in cancer diagnosis. *Cancer Metastasis Rev.* **37**, 691–717.
- Behnammorshedi, M. and Nazem, H. (2015). The Effect of Gold Nanoparticle on Luteinizing Hormone, Follicle Stimulating Hormone, Testosterone and Testis In Male Rat. *Biomed Res* **26**, 5.
- Bergholt, M. S., Zheng, W., Lin, K., Huang, Z., Ho, K. Y., Yeoh, K. G., Teh, M., and So, J. B. Y. (2011). Characterizing variability in in vivo Raman spectra of different anatomical locations in the upper gastrointestinal tract toward cancer detection. *J. Biomed. Opt.* **16**, 037003.
- Birech, Z., Mwangi, P. W., Bukachi, F., and Mandela, K. M. (2017). Application of Raman spectroscopy in type 2 diabetes screening in blood using leucine and isoleucine amino-acids as biomarkers and in comparative anti-diabetic drugs efficacy studies. *PloS One* **12**.
- Birech, Z., Mwangi, P. W., Sehmi, P. K., and Nyaga, N. M. (2019). Raman Spectroscopic comparative study of Oxytocin and Freeze-dried Extract of *Uvariadendron anisatum*

- Verdeck (Annonaceae) and their influence on diet induced obesity in Sprague Dawley rats. *BioRxiv*.
- Birech, Z., Ondieki, A. M., Opati, R. I., and Mwangi, P. W. (2020). Low cost Raman sample substrates from conductive silver paint smear for Raman spectroscopic screening of metabolic diseases in whole blood. *Vib. Spectrosc.*, 103063.
- Bonnier, F. and Byrne, H. J. (2012). Understanding the molecular information contained in principal component analysis of vibrational spectra of biological systems. *Analyst* **137**, 322–332.
- Borio, V. G., Vinha, R., Nicolau, R. A., Oliveira, H. P. M. de, Lima, C. J. de, and Silveira, L. (2012). Quantitative Evaluation of Acetaminophen in Oral Solutions by Dispersive Raman Spectroscopy for Quality Control. *Spectrosc. Int. J.* **27**, 215–228.
- Botting, R. M. (2000). Mechanism of Action of Acetaminophen: Is There a Cyclooxygenase 3? *Clin. Infect. Dis.* **31**, S202–S210.
- Butler, H. J., Ashton, L., Bird, B., Cinque, G., Curtis, K., Dorney, J., Esmonde-White, K., Fullwood, N. J., Gardner, B., Martin-Hirsch, P. L., Walsh, M. J., McAinsh, M. R., Stone, N., and Martin, F. L. (2016). Using Raman spectroscopy to characterize biological materials. *Nat. Protoc.* **11**, 664–687.
- Byers, S. L., Wiles, M. V., Dunn, S. L., and Taft, R. A. (2012). Mouse Estrous Cycle Identification Tool and Images. *PLoS ONE* **7**, e35538.
- Caligioni, C. S. (2009). Assessing Reproductive Status/Stages in Mice. In ‘Current Protocols in Neuroscience’ (J.N. Crawley, C.R. Gerfen, M.A. Rogawski, D.R. Sibley, P. Skolnick, S. Wray, Eds.), ppnsa04is48. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Cheng, J.-H. and Sun, D.-W. (2015). Recent Applications of Spectroscopic and Hyperspectral Imaging Techniques with Chemometric Analysis for Rapid Inspection of Microbial Spoilage in Muscle Foods: Microbial spoilage in muscle foods.... *Compr. Rev. Food Sci. Food Saf.* **14**, 478–490.

