UNIVERSITY OF
LINCOLN

Early Screening and Diagnosis of Diabetic Retinopathy

G Leontidis

Doctor of Philosophy

2016

Early Screening and Diagnosis of Diabetic Retinopathy

Georgios Leontidis

A thesis submitted in partial fulfilment of the requirements of the University of Lincoln for the degree of Doctor of Philosophy

October 2016

Abstract

Diabetic retinopathy (DR) is a chronic, progressive and possibly vision-threatening eye disease. Early detection and diagnosis of DR, prior to the development of any lesions, is paramount for more efficiently dealing with it and managing its consequences. This thesis investigates and proposes a number of candidate geometric and haemodynamic biomarkers, derived from fundus images of the retinal vasculature, which can be reliably utilised for identifying the progression from diabetes to DR. Numerous studies exist in literature that investigate only some of these biomarkers in independent normal, diabetic and DR cohorts. However, none exist, to the best of my knowledge, that investigates more than 100 biomarkers altogether, both geometric and haemodynamic ones, for identifying the progression to DR, by also using a novel experimental design, where the same exact matched junctions and subjects are evaluated in a four year period that includes the last three years pre-DR (still diabetic eye) and the onset of DR (progressors' group). Multiple additional conventional experimental designs, such as non-matched junctions, non-progressors' group, and a combination of them are also adopted in order to present the superiority of this type of analysis for retinal features. Therefore, this thesis aims to present a complete framework and some novel knowledge, based on statistical analysis, feature selection processes and classification models, so as to provide robust, rigorous and meaningful statistical inferences, alongside efficient feature subsets that can identify the stages of the progression. In addition, a new and improved method for more accurately summarising the calibres of the retinal vessel trunks is also presented.

The first original contribution of this thesis is that a series of haemodynamic features (blood flow rate, blood flow velocity, etc.), which are estimated from the retinal vascular geometry based on some boundary conditions, are applied to studying the progression from diabetes to DR. These features are found to undoubtedly contribute to the inferences and the understanding of the progression, yielding significant results,

mainly for the venular network.

The second major contribution is the proposed framework and the experimental design for more accurately and efficiently studying and quantifying the vascular alterations that occur during the progression to DR and that can be safely attributed only to this progression. The combination of the framework and the experimental design lead to more sound and concrete inferences, providing a set of features, such as the central retinal artery and vein equivalent, fractal dimension, blood flow rate, etc., that are indeed biomarkers of progression to DR.

The third major contribution of this work is the new and improved method for more accurately summarising the calibre of an arterial or venular trunk, with a direct application to estimating the central retinal artery equivalent (CRAE), the central retinal vein equivalent (CRVE) and their quotient, the arteriovenous ratio (AVR). Finally, the improved method is shown to truly make a notable difference in the estimations, when compared to the established alternative method in literature, with an improvement between 0.24% and 0.49% in terms of the mean absolute percentage error and 0.013 in the area under the curve.

I have demonstrated that some thoroughly planned experimental studies based on a comprehensive framework, which combines image processing algorithms, statistical and classification models, feature selection processes, and robust haemodynamic and geometric features, extracted from the retinal vasculature (as a whole and from specific areas of interest), provide altogether succinct evidence that the early detection of the progression from diabetes to DR can be indeed achieved. The performance that the eight different classification combinations achieved in terms of the area under the curve varied from 0.745 to 0.968.

Acknowledgements

First and foremost, I shall pay tribute to my two supervisors, professor Andrew Hunter and Dr Bashir Al-Diri, who both made the work of this Ph.D thesis possible, with their constant support, guidance and patience that they showed towards me. My respect to them is unconditional and cannot be expressed with plain words. Throughout this project they were always eager to answer any questions and provide the necessary advice in order to overcome the small problems that always existed. I definitely owe them a lot and I would definitely have not arrived in this level if it were not for them.

In addition, I need to thank my friends and collaborators at the School of Information Engineering at the University of Padova in Italy, Jeffrey Wigdahl and Alfredo Ruggeri, with whom I had a great time working together and they were also eager to help me with providing their expertise in some technical parts. The same goes to my friend and colleague Francesco Calivá, with whom I had the pleasure to also collaborate and work for accomplishing some parts.

Moreover, the financial support of the European Commission through the Marie-Skłodowska Curie Actions was paramount for even starting this project, and made sure to receive all the necessary trainings and conference attendances that is so important for building a complete scientist.

Last but not least, my gratitude also has to go to my wife Francesca Soldati and my daughter Savina, who both made sure to cheer me up and encourage me to arrive at the end of this difficult journey.

Declaration of Authorship

I, Georgios Leontidis, declare that this thesis titled, “Early screening and diagnosis of diabetic retinopathy” and the work presented in it are my own. I confirm that:

- This work was wholly done while in candidature for a research degree at the University of Lincoln.
- No part of this thesis has ever been submitted for any other degree or qualification in this or any other University.
- I have acknowledged all the sources of help and also every source is given.
- Where the work is based on work done by myself jointly with others, I have made clear what my exact contribution is.
- Finally, this thesis is an original work, except where otherwise indicated.

Contents

List of Figures

List of Tables

Abbreviations

AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
AUC	Area Under the Curve
AVR	Arteriovenous Ratio
BC	Branching Coefficient
BIC	Bayesian Information Criterion
CRA	Central Retinal Artery
CRAE	Central Retinal Arterty Equivalent
CRV	Central Retinal Vein
CRVE	Central Retinal Vein Equivalent
DR	Diabetic Retinopathy
FD	Fractal Dimension
FOV	Field of View
FNR	False Negative Rate
FPR	False Postive Rate
ICC	Intraclass Correlation
LASSO	Least Absolute Shrinkage and Selection Operator
LDV	Laser Doppler Velocimetry
LMM	Linear Mixed Model
LRT	Likelihood Ratio Test

ML	Maximum Likelihood
MSA	Maximum Shadow Attribute
OD	Optic Disc
OOB	Out of Bag
PDR	Proliferative Diabetic Retinopathy
RANSAC	Random Sample Consensus
REML	Restricted Maximum Likelihood
RI	Resistance Index
ROC	Receiver Operating Characteristic
RRF	Regularised Random Forests
SD	Standard Deviation
SE	Standard Error
SIFT	Scale Invariant Feature Transform
TNR	True Negative Rate
TPR	True Positive Rate
UK	United Kingdom
VEGF	Vascular Endothelial Growth Factor
WHO	World Health Organisation

Chapter 1

Introduction

This thesis presents a framework for the study of multiple geometric and haemodynamic biomarkers of progression from diabetes to diabetic retinopathy (DR), derived from the retinal vasculature. In addition, an improved estimation of the calibres of the parent vessel trunks with an application to the estimation of the central retinal artery equivalent(CRAE), the central retinal vein equivalent (CRVE) and their quotient, arteriovenous ratio (AVR), is also presented. All of the above studies and implementations aim to identify whether early alterations are indeed present in the retinal vasculature prior to the onset of diabetic retinopathy (DR) and the appearance of any lesions, when still dealing with a diabetic eye. This is paramount, given that the earlier a disease is identified, the more efficient the treatment plan can be (WHO, 2016; AmDiabAss *et al.*, 2010).

DR is a vision-threatening eye disease that affects the retinal blood vessels and is primarily a consequence of the systemic disease of diabetes (WHO, 2016; Group *et al.*, 2004). Diabetes is a chronic disease that occurs when the pancreas does not produce adequate insulin, or when the body cannot actually make use of the insulin it produces. Insulin is an important hormone that regulates the blood sugar (WHO, 2016; Alberti & Zimmet, 1998). Uncontrolled diabetes leads to hyperglycaemia, which is

the raise of the blood sugar above a normal level, and over a period of time can cause severe damage to various body's systems, including kidneys, nerves and blood vessels (Creager *et al.*, 2003). There are two types of diabetes, Type 1 and Type 2. Type 1 diabetes, previously known as insulin-dependent childhood-onset is not preventable with today's knowledge and also requires a daily administration of insulin (Alberti & Zimmet, 1998). Type 2 diabetes, formerly known as non-insulin dependent or adult-onset is a consequence of the body's incapability of effectively using insulin. The majority of people that suffer from diabetes fall into this category and is linked to the lifestyle (body weight, physical inactivity, etc.). Recently this type of diabetes is also found in children (WHO, 2016; Alberti & Zimmet, 1998).

