THE USE OF HAIR TO DETECT AND MONITOR CHRONIC HYPERGLYCAEMIA – A PILOT STUDY

Sian-Ailin da Silva

DSLSIA001

A dissertation presented for the degree MSc(Med) specialising in Trichology and Cosmetic Science, Department of Medicine, University of Cape Town

Supervisor: Dr. Jennifer Van Wyk

Co-Supervisor: Prof. Nonhlanhla Khumalo

DECLARATION

I, *Sian-Ailin da Silva,* hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever after publication of the data in a peer reviewed journal.

Signature: ….. Signed by candidate

Date: ……16/02/2018……

TABLE OF CONTENTS

DECLARATION 1

ACKNOWLEDGMENTS 4

ABSTRACT 5

ABBREVIATIONS 7

LIST OF FIGURES 10

LIST OF TABLES 12

1. LITERATURE REVIEW

1.1. Introduction 13

1.2. Insulin and Diabetes 15

1.3. Role of Hyperglycaemia in the Development of Diabetic Complications 18

1.4. Current Methods for Detecting and Monitoring Hyperglycaemia 20

1.5. Potential Use of Hair to Detect and Monitor Hyperglycaemia 24

1.5.1. Hair Anatomy and Growth 24

1.5.2. Hair as a Diagnostic Tool in Diabetes 25

1.6. Attenuated Total Reflection (ATR) Fourier Transform Infrared (FTIR) Spectroscopy and Multivariate Data Analysis (MVA) 28

1.7. Aims and Objectives 31

2. RESEARCH METHODOLOGY

2.1. Study Participants 31

2.2. Ethical Considerations 31

2.3. Research Site 32

2.4. Blood Sample Collection and HbA1c Measurement 32

2.5. Hair Sample Collection 32

2.6. Hair Sample Preparation and Storage 32

2.7. Attenuated Total Reflection Fourier Transform Infrared Spectroscopy (ATR-FTIR) Analysis of Hair samples 33

2.8. Data Analysis and Statistics for ATR-FTIR Data 34
2.8.1. Principal Component Analysis (PCA-X) 34
2.8.2. Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) 34
2.8.3. Partial Least Squares regression (PLS) 36
2.9. Amino Acid Analysis 36
2.10. Data Analysis and Statistics for Amino Acid Data 37

3. INVESTIGATING THE USE OF ATTENUATED TOTAL REFLECTION FOURIER TRANSFORM INFRARED SPECTROSCOPY (ATR-FTIR) AND MULTIVARIATE DATA ANALYSIS (MVA) TO DETECT HYPERGLYCAEMIA IN HAIR.

3.1. Introduction 38
3.2. Results 40
3.2.1. Analysing Inter- and Intra-Individual Variation in FTIR Hair Spectra 40
3.2.2. Factors Affecting Hair Spectral Classification 43
3.2.2.1. Effect of Age: Adults vs children 43
3.2.2.2. Effect of Chemical Hair Treatment 43
3.2.3. Discrimination Between ATR-FTIR Spectra Obtained from Hair of Controls vs. Diabetics. 47
3.2.4. OPLS-DA Model Validation 52
3.2.5. Investigating the Correlation Between HbA1c Levels and ATR-FTIR Spectra . 53
3.2.6. PLS Model Validation 58
3.2.7. Confirming Classification of Hair into Chemical Treatment Groups is Necessary to Optimally Distinguish Between Hair from Diabetic and Control Groups and to Predict HbA1c Levels from Hair ATR-FTIR Spectra 59

4. COMPARATIVE AMINO ACID ANALYSIS OF HAIR FROM DIABETICS AND HEALTHY CONTROLS.

4.1. Introduction 61
4.2. Results 61

4.2.1. Effect of Age and Hair Dye on Amino Acid Concentrations of Whole Human Hair Fibers 61
4.2.2. Difference in Amino Acid Concentrations in Hair of Diabetic Participants Compared to Healthy Controls 65

5. GENERAL DISCUSSION

5.1. Inter and Intra-Individual Variation Between Hair FTIR Spectra 68
5.2. Factors Affecting Separation of Hair FTIR Spectra 68
5.3. Using Hair FTIR Spectra to Detect Hyperglycaemia 70
5.4. Correlation Between Hair FTIR Spectra and HbA1c Levels 71
5.5. Prospects for Using ATR-FTIR Technology for Diagnosing and Monitoring Diabetes in the Future 72
5.6. Comparative Amino Acid Analysis of Hair from Diabetics and Healthy Controls 74
5.7. Limitations of the Study 75
5.8. Conclusion 77

REFERENCES 78
APPENDIX 1 84
APPENDIX 2 85
APPENDIX 3 86
APPENDIX 4 87

ACKNOWLEDGEMENTS

I would like to thank the following people and institutions for their contributions and support throughout my MSc degree.

Thank you to the National Research Foundation and SARChi for my bursary and funding my Masters.

My supervisor, Dr. Jennifer Van Wyk, for her guidance and mentorship, and for trusting and allowing me the independence to work on my project but always being there to help when it came to bumps in the road. Nothing is too difficult with a little hard work.

My co-supervisor, Prof. Nonhlanhla Khumalo, for her mentorship and support, and for teaching me that if you fight for what you want you will achieve.

Prof. Ian Ross for guidance with regards to many clinical aspects of the project and for allowing me to recruit study participants via his clinic at Groote Schuur Hospital.

Prof. Ariane Spitaels, Prof. Steve Delport and the Sisters at Red Cross Children's hospital's diabetes clinic for allowing us to invade your space to recruit study participants.

Prof. Helene Nieuwoudt for her invaluable mentorship with SIMCA and FTIR data analysis.

To Julie Estherhuizen, the research nurse who assisted in sample collection. Without Julie, I never would have obtained as many samples in such a short period of time.

My colleagues at the Hair and Skin Research Laboratory for their support and friendship throughout my time in the laboratory. I would especially like to thank my fellow students Ernest Mabotha, Relebohile Matobole, Tem Chalwa and Kutlwano Chabaesele for being by my side through thick and thin. I will cherish these friendships forever.

A special little thank you to Brandon Davidson for the tea time chats and words of encouragement.

And finally, to my parents for the constant support and for always believing in me, and my sister and friends for all the encouragement.

ABSTRACT

Background: Diabetes mellitus is a major public health problem resulting in about 5 million deaths per year. This metabolic disorder is characterized by hyperglycaemia, which results in debilitating and life-threatening complications. It is, therefore, vital for diabetics to monitor and control their blood glucose levels in order to keep them below 7mmol/L while fasting and below 9mmol/L after meals. Chronic estimates of glucose control of 8-12 weeks are obtained using glycated haemoglobin A1 (HbA1c). Non-invasive, less expensive methods of monitoring long term glycaemic control may be useful. Since scalp hair consists of about 80% protein, which is subject to non-enzymatic glycation, and growing hair has a rich blood supply exposing it to free glucose, it is likely that hair can be used as an alternative substrate for monitoring chronic hyperglycaemia.

Subjects and Methods: Scalp hair and a blood samples (for HbA1c) were collected from 46 diabetic and 46 healthy control subjects. There were 26 diabetic adults (30-70 years), recruited from the outpatient clinic at Groote Schuur hospital and 20 children (7-18 years) recruited from the diabetic clinic at the Red Cross children's hospital. There were 29 healthy control adults (26-65 years) and 17 children (7- 17 years) recruited from the Groote Schuur and Red Cross hospitals respectively. History of chemical hair treatment was recorded for each participant. Hair samples were washed using 1% sodium dodecyl sulphate and analysed using Fourier transform infrared- attenuated total reflection (ATR-FTIR) spectroscopy. Spectra were analysed using statistical software (SIMCA, Umetrics) to determine whether the hair of diabetics was distinguishable from hair of healthy controls as well as whether spectra correlated with HbA1c levels of participants. Hair amino acid concentrations were also analysed as it is known that circulating amino acid concentrations are altered in people with diabetes.

Results and discussion: The Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) models between spectra obtained from hair of diabetic participants and spectra

obtained from control hair show good separation and predictive ability. When ATR-FTIR spectra were analysed in four groups: children with natural hair, adults with natural hair, adults with dyed hair and adults with relaxed hair, the models yielded predictive abilities (Q2(cum)) ranging from 0.753 to 0.85 and classification abilities ranging from 97.97 to 100%. This demonstrated excellent separation and predictive ability for controls vs. diabetics. Partial least squares (PLS) revealed a good correlation between hair FTIR spectra and participant HbA1c levels (R^2 ranging between 0.8067 and 0.9296). These results demonstrated the possibility to use ATR-FTIR alongside multivariate data analysis to detect hyperglycaemia and monitor blood glucose levels via prediction of HbA1c levels from the hair spectra. Amino acid analysis supported the OPLS-DA classifications, as the largest differences were seen between age and chemically treated hair groups. Amino acid results reinforced the necessity to classify spectra into groups in order to distinguish between hair spectra from diabetics or controls, as well as to predict HbA1c. Twelve amino acids (Asp, Glu, Pro, Gly, Met, Ile, His, Lys, Arg, Amm, Cys, Leu) were significantly different between hair from adults and children, eight amino acids (Ala, Cys, Val, Met, Lys, Amm, Ser, Tyr) were statistically significantly different between natural and dyed hair and only up to four amino acids (Gly, Val, Met, Ile in children or Val, Phe, Amm, Arg in adults) were significantly different between diabetic and control groups.

Conclusion: There is a need for non-invasive means of monitoring chronic hyperglycaemia. This study demonstrated the ability to distinguish between the hair of diabetics and controls as well as the ability to predict HbA1c levels from hair using ATR-FTIR. However, factors such as age and chemical treatment, which affect the chemical properties of hair, like amino acid levels, should be considered first. This would lead to promising prospects for long term blood glucose monitoring, due to the ability to estimate hair growth rate, and greater insights into the timing and development of diabetic complications. ATR-FTIR relatively simple to use, requires minimal sample preparation and does not require the use of expensive consumables. This technology could, potentially, be adapted into a primary health point of care or home screening or monitoring device for long-term hyperglycaemia, which would assist in early detection and preventing the progression of debilitating complications.

ABBREVIATIONS

AGE: Advanced glycation end-product

Amm: Ammonia

Arg: Arginine

Asp: Aspartic acid

ATP: Adenosine triphosphate

ATR: Attenuated total reflection

Cys: Cysteine

EDTA: Ethylenediaminetetraacetic acid

ER: endoplasmic reticulum

FTIR: Fourier Transform Infrared spectroscopy

GC-MS: Gas chromatography – mass spectrometry

GDM: Gestational diabetes mellitus

GIP: Gastric inhibitory polypeptide

GLP-1: Glucagon-like peptide-1

Glu: Glutamic acid

GLUT: Glucose transport proteins

Gly: Glycine

GSIS: Glucose stimulated insulin secretion

HbA1c: Glycated haemoglobin A1c

HbC: Haemoglobin with a genetic variant

HbF: Foetal haemoglobin

HbS: Sickle cell haemoglobin

HCl: Hydrochloric acid

His: Histidine

HPLC: High performance liquid chromatography

IDF: International Diabetes Federation

Ile: Isoleucine

IRS: Inner root sheath

IRS: Insulin receptor substrates

Leu: Leucine

Lys: Lysine

Met: Methionine

MVA: Multivariate data analysis

NCD: Non-communicable disease

NIR: near infrared radiation

OGTT: Oral glucose tolerance test

OPLS-DA: Orthogonal Projections to Latent Structures Discriminant Analysis

ORS: Outer root sheath

PCA: Principal Component analysis

Phe: Phenylalanine

PI3K: Phosphoinositide 3 Kinase

PLS: Partial least squares regression

PRESS: Predicted Residual Sum of Squares

Pro: Proline

ROS: Reactive oxygen species

SDS: Sodium dodecyl sulphate

Ser: Serine

TBA: Thiobarbituric acid

Thr: Threonine

Tyr: Tyrosine

Val: Valine

WHO: Whole Health Organization

LIST OF FIGURES

Figure 1.1. Map of regional prevalence of diabetes mellitus taken from 6th and 8th editions of the IDF Diabetes Atlas (International Diabetes Federation) [6,7] 13

Figure 1.2. Simplified diagram depicting insulin signaling and GLUT4 translocation within a muscle cell 18

Figure 1.3. Flow diagram by The Committee of the Japan Diabetes Society showing the steps taken in clinical diagnosis of diabetes mellitus [60] 21

Figure 1.4. Formation of HbA1c (Marchetti, 2009) [61] 22

Figure 1.5. Diagram showing a hair fiber and the structures necessary for its growth out of the scalp (Buffoli*., et al* (2014)) [71]. 24

Figure 1.6. Diagram showing the stages of hair growth and the vascularization associated with each phase (Buffoli., et al (2014)) [71] 25

Figure 1.7. Schematic diagram showing the principle of ATR-FTIR 28

Figure 1.8. Diagram depicting the basic principle of multivariate data analysis when analysing spectral data (Wiklund, 2008) [118]. 30

Figure 2.1. Diagram showing how FTIR spectral readings were collected for each individual 33

Figure 3.1. Typical ATR-FTIR spectrum for an intact human hair fiber 39

Figure 3.2. Intra-individual variation in hair FTIR spectra 41

Figure 3.3. Inter-individual variation in hair FTIR spectra 42

Figure 3.4. Effect of age: Adults vs. children 45

Figure 3.5. Effect of chemical hair treatment 46

Figure 3.6. OPLS-DA model showing separation between spectra obtained from hair of diabetics vs. controls amongst children with natural hair 48

Figure 3.7. OPLS-DA model showing separation of spectra obtained from hair of diabetics vs. controls amongst adults with natural hair 49

Figure 3.8. OPLS-DA model showing separation between spectra obtained from hair of diabetics vs. controls amongst adults with dyed and relaxed hair 50

Figure 3.9. Permutation plots for the validation of OPLS-DA models 52

Figure 3.10. PLS model displaying correlation between hair spectra and HbA1c for children with natural hair 54

Figure 3.11. PLS model displaying correlation between hair spectra and HbA1c for adults with natural hair 55

Figure 2.12. PLS model displaying correlation between hair spectra and HbA1c for Adults with dyed hair 56

Figure 3.13. PLS model displaying correlation between hair spectra and HbA1c for Adults with relaxed hair 57

Figure 3.14. Permutation plots for the validation of PLS models 58

Figure 3.15. OPLS-DA and PLS models including adults who have natural, dyed and relaxed hair 60

Figure 4.1. Box and whisker plots showing amino acid concentrations in hair of adults vs. children and natural hair vs. dyed hair 63

Figure 4.2. Box and whisker plots showing amino acid concentrations in hair of diabetics vs. controls 66

Figure 5.1. An example of a recommended work flow when attempting to determine whether an individual is diabetic and predict HbA1c via ATR-FTIR analysis of scalp hair..... 73

LIST OF TABLES

Table 1.1. Factors which influence HbA1c measurement taken from the 2012 SEMDSA Guideline for the Management of Type 2 Diabetes [13] 23

Table 3.1. FTIR peak bond vibration and wavenumber assignment 38

Table 3.2. Summary of different treatment types among adults 44

Table 3.3. Classification table for spectra obtained from hair of diabetics vs. controls amongst children with natural hair 50

Table 3.4. Classification table for spectra obtained from hair of diabetics vs. controls amongst adults with natural hair 51

Table 3.5. Classification table for spectra obtained from hair of diabetics vs. controls amongst adults with dyed hair 51

Table 3.6. Classification table for spectra obtained from hair of diabetics vs. controls amongst adults with relaxed hair 51

Table 4.1. Amino acid concentrations in hair of control children and adults (with natural and dyed hair) and concentrations in hair as reported in literature 64

Table 4.2. Amino acid concentrations in hair of diabetic participants compared to healthy controls. 67

CHAPTER 1: LITERATURE REVIEW

1.1. INTRODUCTION

Diabetes mellitus is a major public health problem resulting in about 5 million deaths per year [1]. However, because it is a non-communicable disease (NCD), its impact was, until recently, sorely underestimated. The global prevalence of diabetes is increasing more rapidly than expected. In 1998 the World Health Organization (WHO) and the International Diabetes Federation (IDF) estimated that there were 135 million people with diabetes mellitus, which would increase to 300 million by 2025 [2]. In the year 2000 it was predicted that 366 million people would have diabetes by 2030 [3]. This number was already exceeded in 2013 and is estimated to reach 629 million by the year 2045 (Figure 1.1.) [4-7].

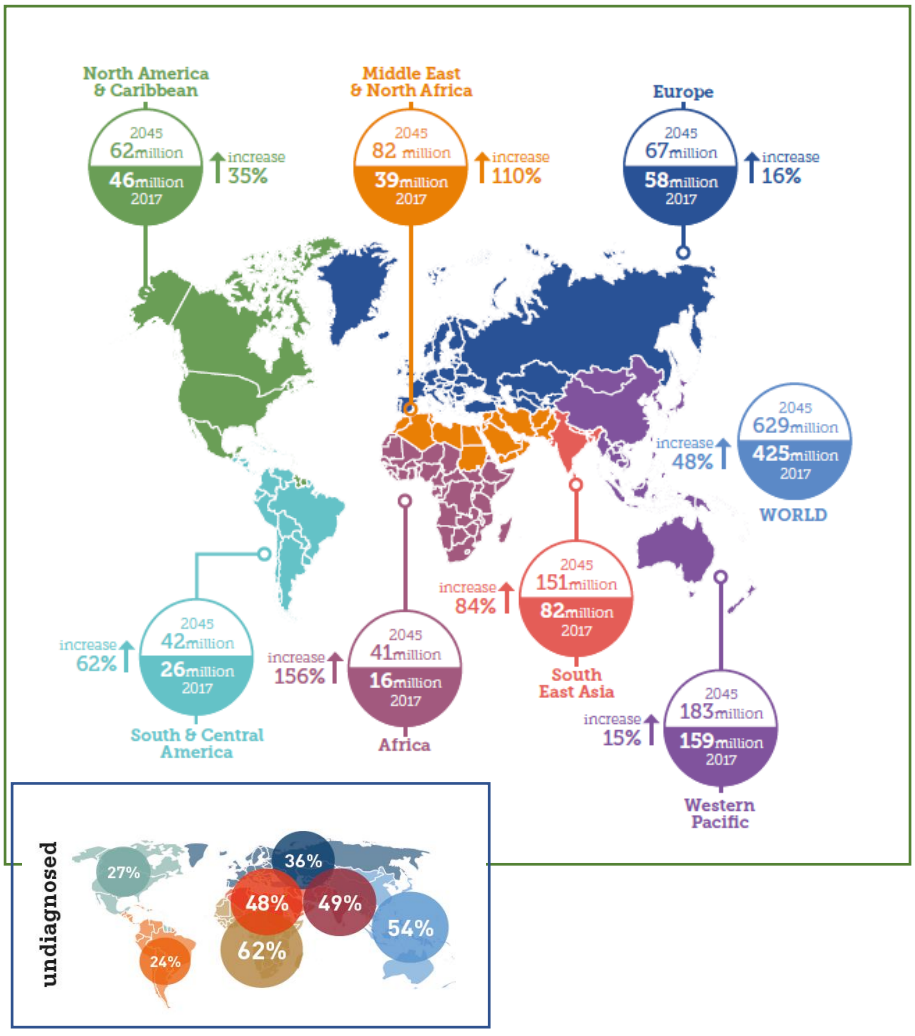


Figure 1.1. *The regional prevalence of diabetes mellitus in 2017 and expected prevalence and increase by the year 2045. The insert shows the estimated proportions of undiagnosed cases worldwide. taken from 6th and 8th editions of the IDF Diabetes Atlas (International Diabetes Federation) [6,7].*

Diabetes is characterised by hyperglycaemia, which is either due to lack of insulin production (type 1 diabetes), or due to resistance to insulin signalling (type 2 diabetes). Chronic hyperglycaemia is associated with several debilitating and life-threatening complications affecting the heart, kidneys, eyes or nerves [7-9]. In order to slow progression of these complications, it is important that diabetics constantly monitor and keep their fasting blood glucose levels below 7mmol/L and below 9mmol/L after meals [10-12]. However, a large proportion of the diabetic population is undiagnosed, especially in low and middle-income countries (up to 62% in Africa) (Figure 1.1.), which could lead to unsuspected development of complications and early death. [6, 7]

There are immediate, methods to measure short term blood glucose levels in bodily fluids (blood, urine and saliva). However, these glucose levels fluctuate vastly throughout the day, so a more stable measure is needed to monitor chronic hyperglycaemia. The percentage of glycated haemoglobin A1c (HbA1c) reflects an average blood glucose level for the past 8-12 weeks and is currently used to monitor chronic hyperglycaemia. An HbA1c value <6% is considered non-diabetic and 6.5-7.5% is the ideal HbA1c for diabetics to avoid complications [10-12]. A 1% increase in HbA1c levels, increases the risk of myocardial infarction by 14%, microvascular complications by 37%, amputation due to diabetic foot by 43% and death by 21% [13].

There are limitations to HbA1c measurement. Obtaining an HbA1c value is invasive, as blood needs to be drawn. It is also an expensive test to perform and HbA1c values are affected by several factors such as vitamin B12 or iron deficiency, abnormal haemoglobin traits, alcoholism [13] and ethnicity [14-17]. Alternative means of monitoring chronic hyperglycaemia would, therefore, be useful. Hair has long been used as a testing substrate for recreational drug abuse and is becoming an increasingly important tool in diagnostics [18-24]. This is primarily due to a consistent growth rate and the stable incorporation of molecules into the hair fibre during growth.

There are multiple reasons to believe that it may also be possible to monitor chronic hyperglycaemia using hair. Firstly, the hair follicle has a rich blood supply, which exposes the hair to free glucose within the blood. Secondly, hair also comes in contact with sweat, which is known to have higher glucose concentrations in people with hyperglycaemia [25]. Hair fibres are keratinized structures, rich in amino acids, which are prone to glycation. It is possible that glucose molecules from blood and sweat covalently bind to the amino acids in the growing hair strand. Since human hair grows at a consistent rate of about 1 cm per month, it may give a chronological indication of glycaemic status for years previously, depending on the length of a person's hair.

Scientists have, for a long time, realised the effect of diabetes on hair growth and structure [26, 27] and have attempted multiple wet laboratory experiments and extraction methods in an attempt to detect hyperglycaemia using hair [28-31]. However, recent advancements in the potential use of Fourier Transform Infrared (FTIR) spectroscopy, which allows for non-invasive study of chemical and structural properties of substances, and multivariate data analysis have opened up new possibilities to research.

1.2. INSULIN AND DIABETES

After food is ingested, the digestive system works to break down large molecules of food into smaller molecules. This process is important as it allows nutrients to be taken up into the blood stream by the small intestine to be transported around the body. When nutrients reach the pancreas, β-cells within the pancreatic islets of Langerhans produce and secrete insulin – a hormone which is required to control blood glucose levels within a strict range (4-7mmol/L)[32] by suppressing glucose production in the liver and stimulating glucose uptake by the cells of the body to use as energy [33]. The pancreas secretes insulin in response to glucose, but the sensing of other nutrients amplifies glucose stimulated insulin secretion (GSIS). For example, gastric inhibitory polypeptide (GIP) and glucagon-like peptide-1 (GLP-1) are hormones produced by the gastrointestinal tract in response to the ingestion of amino acids. These hormones bind to receptors on β-cells directly increasing GSIS [34-36]. Free fatty acids also directly potentiate GSIS via β-cell free fatty acid receptors [37].

Within β-cells, insulin is transcribed and synthesized as preproinsulin and then processed to proinsulin. Proinsulin is then transported into secretory vesicles where it is cleaved into C-peptide and insulin [38], ready to be released into the blood stream via exocytosis in response to nutrient intake. Insulin secretion happens in two phases: a transient first phase, where a small quantity of insulin granules is readily available for quick release after detection of glucose, and a more sustained second phase which replaces the readily available insulin [39, 40].

Once insulin is in the blood stream, it is transported around the body to facilitate uptake of glucose, and other monosaccharides, from the blood stream into tissues. Since monosaccharides are restricted from crossing the cell membrane, cells require glucose transport proteins (GLUTs) in order to take up carbohydrates for energy production [41]. There are fourteen GLUT protein isoforms that transport carbohydrates into cells via facilitated diffusion. GLUT isoforms are expressed in a tissue dependent manner and transport various monosaccharides at varying rates, depending on the glucose requirements of each tissue [41-43]. GLUT4 is the most widely understood GLUT protein and is the predominant isoform in skeletal muscle and adipose tissue. These tissues contribute to the greatest percentage of glucose uptake from the blood and for simplicity; GLUT4 will be the isoform which is referred to.

GLUT4 is contained in vesicles within skeletal muscle and adipose cells. When insulin binds to its receptors on the surface of these cells, it causes recruitment of signalling molecules such as Insulin receptor substrates (IRS) to cause an intracellular signalling cascade, which eventually causes translocation of GLUT4 containing vesicles to the surface of the cell. Fusion of these vesicles causes GLUT4 to become bound to the cell membrane [32, 41, 42]. Glucose then binds to GLUT4, which causes a conformational change in the transporter leading to movement of glucose across the cell membrane into the cell [32, 44]. In skeletal muscle, contraction of the muscle also leads to translocation of GLUT containing vesicles to the

surface of the cell [45]. So, although adipose tissue is more sensitive to insulin signalling, skeletal muscle is responsible for the majority of glucose uptake from the blood.

In diabetes, the above process is compromised due to disruptions in one or more of the steps. Type I diabetes accounts for between 5 and 10% of the diabetic population and occurs either due to autoimmune mediated destruction of pancreatic β-cells (which can be detected through identification of antibodies such as islet cell autoantibodies and autoantibodies to insulin within the patient) or through spontaneous idiopathic destruction of the β-cells, where the patient exhibits insulin deficiency but no signs of autoimmunity. Due to deficiency of insulin, these patients usually depend on insulin for survival. In type 2 diabetes, the pancreas produces insulin, but tissues do not respond to insulin signalling. Type 2 diabetes makes up 90-95% of diabetic cases and usually occurs later in life. It is mainly caused by bad diet, inactivity and obesity [7, 8]. Insulin resistance of this nature occurs when lipids are ectopically deposited on organs such as the pancreas, liver and muscle, either preventing binding of insulin to insulin receptors or by attenuating insulin receptor signalling in downstream pathways [46]. People with type 2 diabetes may also eventually rely on exogenous insulin for survival, as pancreatic β-cells become exhausted in the effort to compensate for high blood glucose concentrations [9].

There are other causes of diabetes, which makes up a minute proportion of the diabetic population, such as genetic abnormalities in the β-cell or insulin receptor/substrates, diseases of the pancreas, endocrinopathies i.e. increased production of insulin inhibiting hormones, certain drugs, infections such as mumps and lastly, gestational diabetes mellitus (GDM) which occurs during pregnancy and tends to retract after birth [8].