- Choquette, S. J., Etz, E. S., Hurst, W. S., Blackburn, D. H., and Leigh, S. D. (2007). Relative Intensity Correction of Raman Spectrometers: NIST SRMs 2241 through 2243 for 785 nm, 532 nm, and 488 nm/514.5 nm Excitation. *Appl. Spectrosc.* **61**, 117–129.
- Cohen, I. V., Cirulli, E. T., Mitchell, M. W., Jonsson, T. J., Yu, J., Shah, N., Spector, T. D., Guo, L., Venter, J. C., and Telenti, A. (2018). Acetaminophen (Paracetamol) Use Modifies the Sulfation of Sex Hormones. *EBioMedicine* **28**, 316–323.
- Cramer, D. W., Liberman, R. F., Hornstein, M. D., McShane, P., Powers, D., Li, E. Y., and Barbieri, R. (1998). Basal hormone levels in women who use acetaminophen for menstrual pain. *Fertil. Steril.* **70**, 371–373.
- Davidson, B., Murray, A. A., Elfick, A., and Spears, N. (2013). Raman Micro-Spectroscopy Can Be Used to Investigate the Developmental Stage of the Mouse Oocyte. *PLoS ONE* **8**, e67972.
- Desimoni, E. and Brunetti, B. (2015). About estimating the limit of detection by the signal to noise approach.
- Díaz-Cruz, M. S., López de Alda, M. J., López, R., and Barceló, D. (2003). Determination of estrogens and progestogens by mass spectrometric techniques (GC/MS, LC/MS and LC/MS/MS): Determination of estrogens and progestogens by MS techniques. *J. Mass Spectrom.* **38**, 917–923.
- Dingari, N. C., Horowitz, G. L., Kang, J. W., Dasari, R. R., and Barman, I. (2012). Raman Spectroscopy Provides a Powerful Diagnostic Tool for Accurate Determination of Albumin Glycation. *PLoS ONE* **7**, e32406.
- Downes, A. and Elfick, A. (2010). Raman Spectroscopy and Related Techniques in Biomedicine. *Sensors* **10**, 1871–1889.
- Duraipandian, S., Zheng, W., Ng, J., Low, J. J. H., Ilancheran, A., and Huang, Z. (2013). Non-invasive analysis of hormonal variations and effect of postmenopausal Vagifem treatment on women using in vivo high wavenumber confocal Raman spectroscopy. *The Analyst* **138**, 4120.

- Farré, M., Kuster, M., Brix, R., Rubio, F., Alda, M.-J. L. de, and Barceló, D. (2007). Comparative study of an estradiol enzyme-linked immunosorbent assay kit, liquid chromatography–tandem mass spectrometry, and ultra performance liquid chromatography–quadrupole time of flight mass spectrometry for part-per-trillion analysis of estrogens in water samples. *J. Chromatogr. A* **1160**, 166–175.
- Ferrari, A. C. (2001). A Model to Interpret the Raman Spectra of Disordered, Amorphous and Nanostructured Carbons. *MRS Proc.* **675**, W11.5.1.
- Fine, P. G. and Rosenfeld, M. J. (2013). The Endocannabinoid System, Cannabinoids, and Pain. *Rambam Maimonides Med. J.* **4**.
- Gadd, S. L., Hobbs, G., and Miller, M. R. (2002). Acetaminophen-induced proliferation of estrogen-responsive breast cancer cells is associated with increases in c-myc RNA expression and NF- κ B activity. *Toxicol. Sci.* **66**, 233–243.
- Galata, Farkas, Könyves, Mészáros, Szabó, Csontos, Pálos, Marosi, Nagy, and Nagy (2019). Fast, Spectroscopy-Based Prediction of In Vitro Dissolution Profile of Extended Release Tablets Using Artificial Neural Networks. *Pharmaceutics* **11**, 400.
- Gates, M. A., Tworoger, S. S., Eliassen, A. H., Missmer, S. A., and Hankinson, S. E. (2010). Analgesic use and sex steroid hormone concentrations in postmenopausal women. *Cancer Epidemiol. Prev. Biomark.* **19**, 1033–1041.
- Gautam, P. and Purvis, T. (2017). *Pharmaceutical Analytical Chemistry*.
- Gautam, R., Vanga, S., Ariese, F., and Umopathy, S. (2015). Review of multidimensional data processing approaches for Raman and infrared spectroscopy. *EPJ Tech. Instrum.* **2**, 8.
- Gava, N., Clarke, C. L., Byth, K., Arnett-Mansfield, R. L., and deFazio, A. (2004). Expression of progesterone receptors A and B in the mouse ovary during the estrous cycle. *Endocrinology* **145**, 3487–3494.