Some important epidemiological facts from the world health organisation (WHO) suggest that between 1980 and 2014 the number of people with diabetes has risen from 108 million to a massive 422 million. Similarly the prevalence of diabetes has risen from 4.7% to 8.5% during the above same period. In 2012, 1.5 million deaths were directly associated with diabetes and another 2.2 million deaths were linked to high blood glucose levels. Moreover, 2.6% of global blindness can be attributed to diabetes (WHO, 2016; Alberti & Zimmet, 1998).

DR is one of the most common causes of preventable blindness and of moderate to severe visual impairment (Bourne *et al.*, 2013). According to 2010 estimations, more than a third of diabetic patients have signs of DR and a third of those suffer from vision-threatening DR (severe non-proliferative DR, or proliferative DR (PDR)), or presents signs of diabetic macular oedema (Lee *et al.*, 2015b).

1.1 Motivation and Objectives

It is imperative to point out, that throughout this thesis, the purpose is to use and combine state of the art computer vision, image processing, machine learning and statistical

analysis techniques, in order to conduct a comprehensive, detailed, robust and complete study, with scientific, technical and clinical importance.

More specifically, the main aim and novelty of this thesis is to utilise and develop different tools, in the areas of image processing, statistical analysis and machine learning, in order to thoroughly investigate the alterations that occur to the retinal vasculature during the progression from diabetes to DR, and also present a framework that accomplishes that. In literature, there is a vast and endless number of implemented algorithms for segmenting the retinal vessels (Oliveira *et al.*, 2016; Li *et al.*, 2016; Christodoulidis *et al.*, 2016; Soares *et al.*, 2006; Staal *et al.*, 2004; Al-Diri *et al.*, 2009) and/or extracting different kind of vascular measurements (Lowell *et al.*, 2004b; Perez-Rovira *et al.*, 2011; Al-Diri *et al.*, 2009, 2010; Grisan *et al.*, 2008; Hart *et al.*, 1999). All of these algorithms justify their existence by stressing the importance that they might have in providing a way of studying different pathological disorders. However, to the best of my knowledge, no previous work exists, that investigates at the same time and in multiple experimental designs, many candidate biomarkers of progression to DR, as well as estimating haemodynamic features from the vascular geometry and also evaluating them both statistically and by using classification models. An additional novelty is that different areas of the retina, as described in (Hove *et al.*, 2004), are also separately investigated in order to find out whether they can offer us more powerful biomarkers. Having said that, all the biomarkers presented in this thesis are evaluated in two complementary ways. First, statistically, in order to understand how each of them changes across the investigated periods of time, and secondly, for their discrimination potential, when utilised for diagnosing the different stages of the disease.

Finally an alternative, more accurate method for summarising the calibre of the parent vessel trunk is also presented (Leontidis *et al.*, 2016b) and compared with the established method in literature (Knudtson *et al.*, 2003), with an application to esti-

mating the arteriovenous ratio (AVR). It should be stressed that the work of the thesis deals with the estimation of the candidate biomarkers and with the evaluation of the changes in the measured biomarkers among the different populations at risk, given the covariates and the condition of the patient, rather than the prediction/assignment of a person to a particular sub-population to stratify its risk of developing retinopathy.

The main objectives of this thesis are the following:

- To provide complete, detailed and valid studies, which evaluate for their statistical significance and robustness more than 100 candidate biomarkers that are extracted from different areas of the retina. In addition a feature selection process will be conducted and classification models will be built, in order to evaluate the discrimination potential of these biomarkers.

- To investigate, primarily from a technical perspective but with keeping in mind the clinical implications as well, whether the retinal vasculature significantly and distinctly changes prior to the onset of DR, as a consequence of the progression of diabetes.

- To present a framework that accommodates the proposed way of analysis, where the same exact segments are studied over a period of time, and also compare it against alternative experimental designs. The proposed framework offers a pipeline that is able not only to automate the process of analysing retinal images and extracting features for studying DR and other retinal diseases, but to also ascertain that the bifurcations are correctly matched (progression studies), the areas of interest are properly identified and the analysis is valid and therefore meaningful.

- To estimate various haemodynamic features, directly from the vascular geometry and to apply them in the study of the progression to DR. This process includes

some constraints, imposed by the boundary conditions, which are used to make these estimations.

- To propose an alternative and more accurate method for summarising the calibre of the parent vessel trunk, separately for healthy, diabetic and DR groups, for estimating the central retinal artery equivalent (CRAE), the central retinal vein equivalent (CRVE) and their quotient, the arteriovenous ratio (AVR). More accurate estimates can help build better performing classification systems, both for the earlier diagnosis of the progression of the disease and the overall improvement of the classification accuracy.

1.2 Statement of Originality

The primary contributions of this thesis are presented below, with all of them elaborated in the chapters to come.

1. The first major contribution is the proposed framework and the experimental design that offer an efficient and accurate way to more robustly study the alterations in the retinal vasculature, for identifying and diagnosing early the progression from diabetes to DR. This has an important clinical implication, because the proliferation of the disease can sometimes be rapid, leading to irreversible damages in the retinal tissue. Providing a framework of automated analysis and diagnosis could help clinicians monitor the progression of the disease more closely and/or adjust the treatment plans of the patients.

   To achieve all these, many different methods already proposed in literature will be combined and some key parts will be implemented, in order to create the feature set that will be used to conduct the statistical analysis, the feature selection process and the multiple binary classifications. In addition, all the classifica-

tion models will include solely geometric and haemodynamic features, with no information extracted from the images per se.

2. The second major original contribution is to introduce a pilot estimation of haemodynamic features, directly from the retinal vascular geometry, and more importantly to evaluate them as candidate biomarkers for identifying the progression to DR. Haemodynamic features' calculation is possible by utilising devices, like laser doppler flowmeter and colour doppler imaging, that can provide in-vivo measurements. However, this is a demanding process, not easily applicable, and it also requires prospective studies, given that these data cannot be retrospectively found. Therefore, in this study, a number of haemodynamic features are estimated, such as blood flow rate, blood flow velocity, wall shear stress, pressures and Reynolds number, by utilising a 0-D mathematical model with some boundary conditions. The estimations are made possible by a combination of image analysis techniques and mathematical modelling, with the former used for extracting geometric parameters needed for the estimations (lengths and widths of the vessel segments), and the latter for simulating some of the conditions that are governed by the functionality of the retina.

3. The third major contribution is the new and more improved method for summarising the width of the parent vessel trunk, with a direct application to the more accurate estimation of the AVR. This method does not only improve the general estimation of the AVR, but also proposes adjusted equations, specifically for healthy, diabetic and DR groups. Any improvement in the above biomarkers, which are vastly used in many studies of different disorders (Golpe *et al.*, 2016; Leontidis *et al.*, 2016a; Triantafyllou *et al.*, 2014; Heitmar *et al.*, 2010; Chew *et al.*, 2016), can lead to more accurate analyses and inferences. The superiority of the proposed method is proven both in terms of the mean absolute

percentage error and the classification performance of the newly estimated features (CRVE, CRAE and AVR), and is also compared with the gold standard method in literature.

1.3 Thesis Overview

The thesis is organised as follows; **Chapter 2** gives an overview of the medical background and previous studies. This chapter presents the necessary background information to understand the functionality and importance of the retina and what happens during the proliferation of diabetes and DR. It also discusses various previous studies, establishing the necessary basis for the subsequent chapters. **Chapter 3** is devoted on describing the proposed framework, with details about its parts, the methods and the techniques that were adopted and combined. It also describes the different areas of interests and how the vessel segments are matched and assigned to each of these areas. Finally it gives important information about the biomarkers under investigation and also the fundamentals for the estimation of the haemodynamic features. **Chapter 4** focuses on the statistical analysis methods, the feature selection processes and the classification models. All the metrics that were utilised to evaluate and validate the models are also elaborated, followed by the feature selection and classification techniques that were adopted, and finishing with the binary classification combinations that were conducted. **Chapter 5** is solely devoted on the new method of summarising the width of the parent vessel trunk, which is also compared with the gold standard method in literature, using statistical and classification approaches for the validation. **Chapter 6** will present the novel results of the statistical analyses of the six different data categories and also the results of the eight different classification combinations. Finally, some critical and concluding remarks are given in **Chapter 7**.