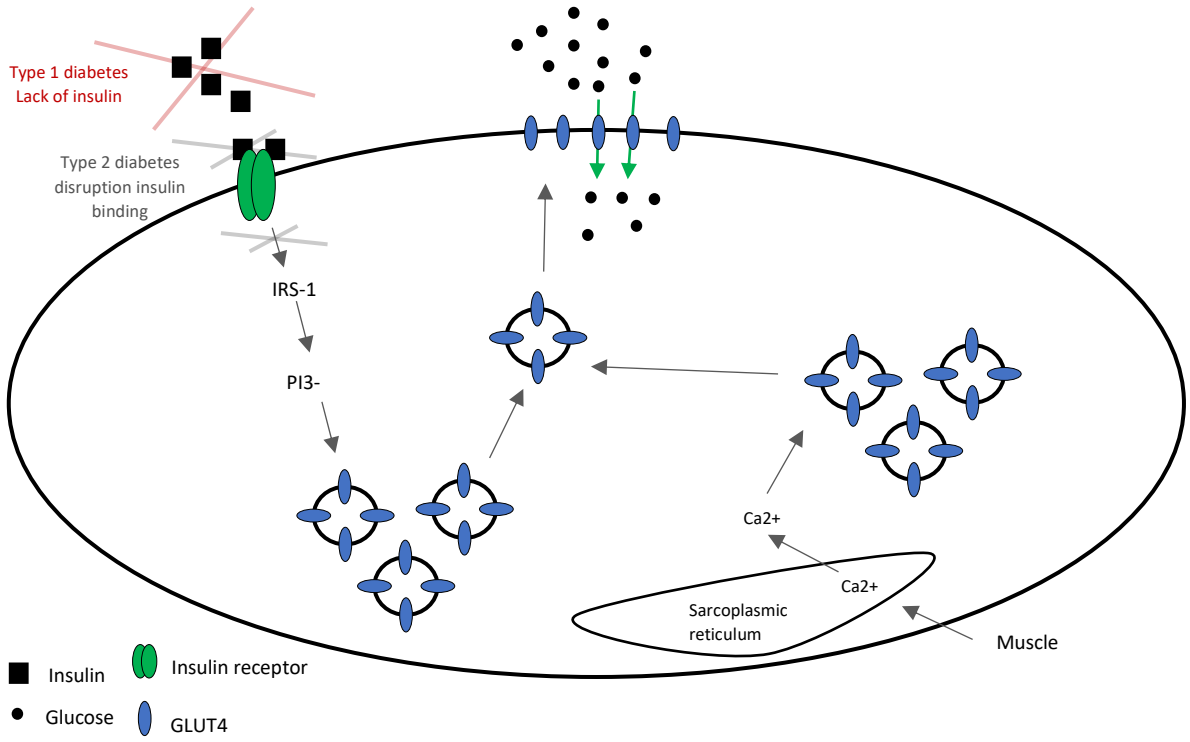


Figure 1.2. *A Simplified diagram depicting insulin signalling and GLUT4 translocation within a muscle cell. Insulin binds to the insulin receptor on the surface of the cell to cause activation of downstream signalling molecules (Insulin receptor substrate 1 (IRS1) and Phosphoinositide (PI) 3-kinase which cause translocation of GLUT4 containing vesicles to the surface of the cell. Muscle contraction causes the sarcoplasmic reticulum, within muscle cells, to release calcium, which also causes translocation of GLUT 4 containing vesicles. When GLUT4 is bound to the cell membrane glucose can enter the cells from the blood stream.*

1.3. ROLE OF HYPERGLYCAEMIA IN THE DEVELOPMENT OF DIABETIC COMPLICATIONS.

Hyperglycaemia is associated with life-threatening complications including diabetic comas, cardiovascular disease (stroke, angina, myocardial infarction etc.) - the most common cause of death amongst people with diabetes, nephropathy and kidney failure, neuropathy which leads to erectile dysfunction as well as problems with digestion and urination, retinopathy and blindness. Damage to nerves and small blood vessels in the extremities causes numbness, unnoticed injuries, slow wound healing and infection. All of these contribute to the development of diabetic foot which eventually results in amputation [7-9].

There are a few major reasons why hyperglycaemia is thought to cause development of diabetic complications. Firstly, cells lose control of energy production [9]. This is either due to

lack of internal glucose, in the case of cells which are dependent on insulin for glucose uptake, or due to excessive intracellular accumulation of glucose, in the case of cells which are insulin independent (cells which express high levels of GLUT1 such as vascular endothelial cells [47]). In the latter scenario, cells cannot control glucose intake, therefore energy production becomes uncontrolled due to excessive glucose accumulation and the ability for these cells to utilize glucose for energy production eventually becomes impaired. Normally, glucose undergoes glycolysis once it has been taken up by the cell. This process leads to the production of pyruvate, which can be used by the mitochondria to produce energy in the form of Adenosine triphosphate (ATP). Pyruvate is the preferred substrate for aerobic ATP production, as it is most effectively broken down by oxidative phosphorylation within the mitochondria. The cells can however switch to using fatty acids and amino acids for ATP production in glucose starved environments, but this leads to the production of ketone bodies by the liver. This is normal in a fasting state, however, excessive accumulation of ketone bodies in the blood leads to ketoacidosis which lowers the pH of the blood and can cause dangerous symptoms such as vomiting, dehydration, a change in mental status, a coma or even death [48, 49].

Secondly, abnormal mitochondrial function (caused by hyperglycaemic loss of control of energy production), leads to the production of reactive oxygen species (ROS), which is thought to be the initial cause of diabetic complications. ROS such as superoxide (O^{-}_{2}) are highly reactive and particularly damaging to cells and tissues when they are produced in excess because antioxidant enzymes, such as superoxide dismutase, cannot keep up with converting superoxide into oxygen and hydrogen peroxide. ROS cause direct damage to cells and tissues but are also thought to contribute to the initiation of intracellular advanced glycation end-product (AGE) formation [50]. Glycosylation is the enzymatic addition of glucose residues onto amino acids and lipids. This post translational modification occurs in a controlled manner and is required for correct protein folding, function, trafficking and degradation[9]. Advanced glycation, however, is the non-enzymatic addition of glycans onto free amino groups of amino acids. This process, which leads to the development of AGEs, is known as the Maillard reaction and it is dependent on pH, temperature, glucose concentrations and time.

The accumulation of AGEs occurs at most sites of diabetic complication and is also thought to be one of the major contributors to many different diabetic complications including diabetic cataract, retinopathy, neuropathy, nephropathy and cardiomyopathy [47, 51]. Glycation affects protein function by causing conformational changes, blocking receptors and enzymatic active sites and changing the half-life of proteins [52]. Normally, misfolded and damaged proteins are processed by the endoplasmic reticulum (ER) stress pathway within cells. In diabetes the ER stress pathways are overwhelmed due to extensive glycation, leading to activation of proapoptotic pathways and cell death [53]. AGEs also interact with cellular receptor, RAGE, to cause inflammation leading to immune mediated damage and production of further ROS [54]. Blockage of this receptor leads to significant attenuation of the inflammatory related tissue damage in diabetes [55]. Since AGEs develop over a period of weeks, they affect long lived proteins [56]. Therefore, the structural components of the connective tissue matrix or the basement membrane, such as collagen, are especially prone to advanced glycation, which causes intermolecular cross-linkages resulting in collagen becoming ridged [57]. This has numerous disadvantages, the prominent one being the stiffening of arterial walls, contributing to diabetic cardiovascular complications. Further, tissue stiffening and poor vascular perfusion contribute to poor wound healing. Glycation of collagen also hinders osteoblast differentiation [58, 59] and the combination of this and decreased flexibility of collagen leads to bone fragility. The major form of haemoglobin (HbA) also becomes glycated, and although regular erythrocyte turnover diminishes adverse health effects of this, it is useful in monitoring chronic hyperglycaemia and helping to prevent complications from arising in diabetic patients (discussed further in the next section).

1.4. CURRENT METHODS FOR DETECTING AND MONITORING HYPERGLYCAEMIA

Most of the existing methods for diabetes detection and monitoring are transient, as these methods involve measuring glucose levels within bodily fluids (blood, saliva and urine) and there are massive glucose level fluctuations in these fluids throughout the day. Currently, diabetes is diagnosed by measuring the fasting plasma glucose level alone or in combination with an oral glucose tolerance test (OGTT)[8]. Fasting plasma glucose levels greater than 7mmol/L after no caloric intake for at least 8 hours, casual plasma glucose levels greater than 11.1mmol/L or glucose levels greater than 11.1mmol/L 2 hours after 75g of oral glucose

administration (OGTT) are the cut of values for diagnosis of diabetes according to the American Diabetes Association [8]. However, these values are referred to as a 'diabetic type' according to the Committee of the Japan Diabetes Society and further steps must be taken to conclusively diagnose diabetes mellitus as shown in Figure 1.3. [60]. Similarly, an HbA1c level that is greater than 6.5% this is also considered to be a 'diabetic type'.

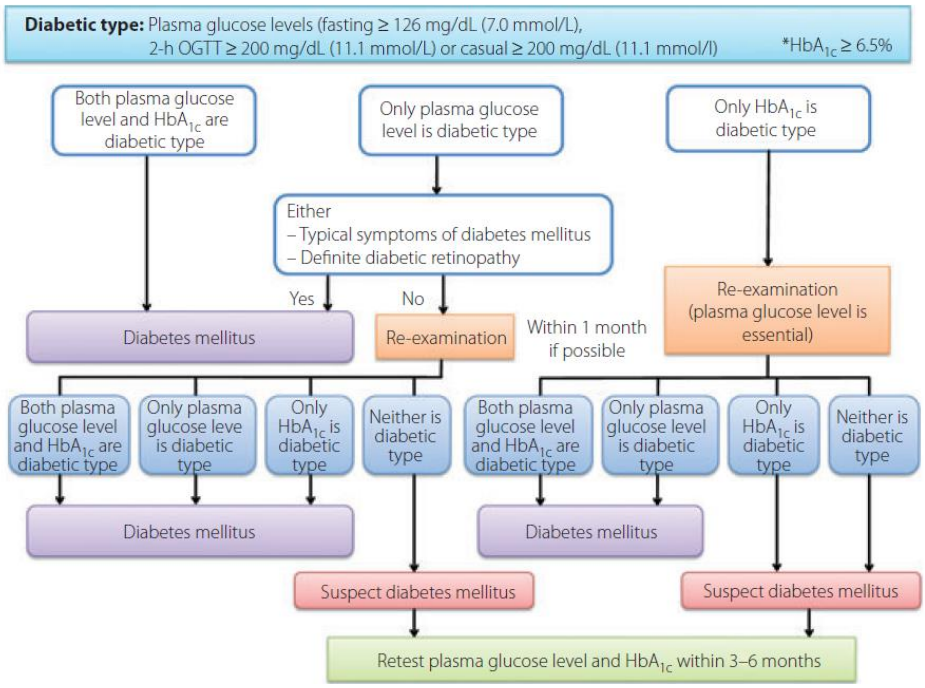


Figure 1.3. *Flow diagram set out by The Committee of the Japan Diabetes Society showing the steps taken in clinical diagnosis of diabetes mellitus [60]*

HbA1 is the major form of haemoglobin within healthy adult red blood cells, comprising about 90%. HbA1 is a tetrameric protein composed of 2 alpha and 2 beta chains ($\alpha_2\beta_2$) attached to a haeme-group. HbA1c refers to the glycated form of the protein, where reactive ends of glucose covalently binds to the N-terminal valine of the haemoglobin β-chain as well as other sites on haemoglobin (see Figure 1.4. [61]). The other two common haemoglobins are HbA2 ($\alpha_2\delta_2$) and HbF ($\alpha_2\gamma_2$), which comprise 2.5% and 0.2% of haemoglobin respectively. Normally about 5% of an adult's haemoglobin is glycated, if this level reaches above 6.5% it is an indicator of chronic hyperglycaemia [60, 62, 63].

In addition to a strict diet, treatment programmes recommend that diabetic patients regularly monitor their glucose levels are required to self-test via finger prick. This is painful and may lead to bruising, infections and nerve damage in the fingertips. There are non-invasive methods of monitoring blood glucose levels through other fluids (urine and saliva) and using near infrared (NIR) radiation through the skin but none of these methods are as accurate as invasive methods. The latter requires regular calibration and results are easily affected blood flow rate, erythrocyte size, site of measurement and density of the capillary network [64].

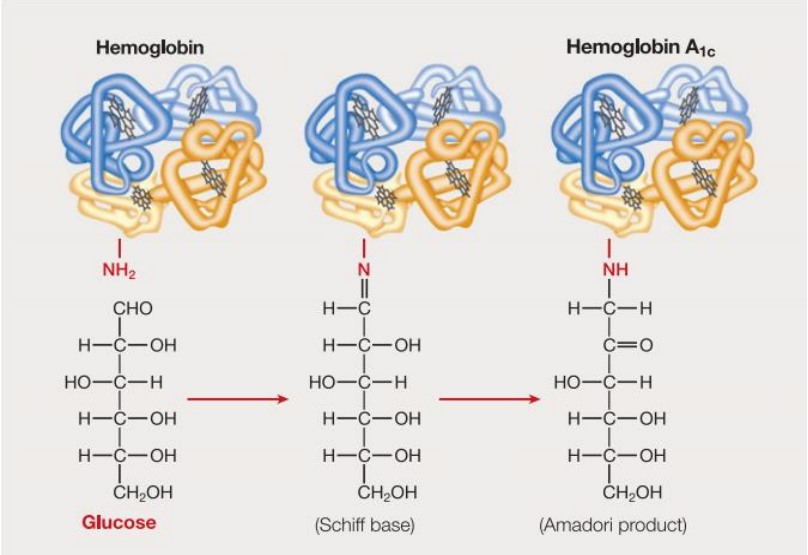


Figure 1.4. *Formation of HbA1c. Each of the α-chains (blue) and β-chains (yellow) is attached to a heme group (Marchetti, 2009) [61]]*

Testing HbA1c level is the 'gold standard' for monitoring chronic hyperglycaemia. It reflects average blood glucose levels for the past 8-12 weeks, as erythrocyte turnover is about 120 days [65, 66]. This is advantageous over testing plasma glucose as is enables better risk assessment for long-term diabetic complications, there is less biological variability as it is not affected by acute blood glucose fluctuations and there is no need for fasting or timed sampling [67].

Testing HbA1c does, however, have limitations. It should not be used on its own to definitively diagnose diabetes, as there is inter-individual variation, particularly among people with diabetes, where the distribution of HbA1c levels is very broad compared to those with normal

blood glucose levels [60]. Further confirmatory plasma glucose tests are, therefore, essential (refer to Figure 1.3.). Further, HbA1c may also be unreliable in the diagnosis of rapidly developing type 1 diabetes as HbA1 glycation is not given sufficient time to 'catch up' to the acute elevation in blood glucose [67]. There are also certain factors which affect HbA1c readings (See Table 1.1.) including abnormal haemoglobin traits, HbS (sickle cells) and HbC [68] or differences in erythrocyte turnover caused by haemolytic anaemia, chronic renal failure, malaria, frequent alcohol consumption, pregnancy, blood loss, or blood transfusions [67, 69]. Further, HbA1c readings also increase with age [70] and vary between races (where African, African-American, Asian and Latino populations have higher HbA1c levels than Caucasian population) [15-17]. A Study by Zemlin *et al.,* found that a cut off level of 6.1% rather than 6.5% was optimal among the mixed race community in Bellville, South Africa, and claimed that different HbA1c cut off values should be established for different population groups [14].

Table 1.1. Factors which influence HbA1c measurement, taken from the 2012 SEMDSA Guideline for the Management of Type 2 Diabetes [13]

Erythropoiesis	Increased HbA1c: Iron deficiency, vitamin B12 deficiency, decreased erythropoiesis
	Decreased HbA1c: Administration of erythropoietin, iron or vitamin B12, reticulocytosis, chronic liver disease
Altered haemoglobin	Genetic or chemical alterations in haemoglobin may increase or decrease HbA1c: Haemoglobinopathies, HbF, methaemoglobin
Glycation	Increased HbA1c: Alcoholism, chronic renal failure, decreased intra-erythrocyte pH
	Decreased HbA1c: Aspirin, vitamins C and E, certain haemoglobinopathies, increased intra-erythrocyte pH
	Variable HbA1c: Genetic determinants
Erythrocyte destruction	Increased HbA1c with increased erythrocyte life span: Splenectomy
	Decreased HbA1c with decreased erythrocyte life span: Haemoglobinopathies, splenomegaly, rheumatoid arthritis, drugs (e.g. antiretrovirals, ribavirin, dapsone)
Assays	Increased HbA1c: Hyperbilirubinaemia, carbamylated haemoglobin, alcoholism, large doses of aspirin, chronic opiate use
	Decreased HbA1c: Hypertriglyceridaemia
	Variable HbA1c: Haemoglobinopathies

1.5. POTENTIAL USE OF HAIR TO DETECT AND MONITOR HYPERGLYCAEMIA

1.5.1. Hair Anatomy and Growth

Human hair is composed of two structures. The dead, fully keratinized hair shaft and the living hair follicle within the skin. The hair shaft consists of the outer cuticle cells, the cortex, which contributes to the bulk of mechanical strength and other physical properties of the hair fibre and the inner inconsistent inner medulla, more commonly present in terminal hair. Human hair grows within a subcutaneous hair follicle originating from the hair bulb (Figure 1.5.)[71]. The follicle is the imperative structure for hair growth, as it contains multipotent keratinocyte and melanocyte stem cells necessary for new growth within the outer root sheath (ORS), as well as mature cells in the inner root sheath (IRS) which produce the keratin and trichohyalin necessary to provide strength for the outward growing hair shaft [71, 72].

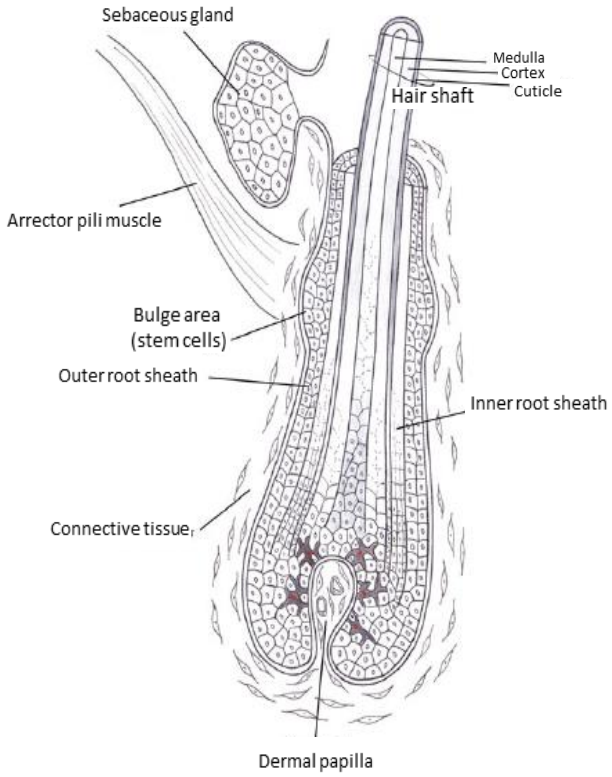


Figure 1.5. *Diagram showing a hair fibre and the structures necessary for its growth out of the scalp (Buffoli., et al., 2018) [71]*

The centre of the hair bulb contains a dermal papilla, which is the main driver in hair growth. The bulb is also the point at which arterioles form a plexus to provide rich vascularization to the hair follicle, which is important in providing nourishment to growing hair fibres [73]. Vascularization of the hair follicle fluctuates according to the different phases of the hair cycle and is seen to be the highest during anagen or growth phase (Figure 1.6.)[71]. The subsequent phase of hair growth is the catagen or transitional phase which lasts several weeks. During catagen the production of the hair shaft is completed and therefore differentiation and proliferation of keratinocytes decreases. The dermal papilla becomes a cluster of quiescent cells and the hair follicle travels upwards towards the dermis/subcutis border forming the club hair. The final phase of the hair growth cycle is the telogen or resting phase, which lasts from a few weeks to 8 months. Between 10% and 15% of all hairs are in telogen at a given time. After this phase, the hair falls out and anagen begins again [71, 72].

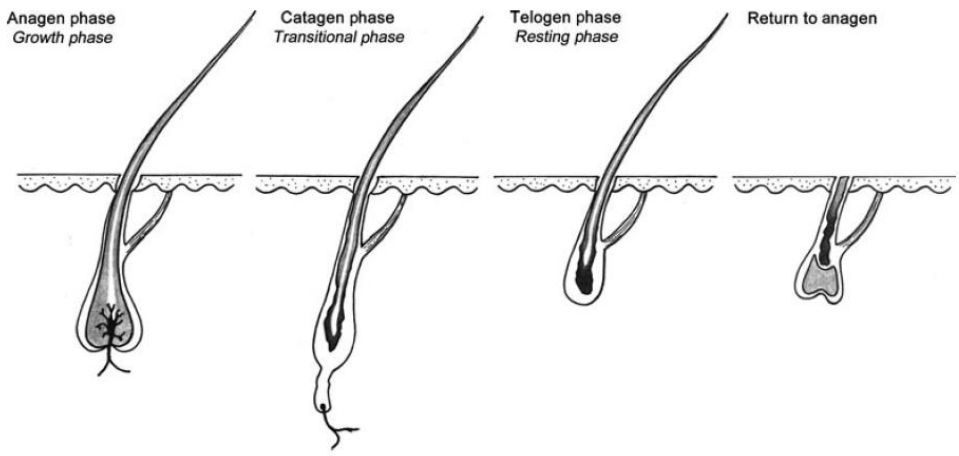


Figure 1.6. *Diagram showing the stages of hair growth and the vascularization associated with each phase (Buffoli., et al., 2014)[71]*

1.5.2. Hair as a Diagnostic Tool in Diabetes.

Due to its rich blood supply from the dermal papilla and the nature of hair growth into a stable keratinized structure, the hair shaft stably incorporates molecules as it grows. The use of hair as a diagnostic tool has, therefore, gained increasing importance over the past few decades. Initially, analysis of hair by mass spectrometry was used to detect drug abuse and poisoning [18]. Today, hair analysis can be used to detect exposure to chemical pollutants [19] and other occupational hazards such as heavy metals [20] as well as exposure to therapeutic drugs, such

as antiretrovirals, in order to predict outcomes of therapy [21]. Hair is also increasingly being used to detect levels of endogenously produced compounds, such as cortisol as a biomarker of chronic stress [22, 23].

There is evidence to suggest that it is possible to detect and monitor a person's glycaemic status using their hair. Diabetic patients may experience unexplained hair loss, breakage and thinning [26]. The discovery that the percentage of glycated haemoglobin A1c (HbA1c) was a reliable indication of diabetic control [74] led to studies done in the 1980s investigating whether other, more stable tissues also became excessively glycated under hyperglycaemic conditions. Since haemoglobin has a reasonably high turnover rate (8-12 weeks), it was of interest to find a longer-term measure of glycation.

A number of methods have been developed to measure glycation in biological samples [28-30, 75-79]. Using a colorimetric assay or using a modified version of the thiobarbituric acid (TBA) reaction, Paisey *et al*., found fructosamine concentration was higher in hair of diabetic patients than in hair of healthy controls and that the concentration of hair glycation positively but loosely correlated with the concentration of glycated haemoglobin [30]. Later, Oimomi *et al*., showed that hair of diabetic patients had increased levels of fructose-lysine, a product of glycation, which is a method more specific to non-enzymatic glycation than the TBA method. Hair fructose-lysine was converted to furosine by hydrolysing hair in Hydrochloric acid (HCl) and furosine levels were measured using high performance liquid chromatography (HPLC) [28, 29, 75]. In 1996, Kobayshi & Igimi, developed a new method to study the glycation index in hair, as they were seeking a method which resulted in hair glycation measurements which correlated more strongly with patient HbA1c levels. Their glycation index was a ratio of glycated protein- to cysteine-induced colouration (A_{390}/A_{412}), which ultimately represented the amount of glycated protein relative to total protein. These glycation indices correlated strongly with HbA1c measurements of both diabetic and control patients [76], however due to excessive sample preparation, this method was not practical to be adopted into standard of care. Immunochemical detection of non-enzymatic glycation products in hair also showed a 3.1 – 4.9-fold increase of glycation in hair of diabetics compared to healthy controls [80].

Several studies also show that there is an increase in glycation of nail proteins among people with diabetes when compared to healthy controls [77-79]. Hair and nails are both keratinized structures which means the mass of both consist of over 80% keratin protein [81]. Keratins contain a significant number of lysine residues, which supports the evidence above, as the preferential target for non-enzymatic glycation is the ε-amino of a lysine residue [78, 82].

It is not only glycation levels that are affected in hair of diabetics. There are changes in levels of trace elements [83] and changes in the molecular structure of the α-keratin, possibly due to binding of glycation end products [27]. There is also evidence that the amino acid content is altered in hair of people with diabetes. Diabetics not only have increased blood glucose levels but also have increased circulating branched chain amino acids [84]. Recently, Rashaid *et al*., used gas chromatography – mass spectrometry (GC-MS) to analyse amino acid levels in hair of diabetic patients compared to hair of healthy controls. The hair amino acid profiles of people with diabetes were significantly different to those of the healthy controls. Particularly glycine, glutamic acid and isoleucine, which were increased in hair of diabetics. This may be a possible means of diagnosing diabetes, although a larger scale study needs to be done to confirm their findings [31]. In contrast, Shimode *et al*., showed that glycation of hair caused a significant reduction in total protein content [85]. However, the hair in this particular study was obtained from healthy volunteers and was glycated in vitro, meaning that the results did not fully represent the hair protein content of diabetics.

It is evident from the above-mentioned research that glycation levels are detectable in hair, which may even correlate with HbA1c levels and that there are other differences between the hair of diabetic patients and healthy controls. There are, however, confounding variables which may have affected results of the above-mentioned studies. Some of these factors include the type of diabetes, anti-diabetic medication [86], age, diet, the growth phase at which a strand of hair is currently in, interindividual variation in hair morphology (curl, colour, shine, thickness) and chemical treatment of hair.

1.6. ATTENUATED TOTAL REFLECTION (ATR) FOURIER TRANSFORM INFRARED (FTIR) SPECTROSCOPY AND MULTIVARIATE DATA ANALYSIS (MVA)

FTIR is a technique that uses infrared light absorption to infer chemical and structural information about a sample. A sample is subjected to infrared light which causes molecular vibrations on the surface of the sample and the light is either absorbed or transmitted depending on the stretching or bending of the covalent bonds within the molecules (See Figure 1.7.), resulting in spectral data [87, 88].

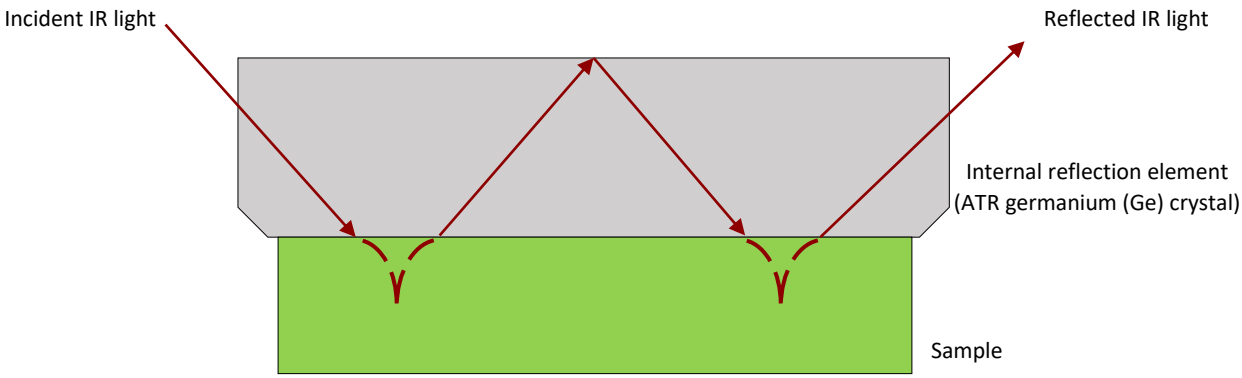


Figure 1.7. *Schematic diagram showing the principle of ATR-FTIR. Solid red line represents infrared light from the source which is reflected and dashed lines represent infrared light which has been absorbed by the sample due to molecular vibrations. (Own interpretation adapted from Lee et al., 2017 [88])*

The use of FTIR to study human hair is not new. Initially FTIR was used to characterise the chemical properties of hair in order to study chemical changes induced in hair by oxidation, photodamage, weathering and various chemical cosmetic treatments [89-104]. The technique has also been used to assess the diffusion of chemicals into hair fibres [105]. FTIR hair analysis also gained interest in forensic sciences to determine whether hair found at a crime scene is synthetic hair, animal hair or human hair and as the technology advances may even be used to determine certain characteristics (chemical treatment, gender and race) of human hair [106-108]. It has also been used in archaeological studies to determine conservation of hair samples after mummification [109, 110].