- Gavrilova, N. and Lindau, S. T. (2009). Salivary Sex Hormone Measurement in a National, Population-Based Study of Older Adults. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* **64B**, i94–i105.
- Ghanem, C. I., Pérez, M. J., Manautou, J. E., and Mottino, A. D. (2016). Acetaminophen from liver to brain: New insights into drug pharmacological action and toxicity. *Pharmacol. Res.* **109**, 119–131.
- Gonzaga, L. W., Botelho, M. A., Queiroz, D. B., Fechine, P. B., Freire, R., Azevedo, E., Morais, A., Ruela, R., Lyra, A., Gomes, S., and Jr, L. Q. (2012). Nanotechnology in Hormone Replacement Therapy: Safe and Efficacy of Transdermal Estriol and Estradiol Nanoparticles after 5 Years Follow-Up Study. *Lat. Am. J. Pharm.*, 9.
- Guiloski, I. C., Ribas, J. L. C., Piancini, L. D. S., Dagostim, A. C., Cirio, S. M., Fávoro, L. F., Boschen, S. L., Cestari, M. M., Cunha, C. da, and Silva de Assis, H. C. (2017). Paracetamol causes endocrine disruption and hepatotoxicity in male fish *Rhamdia quelen* after subchronic exposure. *Environ. Toxicol. Pharmacol.* **53**, 111–120.
- Han, X. X., Pienpinijtham, P., Zhao, B., and Ozaki, Y. (2011). Coupling Reaction-Based Ultrasensitive Detection of Phenolic Estrogens Using Surface-Enhanced Resonance Raman Scattering. *Anal. Chem.* **83**, 8582–8588.
- Herrero, A. M. (2008). Raman spectroscopy for monitoring protein structure in muscle food systems. *Crit. Rev. Food Sci. Nutr.* **48**, 512–523.
- Hombal, P., Gudadappanavar, A., and Kavi, A. (2016). Randomized clinical trial comparing the effect of paracetamol with diclofenac in combination with on request rescue analgesic tramadol: analgesic efficacy, safety and tolerability after abdominal surgery. *Int. Surg. J.*, 557–561.
- Huang, Z., Lui, H., Chen, X. K., Alajlan, A., McLean, D. I., and Zeng, H. (2004). Raman spectroscopy of in vivo cutaneous melanin. *J. Biomed. Opt.* **9**, 1198.
- Hurtado-Gonzalez, P., Anderson, R. A., Macdonald, J., Driesche, S. van den, Kilcoyne, K., Jørgensen, A., McKinnell, C., Macpherson, S., Sharpe, R. M., and Mitchell, R. T. (2018).

- Effects of Exposure to Acetaminophen and Ibuprofen on Fetal Germ Cell Development in Both Sexes in Rodent and Human Using Multiple Experimental Systems. *Environ. Health Perspect.* **126**, 047006.
- Hurtado-Gonzalez, P. and Mitchell, R. T. (2017). Analgesic use in pregnancy and male reproductive development: *Curr. Opin. Endocrinol. Diabetes Obes.* **24**, 225–232.
- Ioele, G., De Luca, M., Dinç, E., Oliverio, F., and Ragno, G. (2011). Artificial neural network combined with principal component analysis for resolution of complex pharmaceutical formulations. *Chem. Pharm. Bull. (Tokyo)* **59**, 35–40.
- Kaingu, C. K., Oduma, J. A., Mbaria, J., and Kiama, S. G. (2017). Effects of *Croton menyharthii* and *Uvariadendron kirkii* extracts on ovarian corpora lutea and reproductive hormones. *Discov. Phytomedicine* **4**, 21.
- Kaingu, C. K. and Odumaa, J. A. (2019). Potential of *Croton menyharthii* and *Uvariadendron kirkii* aqueous extracts as female Novel contraceptives. *Discov. Phytomedicine* **6**, 112–118.
- Kalantri, P. P., Somani, R. R., and Makhija, D. T. (2010). Raman spectroscopy: A potential technique in analysis of pharmaceuticals. 12.
- Kanter, E. M., Vargis, E., Majumder, S., Keller, M. D., Woeste, E., Rao, G. G., and Mahadevan-Jansen, A. (2009). Application of Raman spectroscopy for cervical dysplasia diagnosis. *J. Biophotonics* **2**, 81–90.
- Khatun, S., Nara, S., Tripathi, V., Rangari, K., Chaube, S. K., Kariya, K. P., Kumar, S., and Shrivastav, T. G. (2009). Development of ELISA for Measurement of Progesterone Employing 17- α -OH-P-HRP as Enzyme Label. *J. Immunoass. Immunochem.* **30**, 186–196.
- Klinger-Gratz, P. P., Ralvenius, W. T., Neumann, E., Kato, A., Nyilas, R., Lele, Z., Katona, I., and Zeilhofer, H. U. (2018). Acetaminophen Relieves Inflammatory Pain through CB₁ Cannabinoid Receptors in the Rostral Ventromedial Medulla. *J. Neurosci.* **38**, 322–334.