1.4 List of Publications

The research presented in this thesis has been documented, in part, within the following publications:

1. Leontidis, G., Al-Diri, B., & Hunter, A. (2016). Summarising the retinal vascular calibres in healthy, diabetic and diabetic retinopathy eyes. Computers in biology and medicine,72, (pp.65-74), Elsevier.

2. Leontidis, G., Al-Diri, B., & Hunter, A. (2016). Exploiting the retinal vascular geometry in identifying the progression to diabetic retinopathy using penalized logistic regression and random forests. In Emerging Trends and Advanced Technologies for Computational Intelligence (pp. 381-400). Springer International Publishing.

3. Caliva, F., Leontidis, G., Al-Diri, B.,Hopkins, P., Antiga, L., Hunter, A. Haemodynamics in the retinal vasculature during the progression of diabetic retinopathy. In XXII Biennial Meeting of the International Society for Eye Research (ISER), Abstract.

4. Leontidis, G., Al-Diri, B., Wigdahl, J., & Hunter, A. (2015, August). Evaluation of geometric features as biomarkers of diabetic retinopathy for characterizing the retinal vascular changes during the progression of diabetes. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 5255-5259). IEEE.

5. Leontidis, G., Al-Diri, B., & Hunter, A. (2015, July). Retinal vascular geometry: examination of the changes between the early stages of diabetes and first year of diabetic retinopathy. In Science and Information Conference (SAI), 2015 (pp. 709-713). IEEE.

6. Leontidis, G., Al-Diri, B., & Hunter, A. (2014). Diabetic retinopathy: current and future methods for early screening from a retinal hemodynamic and geometric approach. Expert Review of Ophthalmology, 9(5), 431-442.

7. Leontidis, G., Wigdahl, J., Al-Diri, B., Ruggeri, A., & Hunter, A. (2015). Evaluating tortuosity in retinal fundus images of diabetic patients who progressed to diabetic retinopathy.In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) Group, 101, 95 (Short paper).

8. Leontidis, G., Hunter, A., & Al-Diri, B. (2014). Study of the retinal vascular changes in the transition from diabetic to diabetic retinopathy eye.In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) IEEE (Short paper).

Chapter 2

Medical Background and Previous Studies

2.1 Introduction

This chapter is focusing on providing the necessary background information in order to understand how the human retina is functioning, how is affected by diabetes and diabetic retinopathy (DR) and what impact this has to the retinal vascular geometry and haemodynamics. Diverse and comprehensive research is being conducted in this field, both by clinicians and computer scientists that creates the need for a deeper understanding of the mechanisms underneath the development of diabetic retinopathy (DR). This fact enhances the necessity for a comprehensive approach in quantifying the functional impairments and the retinal vascular alterations (objective of this thesis) during the development and progression of diabetes, until the first lesions appear in the retina (onset of DR). Since DR can be managed upon early diagnosis, it is useful to focus on finding and categorising all the changes that are triggered by the progression of diabetes.

The **first part** of this chapter is devoted on providing some medical background

information about the retinal tissue, and also how diabetes and DR is affecting its functionality. The **second part** presents the findings of some key previous studies, with the **third part** addressing how this thesis is aiming to contribute to this research field with some novel elements.

2.2 Anatomy of the Retina

The vertebrate retina, or simply the retina, is composed by ten distinct layers, which are categorised from closest to farthest from the vitreous body (Villegas, 1960). The retina of the eye is a light-sensitive layer of tissue, which is located in the inner surface of the eye and enables the conversion of the incoming light into a neural signal for further processing in the visual cortex of the brain (Dowling, 1987). It is easily accessible nowadays with many different methods (ophthalmoscope, fluorescein angiogram, optical coherence tomography etc.), to both scientists and clinicians (Ciulla *et al.*, 2003). Its importance is highlighted by the fact that, according to estimations, around 80% of all sensory information in humans originates from the retina (Hildebrand & Fielder, 2011; Oyster, 1999; Riordan-Eva & Whitcher, 2008). Figure 2.1 shows where the retina stands in respect to some of the other anatomical parts of the eye.

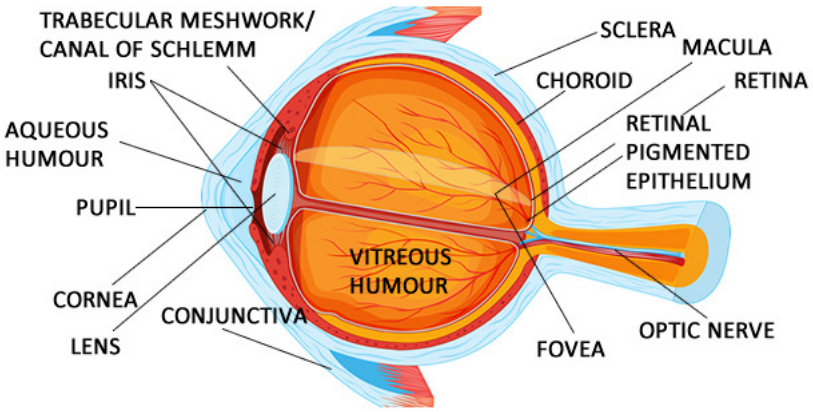

Figure 2.1: A schematic representation of some of the anatomical parts of the eye, including the retina. The diagram shows the position of the retina into this complex organ, the eye. Source:http://www.thomasutton.com/about-your-eyes/eye-anatomy-diagram.html.

As described above, retina is a layered structure of neurons interconnected with synapses with only the photoreceptors cells being light sensitive. There are two main types of photoreceptors: rods, which are responsible for black-and-white vision, mainly in dark light, and cones, which make the eye perceive colour and support daytime vision. A third one is also found, which is a much more rare type of photoreceptor, the intrinsically photosensitive ganglion cell, which responds to light in the absence of all rod and cone photoreceptors (Do & Yau, 2010).

In humans, the entire retina is approximately 24 mm in diameter (Riordan-Eva & Whitcher, 2008). It contains 7 million cones and 75–150 million rods. An important part of the retina is the optic disc or optic nerve head, which is often called the blind spot, since it has no photoreceptors. It is found at the optic papilla, a nasal zone where the optic nerve fibres leave the eye (Navarro, 2009; Oyster, 1999). In a fundus image, figure 2.2, this can be found as a white oval area, extending in an area of 3 mm^2. Temporal to the optic disc is the macula, a small and highly sensitive part of the retina responsible for detailed central vision. In the very centre of the macula is the fovea, which is responsible for sharp central vision, necessary in humans for activities that require visual details (Provis *et al.*, 2005).

In section, the size of the retina is approximately 0.25 mm in thickness, which varies with age (Alamouti & Funk, 2003). It has three layers of nerve cells and two of synapses including the ribbon synapse. The connection between brain and retina is made via the optic nerve head, which carries the ganglion cell axons in the brain, and the blood vessels that are spread inside the retina (Riordan-Eva & Whitcher, 2008).

2.3 Diabetic Retinopathy

Diabetic retinopathy is a complication of diabetes, characterised by lesions and vascular anomalies in the retina, which include micro-aneurysms, haemorrhages, cotton

wool spots, exudates (bright spots), venous beading, intra-retinal microvascular abnormalities, neovascularisation, loop and fibrous proliferation (Donnelly & Horton, 2008). Figure 2.2 shows a normal retina with the blood vessels, the macular area and the optic nerve head, and also a diabetic retinopathy retina one with some apparent lesions. Diabetic maculopathy occurs when diabetic retinopathy starts affecting the central macula, which may also cause loss of vision (Antcliff & Marshall, 1999). Making a proper assessment of the stage of retinopathy requires the ability to be able

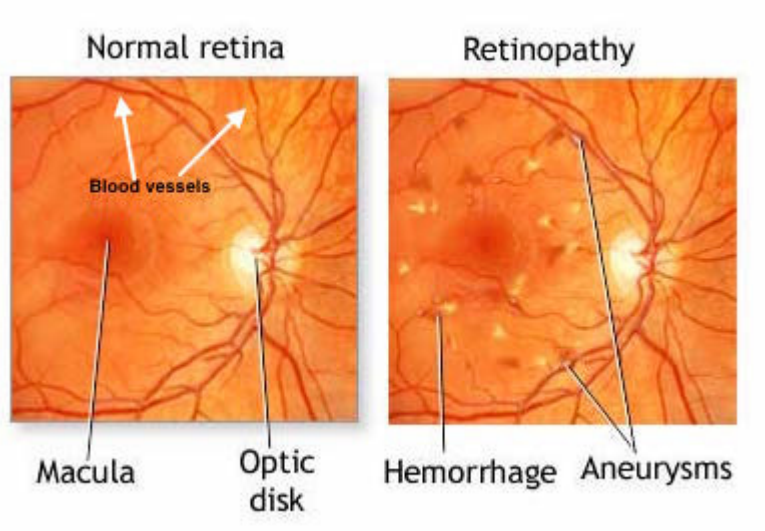

Figure 2.2: Normal retina on the left, showing the blood vessels, the macula and the optic disc. On the right the same retina can be seen, which now includes haemorrhages and micro-aneurysms. Source:http://www.silversteineyecarenj.com/diabetic-eye-care.php.

to successfully identify all of the the following (Donnelly & Horton, 2008; Wilkinson *et al.*, 2003):

Micro-aneurysms represent the earliest visible change of diabetic retinopathy. They appear as round, red dots, mainly in the posterior part of the eye, and usually increase with the progression of DR. It is also common to appear in clusters (Wilkinson *et al.*, 2003).