Investigations into FTIR analysis of human hair revealed that the technique may be used to detect recreational drug abuse [111], as well as diagnosing diseases such as breast cancer and thyroid disease [112-115]. Coopman *et al*. first demonstrated the ability to detect glycation of fingernail clippings (which are, like hair, keratinized structures) using FTIR in patients with diabetes [79]. Their methodology was using a conversion factor of 388 to convert the area under the curve of a single peak (970 to 1140 cm^{-1}) in the infrared spectrum obtained from the nails of each individual to glycated nail protein concentration. This peak corresponds to the carbohydrate region of the spectrum, but it is also influenced by many other factors including molecules with similar chemical bond profiles, such as certain amino acids. Due to the various factors not accounted for and the lack of correlation with an external validated gold-stand (e.g. HbA1c), this methodology is not reliable on its own to deduce glycaemic status.

Human scalp hair samples were been studied using FTIR by Zhang and Hassan to evaluate the differences in hair spectra between individuals with normal blood glucose levels compared to those with high blood glucose levels [116]. This study found differences in hair spectra obtained from people with high blood glucose, indicating that diagnosis and monitoring of diabetes through FTIR analysis of hair may be possible. They did, however, have a small sample size (6 individuals per group) and the transmittance-FTIR technique that was used in this study involved extensive sample preparation (including bleaching the hair samples, mixing them with Potassium Bromide and pulverizing and pressing the mixture into a thin film disc). This may have altered the chemical composition of the hair. ATR-FTIR is advantageous as there is no sample preparation involved, making it inexpensive and it is non-destructive, making investigation into spatial configuration possible and the same samples can be re-analysed or used again in alternative experiments.

Spectral data consist of thousands of variables, making small changes undetectable to the human eye. Traditional univariate analyses (such as t-tests) are not beneficial in the analysis of spectral data, as spectra consist of many variables to one observation i.e. “short and wide” datasets, which would result in false positives [117]. Multivariate data analysis (MVA) is a

highly graphical statistical technique used to analyse data that arises from multiple variables instead of investigating a single variable at a time. It allows for the study of the relationship between these variables as well as the variability in a data set. MVA is useful when analysing spectral data, especially in discrimination between spectra obtained from separate experimental groups (e.g. spectra obtained from hair of diabetics vs. healthy controls) or when carrying out multiple regression analysis between spectra and another variable (e.g. being able to predict HbA1c levels from spectra obtained from hair). MVA is also a useful visualization tool for detecting interesting information from extremely large datasets, as the software assigns a point in space to each observation according to the variable values and projects it onto a 2-dimensional plot (see Figure 1.8.[118]).

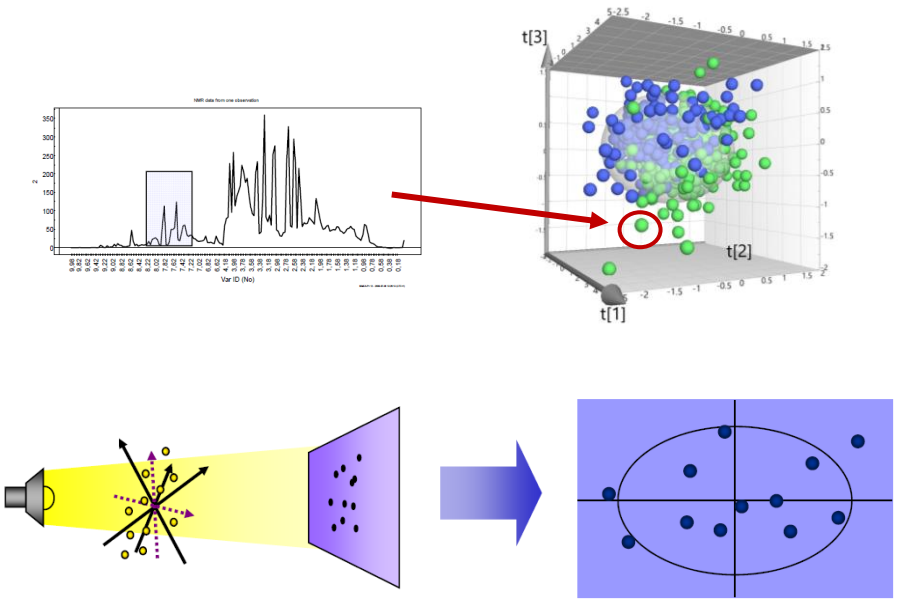


Figure 1.8. *Diagram depicting the basic principle of multivariate data analysis when analysing spectral data. Each spectrum produced by FTIR analysis is a single observation which is assigned as a point in 3D space. These points together are then projected onto a 2D graph for easier visualization. (Adapted from Wiklund, 2008) [118]*

1.7. AIM AND OBJECTIVES

The primary aim of the case control pilot study was to determine whether it is possible to detect and monitor chronic hyperglycaemia using hair.

The key objectives included:

- To use ATR-FTIR alongside multivariate analysis to determine whether it is possible to differentiate between hair taken from participants with chronic hyperglycaemia and hair taken from healthy controls.
- To use multivariate data analysis to determine whether there is a correlation between observed participant HbA1c levels and ATR-FTIR spectra obtained from participant hair.
- To determine whether there is a difference in amino acid levels in hair of participants with chronic hyperglycaemia compared to healthy controls.

CHAPTER 2: RESEARCH METHODOLOGY

2.1. STUDY PARTICIPANTS

The study comprised of 92 participants: 26 diabetic adults (30-70 years old), 29 healthy control adults (26-65 years old), 20 diabetic children (7- 18 years old) and 17 healthy children (7- 18 years old). Information about diabetic type (1 or 2) and chemical hair treatment (Dye, bleach, relaxers etc.) was collected for each participant.

2.2. ETHICAL CONSIDERATIONS

The study was approved by the University of Cape Town's ethics committee (HREC REF: 450/2016). Each participant volunteered to take part in the study and gave written informed consent (an assent and parental consent form were signed in the case of minors). Examples of these consent forms as well as the letter of approval can be seen in appendix 4.

2.3. RESEARCH SITE

Adult and children participants were recruited from the diabetic clinics at Groote Schuur hospital and Red Cross children's hospital in Cape Town, South Africa, respectively. All control participants recruited were escorts of people visiting the relative hospitals or visiting for minor ailments (e.g., routine dermatological issues).

2.4. BLOOD SAMPLE COLLECTION HBA1C MEASUREMENT

A finger prick blood sample was collected from each participant using the Haemoglobin Capillary Collection System (Bio-Rad, Richmond, CA). The kit is specifically designed to be compatible with Bio-Rad analysers for HbA1c testing and involves collecting blood in a plastic capillary which is then placed in a 1ml vial containing ethylenediaminetetraacetic acid (EDTA). Sample vials were stored at room temperature for no longer than one week before HbA1c levels were measured on the Bio-Rad D-10™ Haemoglobin testing system according to the manufacturers specifications. HbA1c has a stability of up to 2 weeks at room temperature but levels should be analysed on the Bio-Rad system within 7 days [119].

2.5. HAIR SAMPLE COLLECTION

Hair samples, wads approximately 5mm in diameter, were collected from the vertex of each participants' scalp. The proximal end of the hair sample was marked with string and the hair was cut as close as possible to the scalp with scissors. Samples were wrapped in tinfoil for storage and labelled with a unique identifier code corresponding to patient information.

2.6. HAIR SAMPLE PREPARATION AND STORAGE

Each hair sample was placed in a beaker and immersed in 1% sodium dodecyl sulphate (SDS) (±20ml) at 37°C and allowed to soak for 30 seconds. The sample was then gently shaken in the 1% SDS solution for 1 minute by hand and finally placed into a sieve to be rinsed under 37°C water running at 4L/min. In cases where the hair was especially dirty (leaving the solution murky or coloured by excess dye) the sample was shaken in fresh 1% SDS solution

for a further 30 seconds. The sample was then placed on a paper towel to air dry for at least 4 hours before various analyses.

2.7. ATTENUATED TOTAL REFLECTION FOURIER TRANSFORM INFRARED SPECTROSCOPY (ATR-FTIR) ANALYSIS OF HAIR SAMPLES

Intact individual hair fibres were mounted onto clean glass slides and subjected to ATR-FTIR analysis using a Bruker LUMOS FTIR Microscope, FTIR spectra were obtained from 5 random separate points within the proximal 3cm from 3 individual hair fibres from each participant resulting in 15 spectra per participant (Figure 2.1). Each spectral recording was performed in the mid-IR range of 4000-450 cm^{-1} for a scan time of 64 scans at $4cm^{-1}$ resolution. The ATR germanium (Ge) crystal was set at medium ATR pressure with the knife edge aperture set to 50x50µm. Collected spectra were min-max normalized using OPUS software (OPUS v.7.5 for Microsoft, Bruker Optics, Ettlingen, Germany) to correct for varying optical densities of different hair samples.

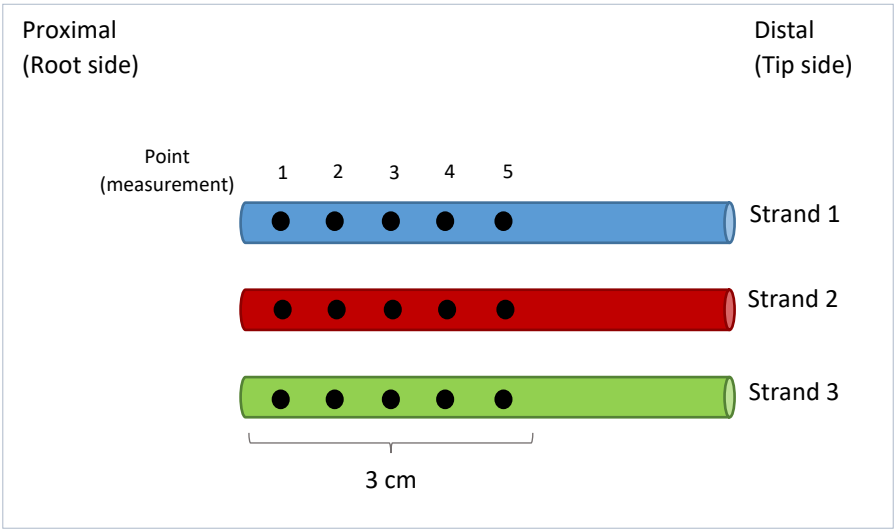


Figure 2.1. *Diagram showing how FTIR spectral readings were collected for each individual*

2.8. DATA ANALYSIS AND STATISTICS FOR ATR-FTIR DATA

FTIR data was analysed with SIMCA (MKS Umetrics, https://umetrics.com/products/simca) multivariate data analysis software. The following methods of multivariate data analysis were used:

2.8.1. Principal Component Analysis (PCA-X)

PCA is an unsupervised method which allows for an overview of the data such as recognising dominating variables (wavenumbers), observations (hair samples), obvious trends in the data and outliers [118, 120, 121]. This method is often used to reduce dimensionality in multidimensional data sets and allows for better visualization the variables that are responsible for the variation. The horizontal and vertical axes on PCA-X scores scatter plots represent model components t[1] and t[2] where t[1] explains the largest variation in the spectra (percentage of explained variance is represented by R2X[1]) followed by t[2] (percentage represented by R2X[2]).

Hoteling's T^2 plots and confidence ellipses (seen in scores scatter plots) were used to exclude extreme outliers. These are multivariate t-tests which estimate the probability that a score assigned to an observation is different from the mean of the distribution [122]. However, due to large variation in the spectra, outliers were excluded sparingly (only if the spectrum corresponding to the outlier looked visibly different).

2.8.2. Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA)

The traditional method for discriminant analysis is partial least squares discriminant analysis (PLS-DA), which is a supervised method based on displaying data in the direction of variation that separates classes rather than the direction of highest variation (as in PCA-X). OPLS-DA is a variation of PLS-DA where orthogonal signal correction is applied. This causes a 'rotation' of the scores scatter plot so that predictive variation is separated from the orthogonal variation [122]. This allows classification or discrimination between two groups (e.g. controls vs

diabetic) by separating discriminant variation in the data from variation that is not correlated to class separation. This separation between predictive and orthogonal variables is useful as it causes these models to be more robust against confounding variables than traditional methods of statistical analysis and allows for better interpretation of plots in terms of within class variation [118, 120-122].

In OPLS-DA two-class separation, the predictive variation in the spectra (i.e. the between class variation which causes the separation of the two groups) is represented on the horizontal axis of the scores scatter plot (percentage represented by R2X[1]). The vertical axis is orthogonal to the predictive component and represents within class or variation which is not related to the separation between the classes (percentage represented by R2X0[1]). If more than two classes are modelled, there will be several predictive components therefore the horizontal and vertical axes could both explain between class separation (the horizontal axis represents t[1] or the variation which is most important for the separation between the classes, followed by t[2], represented by the vertical axis).

R2(cum) and Q2(cum) values are calculated for models by the software. R2(cum) signifies the goodness of separation between the classes and represents the cumulative fraction of the variation in the classes which can be explained by all the components of a specific model. The Q2(cum) is an estimate of the predictive ability of the model and is calculated by cross validation of R2 values. For cross validation, the data is automatically divided into 7 groups by the software. A 7^{th} of the data is removed and a new model is built with the remaining data. The $1/7^{th}$ data that was excluded is then predicted by the new model (spectra are assigned to the nearest class using no threshold). This is repeated with all 7 groups, predicted data is compared with the original data and the sum of squared errors calculated (Predicted Residual Sum of Squares/PRESS). PRESS is then divided by the original sum of squares and subtracted from 1 to generate the Q2 values [109]. For spectroscopic calibration R2(cum) and Q2(cum) values are expected to be 0.99 and 0.98 respectively but considering this study was a pilot study including biological samples, values above 0.5 for both R2(cum) and Q2(cum) represent a good model for separation between classes [105, 109].

2.8.3. Partial Least Squares Regression (PLS)

PLS is also a supervised method of MVA which allows the identification of a relationship between different blocks of variables, e.g. spectral variables obtained from hair (X) and HbA1c (Y) [118, 120, 121]. Spectra (X-variables) are projected onto a PLS scores scatter plot in a similar fashion as a PCA-X plot, but taking the corresponding y-variable (HbA1c) into account. The horizontal and vertical axes represent model components t[1] and t[2], where t[1] explains the largest variation in the spectra followed by t[2].

Observed vs. predicted plots can also be plotted for PLS models. The model uses the leave $1/7^{th}$ of the data cross validation method again, but this time attempts to predict the Y-variables of the left-out data based on the X-variables i.e. it attempts to predict the HbA1c level of a participant based on the FTIR spectra obtained from the hair. The horizontal axis of an observed vs. predicted plot denotes the HbA1c levels predicted by the model and the vertical axis denotes the observed HbA1c levels.

2.9. AMINO ACID ANALYSIS

Amino acid analysis was performed on a subset of samples including, children with natural hair (n=25), adults with natural hair (n=14) and adults with dyed hair (n=20) as these were the groups with the most participants. Only the proximal 3cm of the hair was used. Two 20µg samples of hair from each participant was weighed into screw cap Eppendorf tubes for acid hydrolysis. A 1ml volume of 6N Hydrochloric acid was added to each tube and samples were heated on a heating block at 110°C for 24 hours. Samples were cooled, transferred to 15ml falcon tubes and neutralised (pH ~7.5) using 5% lithium hydroxide. Thereafter, 2ml of 3% 5-sulfosalicylic acid was added to each tube and left to stand for 30 minutes to allow for precipitation of soluble proteins. Each sample was filtered using a 0.2µm filter to remove debris. Each sample was mixed 1:1 with the sodium loading buffer prior to injection into the Biochrom 30 series amino acid analyser (Pharmacia Biochrom Ltd, United Kingdom).

This system utilised ion exchange chromatography with post column derivatization with ninhydrin to analyse amino acids in the hydrolysed hair samples. The samples were injected into a cation-exchange column and the amino acids were separated using a stepwise elution process involving two Sodium Hydrolysate buffers (Buffer 1=pH 3.2 and Buffer 5=pH 6.45) one Sodium Hydro/Oxid buffer (Buffer 2) and one Sodium regeneration buffer (Buffer 6). The varying pH and ionic strengths of the buffers, as well as carefully controlled column temperatures during the elution process, was necessary to produce accurate amino acid separation. The eluent was passed through a high temperature reaction coil along with ninhydrin to produce amino acid derivatives which were monitored at 570 and 440nm using a dual wavelength detector. This resulted in a chromatogram where the amino acid was identified by the peak retention time and the amino acid concentration was represented by the area under the specified peak. Amino acid quantification was based on the amino acids in a protein hydrolysate analytical calibration standard (Sigma-Aldrich) which was run after every 5^{th} sample. After each sample, the column was equilibrated by pumping a strong base followed by buffer 1 which regenerated it for the next analysis. The steps used in the modified protein hydrolysate high performance (NaHP) program are listed in Table A1 in appendix 1.

2.10. DATA ANALYSIS AND STATISTICS FOR AMINO ACID DATA

All amino acid data was analysed for normal distribution and statistical significance using STATA14 statistical analysis software.

CHAPTER 3: INVESTIGATING THE USE OF ATTENUATED TOTAL REFLECTION FOURIER TRANSFORM INFRARED SPECTROSCOPY (ATR-FTIR) TO DETECT HYPERGLYCAEMIA IN HAIR.

3.1. INTRODUCTION

When a hair sample is analysed with FTIR, the sample is exposed to infrared light in mid-IR region (450-4000cm^{-1}) which is either absorbed or reflected depending on the molecular properties of the sample. The result is a spectrum that is representative of the protein secondary structure for the hair fibre. Molecules also absorb the light at different wavenumbers depending on the molecular mass and bond forces geometries [97]. A typical spectrum for intact human hair fibre with labelled peak assignments [110] and the corresponding bond vibrations and wavenumber is shown in Table 3.1 and Figure 3.1, respectively. [107, 110, 123].

Table 3.1. FTIR peak bond vibration and wavenumber assignment

Peak	Bond vibration	Assigned Wavenumber (cm^{-1})
Cystine	SO2, SO & S=O stretching and bending	1121, 1070 and 1040
Amide III	(C=O) & (C-N) stretching and (N-H) & (O=C=N) deformation	1230
CH3 & CH2	(C-H) deformation	1411-1471
Amide II	(S-N) stretching and (N-H) deformation	1480-1575
Amide I	(C=O) & (C-N) stretching and (N-H) deformation	1600-1690
CH3 & CH2	(C-H) stretching	2875-2955
Amide A&B	(N-H) stretching	3070 and 3300

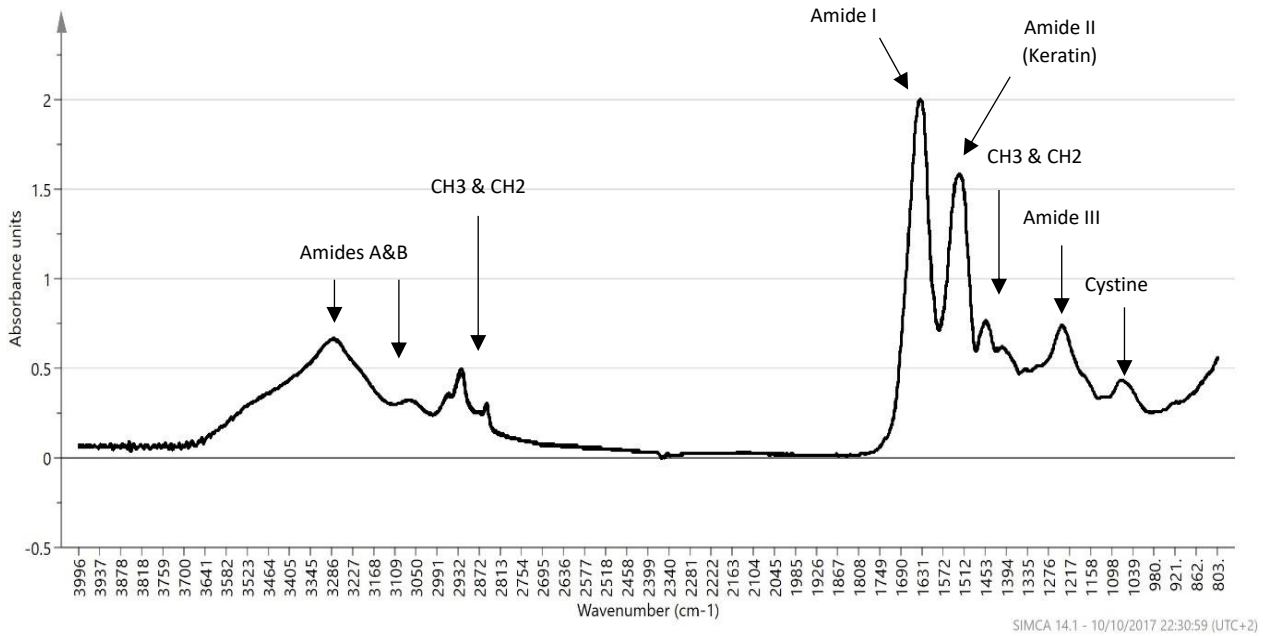


Figure 3.1. *Typical ATR-FTIR spectrum for an intact human hair fibre* (own results)

Since each ATR-FTIR spectrum is composed of approximately 1667 data points representing absorptions at different IR wavenumbers, and small spectral differences are difficult to spot with the human eye, multivariate data analysis (MVA) (SIMCA, MKS Umetrics) was used to extract relevant information from the data. Each spectrum is projected as a point in a 2-dimensional plot and the pattern can be used to interpret the relationships between the observations (participants) and the variables (spectral wavenumbers).

The aim of this chapter was to develop a workflow to differentiate between FTIR spectra of hair from individuals with diabetes and healthy controls and to determine the correlation between these spectra and measured HbA1c levels of participants. This involved several essential steps. Firstly, the raw spectra were assessed to determine whether they captured the relevant chemical information though peak identification. Secondly intra- and inter-individual variation of hair spectra were considered to develop a FTIR sample preparation and spectral collection protocol. Thirdly, potential factors which could affect classification of FTIR spectra into diabetic and control groups, such as age and chemical hair treatments were considered. Finally, all preceding factors were considered to build the best possible models to discriminate between ATR-FTIR hair spectra from diabetics and healthy controls as well as

to determine the correlation of these spectra with HbA1c readings, taking the limitations of the dataset into consideration. The models could then be internally validated through permutation testing.

3.2. RESULTS

3.2.1. Analysing Inter- and Intra-Individual Variation in FTIR Spectra to Determine the Best Method of Preparing and Scanning Hair Samples

Six individuals were chosen at random and their ATR-FTIR hair spectra were analysed to determine intra- and inter-individual variation in hair spectra. This was an important step in creating a FTIR sample collection workflow. Fifteen spectra were collected from each person (i.e., 5 spectra per strand from 3 hair strands) as described in Figure 3.2. This resulted in a total of 1380 spectra (30 of these spectra were excluded as they had low absorption values due to bad contact between the ATR Geranium crystal and the hair). Spectra were min-max normalized using OPUS, the instrument software, for baseline/spectral shift correction and to correct for the different optical densities between hair fibres, as hair colour and thickness differs from person to person and between two separate hairs from the same individual.

PCA-X scores scatter plots of spectra were generated for each individual to get an overview of the trends relating to intra-individual variation (Figure 3.2. (B-G)). Spectra from the same strand grouped more closely than spectra taken from the same point on each strand (points of matching colour group together rather than points with matching labels 1-5), indicating that there is greater variation between different hair strands than there is along a single strand (i.e. proximal vs. distal side). The 90 FTIR spectra from all 6 randomly chosen individuals were then all plotted onto the same PCA-X plot to get an overview of trends relating to inter-individual variation (Figure 3.2. (H)). As expected, spectra from the hair of the same individuals grouped close together, indicating larger inter-individual variation than intra-individual variation.

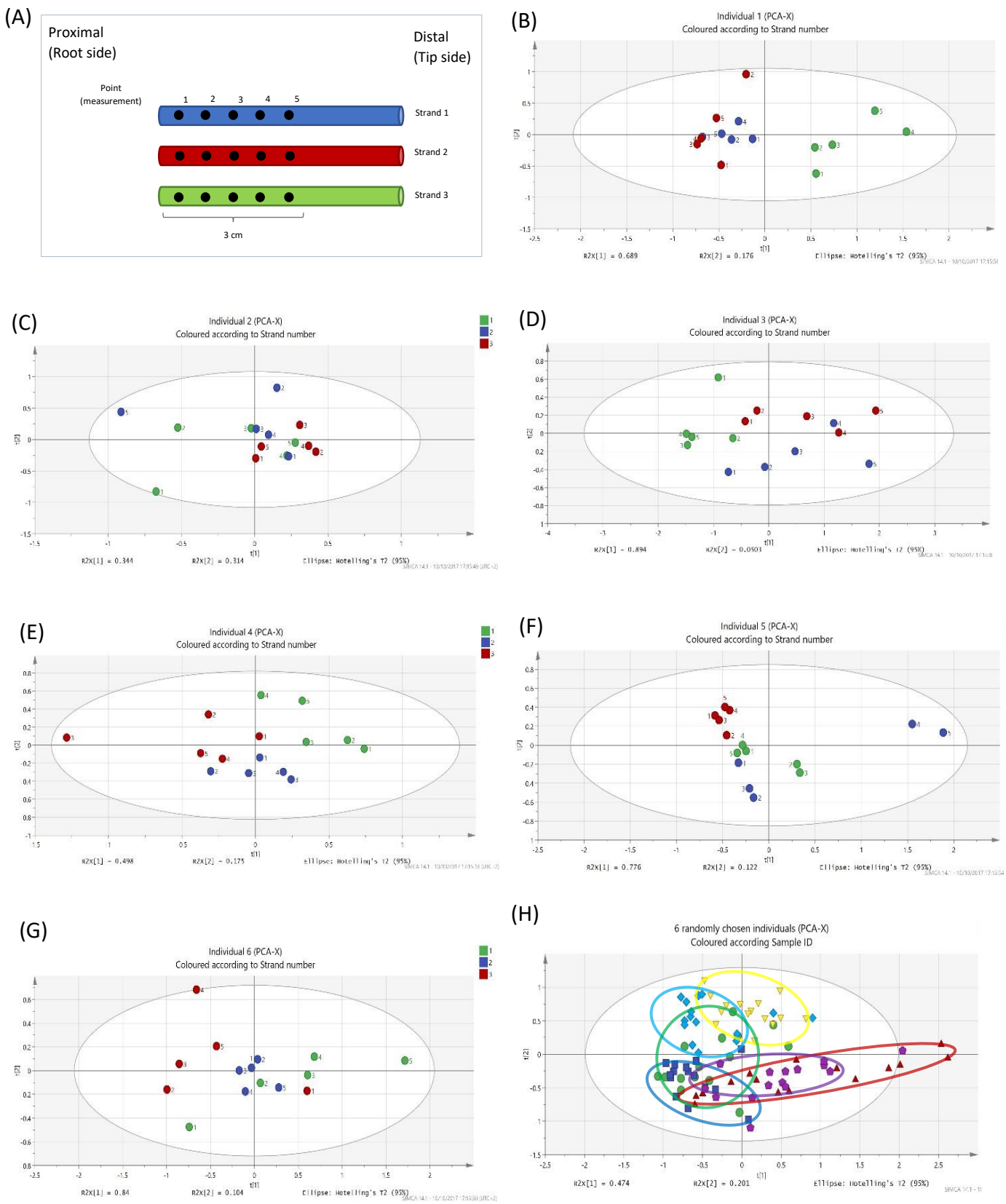


Figure 3.2. (A) *Sketch showing FTIR spectral data collection for each individual* ***(B-G)*** *PCA-X scores scatter plots displaying distribution of the 15 spectra obtained from hair of 6 randomly chosen individuals. Plots are coloured according to strand number and points are labelled according to the point measured on each strand (1=Root side, 5= Tip side)* ***(H)*** *PCA-X scores scatter plot displaying distribution of 6 randomly chosen individuals based on hair spectra, coloured according to Sample ID (coloured circles used to emphasise grouping of individuals).*

OPLS-DA models between individuals were then created including spectra measured from either 1, 3 or 5 points on each of the 3 strands. Separation and predictive ability of OPLS-DA models increased when more points were measured (Figure 3.3.). When only 1 point per strand was measured (A), the predictive ability of the model was poor (Q2(cum)=0.435). Separation between individuals and predictive ability improved when 3 points per strand were measured (B) (Q2(cum)=0.682) and the 6 individuals showed excellent separation when all 5 points from each strand were included in the analysis (C) (Q2(cum)=0.804). This result emphasises that it is necessary to measure at least 5 points on each strand for optimum distinction between groups.