- Konforte, D., Shea, J. L., Kyriakopoulou, L., Colantonio, D., Cohen, A. H., Shaw, J., Bailey, D., Chan, M. K., Armbruster, D., and Adeli, K. (2013). Complex biological pattern of fertility hormones in children and adolescents: a study of healthy children from the CALIPER cohort and establishment of pediatric reference intervals. *Clin. Chem.* **59**, 1215–1227.
- Kumar, R., Malik, A. K., Kabir, A., and Furton, K. G. (2014). Efficient analysis of selected estrogens using fabric phase sorptive extraction and high performance liquid chromatography-fluorescence detection. *J. Chromatogr. A* **1359**, 16–25.
- Lakhwani, G. R., Sherikar, O. D., and Mehta, P. J. (2013). Nondestructive and Rapid Concurrent Estimation of Paracetamol and Nimesulide in Their Combined Dosage Form Using Raman Spectroscopic Technique. *Indian J. Pharm. Sci.*, 6.
- Lauretta, R., Sansone, A., Sansone, M., Romanelli, F., and Appetecchia, M. (2019). Endocrine Disrupting Chemicals: Effects on Endocrine Glands. *Front. Endocrinol.* **10**, 178.
- Li, J., Kim, B. K., Wang, K.-K., Im, J.-E., Choi, H. N., Kim, D.-H., Cho, S. I., Lee, W.-Y., and Kim, Y.-R. (2016). Sensing Estrogen with Electrochemical Impedance Spectroscopy. *J. Anal. Methods Chem.* **2016**, 1–6.
- Li, Y., Du, G., Cai, W., and Shao, X. (2011). Classification and Quantitative Analysis of Azithromycin Tablets by Raman Spectroscopy and Chemometrics. *Am. J. Anal. Chem.* **02**, 135–141.
- Liu, Y., Chen, Y., Zhang, Y., Kou, Q., Zhang, Y., Wang, Y., Chen, L., Sun, Y., Zhang, H., and MeeJung, Y. (2018a). Detection and Identification of Estrogen Based on Surface-Enhanced Resonance Raman Scattering (SERRS). *Molecules* **23**, 1330.
- Liu, Y., Chen, Y., Zhang, Y., Kou, Q., Zhang, Y., Wang, Y., Chen, L., Sun, Y., Zhang, H., and MeeJung, Y. (2018b). Detection and Identification of Estrogen Based on Surface-Enhanced Resonance Raman Scattering (SERRS). *Molecules* **23**, 1330.
- Liu, Y., Zhu, Y., Di, L., Osterberg, E. C., Liu, F., He, L., Hu, H., Huang, Y., Li, P. S., and Li, Z. (2014). Raman spectroscopy as an ex vivo noninvasive approach to distinguish complete