Haemorrhages usually accompany micro-aneurysms in the retina, but they may take different shapes, such as a dot or blot (mostly in the inner layer), and also

flame-shaped (mostly within the superficial layer in the nerve fibre layer), depending on their depth within the retina (Donnelly & Horton, 2008).

Vitreous haemorrhage refers to the bleeding within the vitreous cavity. When the macula is obscured, the incidence of vision loss rises sharply (Donnelly & Horton, 2008).

Preretinal or subhyaloid haemorrhage is bleeding found anterior to the retina and under the posterior vitreous face; it often appears as a capsized shape(Ramsay *et al.*, 1986).

Hard exudates are formed by leaked cellular lipids from abnormal intra-retinal capillaries. They appear as shiny yellow-white deposits with sharp margins, which vary from small spots to larger patches. They may also evolve into rings known as circinates. In case the leaked lipids coalesce into the fovea, then vision can be compromised (Wilkinson *et al.*, 2003).

Cotton wool spots are greyish or white patches of discolouration in the nerve fibre layer, which have indistinct edges. They result from local ischaemia, so multiple cotton wool spots might indicate generalised retinal ischaemia - a feature of pre-proliferative retinopathy(Donnelly & Horton, 2008) .

Venous beading refers to the localised dilation of veins. The degree of venous beading can be a useful sign of proliferative diabetic retinopathy as well as diffuse retinal ischaemia (Wilkinson *et al.*, 2003).

Intraretinal microvascular abnormalities are regions where capillaries appear dilated with new tortuous vessels formed within the retinal layers (Lee *et al.*, 2015a).

Neovascularisation refers to the process of abnormally growing new vessels. These new vessels stem from large veins or major arcade vessels. They initially

appear as fine tussocks on the surface of the retina. These newly formed vessels are fragile and bleed easily (Wilkinson *et al.*, 2003).

Loop is the sudden deviation of a vein from its normal course. The deviation varies from a gentle curve to a nearly omega shaped one (Cogan *et al.*, 1961).

Retinal detachment is an outcome of the neovascular traction. It usually occurs slowly and may remain stable for years, when laser treatment has been applied to control the neovascular process (Wilkinson *et al.*, 2003; Donnelly & Horton, 2008).

DR may progress through four stages (Watkins, 2003; Frank, 1995):

1. **Background or mild non-proliferative retinopathy**, where small balloon-like swellings appear in the retina's blood vessels called micro-aneurysms.

2. **Moderate non-proliferative retinopathy**, where swelling and distortion of the vessels might occur, as well as possibly losing their ability to transport blood.

3. **Severe non-proliferative retinopathy**, where even more blood vessels get blocked, depriving of the blood supply to various areas of the retina. These areas can stimulate the production of growth factors that signal the retina to start growing new blood vessels.

4. **Proliferative retinopathy**. At this stage, the proliferation of new blood vessels occurs, which grow along the inner surface of the retina and the vitreous gel. These vessels are fragile, which makes them more likely to leak and start bleeding. It is also possible that the scar tissue will contract and cause the retina to be detached from the underlying tissue.

An example of the four stages can be seen in figure 2.3.

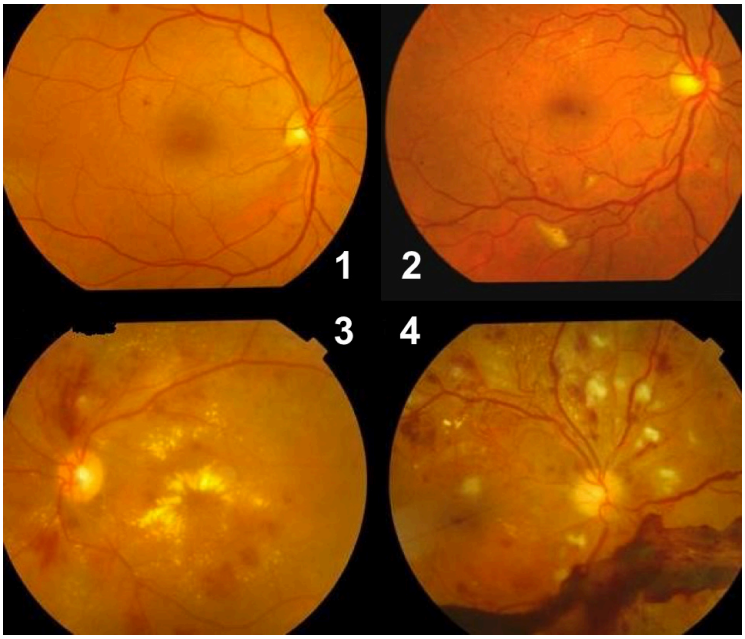

Figure 2.3: Four different fundus images, representing a different stage of DR. Image 1: mild non-proliferative DR, image 2: moderate non-proliferative DR, image 3: severe non-proliferative DR and image 4: proliferative DR. Source: http://sdhawan.com/eye-diseases-diabeteseye.htm.

Although DR is a very common complication of diabetes, still many cases are observed at a late stage, where visual acuity is impaired and irreversible damage has already occurred (Aiello, 2003; Ruta *et al.*, 2013). The prevalence of DR is found to increase with the duration of diabetes. According to some very interesting results from the Australian Diabetes, Obesity and Lifestyle study in patients with diabetes with less than 5 years in diabetic state, the prevalence of DR is less than 10%, but this number becomes more than 50% in those patients having diabetes for 20 or more years (Nguyen *et al.*, 2008a; Tapp *et al.*, 2003).

The main risk factors of DR remain hypertension and hyperglycaemia, but hyperlipidaemia should not be excluded as well (Nguyen *et al.*, 2007). A number of epidemiological studies have highlighted the importance of hyperglycaemia in DR, in addition to two pivotal studies: UK Prospective Diabetes Study in patients with Type 2 diabetes (UKPDS, 1998a) and the Diabetes Control and Complications Trial in patients with Type 1 diabetes (Ohkubo *et al.*, 1995; DiabConCompTr, 1993). In

the UK Prospective Diabetes Study, there is another interesting part, which shows that controlling the blood pressure level reduces the risk of retinopathy regardless of the glycaemic level (UKPDS, 1998b).

An important thing to keep in mind is that the isolated retinopathy signs (micro-aneurysms, haemorrhages and cotton wool spots) are found to be more common now in subjects without diabetes and hypertension, in comparison with what had been previously believed (Nguyen *et al.*, 2008a).

At an early stage, DR affects the endothelial cells and the structure of muscle cells, leading to the loss of pericytes (Eva & Mansour, 1998). The proliferation of the endothelial cells and the thickening of the membrane cause vascular occlusion, whereas pericyte loss is responsible for the formation of micro-aneurysms. The big challenge remains to detect and understand the microvascular haemodynamic abnormalities at a stage when there are no morphological alterations in the retina. In this stage, the clinicians could possibly be able to intervene and manage to some extent the progression of the disease.

Unfortunately, the pathogenesis of DR is not yet fully understood both at the cellular and molecular level, limiting the options for effective therapeutic interventions early, while the disease still develops (Stitt *et al.*, 2005; Stitt & Curtis, 2005). The exact mechanism that triggers the formation of the microvascular lesions remains to be fully understood.