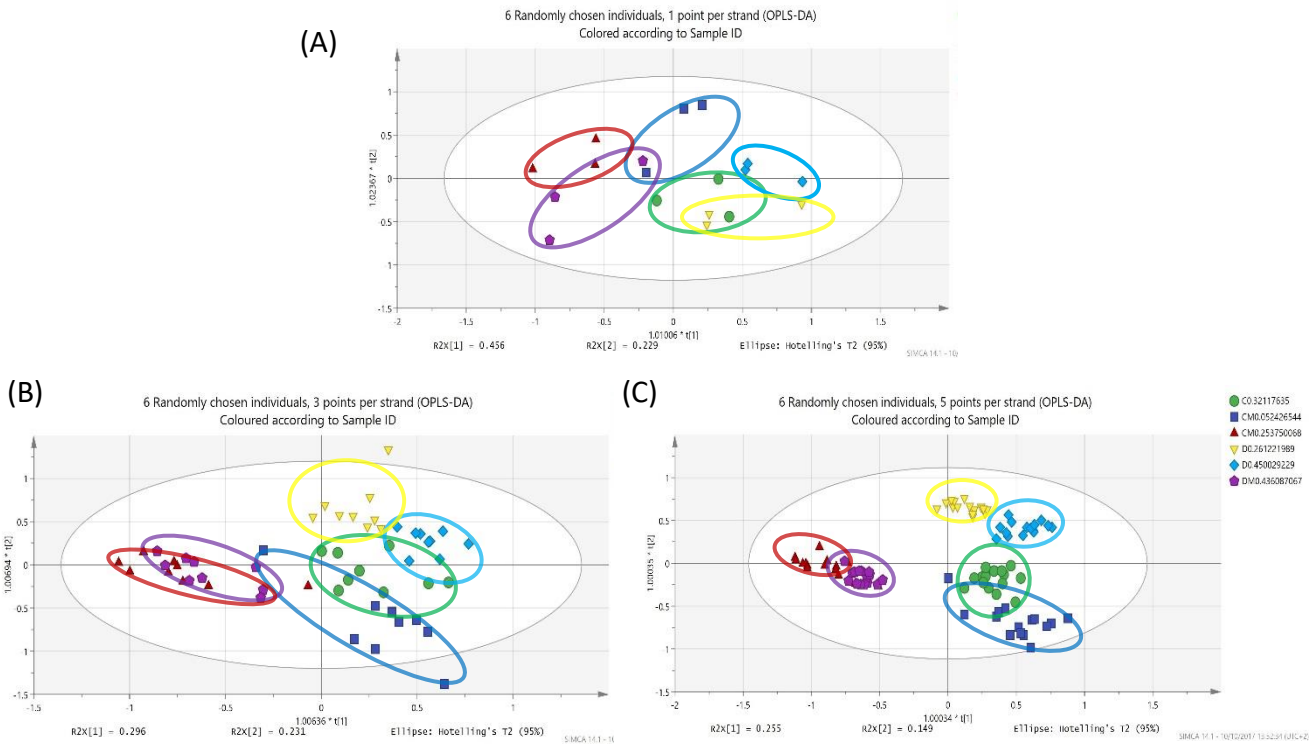


Figure 3.3. (A-C) *OPLS-DA scores scatter plots showing the separation between spectra obtained from hair of 6 randomly chosen individuals.* ***(A)*** *1 spectrum taken from each strand of hair where R2(cum)=0.722 and Q2(cum)=0.435.* ***(B)*** *3 spectra taken from each strand of hair where R2(cum)=0.758 and Q2(cum)=0.682* ***(C)*** *5 spectra taken from each strand of hair where R2(cum)= 0.862 and Q2(cum)=0.804. All plots are coloured according to sample ID (coloured circles used to emphasise grouping of individuals- closer grouping results in tighter circles).*

3.2.2. Factors Affecting Hair Spectral Classification

Principal component analysis (PCA-X) was used as a first-pass method to get an overview of the general trends within the entire dataset and to observe whether there was an obvious chemical difference between the FTIR spectra of the diabetics and healthy controls without any training of the data. All 1350 FTIR spectra were analysed using PCA-X and the scores scatter plots were interrogated to determine whether differences, inherent to the study population, resulted in distinct grouping.

3.2.2.1. Effect of Age: Adults vs. children

When the scores scatter plot was coloured according to different classes the clearest distinction was between spectra taken from hair of adults compared to children (Figure 3.4. (A) – Figure A1 in appendix 2 displays the same scores scatter plot coloured according to other classes i.e. chemical hair treatment type and diabetic status). An OPLS-DA scores scatter plot was produced to confirm the distinction between these two groups (Figure 3.4. (B)). A R2(cum) of 0.753 and Q2(cum) of 0.746 was obtained, indicating excellent predictability and separation between groups. Due to differences in spectra of hair from adults vs. children, spectra were analysed according to the two different classes in future analyses.

3.2.2.2. Effect of Chemical Hair Treatment

OPLS-DA models were created to determine whether chemical treatment affected FTIR spectra of hair (Figure 3.5.). Since most children had natural hair, this was only inspected among the adults. Table 3.3. summarizes the different treatment types within the sample.

Table 3.2. Summary of different treatment types among adults.

Treatment type	Number of adults
Natural	19
Dyed	21
Henna	3
Relaxed	8
Permanent waved	1
Brazilian Keratin treatment	1

When all treatment types were included, spectra obtained from individuals who had Brazilian blow out treatments or permanent waves were the most different than those obtained from people who had another, or no chemical treatment (Figure 3.5. (A)). These individuals (and those who had treated their hair with henna) were, however, excluded due to small sample size. Figure 3.5. (B) and (C) show separation between spectra obtained from individuals who had natural hair, dyed hair and relaxed hair. The goodness of separation and predictive ability increased from R2(cum)=0.441 to R2(cum)=0.597 and Q2(cum)=0.426 to Q2(cum)=0.58 respectively when individuals who had not chemically treated their hair in the past 4 months were excluded. This shows that chemical hair treatment influenced hair spectra and therefore further analysis was performed separately for each treatment group.

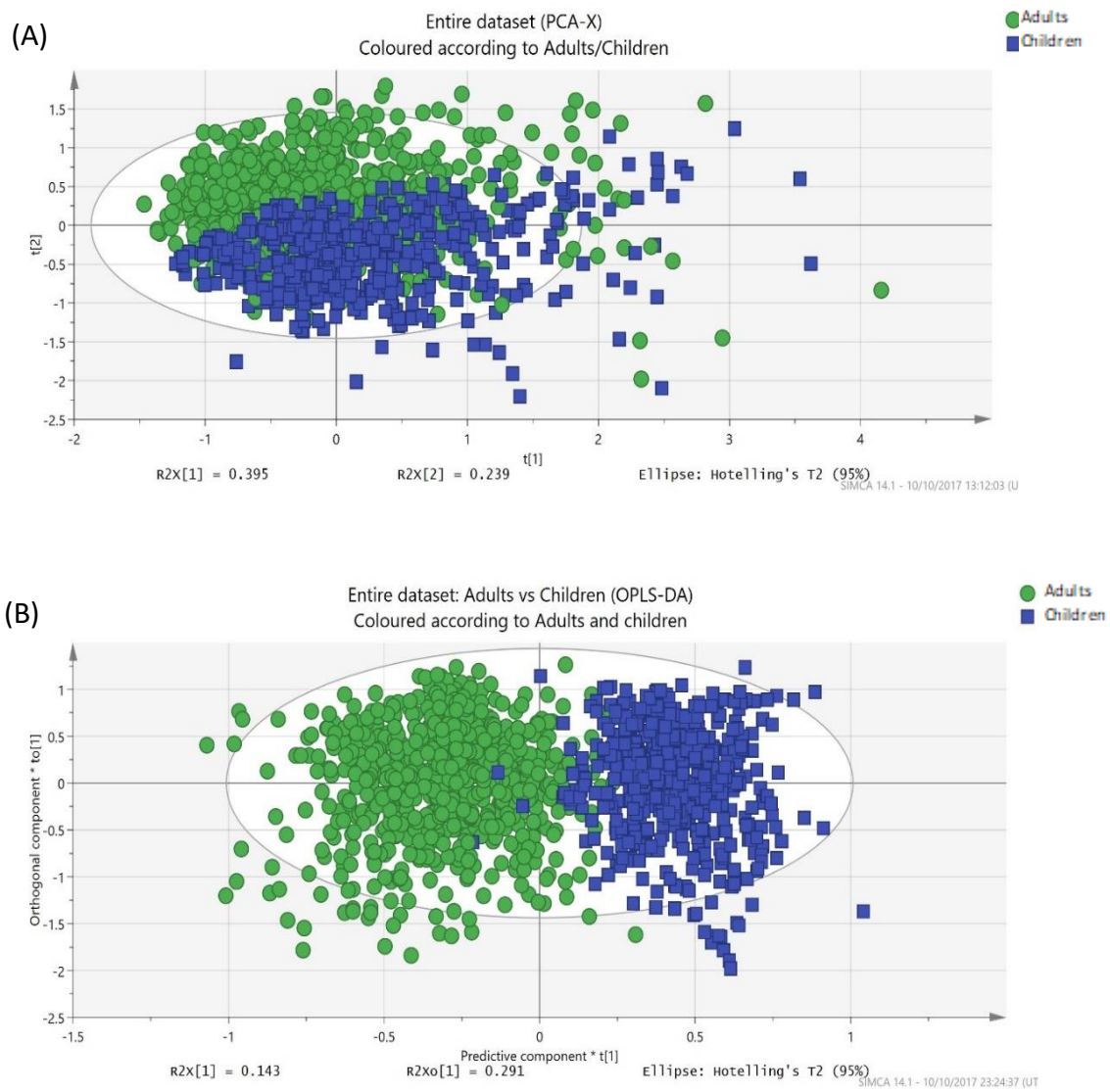
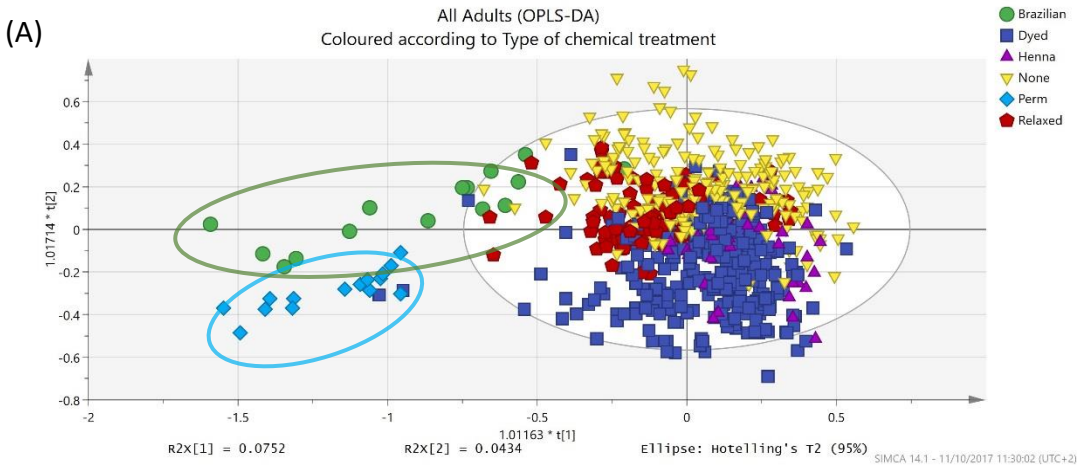
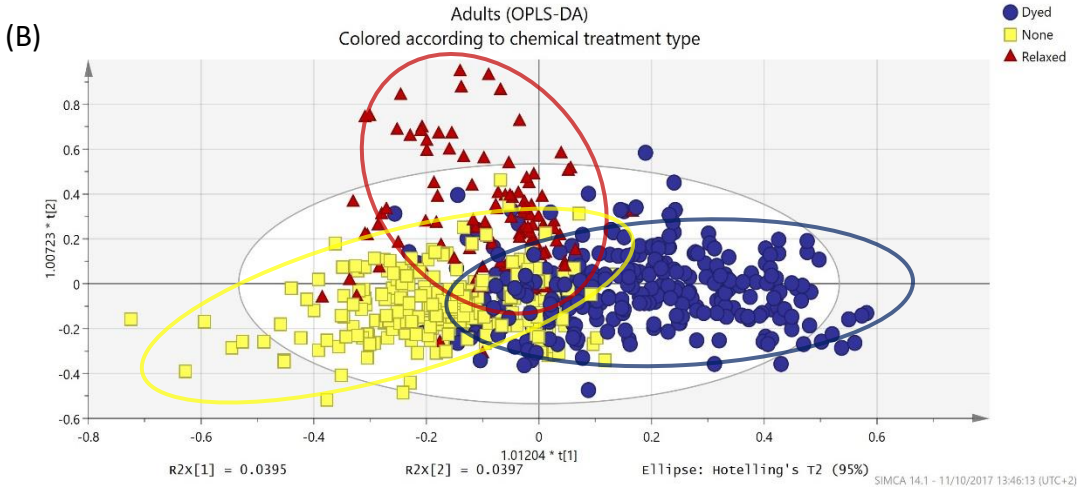


Figure 3.4. (A) *PCA-X scores scatter plot displaying distribution of the entire dataset of participants based on hair spectra.* ***(B)*** *OPLS-DA scores scatter plot of the entire dataset showing the separation between spectra obtained from hair of adults vs. children. R2(cum)=0.753 and Q2(cum)=0.746. Both plots are coloured according to classes Adults (green) and Children (blue).*





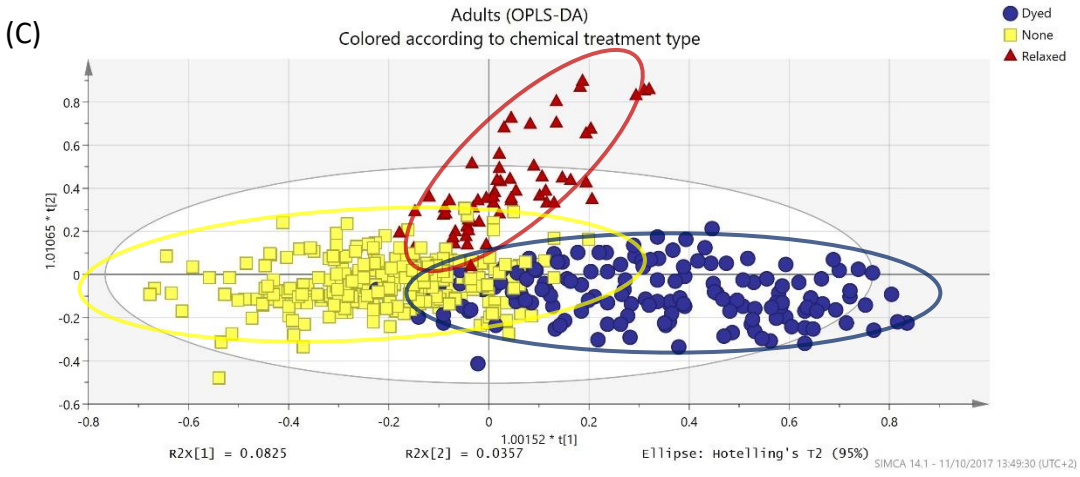


Figure 3.5. *OPLS-DA scores scatter plots showing the separation between spectra obtained from hair of adults with different chemical treatments* ***(A)*** *shows separation between all chemical treatments where R2(cum)=0.389 and Q2(cum)=0.369* ***(B)*** *Shows separation between untreated, dyed and relaxed hair, where R2(cum)=0.441 and Q2(cum)=0.426* ***(C)*** *shows separation between untreated and hair which has been dyed or relaxed within the past 4 months, where R2(cum)=0.597 and Q2(cum)=0.58. All plots are coloured according to treatment type.*

3.2.3. Discrimination Between ATR-FTIR Spectra Obtained from Hair of Controls vs. Diabetics Using Orthogonal Projections to Latent Squares Discriminant Analysis (OPLS-DA)

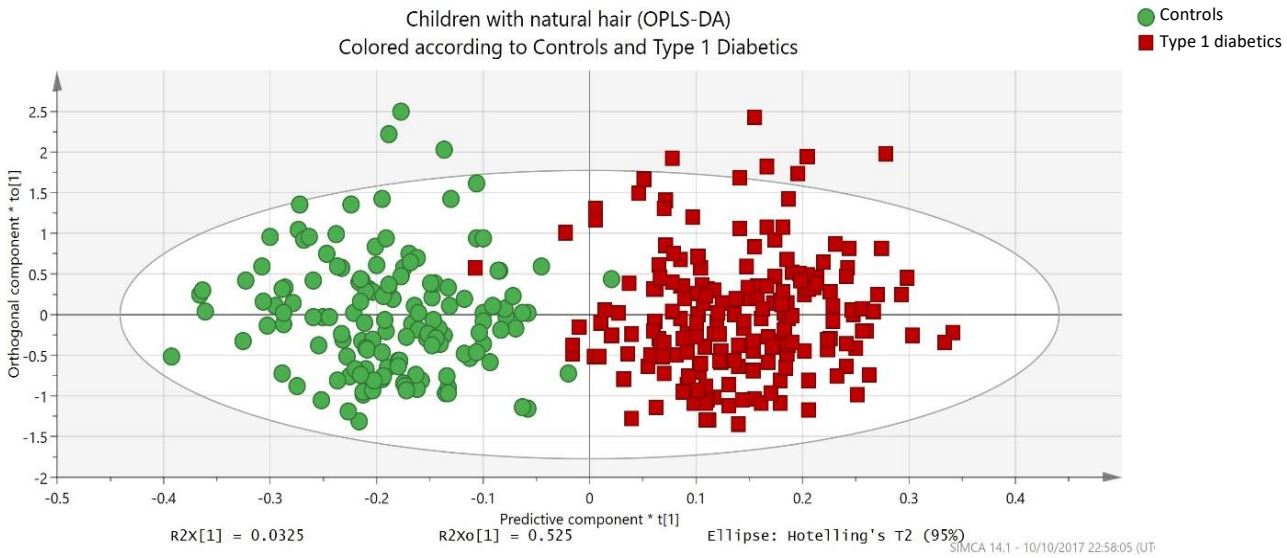
OPLS-DA models were used to determine whether FTIR spectra obtained from hair of diabetic participants were distinguishable from those obtained from hair of healthy controls. All control individuals who had borderline (pre-diabetic) HbA1c levels (HbA1c > 6%) and all well controlled diabetics (HbA1c <8%) were excluded from the analysis so that control and diabetic groups were clearly defined. Adults with type 1 diabetes were also excluded due to small sample size (n=3).

Models were constructed for four groups: children with natural hair (Figure 3.6.), adults with natural hair (Figure 3.7.), adults with dyed hair and adults with relaxed hair (Figure 3.8.). The OPLS-DA scores scatter plots showed good separation and predictive ability between spectra obtained from hair of diabetics compared to those obtained from hair of healthy controls, with the Q2(cum) ranging between 0.736 and 0.85 in these groups.

Corresponding loading plots (below each scores scatter plot) were generated for each model, which display the wavenumbers (horizontal axis) most important in the separation seen between controls and diabetics (p(corr) absorbance units is denoted on the vertical axis). The peak at 1200cm^{-1} which corresponds to the Amide III peak (Refer to Figure 3.1. and Table 3.1.). is commonly important in the discrimination between spectra obtained from hair of diabetics vs. controls in all groups.

Classification tables were constructed for each of the OPLS-DA models. These tables indicate how well the models can classify hair spectra into diabetic groups. Using the same principle of cross-validation as described in section 5.1.1. (1/7th of the data is excluded, then the excluded spectra are assigned to the nearest class using no threshold [121]). The OPLS-DA models could predict spectra as being from either the control or diabetic groups with 96.97–100% accuracy (Tables 3.3 – 3.6).

(A)



(B)

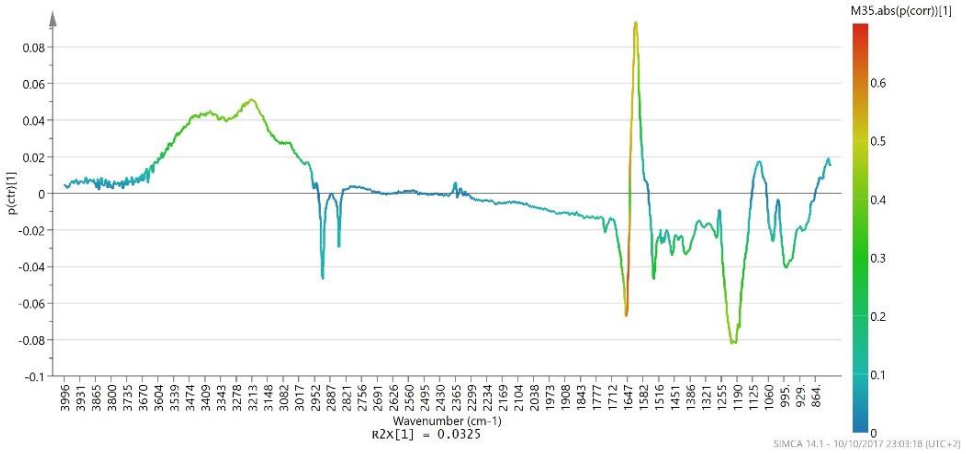
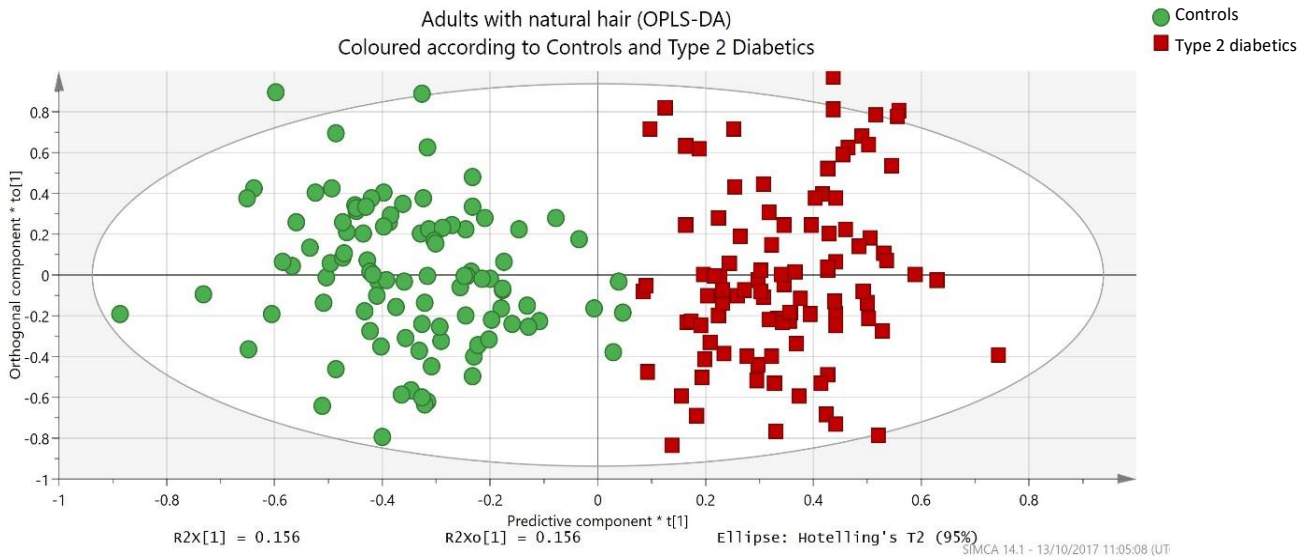


Figure 3.6. (A) *OPLS-DA scores scatter plot for children with natural hair showing the separation between spectra obtained from hair controls (green) vs type 1 diabetics (red) where R2(cum)=0.822 and Q2(cum)=0.753* ***(B)*** *Loading plot showing wavenumbers of spectra which contribute most to the separation seen in* ***(A)****, coloured according to p(corr) absorbance units.*

(A)



(B)

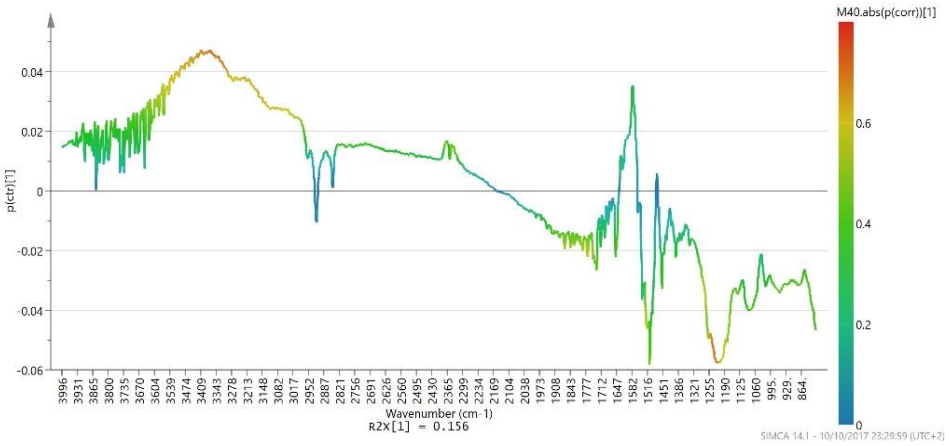


Figure 3.7. (A) *OPLS-DA scores scatter plot for adults with natural hair showing the separation between spectra obtained from hair controls (green) vs type 2 diabetics (red) where R2(cum)=0.841 and Q2(cum)=0.819* ***(B)*** *Loading plot showing wavenumbers of spectra which contribute most to the separation seen in* ***(A)****, coloured according to p(corr) absorbance units.*