- and incomplete spermatogenesis within human seminiferous tubules. *Fertil. Steril.* **102**, 54-60.e2.
- Matyas, R. A., Mumford, S. L., Schliep, K. C., Ahrens, K. A., Sjaarda, L. A., Perkins, N. J., Filiberto, A. C., Mattison, D., Zarek, S. M., Wactawski-Wende, J., and Schisterman, E. F. (2015). Effects of over-the-counter analgesic use on reproductive hormones and ovulation in healthy, premenopausal women. *Hum. Reprod.* **30**, 1714–1723.
- Mouslech, Z. and Valla, V. (2009). Endocannabinoid system: an overview of its potential in current medical practice. *Neuroendocrinol. Lett.* **30**, 153–179.
- Ngeranwa JN, K. K. J. (2015). A Review of the Biochemical, Hematological and Histological Modulations in Acetaminophen Induced Hepatotoxicity and the Potential of Urtica Dioica in the Regeneration of the Liver. *J. Drug Metab. Toxicol.* **06**.
- O'Brien, C. M., Herington, J. L., Brown, N., Pence, I. J., Paria, B. C., Slaughter, J. C., Reese, J., and Mahadevan-Jansen, A. (2017). In vivo Raman spectral analysis of impaired cervical remodeling in a mouse model of delayed parturition. *Sci. Rep.* **7**, 6835.
- Otange, B. O., Birech, Z., Okonda, J., and Rop, R. (2017). Conductive silver paste smeared glass substrates for label-free Raman spectroscopic detection of HIV-1 and HIV-1 p24 antigen in blood plasma. *Anal. Bioanal. Chem.* **409**, 3253–3259.
- Oyedepi, K. O., Bolarinwa, A. F., and Adabanjia, R. B. (2013). Evaluation of haematologica l and reproductive effect of paracetamol (acetaminophen) in Female Albin o Rats. *IOSR J Dent Medic Sci* **3**, 72–75.
- Paccola, C. C., Resende, C. G., Stumpp, T., Miraglia, S. M., and Cipriano, I. (2013). The rat estrous cycle revisited: a quantitative and qualitative analysis. **7**.
- Potcoava, M. C., Futia, G. L., Aughenbaugh, J., Schlaepfer, I. R., and Gibson, E. A. (2014). Raman and coherent anti-Stokes Raman scattering microscopy studies of changes in lipid content and composition in hormone-treated breast and prostate cancer cells. *J. Biomed. Opt.* **19**, 111605.

- Ramírez-González, J. A., Vaamonde-Lemos, R., Cunha-Filho, J. S., Varghese, A. C., and Swanson, R. J. (2016). Overview of the Female Reproductive System. *In* 'Exercise and Human Reproduction' (D. Vaamonde, S.S. du Plessis, A. Agarwal, Eds.), pp19–46. Springer New York, New York, NY.
- Robertson, D. M., Pruysers, E., Stephenson, T., Pettersson, K., Morton, S., and McLachlan, R. I. (2001). Sensitive LH and FSH assays for monitoring low serum levels in men undergoing steroidal contraception. *Clin. Endocrinol. (Oxf.)* **55**, 331–339.
- Rousseau, R., Govaerts, B., Verleysen, M., and Boulanger, B. (2008). Comparison of some chemometric tools for metabonomics biomarker identification. *Chemom. Intell. Lab. Syst.* **91**, 54–66.
- Rudolph, M., Döcke, W.-D., Müller, A., Menning, A., Röse, L., Zollner, T. M., and Gashaw, I. (2012). Induction of Overt Menstruation in Intact Mice. *PLoS ONE* **7**, e32922.
- Schultz, S., DeSilva, M., Gu, T. T., Qiang, M., and Whang, K. (2012). Effects of the Analgesic Acetaminophen (Paracetamol) and its para-Aminophenol Metabolite on Viability of Mouse-Cultured Cortical Neurons: EFFECTS OF THE ANALGESIC ACETAMINOPHEN (PARACETAMOL) AND ITS PARA-AMINOPHENOL METABOLITE. *Basic Clin. Pharmacol. Toxicol.* **110**, 141–144.
- Seli, E., Sakkas, D., Scott, R., Kwok, S. C., Rosendahl, S. M., and Burns, D. H. (2007). Noninvasive metabolomic profiling of embryo culture media using Raman and near-infrared spectroscopy correlates with reproductive potential of embryos in women undergoing in vitro fertilization. *Fertil. Steril.* **88**, 1350–1357.
- Shenglong Zou and Ujendra Kumar (2018). Cannabinoid Receptors and the Endocannabinoid System: Signaling and Function in the Central Nervous System. *Int. J. Mol. Sci.* **19**, 833.
- Sui, H., Wang, Y., Zhang, X., Wang, X., Cheng, W., Su, H., Wang, X., Sun, X., Han, X. X., Zhao, B., and Ozaki, Y. (2016). Ultrasensitive detection of thyrotropin-releasing hormone based on azo coupling and surface-enhanced resonance Raman spectroscopy. *The Analyst* **141**, 5181–5188.