In general, hyperglycaemia appears to be sufficient to initiate the development of DR as revealed by some experiments in animals, which were made hyperglycaemic in the laboratory (Engerman & Kern, 1984; Kador *et al.*, 1990; Kern & Engerman, 1996). At the same time, similar studies have shown that by intensively and sufficiently controlling hyperglycaemia leads to the inhibition of the development of DR (Engerman & Kern, 1993).

Possible geometric alterations in the retina might indicate the existence of a sys-

temic disease. Functional changes can be depicted in the retina vasculature and also be measured using devices such as flow velocity meters, oximeters etc. Assessing these functional and haemodynamic changes in a qualitative way might help to bypass the limitations of studying only the morphological features. Moreover, it can help differentiate between diseases that their structural effect on the retina is ambiguous and prone to misinterpretation.

2.4 Blood Supply

The retinal circulation is directly observable and has its own blood supply, which comes from behind the eye and enters the retina through the optic nerve head. It is an end arterial system without anastomoses. The central retinal artery brings the blood into the retina and the central retinal vein drains the blood out of the eye, leading the blood back to the heart for re-oxygenation (Funk, 1997; Hayreh, 1969).

2.5 Screening of the Retina

Retinal imaging is a non-invasive way of viewing human vessels. Using proper techniques, retina is visible from the outside, which assists in the imaging of the retina and brain tissue non-invasively (Lupascu, 2010; Terai *et al.*, 2014). Moreover, as the retina is a highly metabolically active tissue with a double blood supply, it allows direct non-invasive observation of the circulation (Rice *et al.*, 2002; Kim *et al.*, 2011).

The reason why the retina is so extensively studied is due to the fact that both eye- and other- diseases that affect the circulation and the brain can be illustrated in the retina (Wong *et al.*, 2001, 2006). Macular degeneration and glaucoma are among the most prominent diseases and together with DR the most important causes of blindness in the developed world (Williams *et al.*, 2004).

In addition to the eye diseases, a number of systemic diseases also affect the retina.

Complications of such diseases include DR from diabetes mellitus, hypertension from cardiovascular disease and multiple sclerosis (Hayreh *et al.*, 2001). During the last decade, advances in non-invasive techniques utilised for measuring several features in the retina, have led to the exploration of different aspects about the haemodynamic and geometric features of the retina and blood flow regulation, both in normal and diseased human eyes. Some examples of these techniques are the following: The retinal vessel analyser (Polak *et al.*, 2000), dye dilution technique for arteriovenous passage time (Wolf *et al.*, 1989), blue field simulation for the velocity and the number of leukocytes (Riva & Petrig, 1980), bidirectional laser Doppler velocimetry (LDV) for the red blood cell velocity (Sullivan *et al.*, 1990), laser Doppler flowmetry for the blood flow in the tissue of the optic disc (Nilsson *et al.*, 1980) and colour Doppler imaging for the central retinal arterial blood velocity (Williamson & Baxter, 1994). Regarding the measurement of retinal blood flow, another methodology is the video fluorescein angiography, which relies on the rapid injection of a small bolus fluorescein dye and recording the retinal images with a scanning laser ophthalmoscope in order to distinguish the resultant vascular fluorescence from the passage of fluorescent dye through the retinal circulation (Wolf *et al.*, 1989).

2.6 Blood Flow Mechanisms in Diabetes

Impairment in retinal blood flow is one of the earliest abnormalities to occur in the human body in diabetes (Cunha-Vaz *et al.*, 1978). Only 4% of the blood flow delivered to the eye is distributed into the retina (Besharse & Bok, 2011). Impairment of retinal circulation results in blood flow alterations, which consequently affect the delivery of oxygen and metabolic substrates to the tissue. The maintenance of the function and structure of the retina is highly affected from these abnormalities (Clermont & Bursell, 2007).

Grunwald *et al.* stressed the importance of the blood flow, especially in comparison with perfusion pressure (Grunwald *et al.*, 1996). In their initial study in diabetic patients, they found that both the mean blood pressure and perfusion pressure in all patients who were normotensive were significantly higher in five patients, whose blood flow did not decrease after 5 days and whose retinopathy deteriorated at 6 months. Another observation shows evidence that increased blood pressure is correlated with increased prevalence of retinopathy (Knowler *et al.*, 1980; Klein *et al.*, 1989, 1995). In the Wisconsin study, the authors found that systolic blood pressure is a significant predictor of the incidence of DR, while diastolic blood pressure can be an important predictor of the progression of the disease (Klein *et al.*, 1989).

Perfusion pressure can also be of importance in the early screening of DR, since the normal auto-regulatory response mechanism of the vessels is impaired. Perfusion pressure can be calculated by subtracting the intraocular pressure from the two-thirds of the mean arterial pressure (Sinclair *et al.*, 1982). In another study, using light stimulation in patients with well-controlled Type 1 diabetes and no signs of retinopathy, showed that the functional abnormalities consist of reduced or no dilation of retinal vessels, as well as reduced or no constriction of retinal arteries, as a response to the increase of intraluminal pressure (Garhofer *et al.*, 2004b; Mandecka *et al.*, 2007). A very interesting finding (Hill *et al.*, 2009; Lorenzi *et al.*, 2010) was that flickering light stimulates and activates auto-regulatory mechanisms that dilate the retinal blood vessels and affect the blood flow by enhancing the circulation. In healthy subjects, the response relates to the increase of the calibre of the retinal vessels by 2–4%, whereas in diabetic patients, there is slight or no response to the stimuli at all. An observation of the response of the retinal vessels to increased intraluminal pressure and how the vascular mechanisms manage to regulate this suggests that in normal arteries, an elevation in intraluminal pressure leads to the constriction of the vessels or to the dilation under pressure reduction. All of these occur by using the inherent mechanisms in the vas-

cular smooth muscles cells that are independent of any hormonal, neural or metabolic influences. This mechanism is termed myogenic response (Hill *et al.*, 2009).

It can be inferred from the above studies that blood flow is an important window for understanding and measuring the changes during the progression of the disease. Every single change, either in the vessel structure or in the microcirculation, directly affects the blood flow inside the vessels. It is a fact that most of the studies use different methods to take such measurements and different metric systems. It still remains a very difficult task to associate changes in blood flow with the progression of a disease, since many other factors might influence the blood flow, which have to be excluded in the first place.

2.7 Introduction to Biomarkers

A few tests were run to evaluate the response of the retinal vessel structure to different stimuli, measuring at the same time the changes in the vascular diameter, either with the laser Doppler flowmetry or with the dynamic retinal vessel analyzer (Garhofer *et al.*, 2004a; Lorenzi *et al.*, 2010). In addition, some retinal vascular width changes were identified and associated with early consequences of diabetes, making them candidates to become biomarkers of risk for diabetic complications.

A biomarker can be defined as a feature that is accurately and objectively measurable and is evaluated as an indicator of regular biological and pathogenic processes or responses in a specific drug treatment or disease (Steyerberg *et al.*, 2012). It is clear that some systemic markers such as blood pressure, duration of diabetes, glucose level and lipid levels are definitely relevant factors, but on the other hand, they cannot be used to identify the proliferation of DR (Hove *et al.*, 2006). It is observed that even patients under good glycaemic control can worsen rapidly, in contrast to patients with poor control that might remain in stable condition (Ribeiro *et al.*, 2013). This led to

the identification of different phenotypes of progression, taking into account the characteristics of the retinal lesions.

Time of testing plays a crucial role in accurately and reliably measuring the haemodynamic features at fixed state. Diverse indications of the abnormalities in diabetic patients are observed, which are assumed to be attributed to the influence of blood glucose level during the testing (Ikram *et al.*, 2013; Pournaras *et al.*, 2008). Applying some stimuli in the retina can cause abnormal auto-regulation of the blood flow in DR patients, with a simultaneous increase of the flow, in parallel with the proliferation of DR (Kohner *et al.*, 1995). Changes in the retinal architecture may result in impaired space filling and microcirculatory transport of non-uniform shear distribution in branches and bifurcations. Moreover, it might cause reduced energy efficiency in blood flow, giving a strong indication of early disease state (Ikram *et al.*, 2013).