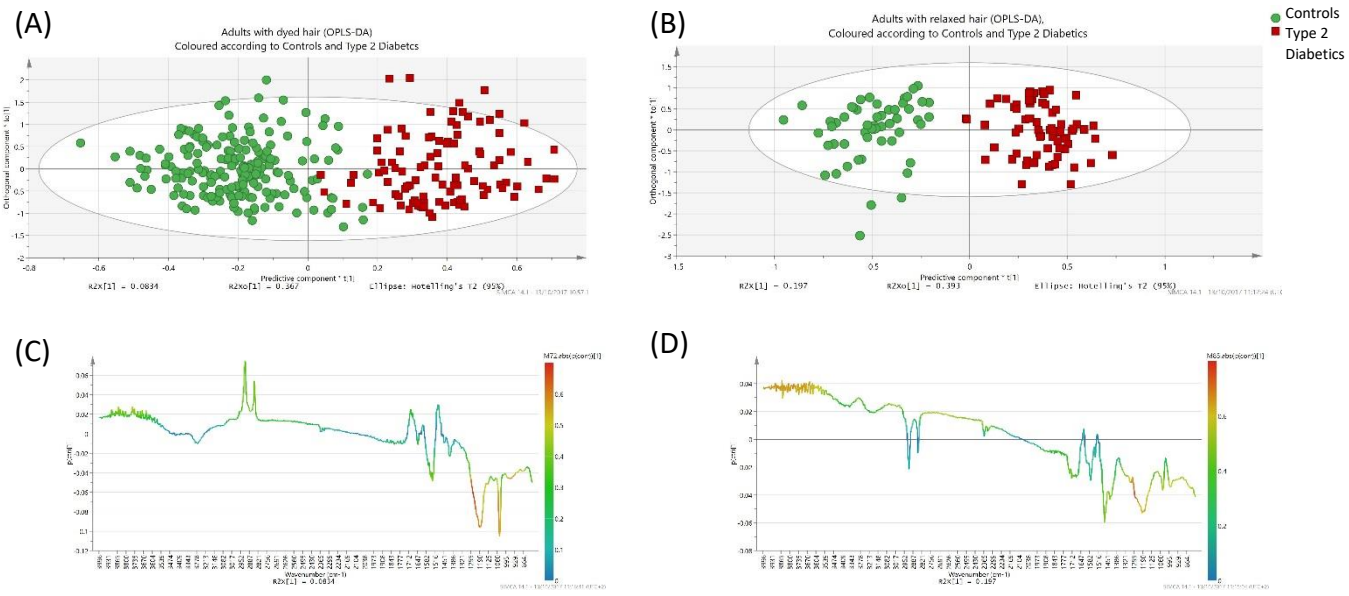


Figure 3.8. (A-B) *OPLS-DA scores scatter plots showing the separation between spectra obtained from hair of type 2 diabetics (red) vs control (green) adults with* ***(A)*** *dyed hair, where R2(cum)=0.791 and Q2(cum)=0.736 and* ***(B)*** *relaxed hair, where R2(cum)=0.88 and Q2(cum)=0.85* ***(C-D)*** *Loading plots showing wavenumbers of spectra which contribute most to the separation seen in* ***(A-B)*** *respectively, coloured according to p(corr) absorbance units.*

Table 3.3. Classification table for spectra obtained from hair of diabetics vs. controls amongst children with natural hair

	Members	Correct	Controls	Type 1 diabetics	No class (YPred <= 0)
Controls	141	98.58%	139	2	0
Type 1 diabetics	188	99.47%	1	187	0
No class	0		0	0	0
Total	329	99.09%	140	189	0
Fisher's prob.	0				

Table 3.4. Classification table for spectra obtained from hair of diabetics vs. controls amongst adults with natural hair

	Members	Correct	Controls	Type 2 diabetics	No class (YPred <= 0)
Controls	99	96.97%	96	3	0
Type 2 diabetics	99	100%	0	99	0
No class	0		0	0	0
Total	198	98.48%	96	102	0
Fisher's prob.	0				

Table 3.5. Classification table for spectra obtained from hair of diabetics vs. controls amongst adults with dyed hair

	Members	Correct	Controls	Type 2 diabetics	No class (YPred <= 0)
Controls	192	98.44%	189	3	0
Type 2 diabetics	103	97.09%	3	100	0
No class	0		0	0	0
Total	295	97.97%	192	103	0
Fisher's prob.	0				

Table 3.6. Classification table for spectra obtained from hair of diabetics vs. controls amongst adults with relaxed hair

	Members	Correct	Controls	Type 2 diabetics	No class (YPred <= 0)
Controls	54	100%	54	0	0
Type 2 diabetics	73	100%	0	73	0
No class	0		0	0	0
Total	127	100%	54	73	0
Fisher's prob.	3.4e-37				

3.2.4. OPLS-DA Model Validation

It is possible that the previously constructed OPLS-DA models may fit the training set well but may not be able to predict whether new spectra belong in the control or diabetic groups. It is also possible that models were overfit. Therefore, each OPLS-DA model was validated using permutation plots to ensure that the models are not spurious and would be able to classify a new spectrum correctly to either the control or diabetic groups. To create a permutation plot, classes were randomly permuted (reshuffled) 20 times and the R2 and Q2 values were recalculated for each of the 20 'new' models. The new R2 and Q2 values were then plotted along with the R2 and Q2 values for the original model, where the y-axis denotes the R2 and Q2 values and the x-axis denotes the correlation between the original values and permuted values. The permutation plots for the 4 groups are shown in Figure 2.9. A model is valid if the regression line for the Q2 points (blue) intersect the y-axis at or below zero and the regression line for the R2 points (green) intersect the y-axis at or below 0.3 [109, 110]. The permutation plots for all OPLS-DA models to discriminate between spectra from hair of diabetics vs. healthy controls fit the above criteria, indicating that these models were valid and are not overfit (Figure 3.9.).

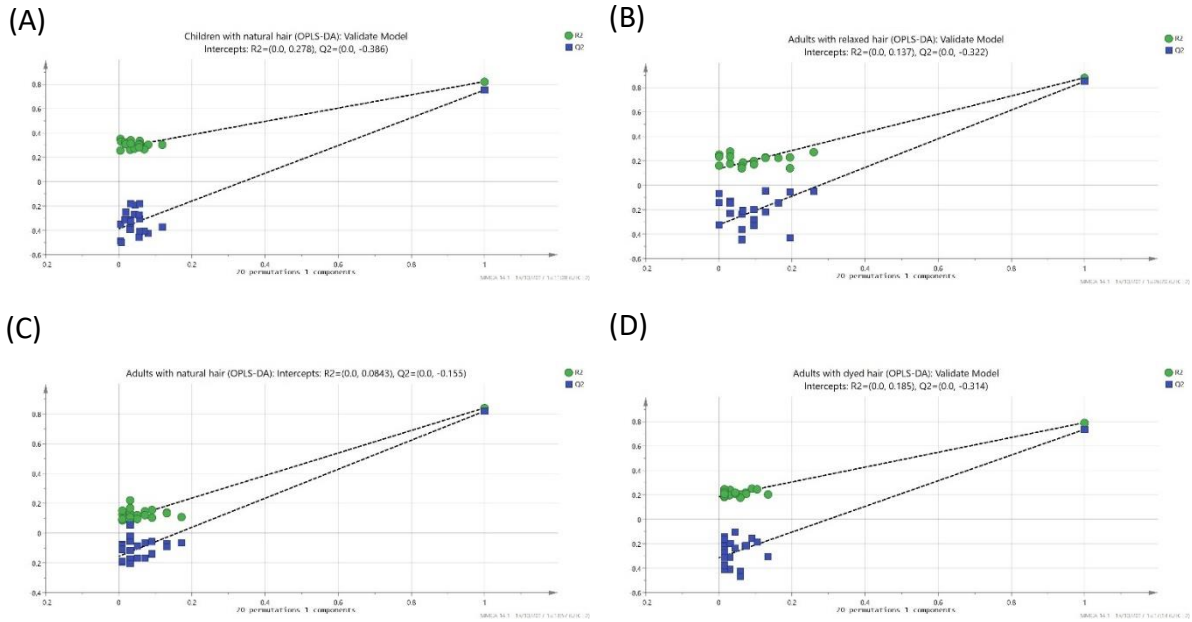


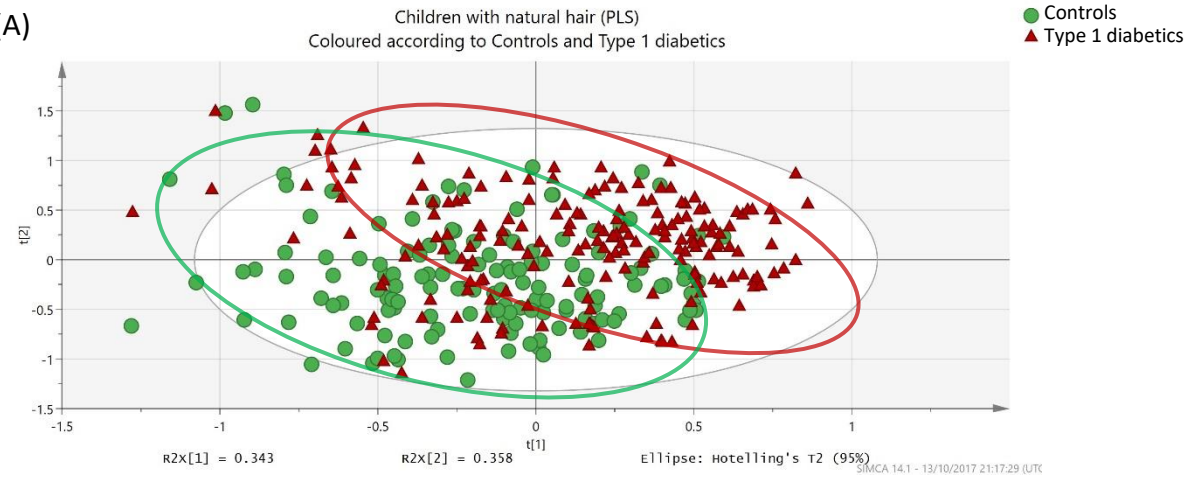
Figure 3.9. *Permutation plots for the internal validation of OPLS-DA models for controls vs. diabetics in the following groups:* ***(A)*** *Children with natural hair* ***(B)*** *Adults with natural hair* ***(C)*** *Adults with dyed hair* ***(D)*** *Adults with relaxed hair.*

3.2.5. Investigating the Correlation Between HbA1c Levels and ATR-FTIR Spectra from Hair Using Partial Least Square Regression (PLS)

PLS can be used to determine whether there is a relationship between different blocks of variables i.e. between ATR-FTIR spectra obtained from hair and HbA1c levels of participants. Separate PLS models were constructed for each of the groups i.e. children with natural hair (Figure 3.10.), adults with natural hair (Figure 3.11.), adults with dyed hair (Figure 3.12.) and adults with relaxed hair (Figure 3.13.), to determine if there was a correlation between hair FTIR spectra and HbA1c readings for each participant.

In each case, although the scores scatter plots show overlap between groups in PCA-X models (Figure A1 in appendix 2), distinct grouping of diabetics and controls was observed in scores scatter plots when the HbA1c measurements were included in the model to create PLS models (Figures 3.10-3.13A). All PLS observed vs. predicted plots showed that there was a strong correlation between FTIR hair spectra and HbA1c levels as indicated by the linear regression coefficient (R^2 values). These values were as follows: 0.8067 for children with natural hair; 0.8481 for adults with natural hair; 0.895 for adults with dyed hair and 0.9296 for adults with relaxed hair. The predicted ranges were large and showed a considerable amount of overlap (refer to x-axis of PLS observed vs. predicted plots in (B) of Figures 3.10 – 3.13). This once again emphasises that multiple spectral readings from multiple strands of hair for each individual would be necessary to calculate a reliable final predicted HbA1c from the mean or the median predicted level. For example, referring to Figure 3.10(A), there is one child who has an observed HbA1c of 14.3%. The lowest predicted HbA1c value for this child (obtained from a single spectrum) is 12.18% and the highest predicted value is 17.69%. However, the correlation line intersects the prediction axis at 14.22% (median) and the mean of all fifteen predicted HbA1c readings for this child is 14.20%.

(A)



(B)

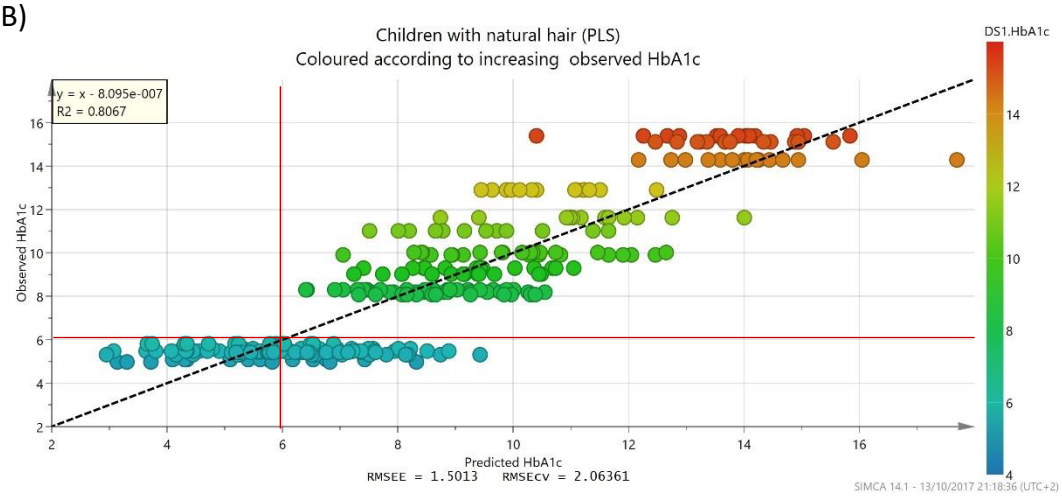


Figure 3.10. *PLS models for children with natural hair* ***(A)*** *PLS scores scatter plot showing grouping of spectra from hair of type 1 diabetics and controls based on participant HbA1c readings (Circles emphasise grouping of controls and diabetics)* ***(B)*** *Observed vs. predicted plot displaying observed HbA1c readings for each participant versus HbA1c predicted by the model using hair spectra (R^2=0.8067). Red lines denote healthy HbA1c threshold.*

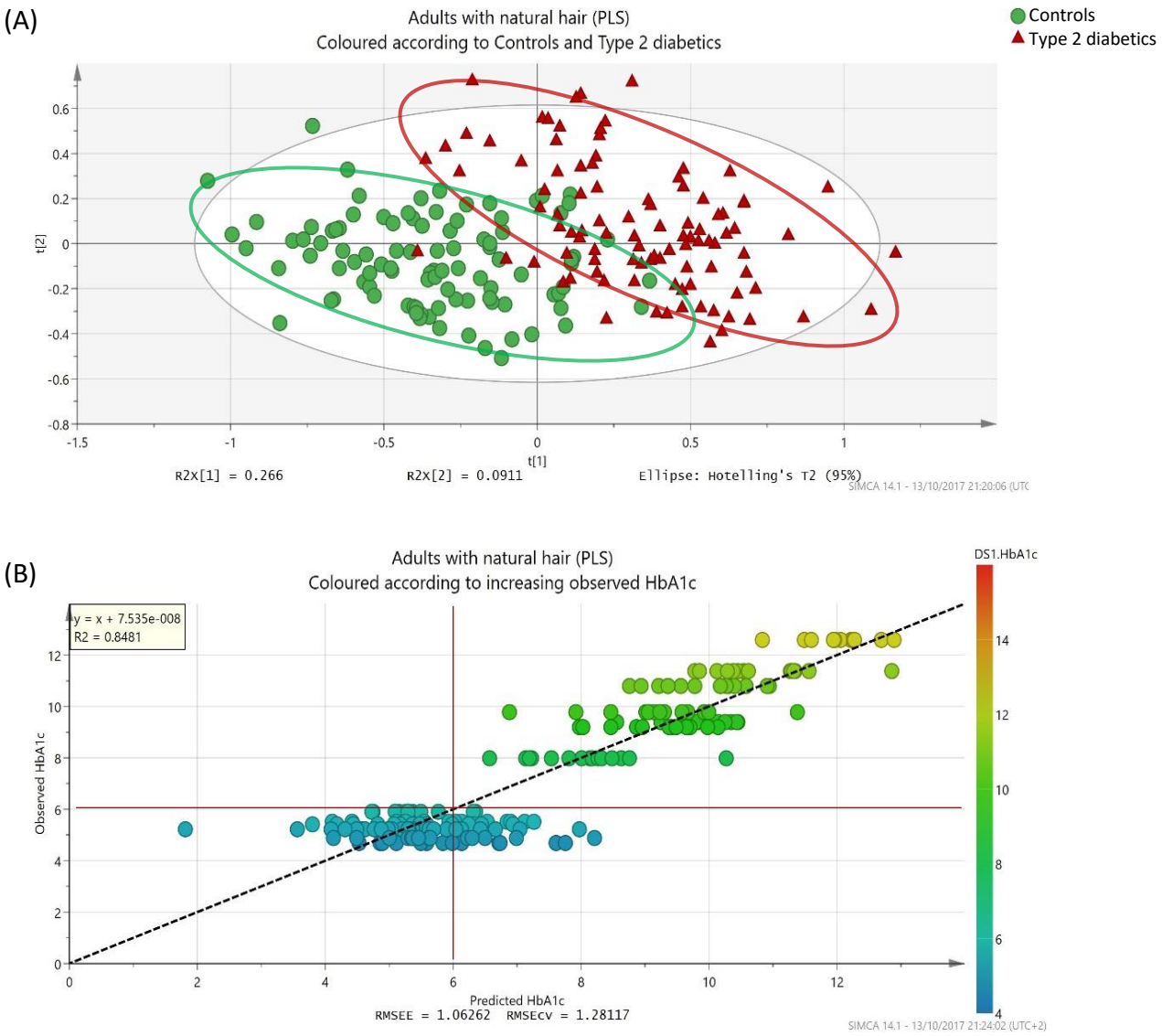
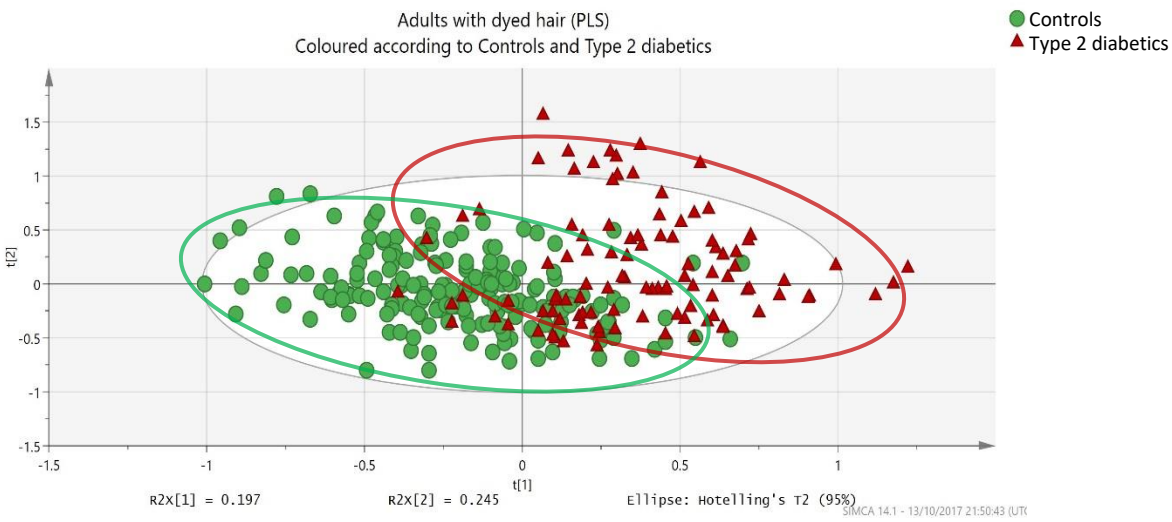


Figure 3.11. *PLS models for adults with natural hair* ***(A)*** *PLS scores scatter plot showing grouping of spectra from hair of type 2 diabetics and controls based on participant HbA1c readings (Circles emphasise grouping of controls and diabetics)* ***(B)*** *Observed vs. Predicted plot displaying observed HbA1c readings for each participant versus HbA1c predicted by the model using hair spectra (R^2=0.8481). Red lines denote healthy HbA1c threshold.*

(A)



(B)

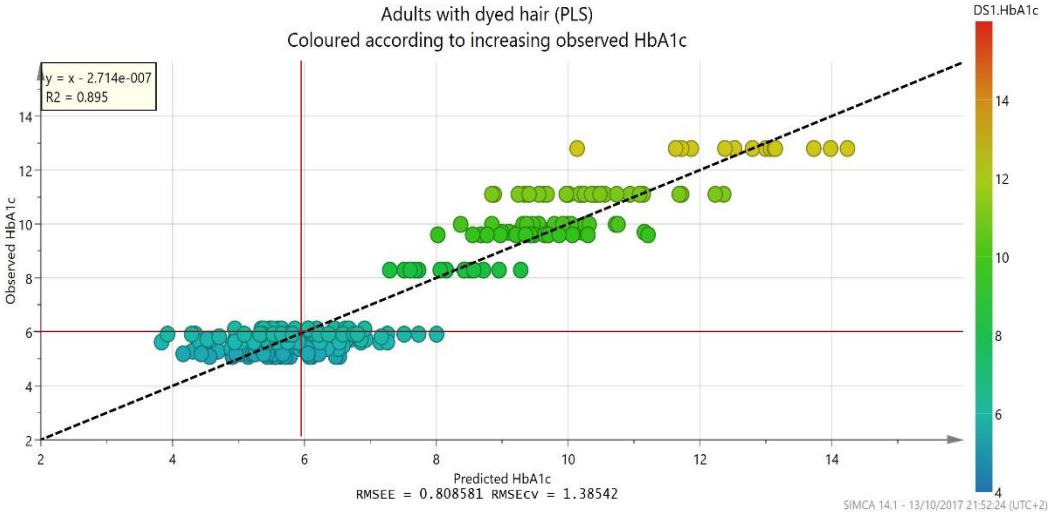
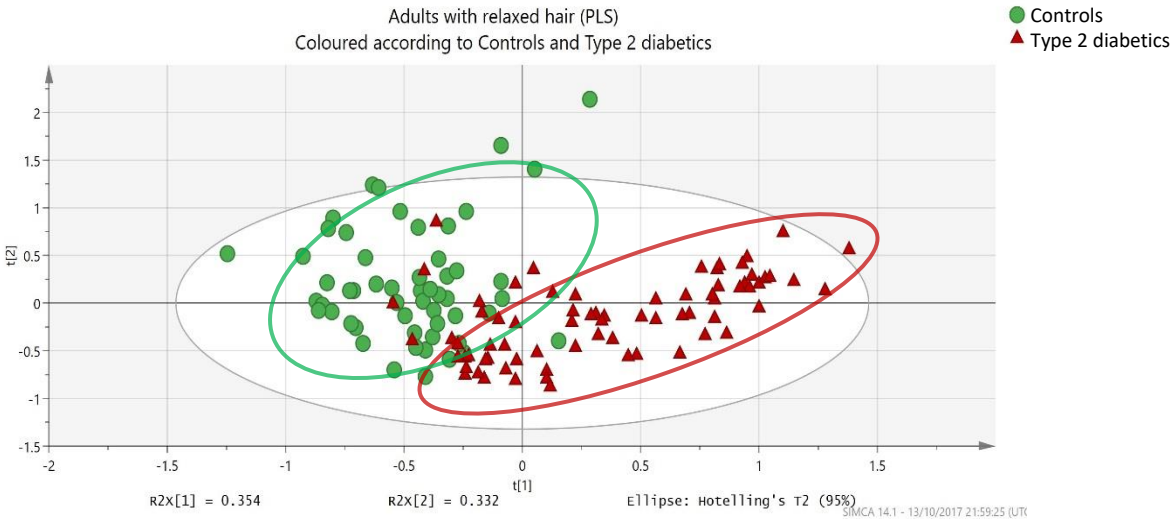


Figure 3.12. *PLS models for Adults with dyed hair* ***(A)*** *PLS scores scatter plot showing grouping of spectra from hair of type 2 diabetics and controls based on participant HbA1c readings (Circles emphasise grouping of controls and diabetics)* ***(B)*** *Observed vs. predicted plot displaying observed HbA1c readings for each participant versus HbA1c predicted by the model using hair spectra (R^2=0.895). Red lines denote healthy HbA1c threshold.*

(A)



(B)

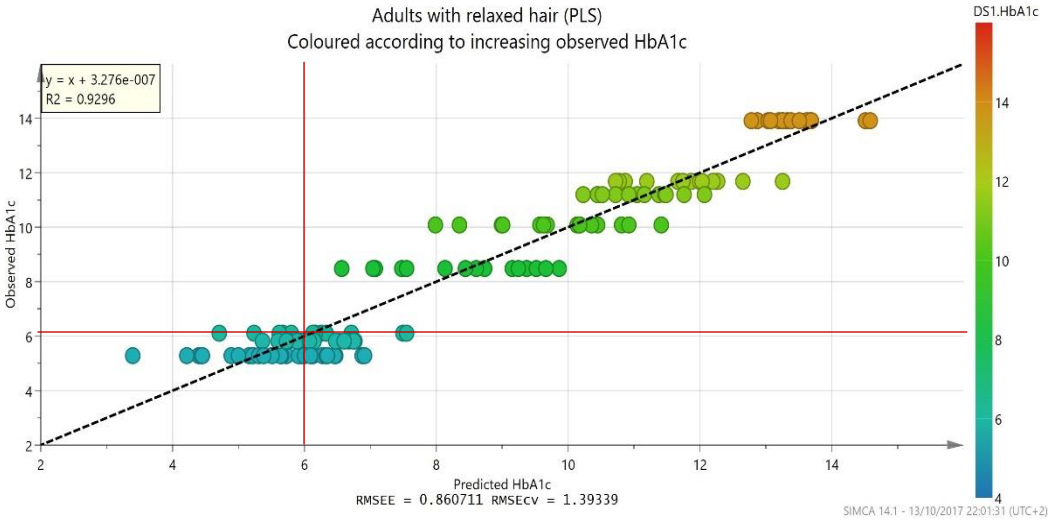


Figure 3.13. *PLS models for Adults with relaxed hair* ***(A)*** *PLS scores scatter plot showing grouping of spectra from hair of type 2 diabetics and controls based on participant HbA1c readings (Circles emphasise grouping of controls and diabetics)* ***(B)*** *Observed vs. predicted plot displaying observed HbA1c readings for each participant versus HbA1c predicted by the model using hair spectra (R^2= 0.9296). Red lines denote healthy HbA1c threshold.*

3.2.6. PLS Model Validation

Permutation plots were used to validate the above PLS models (Figure 3.14.). To create a permutation plot, HbA1c levels predicted by the models were randomly permuted 20 times and the R2 and Q2 values for the PLS scores scatter plot are recalculated for each of the 20 'new' models. The new R2 and Q2 values were then plotted along with the R2 and Q2 values for the original model, where the y-axis denotes the R2 and Q2 values and the x-axis denotes the correlation between the original values and permuted values. New R2 and Q2 values were calculated and plotted along with the original R2 and Q2 values. PLS permutation plots indicate that the above models were valid and not overfit (i.e. the models would be able to predict HbA1c levels for new spectra).