- Suzuki, S. and Yoshimura, M. (2017). Chemical stability of graphene coated silver substrates for surface-enhanced Raman scattering. *Sci. Rep.* **7**, 1–7.
- Toda, K. (2018). Acetaminophen is not Safe in Pregnancy. **1**, 6.
- Ullah, R., Khan, S., Farman, F., Bilal, M., Krafft, C., and Shahzad, S. (2019). Demonstrating the application of Raman spectroscopy together with chemometric technique for screening of asthma disease. *Biomed. Opt. Express* **10**, 600.
- Vargis, E., Byrd, T., Logan, Q., Khabele, D., and Mahadevan-Jansen, A. (2011). Sensitivity of Raman spectroscopy to normal patient variability. *J. Biomed. Opt.* **16**, 117004.
- Vedad, J., Mojica, E.-R. E., and Desamero, R. Z. B. (2018). Raman spectroscopic discrimination of estrogens. *Vib. Spectrosc.* **96**, 93–100.
- Wahadoszamen, Md., Rahaman, A., Hoque, N. Md. R., I Talukder, A., Abedin, K. M., and Haider, A. F. M. Y. (2015). Laser Raman Spectroscopy with Different Excitation Sources and Extension to Surface Enhanced Raman Spectroscopy. *J. Spectrosc.* **2015**, 1–8.
- Wang, Q., Zheng, N., Li, Z., and Ma, Z. (2015). Quantitative analysis of glucose in whole blood using FT-Raman spectroscopy and artificial neural network. In 'Proceedings of the 2015 International Conference on Computational Science and Engineering'. Atlantis Press, Qingdao, China.
- Winnard, P. T., Zhang, C., Vesuna, F., Kang, J. W., Garry, J., Dasari, R. R., Barman, I., and Raman, V. (2017). Organ-specific isogenic metastatic breast cancer cell lines exhibit distinct Raman spectral signatures and metabolomes. *Oncotarget* **8**.
- Zhang, X., Young, M. A., Lyandres, O., and Van Duyne, R. P. (2005). Rapid Detection of an Anthrax Biomarker by Surface-Enhanced Raman Spectroscopy. *J. Am. Chem. Soc.* **127**, 4484–4489.
- Zheng, F., Qin, Y., and Chen, K. (2007). Sensitivity map of laser tweezers Raman spectroscopy for single-cell analysis of colorectal cancer. *J. Biomed. Opt.* **12**, 034002.

Zhu, S., Song, Z., Shi, S., Wang, M., and Jin, G. (2019). Fusion of Near-Infrared and Raman Spectroscopy for In-Line Measurement of Component Content of Molten Polymer Blends. *Sensors* **19**, 3463.