2.8 Pathology in the Diabetic Retina

The pathological processes during diabetes are initially subtle but affect the whole haemodynamic functionality of the retina. During normal state, the auto-regulation mechanisms keep the blood flow constant in the whole range of systemic blood pressures and intraocular ones (Riva *et al.*, 1981). The vessel responses are locally regulated, by targeting the smooth muscle cells in arterioles and capillary pericytes (Burgansky-Eliash, 2012; Sims, 1986). On the other hand, during diabetes, there are changes in local vasoactive factors, and the response of pericytes to these factors is altered as well (Joussen *et al.*, 2002; King *et al.*, 1994; Bursell *et al.*, 1997).

Tight junction complex proteins help in the creation of the blood–retinal barrier. In order to maintain normal neural function, the tight junction is responsible for the connection of the endothelial cells in the brain and retina (Harhaj & Antonetti, 2004). In some diseases such as DR, the actions of the vascular endothelial growth factor

(VEGF) and cytokines on the tight junction proteins affect the vascular permeability and cause changes in the blood–retinal barrier. The importance of blood–retinal barrier can be highlighted from the fact that it is responsible for preventing certain substances from entering the tissues of the retina. The blood–retinal barrier is formed from tight junctions between retinal epithelial cells and non-fenestrated capillaries of the retinal circulation (Harhaj & Antonetti, 2004).

According to emerging evidence, neurodegeneration has been found to occur early in the pathogenesis of DR. In addition to neural apoptosis, some changes in glial cells (non-neurons) occur as well, with the process known as reactive gliosis. It is still unclear, which of the two processes occurs first in the degeneration process. The most important mechanisms that mediate the neurodegeneration process are: oxidative stress, extracellular glutamate accumulation and reduction of neuroprotective factors (Simo & Hernandez, 2014).

The term oxidative stress, describes the imbalance that occurs between the reactive oxygen species and the antioxidant defenses of a living system (Baynes, 1991). Tissue damage and patho-physiology is triggered by oxidants such as reactive oxygen species and reactive nitrogen species. The oxidative stress, which is caused by hyperglycaemia, is considered an important pathway of diabetic microvascular complications (Cui *et al.*, 2005). There is strong evidence that the correlation between hyperglycaemia, redox homoeostasis and oxidative stress is responsible for the pathogenesis of DR (Kowluru & Chan, 2007; El-Remessy *et al.*, 2005).

The Hoorn study reported the significance of subclinical inflammation to the development of DR (Van Hecke *et al.*, 2005). DR is also considered a low-grade inflammatory disease affecting the rolling and adhesion of leukocytes (Miyamoto & Ogura, 1999). Nowadays, the role of inflammation has been highlighted and it is considered very important, though complex and unclear. Inflammation is triggered by factors such as hyperglycaemia, oxidative stress, hypertension etc., but this creates a chain reaction

since inflammation propagates these pathways further through cytokines, VEGF signalling, adhesion molecules, enhanced receptors for advanced glycation end- products expression, nitric oxide regulation and NF-kb signalling. The subclinical inflammation via endothelial nitric oxide synthase leads to increased intraocular pressure (Adamis, 2002).

2.9 Effect of Diabetes in Vessel Structure

Hyperglycaemia leads to intramural pericyte death and thickening of the basement membrane, which contributes highly to the alterations in the integrity of the retinal blood vessels. This fact causes changes to the blood–retinal barrier and vascular permeability (Kitabchi *et al.*, 2009; Qaum *et al.*, 2001). During the hyperglycaemic state, the endothelial cells align and elongate in the direction of shear stress, modifying some of their functions at the same time. Shear stress is defined as the component of stress coplanar with a material cross-sectionally (Libby P, 2002). The endothelial cells respond to the increased shear stress and produce more a vasodilator (nitric oxide), which causes the expansion/dilation of the blood vessels. This homoeostatic reaction of the vessels occurs in order to restore the normal shear stress by decreasing the blood flow velocity (Gross *et al.*, 2003; Beckman *et al.*, 2001).

Another major issue is the development of atherosclerotic plaque, namely the hardening and thickening of arterial wall due to the reactive oxygen species and inflammation. These two factors can be suppressed by the nitric oxide. If the endothelial cells do not produce enough nitric oxide in response to shear stress in a diabetic state, it can lead to the development of atherosclerosis in diabetic patients (Creager *et al.*, 2003).

The pulsatile flow of the blood through the vessels activates the endothelial nitric oxide (Lu & Kassab, 2004). The shear stress, caused by the blood pressure in every heartbeat, makes the vessels stretch and relax, since the blood flow inside the vessels

is not steady. The role of the endothelial nitric oxide is to maintain the diameter of the blood vessels so as to preserve the perfusion of tissues at optimal levels. VEGF mediates the release of nitric oxide from human umbilical venous endothelial cells (Toda *et al.*, 2010; Calles-Escandon & Cipolla, 2001).

Metabolic abnormalities are a main characteristic of diabetes, which include hyperglycaemia, free fatty acids and insulin resistance. These three factors provoke molecular mechanisms, which in turn alter the function and the structure of the vessels. Oxidative stress is one of those affected, in addition to the malfunction of the intracellular signal transduction (Hartnett *et al.*, 2000; Beckman *et al.*, 2002).

One very important concept defines that hyperglycaemia-induced oxidative stress, very common in DR, mediates the endothelial malfunction in diabetic patients. This is proven by the observations that intra-arterial infusion of ascorbic acid restores endothelium-dependent vasodilation in healthy subjects, exposed to a hyperglycaemic clamp as well as in patients with Type 1 and Type 2 diabetes (O'Driscoll *et al.*, 1997).

Vascular muscle cell apoptosis in atherosclerotic lesions is also increased in such a way that patients with diabetes are prone to having fewer smooth muscle cells in the arterial lesions, which increases the tendency of plaque rupture. The signalling pathways in the cells are affected by the way the cells adhere to the substrate proteins. Moreover, since the cells attach to the glycated collagen in a completely different way than normal collagen, it leads to a different way of responding to the mechanical forces. Although the high blood sugar state is more frequent, the cell response to blood flow is altered even in low blood sugar levels (Creager *et al.*, 2003).

The arterial walls tend to become stiffer with age, but diseases such as diabetes can accelerate this natural process. The arterial muscle cells contraction and relaxation are affected by the impaired endothelial cells, which in turn affect the wall stiffness by the modification of the isometric tone (Bank *et al.*, 1995).

2.10 Oxygen Perfusion

The exact cause of the elevated oxygen saturation in DR is still uncertain and unclear, but it is evident that the normal response of the retinal circulation in preventing hyperoxygenation is impaired by hyperglycaemia (Khoobehi *et al.*, 2013).

Oxygen distribution is another major factor in the vessel wall oxygen perfusion. The process of the oxygen distribution can be affected by some factors, including capillary non-perfusion and shunting, thickening of capillary basement membranes and oxygen affinity of haemoglobin in diabetic patients (Konno *et al.*, 1996).

If the capillary network is shunted at some point, the blood may bypass it through dilated channels. During this state, some of the capillaries are closed and some dilated. Some studies in the retina, using fluorescein angiography, have shown that dilated capillaries force the blood to pass from arterioles to venules, leading to capillary non-perfusion (Hardarson & Stefánsson, 2012). This ”unhealthy” process makes the transport of blood faster than normal, making the venular blood hyperoxic and the rest of the blood hypoxic in these non-perfused areas. The capillary non-perfusion in addition to shunting disturbs the normal blood flow, altering the normal oxygenation that may lead to different pathologies. As a consequence, the non-perfused areas do not extract oxygen from the haemoglobin, making the tissues hypoxic and thus ischemic.

One of the observations during DR is the thickening of capillary walls, which can lower oxygen delivery levels. It is assumed that in this situation oxygen is inhibited from efficiently diffusing and perfusing vessels, contributing probably to the elevated oxygenation of the blood. Knowing the high importance of the oxygen for the functionality and preservation of tissues, it can be inferred that the mal-distribution during the disease state makes the tissues hypoxic, elevating the demand of oxygen, and thus increasing the blood flow to deliver more oxygen. The above way is one of the reasons that total blood flow can be increased in DR (Hardarson & Stefánsson, 2012) .