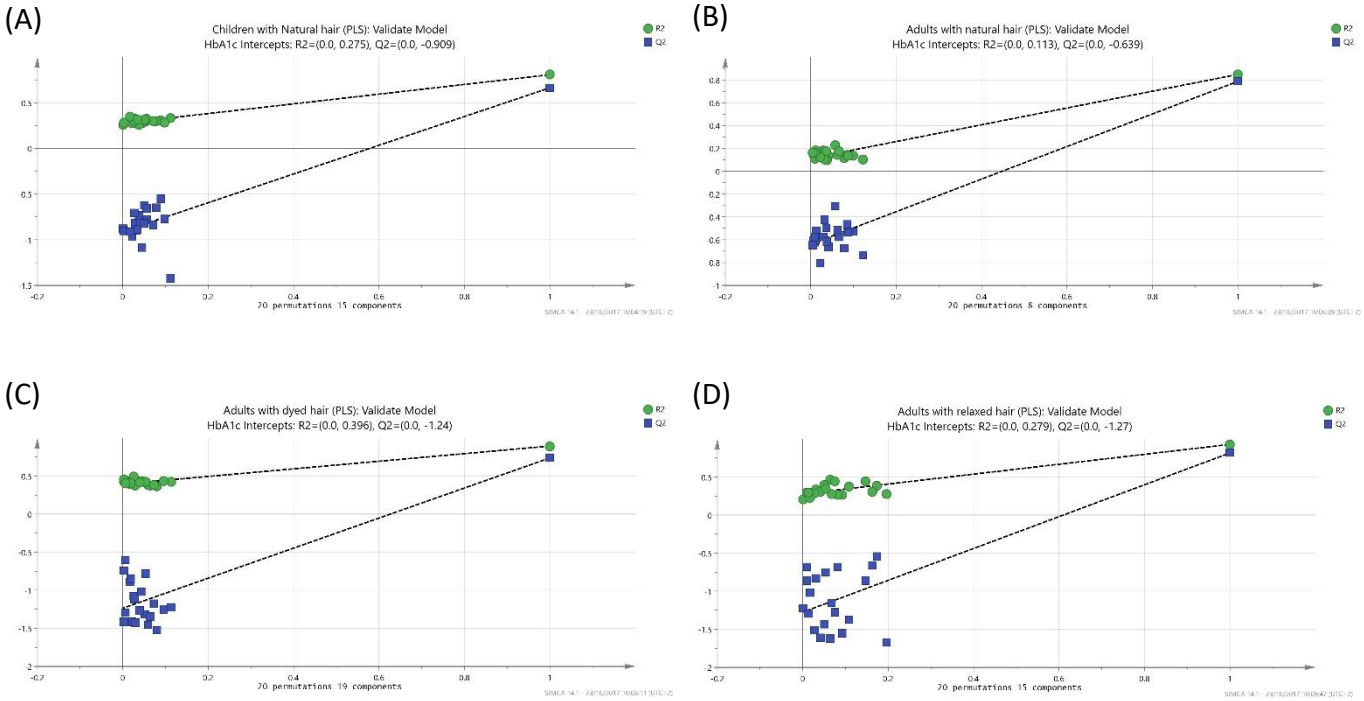
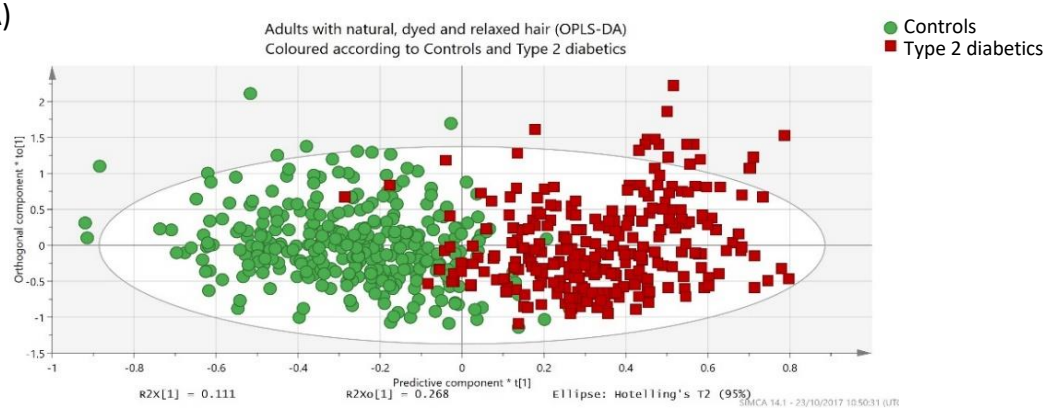


Figure 3.14. *Permutation plots for the validation of PLS models for controls vs. diabetics in the following groups:* ***(A)*** *Children with natural hair* ***(B)*** *Adults with natural hair* ***(C)*** *Adults with dyed hair* ***(D)*** *Adults with relaxed hair.*

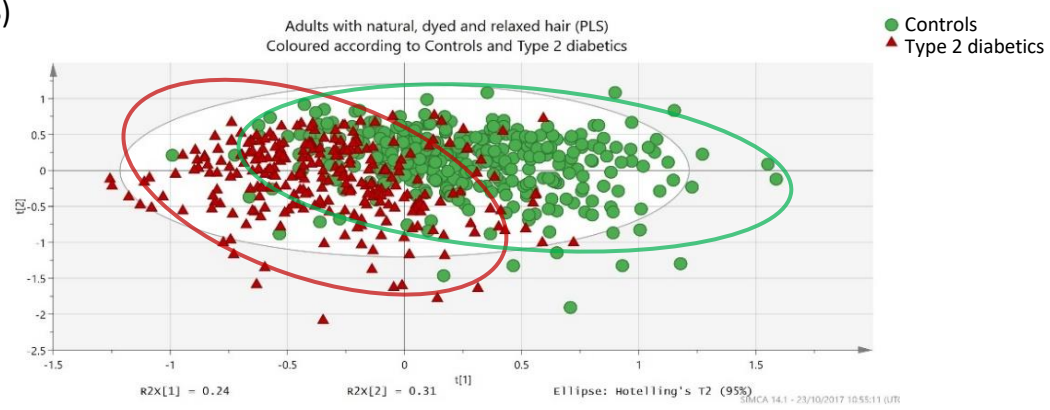
3.2.7. Confirming Classification of Hair into Chemical Treatment Groups is Necessary to Optimally Distinguish Between Hair from Diabetic and Control Groups and to Predict HbA1c Levels from Hair ATR-FTIR Spectra

OPLS-DA and PLS models were created including adults with natural, dyed and relaxed hair within the same model (Figure 3.15). The OPLS-DA model (A) showed less distinct separation between diabetic and control groups and had a lower predictive ability (Q2(cum)=0.675) than when the hair treatment groups were separated (Q2(cum)=0.819; Q2(cum)=0.736 and Q2(cum)=0.85 respectively). The PLS observed vs. predicted plot also had a weaker correlation (R^2=0.6165 compared to R^2=0.8481; R^2=0.895 and R^2=0.9296 respectively). These results emphasise that classification of hair into chemical treatment groups prior to the prediction of glycaemic status from FTIR spectra would be necessary.

(A)



(B)



(C)

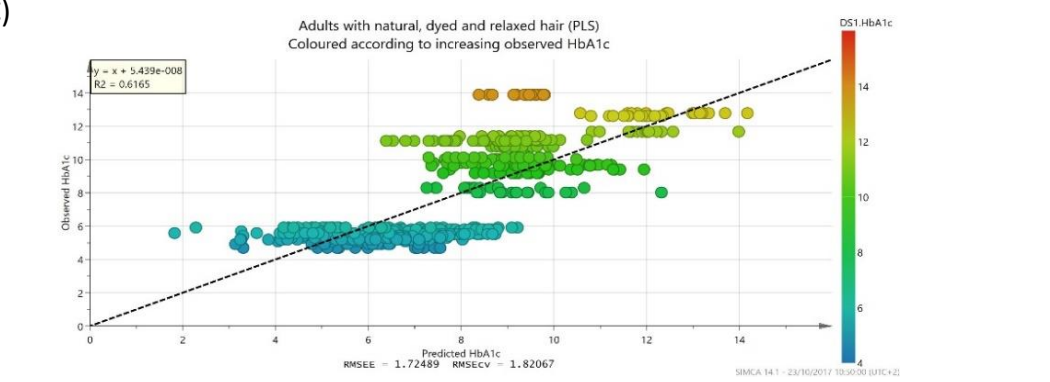


Figure 3.15. *OPLS-DA and PLS models including adults who have natural, dyed and relaxed hair* ***(A)*** *OPLS-DA plot of controls (green) vs. type 2 diabetics (red) where R2(cum)=0.701 and Q2(cum)=0.675* ***(B)*** *PLS scatter plot (Circles emphasise grouping of controls and diabetics) and* ***(C)*** *PLS observed vs. predicted HbA1c plot (R^2=0.6165).*

CHAPTER 4: COMPARATIVE AMINO ACID ANALYSIS OF HAIR FROM DIABETICS AND HEALTHY CONTROLS.

4.1. INTRODUCTION

It is known that circulating amino acid concentrations are altered in people with diabetes [31, 84] and that amino acids make up a large proportion of the chemical structure of hair (about 80%) [31]. Most recently, Rashaid *et al.* investigated amino acid concentrations in hair as a possible tool for diagnosing diabetes using GC-MS. The amino acids glycine (Gly), glutamic acid (Glu) and Isoleucine (Ile) were significantly more abundant in scalp hair of type 2 diabetics than healthy controls [31]. Since little research has been done on amino acid levels in hair of diabetics compared to controls, amino acid concentrations in the hair of participants were investigated. Amino acid concentrations of 17 amino acids, plus ammonia, were analysed in the hair of children with natural hair, adults with natural hair and adults with dyed hair.

4.2. RESULTS

4.2.1. Effect of Age and Hair Dye Treatment on Amino Acid Concentrations of Whole Human Hair Fibres

Before determining whether it was possible to distinguish between amino acid levels in hair of diabetics compared to healthy controls, it was important to consider any other factors (such as age and chemical hair treatment) which could contribute to observed differences and affect these results. The amino acid concentrations in hair of control children and adults were compared to determine if age was an important factor when measuring amino acid levels in hair. Most amino acid concentrations were within the same range or only slightly lower than what was reported in literature (probably due to incomplete digestion), leucine (Leu) and valine (Val) concentrations were however, not detected or were unusually lower than concentrations reported in literature (Table 4.1.). Of the eighteen amino acids analysed, twelve were significantly different between hair of adults and children. Aspartic acid (Asp), glutamic acid, proline (Pro), glycine, methionine (Met), isoleucine, histidine (His), lysine (Lys),

arginine (Arg) and ammonia (Amm) were all higher in hair of adults where cysteine (Cys) and leucine (Leu) were higher in hair of children (Figure 4.1. (A)).

The effect of treatment with hair dye on amino acid concentration was also determined by comparing natural and dyed hair of control adults (very few children has chemically treated hair). Eight amino acids were significantly different between natural and dyed hair. Alanine (Ala), Cys, Val, Met, Lys and Amm were higher in natural hair whereas serine (Ser) and tyrosine (Tyr) were higher in dyed hair (Figure 4.1. (B)). Table 4.1. summarizes these results showing confidence intervals for amino acid concentration in each group, including amino acids concentrations reported in literature by previous studies.

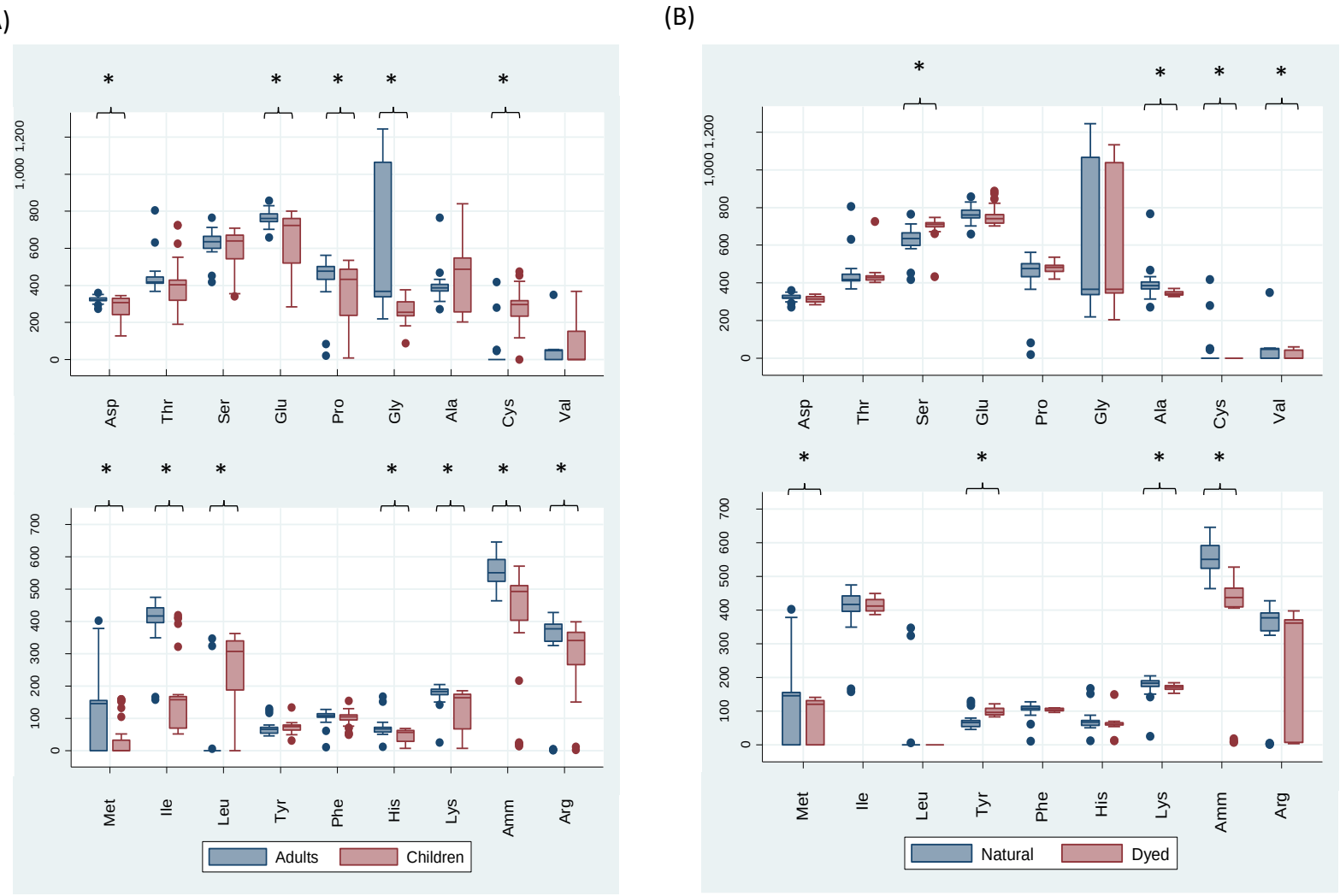


Figure 4.1. *Box and whisker plots showing min, 25th percentile, median 75th percentile and max concentrations for each amino acid (including outliers)* ***(A)*** *Amino acid concentrations in hair of adults vs. children (controls with natural hair)* ***(B)*** *Amino acid concentrations in natural vs. dyed hair (control adults). Asterisks show significant difference ($p<0.05$).*

Table 4.1. A summary of the confidence intervals for amino acid concentrations in hair of control children and adults (with natural and dyed hair) and concentrations in hair as reported in literature [72].

Amino Acid	Concentration (µmol/g hair) [95% confidence interval]			
	Control Children (w/ natural hair)	Control adults (w/ natural hair)	Control adults (w/ dyed hair)	As reported in literature (whole unaltered human hair)
Aspartic acid	273-306	315-332*	304-320	292-578
Threonine	347-427	407-479	411-469	558-714
Serine	571-636	602-660	669-722[a]	705-1013
Glutamic acid	610-703	745-782*	730-778	930-1036
Proline	289-405	398-495*	464-489	374-708
Glycine	250-287	493-793*	479-776	463-560
Alanine	376-487	363-434[a]	341-351	314-384
Cysteine	253-309*	0-66[a]	0	17-70
Valine	39-119	16-67[a]	11-30	470-513
Methionine	15-53	77-166*[a]	42-98	47-67
Isoleucine	128-226	345-439*	404-428	244-366
Leucine	207-285*	0-58	0	489-529
Tyrosine	65-79	61-78	94-104[a]	121-195
Phenylalanine	93-108	94-112	102-106	132-226
Histidine	38-52	59-81*	52-72	40-86
Lysine	104-147	163-189*[a]	167-174	130-222
Ammonia	356-472 ~8.5%	535-570*[a] ~9.9%	307-451 ~7.2%	14.5-15.9%
Arginine	272-333	270-375*	171-340	499-620
Total	4289-5433	4926-6176	4747-5708	-

*amino acid content significantly higher when comparing control adults and children with natural hair

[a] amino acid content significantly higher when comparing control adults with natural hair and control adults with dyed hair ($p<0.05$)

4.2.2. Difference in Amino Acid Concentrations in Hair of Diabetic Participants Compared to Healthy Controls

Due to previously observed differences, amino acid concentrations in control and diabetic participants were analysed separately for adults and children. Amino acid concentrations for adults with natural and with dyed hair were also analysed separately. Figure 4.2. (A) displays amino acid concentrations in children with natural hair showing 4 amino acids (Gly, Val, Met and Ile) which are significantly higher in the hair of diabetics compared to control children. In adults with natural hair Val, phenylalanine (Phe) and Amm were all lower in diabetics compared to controls (Figure 4.2. (B)) and in adults with dyed hair Val, Arg, and Amm were lower in diabetics compared to controls (Figure 4.2. (C)). Table 4.2. summarizes these results showing confidence intervals for amino acid concentration in each group.

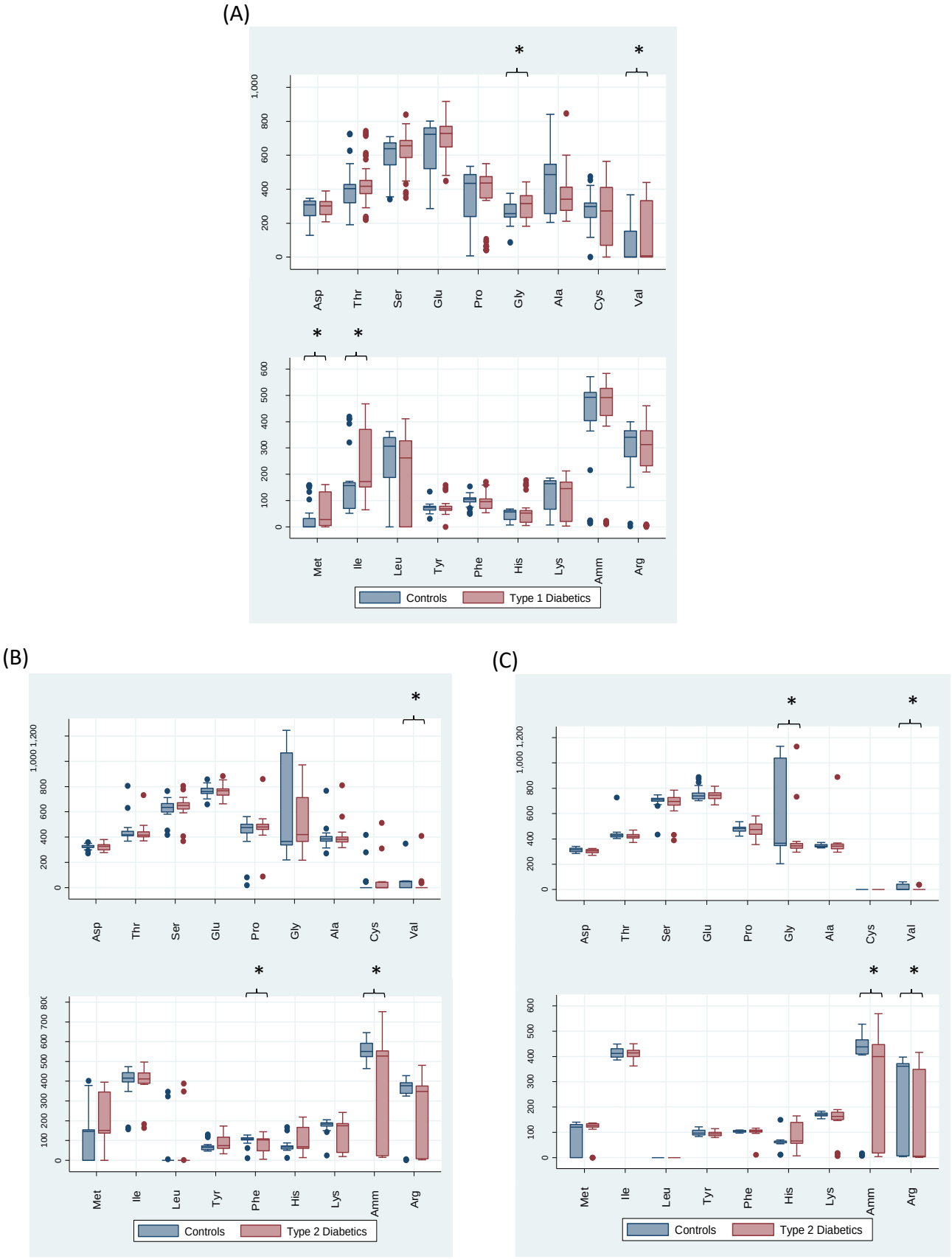


Figure 4.2. *Box and whisker plots showing min, 25th percentile, median 75th percentile and max concentrations for each amino acid (including outliers)* ***(A)*** *Amino acid concentrations in hair of control children vs. type 1 diabetics (controls with natural hair)* ***(B)*** *Amino acid concentrations in control adults vs. type 2 diabetics with natural hair.* ***(C)*** *Amino acid concentrations in control adults vs. type 2 diabetics with dyed hair. Asterisks show significant difference ($p<0.05$).*

Table 4.2. Table summarizing confidence intervals for amino acid concentrations in hair of diabetic participants compared to healthy controls.

Amino Acid	Concentration (µmol/g hair) [95% confidence interval]					
	Control children (w/ natural hair)	Type 1 diabetic children (w/natural hair)	Control adults (w/ natural hair)	Type 2 diabetic adults (w/ natural hair)	Control adults (w/ dyed hair)	Type 2 diabetic adults (w/ dyed hair)
Aspartic acid	273-306	284-308	315-332	314-337	304-320	293-312
Threonine	347-427	397-460	407-479	403-460	411-469	408-433
Serine	571-636	594-654	602-660	605-677	669-722	640-719
Glutamic acid	610-703	664-733	745-782	747-784	730-778	726-758
Proline	289-405	315-408	398-495	438-525	464-489	435-508
Glycine	250-287	282-322[b]	493-793	457-635	479-776[d]	316-466
Alanine	376-487	331-339	363-434	366-439	341-351	314-410
Cysteine	253-309	204-297	0-66	1-89	0	0
Valine	39-119	95-185[b]	16-67[c]	0-53	11-30[d]	0-8
Methionine	15-53	42-75[b]	77-166	141-247	42-98	93-130
Isoleucine	128-226	206-280[b]	345-439	336-442	404-428	401-421
Leucine	207-285	140-225	0-58	0-66	0	0
Tyrosine	65-79	65-88	61-78	70-99	94-104	90-100
Phenylalanine	93-108	87-102	94-112[c]	64-101	102-106	92-110
Histidine	38-52	40-64	59-81	82-129	52-72	68-107
Lysine	104-147	86-126	163-189	106-169	167-174	109-173
Ammonia	356-472 ~8.5%	384-477 ~8.7%	535-570[c] ~9.9%	298-502 ~7.5%	307-451[d] ~7.2%	172-356 ~5.2%
Arginine	272-333	222-301	270-375	160-311	171-340[d]	48-216
Total	4289-5433	4438-5504	4926-6176	4571-6065	4747-5708	4203-5226

[b] amino acid content significantly higher when comparing control and diabetic children

[c] amino acid content significantly higher when comparing control and diabetic adults with natural hair

[d] amino acid content significantly higher when comparing control and diabetic adults with dyed hair ($p<0.05$)

CHAPTER 5: GENERAL DISCUSSION

This study used ATR-FTIR spectroscopy and amino acid analysis to determine whether the scalp hair from people with chronic hyperglycaemia was distinguishable from hair of healthy controls. 1380 spectra were collected from hair of 92 participants and analysed using SIMCA software, looking for trends among the participants relating to aspects other than control and diabetic groups. OPLS-DA was used to distinguish between hair of diabetics and healthy controls. Each participants' HbA1c was also measured and PLS was used to determine whether there was a correlation between ATR-FTIR hair spectra and measured HbA1c levels to determine whether it was possible to predict a patient's HbA1c based on their hair spectra. Amino acid concentrations of the hair were then analysed in order to find any further distinguishing features between hair of diabetics and healthy controls.

5.1. INTER AND INTRA INDIVIDUAL VARIATION BETWEEN HAIR FTIR SPECTRA

Before attempting to discriminate between spectra taken from hair of diabetic participants and healthy controls, it was important to consider factors that may have resulted in spectral differences. Firstly, 6 randomly chosen individuals were used and OPLS-DA models were created to demonstrate the importance of taking multiple spectral measurements on three different strands for each participant in order to reduce the effects of intra-individual variation and optimise separation and predictive ability between individuals.

5.2. FACTORS AFFECTING SEPARATION OF HAIR FTIR SPECTRA

A crucial step in analysing ATR-FTIR data was to get an overview of the general trends within the entire dataset in order to identify potential factors which could have affected classification of FTIR spectra into diabetic and control groups. There was distinct separation between hair FTIR spectra from adults and children, which was even visible in the untrained PCA-X plot, meaning that these groups had significantly different spectral fingerprints. When comparing ATR-FTIR hair spectra of children to adults, the OPLS-DA loading plot displayed intensities which were lower at the position of amide A (3300cm^{-1}) and slightly lower around the position of amide II (1480-1575cm-1). It also displayed significantly increased intensities

from 1260-920cm^{-1}, wavenumbers which coincide with the Amide III peak (1230cm^{-1}) as well as cystine peaks (1040-1121cm-1) (Figure A2 (A) in appendix 3). A study by Kim *et al*., also found numerous significant differences in ATR-FTIR spectra among people of different ages. They reported a slight shift in position of the amide A, CH_3 and amide I peaks as well as change of intensity of the amide A, amide II and cysteine peaks [124]. It is well known that physical properties of hair change with age, particularly when a person goes through hormonal changes such as puberty. Post-pubertal hair is coarser and thicker than pre-pubertal hair, explaining the differences seen in the hair spectra with age [72].

The second largest factor affecting variation in hair FTIR spectra when looking at the general trends was chemical treatment, as OPLS-DA plots displayed grouping of hair spectra according to treatment type. The OPLS-DA loading graph which showed wavenumbers most important in discriminating between natural and dyed hair (Figure A2 (B) in appendix 3) displayed intense peaks at 1040cm^{-1} and 1175cm^{-1}. These peaks correspond to sulfonates, which are formed due to oxidation of cystine, and have been seen in multiple studies on the effects of oxidative treatments on human hair [101-104]. There was another intense peak at 1550cm^{-1}. This peak was also seen in the OPLS-DA loading plot showing wavenumbers most important for discrimination between natural and relaxed hair (Figure A2 (C) in appendix 3), along with a peak at 1655cm^{-1}. These peaks correspond to amino acid bands, amide I and amide II. Oxidative hair treatments, especially relaxer treatments such as guanidine hydroxide, cause amino acid chains to break producing amides which result in higher intensity FTIR peaks at these wavenumbers [104].

The above results demonstrate that, before attempting to classify hair spectra into ‘control’ or ‘diabetic’ groups or predicting HbA1c levels, hair from a patient with unknown glycaemic status would need to be classified into the above groups first. Figure 5.1. shows a diagram of an example of the workflow which would need to be followed in diagnosis and monitoring of hyperglycaemia using ATR-FTIR on hair.

5.3. USING HAIR FTIR SPECTRA TO DETECT HYPERGLYCAEMIA

Based on the preliminary data analysis considering age and hair treatment, ATR-FTIR spectra were divided and analysed in 4 groups: children with natural hair, adults with natural hair, adults with dyed hair and adults with relaxed hair. OPLS-DA was then used to discriminate between the spectral data from hair diabetics and healthy controls within the four groups. Each group showed excellent separation and predictive ability for controls vs. diabetics with Q2(cum) ranging from 0.753 to 0.85 and classification ability ranging from 97.97% to 100%. This demonstrates that it is possible to detect hyperglycaemia using hair using ATR-FTIR. Little research has been done on using this method to monitor hyperglycaemia. Zhang and Hassan found differences in hair spectra obtained from people with high blood glucose using the transmittance-FTIR technique [116] and Coopman *et al.*, were able to use FTIR to detect increased glycation in fingernail clippings (keratinized structures similar to hair) of people with diabetes [79]. None of these studies, however, employed multivariate data analysis to analyse their spectral data which, in this study, allowed for discrimination between diabetic and control groups via detection of changes in the spectra which would usually have been undetectable by the human eye.