APPENDICES

Appendix 1: Neural Network script in R

```
#### loading libraries

library(neuralnet)

library(devtools)

library(caret)

library(ggplot2)

library(boot)

library(plyr)

#### reading data

mydata=read.csv('data.csv',sep="," ,header=TRUE)

#### normalisation of data

normalize <- function(x) {

  x <- as.numeric(x)

  return((x - min(x)) / (max(x) - min(x)))

}

Data <- as.data.frame(apply(mydata, 2, normalize))

#### data partitioning

indexs <- sample(1:nrow(Data), round(0.70*nrow(Data)))

train_data <- Data[indexs, ]

test_data <- Data[-indexs, ]

#### plotting the model
```

```

model=neuralnet(conc~.,Data, hidden=10, act.fct = "logistic", threshold = 0.01, algorithm =
"rprop+", err.fct = "sse", linear.output = T)

model$net.result

plot(model)

### predicting the model

predcal_data = compute(model,train_data[,c(1:26)])

predcal_data$net.result

pred_trainNN = predcal_data$net.result

actual_train = train_data$conc

actual_train = data.frame(pred = predcal_data$net.result,

                           actual = train_data$conc)

actual_train

write.csv(actual_train,"est_train_regression.csv")

### rms(error of prediction obtained)

RMSEcal_model = (sum((train_data$conc-pred_trainNN)^2)/nrow(train_data)) ^ 0.5

### Rsquared_model

Rsquaredcal_model = 1 - sum((train_data$conc-pred_trainNN)^2)/sum(train_data$conc-
(sum(train_data$conc)/nrow(train_data))^2)

### Mean absolute error obtained

MAEcal_model = (sum((train_data$conc-pred_trainNN)^2)^0.5 / nrow(train_data))

### predicting the model

pred_data = compute(model,test_data[,c(1:26)])

```

```

pred_data$net.result

actual_pred = data.frame(pre = pred_data$net.result,

                          actual = test_data$conc)

actual_pred

write.csv(actual_pred,"est_pred_regression.csv")

plot(actual_pred, main="Predicted concentration vs Actual concentration", col='blue' ,
      xlab="Actual concentration", ylab="Predicted concentration", pch=16, type="p")

abline(0,1, col="black")

##calculating the root mean square error

cor(actual_pred)

pred_testNN = (pred_data$net.result)

##rms(error obtained)

RMSEpred_model = (sum((test_data$conc-pred_testNN)^2)/nrow(test_data)) ^ 0.5

##Rsquared_model = 1 - sum((datatest$conc-pred_testNN)^2)/sum(datatest$conc-
(sum(datatest$conc)/nrow(datatest))^2)

Rsquaredpred_model = 1 - sum((test_data$conc-pred_testNN)^2)/sum(test_data$conc-
(sum(test_data$conc)/nrow(test_data))^2)

### Mean absolute error obtained

MAEpred_model = (sum((test_data$conc-pred_testNN)^2)^0.5 / nrow(test_data))

### ANN model using caret (model validation)

##loading libraries

library(caret)

```

```

##library(mlbench)

##set seed for reproducible results

##read input

set.seed(1)

input_data <- read.csv("data.csv",
                      sep=',', header = T)

n <- nrow(input_data)

##normalisation of data

normalize <- function(x) {

  x <- as.numeric(x)

  return((x - min(x)) / (max(x) - min(x)))

}

normalisedData <- as.data.frame(apply(input_data, 2, normalize))

##number of fold (if k=n, leave one out cross validation)

k <- n

folds <- createFolds(1:n, k, list = TRUE,returnTrain = FALSE)

predValue <- NULL

nn_result <- NULL

for(fold in folds){

trainNN <- normalisedData[-fold, ]

testNN <- normalisedData[fold, ]

```

```

form <- as.formula("conc~.")

NeuralNet <- neuralnet(

  form, trainNN, hidden = 10, act.fct = "logistic", threshold = 0.01,

  stepmax = 1e+05, rep = 1, startweights = NULL,

  learningrate.limit = NULL,

  learningrate.factor = list(minus = 0.5, plus = 1.2),

  learningrate = NULL, lifesign = "none",

  lifesign.step = 1000, algorithm = "rprop+",

  err.fct = "sse", linear.output = TRUE, exclude = NULL,

  constant.weights = NULL, likelihood = FALSE)

}

plot(NeuralNet)

###predicting the model

pred_test <- compute(NeuralNet, testNN)

pred_test$net.result

actual_pred = data.frame(pre = pred_test$net.result, actual = testNN$conc)

plot(actual_pred, main="prediction vs actual", col='blue' , xlab="Actual", ylab="Predicted",
pch=15, type="p")

abline(0,1, col="black")

##Calculating RMSE, R_squared and MAE (validation)

predict_testNN = (pred_test$net.result)