It is easily understood that dead tissue cannot consume oxygen. Tissue degenera-

tion lowers the total amount of oxygen extracted by blood vessels, increasing simultaneously the venous oxygen saturation, as mentioned in (Khoobehi *et al.*, 2013). It is speculated that the affected microcirculation in diabetes has an effect on retinal vessel saturation (Hammer *et al.*, 2009). Pathogenesis in DR has been linked to retinal hypoxia, which triggers neovascularization and retinal oedema (Cai & Boulton, 2002).

Khoobehi *et al.* conducted a comprehensive experiment in order to determine the retinal oxygen saturation trend with the onset of diabetes and increasing severity of DR, by comparing diabetic groups with and without retinopathy (Khoobehi *et al.*, 2013). For this purpose, they used a fundus camera-based dual-wavelength snapshot oximeter to take images of the retina for the whole recruited group and analyse them to determine oxygen saturation in the major arteries and veins. It was found that in normal subjects, the saturation was 92.3 ± 4.2% in arteries and 57.2 ± 6% in veins; in diabetic patients without DR, it was 96.3 ± 8.6% in arteries and 58.7 ± 7.5% in veins; in diabetic patients with mild-to-moderate non-proliferative DR, it was 97.7 ± 5.8% in arteries and 61.1 ± 7.6% in veins. The saturation in patients with diabetes with severe non-proliferative DR was 102 ± 10.2% in arteries and 66.8 ± 8.4% in veins; in patients with proliferative DR, it was 103.6 ± 8.7% in arteries and 66.6 ± 10.2% in veins; and finally, in all diabetics with DR combined, it was 100.4 ± 7.6% in arteries and 64.2 ± 8.4% in veins. From this study, it is clear that there is a trend of increasing retinal oxygen saturation from healthy subjects to non-DR group and to DR patients. In DR, the oxygen perfusion might be influenced by the impaired blood flow. In a study by Jorgensen *et al.*, 156 diabetic patients were recruited, 48 with Type 1 and 108 with Type 2 diabetes, in addition to 80 age-matched normal subjects (Jorgensen *et al.*, 2014). For the normal controls, any other diseases such as ocular diseases, diabetes or hypertension were excluded. They used a retinal oximetry device in order to measure the oxygen saturation in veins and arteries. As observed in proliferative DR patients, the arterial saturation was significantly higher than normal subjects and

diabetic patients with retinopathy not requiring treatment, whereas there was no significant difference in diabetic patients without retinopathy. Regarding the veins, the diabetic patients with or without retinopathy presented significantly higher saturation than normal subjects. Another important observation indicates that the oxygen extraction decreases with increasing severity of retinopathy, i.e. from normal subjects until the last stage of proliferative DR.

2.11 Haemodynamic and Geometric Alterations in Diabetes & DR – Experimental Results in Literature

A vast number of studies have focused on the vascular changes in healthy, diabetic and/or DR subjects. Studies since 1921 can be found introducing the vascular changes due to DR (Wagener & Wilder, 1921), with also other studies (Wagener & Keith, 1939; Scheie, 1953; Leishman, 1957; Stokoe & Turner, 1966) introducing general vessel appearance, such as tortuosity, branching patterns and general retinal features (Ashton, 1949). Regarding the more recent and comprehensive studies, Burgansky *et al.* studied the effect of DR on the arterial blood flow velocity in 42 diabetic patients and 38 healthy subjects and found that the velocity was slower in DR patients than healthy subjects (3.74 mm/s ± 1.09 for DR and 4.19mm/s ± 0.99 for healthy) (Burgansky-Eliash *et al.*, 2012, 2010). Moreover, they investigated the early haemodynamic alterations in patients with diabetes before the first lesions appear in the retina. In this study, the blood flow velocity in the retinal vasculature of adult-onset diabetic patients with no signs of DR (23 eyes) was compared with that of age-matched healthy subjects (51 eyes). For all the participants, measurements of blood glucose level, glycosylated haemoglobin (HbA1c), body mass index, intraocular pressure, systemic blood pressure and heart rate were taken. According to their results, the blood flow velocity in arteries was 4.7 ± 1.7 mm/s in the diabetic group, which is significantly higher than

that of the healthy group 4.1 ± 0.9 mm/s. The velocity in the veins in both groups was slower than arteries (3.8 mm/s in diabetes mellitus group and 2.9 mm/s in healthy group). It is worth mentioning that in the diabetic group, the velocity values in both arteries and veins were not correlated with the duration of diabetes, the level of glucose, HbA1C or body mass index (Burgansky-Eliash, 2012; Burgansky-Eliash *et al.*, 2010).

It is apparent that the increased velocity in diabetic patients compared to healthy subjects is opposed to the findings in DR patients (Libby P, 2002). This means that the relationship between the disease and healthy blood flow velocity reverses during the development of morphological alterations in the retina, as arteries reach their limits and capillary resistance defines the flow volume. In two studies (Rimmer *et al.*, 1989; Konno *et al.*, 1996), it was found that blood flow velocity decreases over time in some but not all diabetic patients.

Takahiko *et al.* assessed the blood flow in the bilateral central arteries in 50 insulin-dependent diabetic patients without any signs of DR and they used 20 sex- and age-matched normal subjects as a comparison. For the measurements, they used duplex Doppler sonography (Kawagishi *et al.*, 1995). The parameters that they measured were peak-systolic velocity, end-diastolic velocity, time-averaged velocity, resistance index (RI) and pulsatility index. As was expected, the results were different between diabetic patients and normal subjects. Peak-systolic velocity, end-diastolic velocity and time-averaged velocity were lower in diabetic patients than in normal subjects. Regarding the RI index, it appeared to be higher in diabetic patients than in normal subjects. Finally, for the pulsatility index, there was no significant difference between diabetic patients and normal subjects. Another interesting finding is that RI is correlated with the levels of glucose (r = 0.310) but not with haemoglobin levels (r = 0.184), blood pressure or duration of diabetes. The outcome shows that retinal artery blood flow velocities were decreased, whereas vascular resistance was increased in diabetic

patients without clinical signs of DR.

Grunwald *et al.* recruited 19 diabetic patients with less than 4 years of diabetes and 16 age-matched normal subjects. In their experiment, they measured different venous segments in diabetic patients and normal subjects respectively. Haemodynamic parameters such as blood glucose, mean blood pressure, intraocular pressure, perfusion pressure, haemoglobin and the duration of diabetes were taken into account. The total measured blood flow rate in the diabetic patients was 43.3 (SD8.9) ml/min, which is significantly higher than normal subjects 38.5 (SD 4.7) ml/min by about 12% (Grunwald *et al.*, 1996). A positive correlation was also observed between blood flow and disease duration. The total venous cross-section in diabetic patients was much higher than that of normal subjects by about 12%. Again a correlation was observed between the total venous cross-section and duration of diabetes (r = 0.34). The blood flow velocity in the largest retinal vein in both eyes appeared not to be significantly different between normal subjects and diabetic patients and no significant correlation was observed between the velocity and duration of the disease(r = 0.05). A very interesting part of this experiment was the measurement of the retinal-vascular regulatory response in hypoxia, which was defined as the percentage decrease in blood flow rate, blood flow velocity and large venous diameter between normal room air breathing and 100% oxygen breathing provided externally. This magnitude was found to be –11.6% (SD 4.5%) for venous diameter, –35.2% (SD 8.4%) in blood flow velocity and –49.2% (SD 7.8%) in blood flow rate but showing no significant difference from those of the normal subjects (-12.6%(SD 4.1%), -38.2 (SD 10%) and -53% (SD 8.8%) respectively). Finally they found no significant correlation between any of the haemodynamic variables measured.

In another study (Grunwald *et al.*, 1992), 12 normal subjects and 18 diabetic patients with background retinopathy were used for taking measurements of the total retinal volumetric blood flow and venous diameter in a similar way as the above-

mentioned experiment, using the same haemodynamic parameters. They found a positive correlation between the venous diameter measurements and the maximum erythrocyte velocity in four to five major retinal veins, both in normal subjects and in diabetic patients, measured by using bidirectional laser Doppler. The results showed a significantly larger blood flow and venous cross-section in the diabetic patients than normal subjects. Blood flow was found significantly larger in temporal retina than in nasal retina in normal subjects and diabetic patients. The measured blood flow was also significantly different between the superior and inferior retina in diabetic patients but not in normal subjects.