The loading plots were evaluated to determine which wavenumbers in the spectra contributed to the separation between the groups. In each case there was a strong negative loading at the wavenumbers which correspond to the Amide III peak (1230cm^{-1}) meaning that there was lower absorbance in hair of diabetic participants in the regions of (C=O) & (C-N) stretching and (N-H) & (O=C=N) deformation. This means that there were changes in the masses of atoms and their geometries as well as force constants which correspond to the above bonds. Glycation is when a reactive carbohydrate molecule (such as glucose or fructose) covalently binds to the N-terminal of amino acids, resulting in an Amadori product containing a C-N bond between the amino acid and the carbohydrate, a new N-H bond on the terminal N as well as a new C=O bond within the attached carbohydrate [125] (refer to Figure 1.4.). This is strong evidence to suggest that increased glycation of the amino acids in hair of diabetic participants is what contributed most to the separation between the hair ATR-FTIR spectra of controls vs. diabetic participants – which corresponds with the hypothesis that amino acids in hair keratin become glycated as the fibre grows from the scalp due to the rich

blood supply at the root, and possibly also through sweat. This result also agrees with studies done in the 1980s and 1990s, which utilized wet chemistry methods in attempt to measure glycation levels in hair [28-30, 75, 76]. These studies all reported an increase in glycation of hair from diabetics compared to healthy controls, but only a weak positive correlation with patient HbA1c levels (probably due to the use of indirect methods to measure glycation or due to the sample preparation steps employed).

5.4. CORRELATION BETWEEN HAIR FTIR SPECTRA AND HBA1C LEVELS

Hair ATR-FTIR spectra were subjected to PLS analysis to determine whether there was a relationship between spectra and HbA1c levels measured for each participant. A different PLS model was created for each of the above-mentioned groups, matching the grouping used in OPLS-DA analysis. Plots showing observed HbA1c against HbA1c predicted using hair FTIR spectra displayed excellent correlation, where R^2 values ranged between 0.8067 and 0.9296. This showed there was a strong relationship between hair spectra and HbA1c readings, demonstrating the prospect to accurately predict a patient's long-term glycaemic status using their hair instead of blood.

It is already well known that structural proteins become excessively glycated in people with chronic hyperglycaemia, leading to the onset of numerous diabetic complications [125]. Multiple studies have reported the ability to detect glycation in keratinized proteins, such as hair and nails, in an attempt to find a non-invasive substrate for monitoring chronic hyperglycaemia [28-30, 76-79]. Evidence of glycation has also been found in hair which has been incubated in glucose [30, 126], meaning that glycation of hair happens spontaneously and non-enzymatically. In-vivo, this process most likely occurs due to contact with glucose in the blood as the hair grows and keratinizes at the root as well as through contact with glucose in sweat.

Predicted ranges in PLS models were, however large and overlapped considerably which emphasised the need to have at least five FTIR measurements on multiple strands of hair for

each patient so average predicted HbA1c can be calculated. Shimode *et al*., also reported higher concentration of AGEs on the distal or tip sides of hair fibres compared to proximal or root sides, probably due to UV exposure, which promotes free radical production and in turn AGE formation [126]. Therefore, further research will need to be done on how far along the hair fibre glycation levels correlate with HbA1c measurements.

5.5. PROSPECTS FOR USING ATR-FTIR TECHNOLOGY IN DIAGNOSING AND MONITORING DIABETES IN THE FUTURE

It is evident that it is possible to distinguish between the hair of diabetics and healthy controls using ATR-FTIR spectroscopy and that hair spectra correlate with patient HbA1c levels. These results are promising as they imply that it would be possible to diagnose diabetes as well as monitor the patients' glycaemic status using their hair. However, in order to obtain optimal separation between ATR-FTIR hair spectra from diabetics and controls in OPLS-DA models and optimal correlation between ATR-FTIR hair spectra and measured HbA1c readings in PLS models, each participant first had to be classified into groups according to their age and chemical hair treatment. Therefore, in order to use ATR-FTIR to detect and monitor diabetes, a patient would have to be classified, first as an adult or a child (under 18 years old) and then into a chemical hair treatment group (i.e., natural, dyed, relaxed etc.). This could be done either by observation or by OPLS-DA classification. Once a patient is classified as diabetic, PLS could then be used to monitor chronic glycaemic control i.e. HbA1c levels, more long-term and non-invasively (Figure 5.1.). Once a large enough database of hair spectra for each group is collected, calibration sets could be built into ATR-FTIR devices which are adapted to be handheld or point of care devices for detecting and monitoring hyperglycaemia.

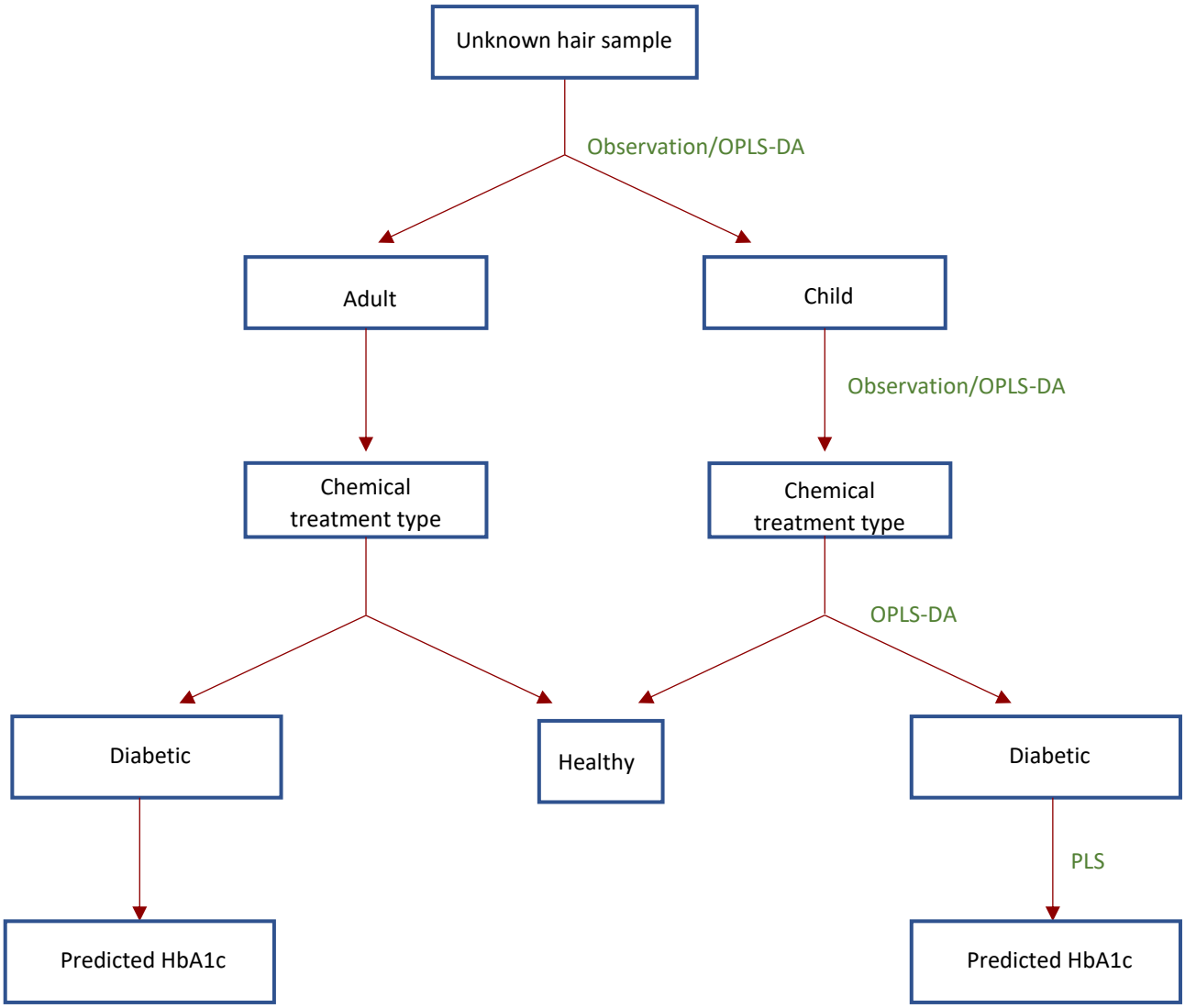


Figure 5.1. *An example of a recommended work flow when attempting to determine whether an individual is diabetic and predict HbA1c via ATR-FTIR analysis of scalp hair.*

5.6. COMPARTIVE ANALYSIS OF AMINO ACID LEVELS IN HAIR FROM DIABETICS AND HEALTHY CONTROLS

It is known that circulating amino acid concentrations are altered in people with diabetes with increased blood concentrations of branched chain amino acids: Leu, Ile and Val [31, 84, 127]. Since amino acids make up a large proportion of the chemical structure of hair (about 80%) [24] it is possible that circulating amino acid levels reflect in the growing hair strand. Hair amino acid concentrations were, therefore, determined in an attempt to discriminate between hair of diabetic participants and healthy controls, however, factors (such as age and chemical hair treatment) which may have affected these results were considered first.

Twelve out of eighteen amino acids were significantly different between hair from adults and children, which most likely explains the distinct separation observed between the ATR-FTIR spectra of these two groups, as amino acids are the building blocks of protein, which is the primary component of hair. Ten of these amino acids were significantly higher in adult hair where only leucine and cysteine seemed to be higher in hair of children. Since these two amino acids had significantly different concentrations in at least one of the groups than the values reported in literature this may not be a reliable result. Pre-pubertal hair is known as primary terminal hair which is known to be shorter (due to a shorter anagen phase) and finer compared to secondary terminal hair of adults which is longer and coarser with a larger maximum diameter containing more keratin than hair of children [72].

Eight amino acids had significantly different concentrations in dyed hair when compared to natural hair. This result, once again, supported the differences seen in hair FTIR spectra. Six of these amino acids were significantly lower in dyed hair, probably because hair dye (specifically permanent hair dye) may cause oxidative damage to proteins and breaking of amino acid side chains within the hair, leading to reduction in amino acid levels [72, 104]. This result also corresponds to the FTIR data, where the peaks which corresponded to the breaking of amino acid side chains had higher intensities in dyed hair compared to natural hair. The two amino acids which were increased in dyed hair were serine and tyrosine. This was probably due to oxidation of other amino acids within the hair as serine and tyrosine are

similar in structure to alanine and phenylalanine respectively, but with an additional OH group [72].

When comparing amino acid concentrations in hair of diabetic participants and healthy controls, there were a few slight differences. These differences did not coincide between groups, as the amino acids which were different between children with diabetes and controls were not the same as the differences seen between adults with diabetes and controls. Further research will have to be done, including adults with type 1 diabetes, to determine whether these differences were due to the type of diabetes or if the result was affected by the age of the participant.

Rashaid *et al*., reported increased Gly, Glu and Ile in hair of diabetics using GC-MS to quantify amino acid concentrations [31]. Using an amino acid analyser (HPLC), this study also found increased concentrations of Gly and Ile, but only in hair of diabetic children, not the adults. Val (known to be elevated in circulating blood of diabetics) was elevated in hair of all diabetic groups compared to healthy controls and circulating amino acids may be incorporated into the growing hair strand. It is, however, known that changes in circulating amino acids are caused by a vast number of factors including aspects as simple as menstrual cycle, diet, vitamin deficiency, infection and certain medications[127]. These factors (which have not been accounted for in this study) could also affect amino acid concentrations in hair. The discrimination between diabetic and control hair FTIR spectra and prediction of HbA1c levels from these spectra by MVA is most likely based on increased glycation of keratin proteins in hair of participants with hyperglycaemia and less likely to be based on changes in amino acid levels.

5.7. LIMITATIONS OF THE STUDY

This study did not include adults with type 1 diabetes and research would have to be done to determine whether all MVA models would yield the same results if adults with type 1 diabetes were to be included or if adults with type 1 diabetes and adults with type 2 diabetes would

have to be classified separately. It is unknown whether the results seen in hair ATR-FTIR spectra are only a consequence of hyperglycaemia (which is observed in all types of diabetes) or if the results would be affected by diabetic type.

All MVA models were internally validated using permutation plots which verified that the models were not overfit or spurious. In an ideal scenario, calibration models would be created and have been externally validated by a test set of new experiments made up of entirely new participants to ensure classification and 'biomarkers' which caused separation were the same. There were not enough participants in this study to split each group into calibration and test sets without the calibration sets becoming less effective or overfit. However, this study did show proof of concept that hair can be used to identify whether someone is diabetic or not and be used to predict long-term blood glucose levels when analysed with ATR-FTIR. Much larger calibration sets of individuals with known glycaemic status can be built in the future, including calibration sets for people with all different chemical treatment types so that test sets of people with unknown glycaemic status can be accurately predicted.

More research needs to be performed on amino acid levels in hair and whether factors that affect circulating amino acids (such as disease, diet, medication and infection) also affect hair amino acid concentrations. It is known that diabetics have increased circulating branched chain amino acids (leucine, isoleucine and valine) [127]. Rashaid *et al.*, found increased glycine, glutamic acid and isoleucine (one of which was a branched chain amino acid) in hair of adults with type 2 diabetes [31]. The results in this study did not fully agree with the result by Rashaid *et al.*, however, different techniques were also employed to analyse amino acid concentrations in each study and subsequent research would need to be done. There has been little research on hair amino acid content and there is a large variation of values that are reported in literature. Therefore, methods of amino acid analysis in hair which reduce destruction of amino acids need to be developed to accurately compare amino acid levels in hair off diabetics and healthy controls.

5.8. CONCLUSION

Diabetes mellitus is a rapidly growing global burden, which leads to debilitating complications if not closely monitored. One of the key problems with monitoring diabetes is lack of non-invasive methods of monitoring chronic hyperglycaemia, as HbA1c measurement requires blood to be drawn. An HbA1c reading also reflects only an average estimate of blood glucose levels for the past 8-12 weeks, is affected by multiple factors and is an expensive test to perform. There is a need for non-invasive means of monitoring chronic hyperglycaemia. This need, coupled with the growing interest in the use of hair as a diagnostic tool, has led to the idea that hair may be a suitable medium to monitor chronic hyperglycaemia. This study demonstrated the ability to distinguish between the hair of diabetics and controls and well as the ability to predict HbA1c levels using ATR-FTIR (R^2 ranging between 0.8067 and 0.9296). Further research needs to be done with regards to whether models can be externally validated and whether calibration sets can be built and used to predict whether a person with an unknown glycaemic status is diabetic as well as to predict their HbA1c level. However, the results of this study show proof of concept that hyperglycaemia can be monitored using a persons' hair through ATR-FTIR, if factors such as age and chemical treatment which affect the chemical properties, like amino acid levels, are considered. Hair grows at a constant rate (about 1cm per month depending on the degree of curl) and it is fairly inert once it has emerged from the scalp, meaning molecules can be stably incorporated into a growing strand. This means that hair may even be used as a timeline of glycaemic status for up to years previously, depending on the length of an individual's hair. This would lead to promising prospects for long term blood glucose monitoring and greater insights into the timing and development of diabetic complications. ATR-FTIR is inexpensive as it uses no additional consumables. This technology can, potentially, be adapted into a primary health point of care or home device for scanning and monitoring for long-term hyperglycaemia, which would assist in preventing the progression of debilitating complications as well as reduce the high proportion of undiagnosed diabetes due to easy, painless scanning.

REFERENCES

1. Zimmet, P., et al., *Diabetes mellitus statistics on prevalence and mortality: facts and fallacies.* Nat Rev Endocrinol, 2016. **12**(10): p. 616-22.
2. King, H., R.E. Aubert, and W.H. Herman, *Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections.* Diabetes Care, 1998. **21**(9): p. 1414-31.
3. Wild, S., et al., *Global prevalence of diabetes: estimates for the year 2000 and projections for 2030.* Diabetes Care, 2004. **27**(5): p. 1047-53.
4. Guariguata, L., et al., *Global estimates of diabetes prevalence for 2013 and projections for 2035.* Diabetes Res Clin Pract, 2014. **103**(2): p. 137-49.
5. Ogurtsova, K., et al., *IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040.* Diabetes Res Clin Pract, 2017. **128**: p. 40-50.
6. International Diabetes Federation., *IDF Diabetes Atlas, 8th edn.* Brussels, Belgium: International Diabetes Federation, 2017.
7. Aguiree, F., et al., *IDF Diabetes Atlas : sixth edition*. 6 ed. IDF Diabetes Atlas, ed. L. Guariguata, et al. 2013, Basel, Switzerland: International Diabetes Federation. 155.
8. American Diabetes Association., *Diagnosis and classification of diabetes mellitus.* Diabetes Care, 2010. **33 Suppl 1**: p. S62-9.
9. Forbes, J.M. and M.E. Cooper, *Mechanisms of diabetic complications.* Physiol Rev, 2013. **93**(1): p. 137-88.
10. National Institute for Health and Clinical Excellence., *Type 1 diabetes in adults: diagnosis and management.* 2015. **NICE guideline (NG17)**.
11. National Institute for Health and Clinical Excellence., *Type 2 diabetes in adults: management.* 2015. **NICE guideline (NG28)**.
12. National Institute for Health and Clinical Excellence., *Diabetes (type 1 and type 2) in children and young people: diagnosis and management.* 2015(NICE guideline (NG18)).
13. Amod A, M.A., Levitt N, Berg J, Young M, Grobler N, Heilbrunn A, Distiller L, Pirie F, Dave J, Huddle K, Jivan D, Paruk I, May W, Raal D,, A.-E.B. Blom D, Brown S, Mollentze W, Rheeder P, Tudhope L, Van Rensburgh G, Ganie Y, Carrihill M, Rauff S, Van Zyl D, Randeree, and K.D. H, Joshi P, Raubenheimer P, *The 2012 SEMDSA Guideline for the Management of Type 2 Diabetes* JEMDSA, 2012. **17**(1): p. S1-S94.
14. Zemlin, A.E., et al., *HbA1c of 6.5% to diagnose diabetes mellitus -- does it work for us? -- the Bellville South Africa study.* PLoS One, 2011. **6**(8): p. e22558.
15. Herman, W.H., et al., *Differences in A1C by race and ethnicity among patients with impaired glucose tolerance in the Diabetes Prevention Program.* Diabetes Care, 2007. **30**(10): p. 2453-7.
16. Kirk, J.K., et al., *Disparities in A1C levels between Hispanic and non-Hispanic white adults with diabetes: a meta-analysis.* Diabetes Care, 2008. **31**(2): p. 240-6.
17. Brown, A.F., et al., *Race, ethnicity, socioeconomic position, and quality of care for adults with diabetes enrolled in managed care: the Translating Research Into Action for Diabetes (TRIAD) study.* Diabetes Care, 2005. **28**(12): p. 2864-70.
18. Pragst, F. and M.A. Balikova, *State of the art in hair analysis for detection of drug and alcohol abuse.* Clin Chim Acta, 2006. **370**(1-2): p. 17-49.
19. Covaci, A., et al., *Hair analysis: another approach for the assessment of human exposure to selected persistent organochlorine pollutants.* Chemosphere, 2002. **46**(3): p. 413-418.
20. Wright, R.O., et al., *Neuropsychological correlates of hair arsenic, manganese, and cadmium levels in school-age children residing near a hazardous waste site.* Neurotoxicology, 2006. **27**(2): p. 210-6.
21. Gandhi, M., et al., *Atazanavir concentration in hair is the strongest predictor of outcomes on antiretroviral therapy.* Clin Infect Dis, 2011. **52**(10): p. 1267-75.

22. Gow, R., et al., *An assessment of cortisol analysis in hair and its clinical applications.* Forensic Sci Int, 2010. **196**(1-3): p. 32-7.
23. Russell, E., et al., *Hair cortisol as a biological marker of chronic stress: current status, future directions and unanswered questions.* Psychoneuroendocrinology, 2012. **37**(5): p. 589-601.
24. Schrader, J., M. Rothe, and F. Pragst, *Ethyl glucuronide concentrations in beard hair after a single alcohol dose: evidence for incorporation in hair root.* Int J Legal Med, 2012. **126**(5): p. 791-9.
25. Moyer, J., et al., *Correlation between sweat glucose and blood glucose in subjects with diabetes.* Diabetes technology & therapeutics, 2012. **14**(5): p. 398-402.
26. Gude, D., *Hair - a yardstick for diabetes.* Int J Trichology, 2011. **3**(2): p. 131.
27. James, V.J., D.K. Yue, and S.V. McLennan, *Changes in the molecular structure of hair in insulin-dependent diabetes.* Biochem Biophys Res Commun, 1997. **233**(1): p. 76-80.
28. Oimomi, M., et al., *Hair protein glycation as a long-term index of blood glucose in diabetics.* Diabetes Res Clin Pract, 1988. **5**(4): p. 305-8.
29. Oimomi, M., et al., *Increased fructose-lysine of hair protein in diabetic patients.* Klin Wochenschr, 1985. **63**(15): p. 728-30.
30. Paisey, R.B., et al., *Glycosylation of hair: possible measure of chronic hyperglycaemia.* Br Med J (Clin Res Ed), 1984. **288**(6418): p. 669-71.
31. Rashaid, A.H., B. Harrington Pde, and G.P. Jackson, *Profiling Amino Acids of Jordanian Scalp Hair as a Tool for Diabetes Mellitus Diagnosis: A Pilot Study.* Anal Chem, 2015. **87**(14): p. 7078-84.
32. Saltiel, A.R. and C.R. Kahn, *Insulin signalling and the regulation of glucose and lipid metabolism.* Nature, 2001. **414**(6865): p. 799-806.
33. Fu, Z., E.R. Gilbert, and D. Liu, *Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes.* Curr Diabetes Rev, 2013. **9**(1): p. 25-53.
34. Tang-Christensen, M., et al., *[Glucagon-like peptide 2, a neurotransmitter with a newly discovered role in the regulation of food ingestion].* Ugeskr Laeger, 2001. **163**(3): p. 287-91.
35. MacDonald, P.E., et al., *The multiple actions of GLP-1 on the process of glucose-stimulated insulin secretion.* Diabetes, 2002. **51 Suppl 3**: p. S434-42.
36. MacDonald, P.E., et al., *Inhibition of Kv2.1 voltage-dependent K+ channels in pancreatic beta-cells enhances glucose-dependent insulin secretion.* J Biol Chem, 2002. **277**(47): p. 44938-45.
37. Itoh, Y., et al., *Free fatty acids regulate insulin secretion from pancreatic beta cells through GPR40.* Nature, 2003. **422**(6928): p. 173-6.
38. Kemmler, W., et al., *On the biosynthesis, intracellular transport and mechanism of conversion of proinsulin to insulin and C-peptide.* Diabetes, 1972. **21**(2 Suppl): p. 572-81.
39. Kashyap, S., et al., *A sustained increase in plasma free fatty acids impairs insulin secretion in nondiabetic subjects genetically predisposed to develop type 2 diabetes.* Diabetes, 2003. **52**(10): p. 2461-74.
40. Olofsson, C.S., et al., *Fast insulin secretion reflects exocytosis of docked granules in mouse pancreatic B-cells.* Pflugers Arch, 2002. **444**(1-2): p. 43-51.
41. Cura, A.J. and A. Carruthers, *Role of monosaccharide transport proteins in carbohydrate assimilation, distribution, metabolism, and homeostasis.* Compr Physiol, 2012. **2**(2): p. 863-914.
42. Bogan, J.S., *Regulation of glucose transporter translocation in health and diabetes.* Annu Rev Biochem, 2012. **81**: p. 507-32.
43. Thorens, B. and M. Mueckler, *Glucose transporters in the 21st Century.* Am J Physiol Endocrinol Metab, 2010. **298**(2): p. E141-5.
44. Naftalin, R.J., *Reassessment of models of facilitated transport and cotransport.* J Membr Biol, 2010. **234**(2): p. 75-112.
45. Lund, S., et al., *Contraction Stimulates Translocation of Glut-4 in Skeletal-Muscle through a Mechanism Distinct from That of Insulin.* Diabetologia, 1995. **38**: p. A10-A10.

46. Yki-Jarvinen, H. and J. Westerbacka, *The fatty liver and insulin resistance.* Curr Mol Med, 2005. **5**(3): p. 287-95.
47. Singh, V.P., et al., *Advanced glycation end products and diabetic complications.* Korean J Physiol Pharmacol, 2014. **18**(1): p. 1-14.
48. Kitabchi, A.E., et al., *Hyperglycemic crises in adult patients with diabetes.* Diabetes Care, 2009. **32**(7): p. 1335-43.
49. Wolfsdorf, J., et al., *Diabetic ketoacidosis in infants, children, and adolescents: A consensus statement from the American Diabetes Association.* Diabetes Care, 2006. **29**(5): p. 1150-9.
50. Nishikawa, T., et al., *Normalizing mitochondrial superoxide production blocks three pathways of hyperglycaemic damage.* Nature, 2000. **404**(6779): p. 787-90.
51. Hammes, H.P., et al., *Differential accumulation of advanced glycation end products in the course of diabetic retinopathy.* Diabetologia, 1999. **42**(6): p. 728-36.
52. Hsieh, C.L., et al., *Kinetic analysis on the sensitivity of glucose- or glyoxal-induced LDL glycation to the inhibitory effect of Psidium guajava extract in a physiomimic system.* Biosystems, 2007. **88**(1-2): p. 92-100.
53. Kim, R., et al., *Role of the unfolded protein response in cell death.* Apoptosis, 2006. **11**(1): p. 5-13.
54. Hofmann, M.A., et al., *RAGE mediates a novel proinflammatory axis: a central cell surface receptor for S100/calgranulin polypeptides.* Cell, 1999. **97**(7): p. 889-901.
55. Hudson, B.I., et al., *Blockade of receptor for advanced glycation endproducts: a new target for therapeutic intervention in diabetic complications and inflammatory disorders.* Arch Biochem Biophys, 2003. **419**(1): p. 80-8.
56. Singh, R., et al., *Advanced glycation end-products: a review.* Diabetologia, 2001. **44**(2): p. 129-46.
57. Avery, N.C. and A.J. Bailey, *The effects of the Maillard reaction on the physical properties and cell interactions of collagen.* Pathol Biol (Paris), 2006. **54**(7): p. 387-95.
58. Sanguineti, R., et al., *Pentosidine effects on human osteoblasts in vitro.* Ann N Y Acad Sci, 2008. **1126**: p. 166-72.
59. Alikhani, M., et al., *Advanced glycation end products stimulate osteoblast apoptosis via the MAP kinase and cytosolic apoptotic pathways.* Bone, 2007. **40**(2): p. 345-53.
60. Committee of the Japan Diabetes Society on the Diagnostic Criteria of Diabetes, M., et al., *Report of the committee on the classification and diagnostic criteria of diabetes mellitus.* J Diabetes Investig, 2010. **1**(5): p. 212-28.
61. Marchetti, P., *Advanced glycation end products (AGEs) and their receptors (RAGEs) in diabetic vascular disease.* Medicographia, 2009. **31**(3): p. 257-265.
62. Bunn, H.F., K.H. Gabbay, and P.M. Gallop, *The glycosylation of hemoglobin: relevance to diabetes mellitus.* Science, 1978. **200**(4337): p. 21-7.
63. Little, R.R. and D.B. Sacks, *HbA1c: how do we measure it and what does it mean?* Curr Opin Endocrinol Diabetes Obes, 2009. **16**(2): p. 113-8.
64. Yadav, J., et al., *Prospects and limitations of non-invasive blood glucose monitoring using near-infrared spectroscopy.* Biomedical Signal Processing and Control, 2015. **18**: p. 214-227.
65. Sacks, D.B., *A1C versus glucose testing: a comparison.* Diabetes Care, 2011. **34**(2): p. 518-23.
66. Nathan, D.M., H. Turgeon, and S. Regan, *Relationship between glycated haemoglobin levels and mean glucose levels over time.* Diabetologia, 2007. **50**(11): p. 2239-44.
67. International Expert, C., *International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-34.
68. Roberts, W.L., et al., *Effects of hemoglobin C and S traits on glycohemoglobin measurements by eleven methods.* Clin Chem, 2005. **51**(4): p. 776-8.
69. Nitin, S., *HbA1c and factors other than diabetes mellitus affecting it.* Singapore Med J, 2010. **51**(8): p. 616-22.