```



```

RMSError_Val = (sum((testNN$conc-predict_testNN)^2) / nrow(testNN)) ^ 0.5

Rsquared_model = 1 - sum((testNN$conc-pred_testNN)^2)/sum(testNN$conc-
(sum(testNN$conc)/nrow(testNN))^2)

##saving model

save(model, file = "model")

### loading model

load(file = "model")

### read input and prediction for new input

set.seed(1)

indata <- read.csv("micedata.CSV", sep=',', header = T)

n <- nrow(indata)

normalize <- function(x) {

  x <- as.numeric(x)

  return((x - min(x)) / (max(x) - min(x)))

}

normalisedinData <- as.data.frame(apply(indata, 2, normalize))

val_data_model_group2_est <- compute(model, normalisedinData[,c(2:26)])

val_data_model_group2_est$net.result

denormalizedval_model_data_group2_est <- val_data_model_group2_est$net.result *
(max(input_data$conc) - min(input_data$conc)) + min(input_data$conc)

denormalizedval_model_data_group2_est

tmp <- cbind(indata$conc, denormalizedval_model_data_group2_est)

```

```
colnames(tmp) <- c("group", "conc")  
  
##tmp <- as.data.frame(tmp)  
  
write.csv(tmp,"Group2_est_predconc.csv")
```

Appendix 2: Principal Component Analysis (PCA) script in R

```
### Importing libraries
```

```
library(ChemoSpecUtils)
```

```
library(ChemoSpec)
```

```
library(knitr)
```

```
library(R.utils)
```

```
library(utils)
```

```
### Reading a matrix data file stored in the working directory
```

```
spc <- matrix2SpectraObject(gr.crit = c("A", "B", "C", "D"),  
                           gr.cols = c("blue", "green", "purple", "red"),  
                           freq.unit = "Raman shift (/cm)",  
                           int.unit = "intensity",  
                           descrip = "Study",  
                           in.file = "mydata.csv",  
                           out.file = "horm", chk = TRUE,  
                           sep = ",", dec = ".")
```

```
### Summarizing the data
```

```
sumSpectra(spc)
```

```
### Baseline correction
```

```
spc1 <- baselineSpectra(spc, int = FALSE, method = "modpolyfit", retC = TRUE)
```

```
### Removing bad samples
```

```

spc2 <- removeSample(spc1, rem.sam = c("p", "q", "r ", "x ", "y ", "z"))

### normalizing data

spc3 <- normSpectra(spc2)

### Binning data

spec <- binSpectra(spc3, bin.ratio = 4)

### robustscaling data

robust <- r_pcaSpectra(spec, choice = "noscale")

### determining the number of pcs to be used in robustpca

plotScree(robust)

### robust2D Scoreplot

plotScores(spec, robust,
            main = "PCA",
            pcs = c(1,2),
            ellipse = "rob",
            tol = 0.5)

abline(h=0, v=0)

plotScores(spec, robust,
            main = "PCA",
            pcs = c(1,3),
            ellipse = "rob",
            tol = 0.5)

```

```
abline(h=0, v=0)
```

```
plotScores(spec, robust,
```

```
  main = "PCA",
```

```
  pcs = c(2,3),
```

```
  ellipse = "rob",
```

```
  tol = 0.6)
```

```
abline(h=0, v=0)
```

```
### 3D plot for PCA
```

```
plotScores3D(spec, robust, ellipse = TRUE, tol = 0.01, axes = "float")
```

```
### robustpotential outliers
```

```
diagnostics<-pcaDiag(spec, robust, pcs = 3, plot = c("SD","OD"))
```

```
### robustpca loadings
```

```
plotLoadings(spec, robust, loads = c(1:3), ref = 1, lwd = 3)
```

Appendix 3: Raman Spectrometer set-up at the Physics department (University of Nairobi)



Appendix 4: (a) Swiss Albino Mouse and (b) Swiss Albino Mice in a cage