70. Pani, L.N., et al., *Effect of aging on A1C levels in individuals without diabetes: evidence from the Framingham Offspring Study and the National Health and Nutrition Examination Survey 2001-2004.* Diabetes Care, 2008. **31**(10): p. 1991-6.
71. Buffoli, B., et al., *The human hair: from anatomy to physiology.* Int J Dermatol, 2014. **53**(3): p. 331-41.
72. Robbins, C.R. and SpringerLink (Online service), *Chemical and Physical Behavior of Human Hair*. 2012, Springer Berlin Heidelberg,: Berlin, Heidelberg. p. 1 online resource.
73. Hordinsky, M.K. and M. Ericson, *Hair innervation and vasculature.* Exp Dermatol, 1999. **8**(4): p. 314.
74. Gabbay, K.H., et al., *Glycosylated hemoglobins and long-term blood glucose control in diabetes mellitus.* J Clin Endocrinol Metab, 1977. **44**(5): p. 859-64.
75. Oimomi, M., et al., *Protein glycosylation as an indicator of blood glucose control using furosine determination.* Kobe J Med Sci, 1986. **32**(2): p. 71-6.
76. Kobayashi, K. and H. Igimi, *Glycation index of hair for non-invasive estimation of diabetic control.* Biol Pharm Bull, 1996. **19**(4): p. 487-90.
77. Kishabongo, A.S., et al., *Glycated nail proteins: a new approach for detecting diabetes in developing countries.* Trop Med Int Health, 2014. **19**(1): p. 58-64.
78. Kishabongo, A.S., et al., *Glycation of nail proteins: from basic biochemical findings to a representative marker for diabetic glycation-associated target organ damage.* PLoS One, 2015. **10**(3): p. e0120112.
79. Coopman, R., et al., *Glycation in human fingernail clippings using ATR-FTIR spectrometry, a new marker for the diagnosis and monitoring of diabetes mellitus.* Clin Biochem, 2017. **50**(1-2): p. 62-67.
80. Nissimov, J., et al., *Method for chronological recording of antigen appearance in human head-hair shafts and its use for monitoring glycation products in diabetes.* J Immunol Methods, 2007. **320**(1-2): p. 1-17.
81. Goldsmith, L.A., *Biochemistry of the skin: 50 years in The Journal of Investigative Dermatology.* J Invest Dermatol, 1989. **92**(4 Suppl): p. 113S-120S.
82. Ansari, N.A., Moinuddin, and R. Ali, *Glycated lysine residues: a marker for non-enzymatic protein glycation in age-related diseases.* Dis Markers, 2011. **30**(6): p. 317-24.
83. Skalnaya, M.G. and V.A. Demidov, *Hair trace element contents in women with obesity and type 2 diabetes.* J Trace Elem Med Biol, 2007. **21 Suppl 1**: p. 59-61.
84. Anuradha, C.V., *Aminoacid support in the prevention of diabetes and diabetic complications.* Curr Protein Pept Sci, 2009. **10**(1): p. 8-17.
85. Shimode A, Y.M., Naito J, *Hair protein glycation reduces hair breaking strength.* Glycative stress research, 2014. **1**(2): p. 37-45.
86. Binz, T.M., et al., *Rapid extraction, identification and quantification of oral hypoglycaemic drugs in serum and hair using LC-MS/MS.* Forensic Sci Int, 2012. **223**(1-3): p. 119-24.
87. Glassford, S.E., B. Byrne, and S.G. Kazarian, *Recent applications of ATR FTIR spectroscopy and imaging to proteins.* Biochim Biophys Acta, 2013. **1834**(12): p. 2849-58.
88. Lee, L.C., C.-Y. Liong, and A.A. Jemain, *A contemporary review on Data Preprocessing (DP) practice strategy in ATR-FTIR spectrum.* Chemometrics and Intelligent Laboratory Systems, 2017.
89. Brenner, L., et al., *A measurement of human hair oxidation by Fourier transform infrared spectroscopy.* Journal of Forensic Science, 1985. **30**(2): p. 420-426.
90. HILTERHAUS-BONG, S. and H. ZAHN, *Contributions to the chemistry of human hair. 1. Analyses of cystine, cysteine and cystine oxides in untreated human hair.* International journal of cosmetic science, 1987. **9**(3): p. 101-110.
91. Joy, M. and D. Lewis, *The use of Fourier transform infra-red spectroscopy in the study of the surface chemistry of hair fibres.* International journal of cosmetic science, 1991. **13**(5): p. 249-261.

92. Signori, V. and D. Lewis, *FTIR investigation of the damage produced on human hair by weathering and bleaching processes: implementation of different sampling techniques and data processing.* International journal of cosmetic science, 1997. **19**(1): p. 1-13.
93. Singh, B. and S. Umapathy, *Effect of SDS on human hair: Study on the molecular structure and morphology.* Journal of biophotonics, 2011. **4**(5): p. 315-323.
94. Zhou, Y., et al., *The effect of various cosmetic pretreatments on protecting hair from thermal damage by hot flat ironing.* Journal of cosmetic science, 2011. **62**(2): p. 265.
95. Kim, K.S., M.K. Shin, and H.K. Park, *Effects of ultraviolet B radiation on physicochemical properties of human hair shaft.* Microscopy research and technique, 2012. **75**(7): p. 949-954.
96. Richena, M. and C. Rezende, *Effect of photodamage on the outermost cuticle layer of human hair.* Journal of Photochemistry and Photobiology B: Biology, 2015. **153**: p. 296-304.
97. Mkentane, K., *Human scalp hair: geometry, biochemistry, growth parameters and mechanical characteristics*. 2016, University of Cape Town.
98. Ryu, S.R., et al., *FT-IR microspectroscopic imaging of cross-sectioned human hair during a bleaching process.* J Cosmet Dermatol Sci Applic, 2016. **6**: p. 181-190.
99. Coderch, L., et al., *Exogenous and endogenous lipids of human hair.* Skin Research and Technology, 2017.
100. Leite, M. and P. Maia Campos, *Mechanical characterization of curly hair: Influence of the use of nonconventional hair straightening treatments.* Skin Research and Technology, 2017.
101. Alter, H. and M. Bit-Alkhas, *Infrared analysis of oxidized keratin.* Textile Research Journal, 1969. **39**(5): p. 479-481.
102. Signori, V. and D.M. Lewis. *FTIR analysis of cysteic acid and cysteine-S-thiosulphate on untreated and bleached human hair*. in *Macromolecular Symposia*. 1997. Wiley Online Library.
103. Strassburger, J. and M.M. Breuer, *Quantitative Fourier transform infrared spectroscopy of oxidized hair.* J. Soc. Cosmet. Chem, 1985. **36**(1): p. 61-74.
104. Dankers, L.M., *Physical analysis of human hair.* 2007.
105. Chan, K.A., et al., *A novel approach for study of in situ diffusion in human hair using Fourier transform infrared spectroscopic imaging.* Applied spectroscopy, 2008. **62**(9): p. 1041-1044.
106. Hopkins, J., L. Brenner, and C. Tumosa, *Variation of the amide I and amide II peak absorbance ratio in human hair as measured by Fourier transform infrared spectroscopy.* Forensic science international, 1991. **50**(1): p. 61-65.
107. Barton, P.M.J., *A forensic investigation of single human hair fibres using FTIR-ATR spectroscopy and chemometrics*. 2011, Queensland University of Technology.
108. Manheim, J., et al., *Forensic Hair Differentiation Using Attenuated Total Reflection Fourier Transform Infrared (ATR FT-IR) Spectroscopy.* Applied spectroscopy, 2016. **70**(7): p. 1109-1117.
109. Mansilla, J., et al., *Archaeological and contemporary human hair composition and morphology.* Chungara, Revista de Antropología Chilena, 2011. **43**(2).
110. Fresnais, M., et al., *Recent advances in the characterization of hair of mummies from the Chilean Andean coast.* Forensic Sci Int, 2015. **249**: p. 25-34.
111. Lehtinen, J., et al., *Human Hair in the Identification of Cocaine Abuse with Cantilever-Enhanced Photoacoustic Spectroscopy and Principal Component Analysis.* Applied spectroscopy, 2013. **67**(8): p. 846-850.
112. Lin, S.-Y., et al., *Diagnosis of congenital hypothyroidism from human anagen scalp hair by infrared microspectroscopy.* Ultrastructural pathology, 2001. **25**(5): p. 357-360.
113. Lyman, D.J. and J. Murray-Wijelath, *Fourier transform infrared attenuated total reflection analysis of human hair: comparison of hair from breast cancer patients with hair from healthy subjects.* Applied spectroscopy, 2005. **59**(1): p. 26-32.
114. Lyman, D.J. and S.G. Fay, *The effect of breast cancer on the Fourier transform infrared attenuated total reflection spectra of human hair.* ecancermedicalscience, 2014. **8**.

115. Kamatchi, S., et al., *FTIR-ATR spectroscopic technique on human single intact hair fibre-A case study of thyroid patients.* International Journal of Advanced Scientific Technologies in Engineering and Management Sciences, 2016. **2**(5): p. 1-6.
116. ZHANG, L. and S. HASSAN, *Fourier transform infrared (FTIR) in diabetes research by analyzing hair samples.* Journal of the North Carolina Academy of Science, 2007: p. 163-166.
117. Wheelock, A.M. and C.E. Wheelock, *Trials and tribulations of 'omics data analysis: assessing quality of SIMCA-based multivariate models using examples from pulmonary medicine.* Mol Biosyst, 2013. **9**(11): p. 2589-96.
118. Wiklund, S., *Multivariate data analysis for Omics.* Umetrics, Umeå, Sweden, 2008.
119. Little, R.R., et al., *Effects of sample storage conditions on glycated hemoglobin measurement: evaluation of five different high performance liquid chromatography methods.* Diabetes Technol Ther, 2007. **9**(1): p. 36-42.
120. Umetrics, M., *Multivariate Data Analysis and Modelling Basic Course.* SIMCA-P and Multivariate Analysis, 2002.
121. Umetrics, M., *User guide to SIMCA.* Malmö (Sweden): MKS Umetrics AB, 2013.
122. Bylesjö, M., et al., *OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification.* Journal of Chemometrics, 2006. **20**(8-10): p. 341-351.
123. Coates, J., *Interpretation of infrared spectra, a practical approach.* Encyclopedia of analytical chemistry, 2000.
124. Kim, K.S. and H.K. Park, *Analysis of aging effects on chemical property of human hair by Fourier transform infrared spectroscopy.* Skin Research and Technology, 2013. **19**(1).
125. John, W.G. and E.J. Lamb, *The Maillard or browning reaction in diabetes.* Eye, 1993. **7**(2): p. 230.
126. Shimode, A., et al., *Hair protein glycation reduces hair breaking strength.* Glycative Stress Research, 2014. **1**: p. 37-45.
127. Shih, V.E., *Amino acid analysis*, in *Physician's guide to the laboratory diagnosis of metabolic diseases*. 2003, Springer. p. 11-26.

APPENDIX 1: METHODS

Table A1. Steps in modified Sodium High Performance (NaHP) program

Step	Time	Temp (℃)	Buffer	Pump (ml/h)	Ninhydrin (ml/h)	Commands
1	01:00	49	1	35	25	
2	00:00	49	1	35	25	Reset
3	01:00	49	1	35	25	Load
4	01:00	49	1	35	25	
5	00:00	49	1	35	25	Reset
6	06:00	49	1	35	25	
7	13:00	47	2	35	25	
8	00:30	47	5	35	25	
9	25:00	98	5	35	25	
10	04:00	98	6	35	25	
11	04:00	98	1	35	25	
12	02:00	50	0	OFF	OFF	
13	12:30	50	1	35	OFF	
14	02:00	49	1	35	25	

APPENDIX 2: RESULTS

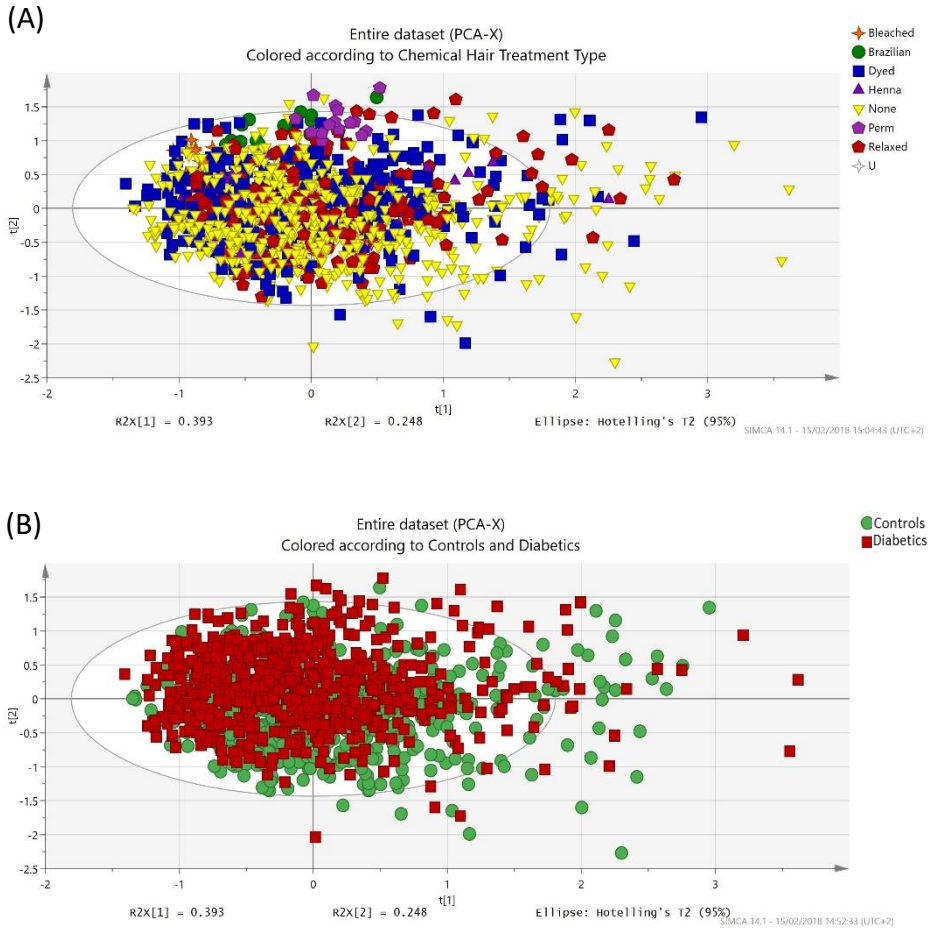


Figure A1. *PCA-X scores scatter plots displaying distribution of the entire dataset and coloured according to* ***(A)*** *Chemical hair treatment* ***(B)*** *Controls and diabetics.*

APPENDIX 3: GENERAL DISCUSSION

Figure A2. *OPLS-DA loading plot showing wavenumbers of spectra which contribute most to the separation between ATR_FTIR hair spectra of* ***(A)*** *Adults and children* ***(B)*** *Adults with natural hair and adults with dyed hair* ***(C)*** *Adults with natural hair and adults with relaxed hair. coloured according to p(corr) absorbance units.*

APPENDIX 4: ETHICS APPROVAL LETTER AS WELL AS EXAMPLES OF INFORMED CONSENT FORM, ASSENT FORM, PARENTAL CONSENT FORM

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee
Room E53-46 Old Main Building
Groote Schuur Hospltal
Observatory 7925
Telephone [021] 406 6492
Email: sumayah.ariefdien@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

02 August 2016

HREC REF: 450/2016

Prof N Khumalo
Division of Dermatology
G-23
NGSH

Dear Prof Khumalo

PROJECT TITLE: THE USE OF HAIR TO DETECT (AND MONITOR) CHRONIC HYPERGLYCAEMIA (MSc-candidate- Ms S da Silva)

Thank you for your response letter dated 29 July 2016, addressing the issues raised by the Human Research Ethics Committee (HREC).

It is a pleasure to inform you that the HREC has formally approved the above-mentioned study.

Approval is granted for one year until the 30 August 2017.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

Please quote the HREC REF in all your correspondence.

We acknowledge that the student, S da Silva will also be involved in this study.
Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator must obtain appropriate institutional approval before the research may OCCUr,

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON. FHS HUMAN RESEARCH ETHICS COMMITTEE

Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

HREC 450/2016

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DOH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI),and Declaration of Helsinki (2013) guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

The use of hair to detect (and monitor) chronic hyperglycaemia
Informed Consent form

Dear Volunteer

We, at the Division of Dermatology at the University of Cape Town have embarked on a study to determine whether it may be possible to detect longstanding high blood sugar levels using hair. The reason for this is that current methods for testing high blood sugar levels mostly include blood tests which are not accurate measures of long standing (chronic) high blood sugar. Using hair instead of blood in order to detect and monitor blood sugar levels would be less painful than blood testing, especially for small babies and elderly people, and the results would show blood sugar levels for months previously depending on the length of your hair. Since hair is not affected by conditions such as anaemia, kidney failure etc. it may be a better measure in these patients than measuring HbA1c levels, which is the current method for monitoring chronic high blood sugar. Hair generally grows about 1cm per month which may give doctors an indication as to when blood sugar levels were high (and to what extent) and in turn help to better prevent complications such as kidney failure, blindness and amputations associated with diabetes by better controlling blood sugar levels.
The reason a small amount of hair is needed from patients with diabetes and from healthy controls is that diabetes causes high blood sugar levels and if we can find a difference between hair from people with diabetes compared to hair from people with normal blood sugar levels, we may be able to eventually find a definitive test for high blood sugar using hair alone.
A small amount of finger prick blood will also be required in order to test for HbA1c levels (average blood sugar levels for the past 3 months) and correlate this with data found in hair for each participant.

Do I have to take part? No. Whether you agree to take part or not, or even change your mind after starting in the study; the health care you receive will not be affected.
Will I receive payment for taking part? No there is no payment for participating in the study as you will be seen on your normal clinic visit.
What will I benefit by participating? There will be no immediate benefit. However, confirmation that testing hair for high sugar levels is reliable could in future help doctors and patients gain insight on how well the sugar control is not just on the day but over previous

months depending on hair length. Such a test could encourage lifestyle changes and treatment adjustment to reduce the risk of complications.

What's involved? On your regular clinic visit, a small (invisible and painless) hair sample will be taken from the crown of your head (or in the case of bald men - from the underarm). In the case of diabetic participants clinical information from your hospital records such as HbA1c, blood sugar levels, development of diabetic complications, type of diabetes etc. may be used for the purpose of the study. A small prick will also be made on your finger in order to obtain a drop of blood which will allow us to measure HbA1c levels.

We only need 10 minutes of your time for this study (this includes collecting hair and finger prick samples) and can be before or after your routine visit at the clinic. Minors involved in this study should be between the ages of 10 and 18 and will need parental consent in order to take part.

Will my taking part be kept confidential? Yes, only the nurses and doctors conducting the study will know your personal information which will be kept in a password protected computer. Your hair samples will be stored securely and only be used for this study; no other tests will be conducted without your permission.

We wish to invite you to be part of this study. If you have any queries or questions, please do not hesitate to contact the numbers listed below.

Yours Sincerely,
Professor Nonhlanhla P Khumalo,
Division of Dermatology
New Groote Schuur Hospital
Observatory 7925, Cape Town

The use of hair to detect (and monitor) chronic hyperglycaemia

I (full name)……
understand the project and hereby give consent to the investigators hair and finger prick samples to be taken and used scientifically for the purpose of furthering knowledge in the testing and monitoring of hyperglycaemia. I understand that all my details will remain completely confidential.

I further understand that any results obtained will be entirely confidential.

I have read all of the above and anything I have not understood has been fully explained to me. I also understand that I may withdraw my participation at any stage from the study without repercussion.

Volunteer's signature:……Date:…….

Place:…….

Witness (full name):……

Witness's signature:…… Date:…….

This study has been reviewed by the University of Cape Town's Research Committee and the research committee can be contacted on the following contact details:

Human Research Ethics Committee
Faculty of Health Science Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory, 7925
Email: sumayah.ariefdien@uct.ac.za

If I have questions, I may contact Sr Anthea Ndyenga at Tel: 021 4045269 or 4043376 Cell: 0761036437
OR – UCT's Faculty of Health Sciences, Human Research Ethics Committee at: 0214066338
OR – UCT's Faculty of Health Sciences, Human Research Ethics Committee at: 0214066338

The use of hair to detect (and monitor) chronic hyperglycaemia
Assent form - Minors

Dear Volunteer

We, at the Division of Dermatology at the University of Cape Town have started a study to see if it's possible to detect long term high blood sugar levels using hair. The reason for this is that the methods at the moment for testing high blood sugar levels mostly involve blood tests which do not perfectly measure long term high blood sugar. If we could use hair instead of blood to look at blood sugar levels, it would be less painful than blood testing and the results would show blood sugar levels from long ago depending on the length of your hair. Your hair grows about 1cm per month which might show doctors when someone's blood sugar levels were high and how high it was. This can help doctors to stop the diabetes making people sick. The reason a small amount of hair is needed from people with diabetes and from healthy people is that diabetes causes high blood sugar levels and if we can find a difference between hair from people with diabetes compared to hair from people with normal blood sugar levels, we may be able to eventually find a test for high blood sugar using only hair.
We also need to make a small prick on your finger so we can get a little bit of blood so we can see if the sugar levels in your blood matches

Do I have to take part? No. You can decide if you want to take part or even change your mind later if you want, the health care you receive will not be affected.
Will I receive payment for taking part? No there is no payment for participating in the study as you will be seen on your normal clinic visit.
What will I benefit by participating? There will be no benefit straight away, but if we can prove that it is possible test hair for sugar levels it will help doctors and patients on how well people are controlling their sugar levels over a long time. This could help the doctors to help patients to not get sick.
What's involved? When you come in for your clinic visit, a small bit of hair will be taken from the back of your head. It doesn't hurt and you won't be able to see where there hair was taken. If you are diabetic, information from your hospital records such as HbA1c, blood sugar levels and the type of diabetes you have etc. may be used in this study. We will also need to make a small prick on your finger so we can get a drop of blood so we can measure you HbA1c levels (which shows your average blood sugar levels over the past 3 months)

We only need 10 minutes of your time for this study (this includes collecting hair and finger prick samples) and can be before or after your visit at the clinic.
Who will know my personal information? Only the nurses and doctors will know your personal information which will be kept in a password protected computer. Your hair and blood samples will be stored safely and only be used for this study; no other tests will be done with the samples you gave us without your without your permission.

We wish to invite you to be part of this study. If you have any questions, please do not hesitate to contact the numbers listed below.

Yours Sincerely,
Professor Nonhlanhla P Khumalo,
Division of Dermatology
New Groote Schuur Hospital
Observatory 7925, Cape Town

The use of hair to detect (and monitor) chronic hyperglycaemia

Assent form

I (full name of minor)……
understand the project and hereby give assent to the investigators for hair and finger prick samples to be taken and used scientifically to help increase the knowledge in the testing and monitoring of high blood sugar. I understand that all my details will remain completely private.

I also understand that any results from the study will be private.

I have read all of the above and anything I have not understood has been explained to me. I also understand that I may change my mind about being part of this study at any stage from the study without consequences.

Volunteer *(minor's*) signature:……Date:…….

Place:…….

Witness (full name):……

Witness's signature:…… Date:…….

Human Research Ethics Committee
Faculty of Health Science Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory, 7925
Email: sumayah.ariefdien@uct.ac.za

If I have questions, I may contact Sr Anthea Ndyenga at Tel: 021 4045269 or 4043376 Cell: 0761036437
OR – UCT's Faculty of Health Sciences, Human Research Ethics Committee at: 0214066338
OR – UCT's Faculty of Health Sciences, Human Research Ethics Committee at: 0214066338

The use of hair to detect (and monitor) chronic hyperglycaemia
Informed parent/guardian consent form for minors between 10 and 17.

Dear Parent/Gaurdian

We, at the Division of Dermatology at the University of Cape Town have embarked on a study to determine whether it may be possible to detect longstanding high blood sugar levels using hair. The reason for this is that current methods for testing high blood sugar levels mostly include blood tests which are not accurate measures of long standing (chronic) high blood sugar. Using hair instead of blood in order to detect and monitor blood sugar levels would be less painful than blood testing, especially for small babies and elderly people, and the results would show blood sugar levels for months previously depending on the length of the hair. Since hair is not affected by conditions such as anaemia, kidney failure etc. it may be a better measure in these patients than measuring HbA1c levels, which is the current method for monitoring chronic high blood sugar. Hair generally grows about 1cm per month which may give doctors an indication as to when blood sugar levels were high (and to what extent) and in turn help to better prevent complications such as kidney failure, blindness and amputations associated with diabetes by better controlling blood sugar levels.

The reason a small amount of hair is needed from patients with diabetes as well as healthy controls is that diabetes causes high blood sugar levels and if we can find a difference between hair from people with diabetes compared to hair from people with normal blood sugar levels, we may be able to eventually find a definitive test for high blood sugar using hair alone.

The reason children's hair is needed for this study is that children are unlikely to have any other illnesses, ageing effects or hair treatments which may affect result of the project. It is also important to include both Type 1 and Type 2 diabetic patients in this study in order to ensure reliability of results for different clinical settings. Type 2 diabetes is, by far, the most common Type in adults and so including children and adults in the study ensures investigation of both phenotypes of diabetes.

A small amount of finger prick blood will also be required in order to test for HbA1c levels (average blood sugar levels for the past 3 months) and correlate this with data found in hair for each participant.

Does my child have to take part? No. Whether he/she agrees to take part or not, or even changes his/her mind after starting in the study; the health care he/she receives will not be affected.

Will I receive payment if my child takes part? No there is no payment for participating in the study as your child will be seen on their normal clinic visit.

What will my child benefit by participating? There will be no immediate benefit. However, confirmation that testing hair for high sugar levels is reliable could in future help doctors and patients gain insight on how well the sugar control is not just on the day but over previous months depending on hair length. Such a test could encourage lifestyle changes and treatment adjustment to reduce the risk of complications.

What's involved? On your child's regular clinic visit, a small (invisible and painless) hair sample will be taken from the crown of their head. In the case of diabetic patients, clinical information from your child's hospital records such as HbA1c, blood sugar levels, development of diabetic complications, type of diabetes etc. may be used for the purpose of the study. A small prick will also be made on your child's finger in order to obtain a drop of blood which will allow us to measure HbA1c levels.

We only need 10 minutes of you and your child's time for this study (this includes collecting hair and finger prick samples) and can be before or after your child's routine visit at the clinic. Minors involved in this study should be between the ages of 10 and 18.

Will my child's taking part be kept confidential? Yes, only the nurses and doctors conducting the study will know your child's personal information which will be kept in a password protected computer. Your child's hair samples will be stored securely and only be used for this study; no other tests will be conducted without your permission.

We wish to invite your child to be part of this study. If you have any queries or questions, please do not hesitate to contact the numbers listed below.

Yours Sincerely,

Professor Nonhlanhla P Khumalo,
Division of Dermatology
New Groote Schuur Hospital
Observatory 7925, Cape Town

The use of hair to detect (and monitor) chronic hyperglycaemia

Parental consent form

I (full name)…… am the parent/legal guardian of (full name of minor)……
I understand the project and hereby give consent to the investigators to approach my child to ask for assent for hair and finger prick samples to be taken and used scientifically for the purpose of furthering knowledge in the testing and monitoring of hyperglycaemia. I understand that all his/her details will remain completely confidential.

I further understand that any results obtained will be entirely confidential.

I have read all of the above and anything I have not understood has been fully explained to me. I also understand that I may withdraw my participation at any stage from the study without repercussion.

Parent/Gaurdian's
signature:……Date:……
Place:……

Witness (full name):……

Witness's signature:…… Date:…….

This study has been reviewed by the University of Cape Town's Research Committee and the research committee can be contacted on the following contact details:

Human Research Ethics Committee
Faculty of Health Science Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory, 7925
Email: sumayah.ariefdien@uct.ac.za

If I have questions, I may contact Sr Anthea Ndyenga at Tel: 021 4045269 or 4043376 Cell: 0761036437

OR – UCT's Faculty of Health Sciences, Human Research Ethics Committee at: 0214066338

OR – UCT's Faculty of Health Sciences, Human Research Ethics Committee at: 0214066338