Bursell *et al.* conducted a comprehensive experiment investigating the retinal blood flow changes in patients with Type 1 diabetes and age-matched normal subjects (Bursell *et al.*, 1996). It was also investigated whether blood glucose levels can alter the retinal blood flow and also whether this can influence blood flow data in cross-sectional studies. Fluorescein angiography was used and blood glucose levels were adjusted in three levels using a glucose clamp in order to achieve values at 100, 200 and 300 mg/dl, taking blood flow measurements in each of these levels. Retinal blood flow was found to be significantly decreased in diabetic patients in comparison with the blood flow in normal subjects. During the glucose clamp in diabetic patients, the retinal blood flow increased at 200 mg/dl and 300 mg/dl in comparison with the 100 mg/dl level. In addition to the blood flow, they took measurements of arterial and venous diameters but the results presented no significant differences across the three glucose clamps and between the two groups.

The rate of retinal blood flow depends on several factors, which determine the perfusion pressure and the vascular resistance. Perfusion pressure is the pressure that drives the blood into the retinal vasculature. Vascular resistance is generated by the combination of the retinal vessels and the blood viscosity (Pournaras *et al.*, 2008).

Blood flow rate using laser doppler velocimetry (LDV) method has been found to

vary with a junction exponent x = 2.76 (SD 0.16) in the arteries (values of diameter (D) between 39 and 145 μm) and x = 2.84 (SD 0.12) for the veins (values of D between 64 and 177 μm) (Riva *et al.*, 1985). These values are in close agreement with Murray's law (Murray, 1926b), which calculates a junction exponent value of 3 for a vascular system that seeks an optimum compromise between blood volume and vascular resistance or in other words that minimises its resistance for a given volume. Feke *et al.* found an exponent of 4.1 for D >100 μm (Feke *et al.*, 1989) and Garcia *et al.* 3.35 for D between 84–177 μm (Garcia Jr *et al.*, 2002). Another study using dye delivery technique found an exponent of 2.9 for retinal arterioles and arteries with D between 20 and 80 μm (Guran *et al.*, 1990).

In literature, as far as normal subjects are concerned, the values of the blood flow rate between studies vary between 30 and 38 $\mu l/min$ (Grunwald *et al.*, 1992; Riva *et al.*, 1985) and from 65 to 80 $\mu l/min$ (Konno *et al.*, 1996; Feke *et al.*, 1989; Garcia Jr *et al.*, 2002). These differences in measurements are attributed to the fact of different methodological approaches.

As outlined in (Burgansky-Eliash, 2012), the correlation of blood flow velocity to physiological parameters is very important in order to understand the effects that diabetes can have to the human body functionality. In figure 2.4, it is shown that for a healthy group the flow velocity in arterioles, but not in veins, is positively correlated with mean arterial pressure (r = 0.29) (Burgansky-Eliash, 2012). The interesting part in this is that there was no significant correlation observed between blood flow velocity and mean arterial pressure in both pre-retinopathy and DR patients. This indication does not exclude the fact that there might be reduced correlation since it could be impaired by other factors as well, while diabetes develops and progresses.

In the epidemiologic study of Wisconsin (Klein *et al.*, 1984a,b,c,d, 1998, 2008), which spans across a twenty five year period, they are focusing on the prevalence of progression and regression to diabetic retinopathy. For that purpose they recruited

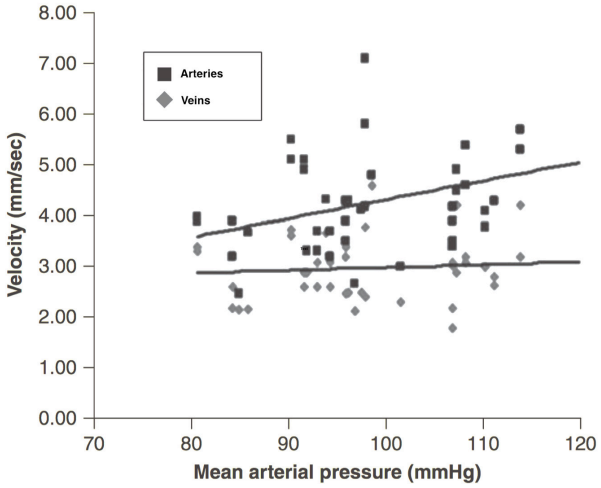

Figure 2.4: Blood flow velocity correlation with the mean arterial pressure using the Retinal Function Imager. Plot taken from (Burgansky-Eliash, 2012).

more than 1000 young and old diabetic patients, for conducting a prospective study. A number of medical, demographic, ocular and other covariates were investigated, e.g. duration of diabetes, glycosylated haemoglobin, body mass index, vessel widths and arteriovenous ratio (AVR). Their studies found multiple associations with the above covariates, regarding the progression risk and incidence of diabetic retinopathy, like duration of diabetes and age. In the particular study that targeted the geometric biomarkers, they found that within four years from the baseline examination the central retinal artery and vein equivalents were associated with greater progression of retinopathy. In addition, larger CRVE and smaller AVR were strongly associated with greater 4-,10- and 14-year incidence of proliferative retinopathy. Larger arteriolar and venular widths, independent of the retinopathy severity level, is associated with the progression of retinopathy and also larger venular caliber is associated with the 4-year incidence of proliferative retinopathy. In contrast, according to (Yang *et al.*, 2016) wider venular but not arteriolar widths were found to be associated with the development and increase severity of diabetic retinopathy in a chinese population. Similarly in (Kifley *et al.*, 2007) they observed that an increasing severity of DR in persons with diabetes is associated with the widening of retinal venular width. In (Tsai *et al.*, 2011) was also

suggested that wider retinal arteriolar width is associated with diabetes and hyperglycaemia, whereas wider retinal venular width is associated with diabetic retinopathy. In (Islam *et al.*, 2009) was found that arteriolar calibre was significantly wider in persons with diabetes compared to DR ($p<0.0001$), but not for venular widths. Other findings in (Nguyen *et al.*, 2008b) suggest that both CRVE and CRAE are larger in individuals with diabetes compared to normal subjects, observing also further dilation of veins with increasing severity of diabetic retinopathy.

In (Habib *et al.*, 2014) they compared subjects with different stages of diabetic retinopathy, including subjects that progressed to proliferative DR and subjects with only diabetes. They investigated only the vessel widths and branching angles. They observed significant changes among no-DR, mild-DR, severe-DR and PDR independent groups for both venular and arterial widths. In their predictive study, which included the same subjects in three periods, i.e. first, penultimate and final visit, with the first being no-DR and the other two mild- and Proliferative- DR respectively, no significant results were observed whatsoever. Tortuous or twisted vessels are regularly seen in humans (Weibel & Fields, 1965); mild tortuosity is usually a common anomaly with no clinical symptons, but severe tortuosity may lead to serious symptoms. Not specifically in the retina, tortuosity has been linked to ageing, atherosclerosis, hypertension, genetic defects and diabetes (Del Corso *et al.*, 1998; Hiroki *et al.*, 2002; Pancera *et al.*, 2000; Callewaert *et al.*, 2008; Owen *et al.*, 2008). The underlying mechanisms that provoke the vessels to become tortuous are poorly understood (Han, 2012). In (Lotmar *et al.*, 1979) vessel tortuosity quantification passed into a more objective assessment, where they subdivided a tortuous vessel into single arcs and measured the chord length and the arrow height (Kalitzeos *et al.*, 2013). A few studies have observed a significant increase of tortuosity in diabetic and DR subjects (Sasongko *et al.*, 2011; Weiler *et al.*, 2015; Cheung *et al.*, 2012; Leontidis *et al.*, 2015b), suggesting that tortuosity might be one of the early signs of vascular changes during the proliferation and progression

of diabetes and DR.

In addition to the above, another investigated candidate biomarker is fractal dimension (FD), which is a measure of the complexity of a pattern (more in chapter 3). Increased FD in proliferative diabetic retinopathy patients in comparison to normal subjects has been observed (Daxer, 1993). In (Kunicki *et al.*, 2009) no statistical difference was found between a diabetic and DR eye, suggesting that the complexity of the retina vessel network remains statistically the same. Moreover higher FD in diabetic subjects compared to normal ones has also been observed, and the associations remained the same even after adjusting for the presence of retinopathy (Yau *et al.*, 2010). It needs to be stressed that FD can be affected from the quality of the segmentation and the amount of segmented vessels, supporting the main approach of the studies of this thesis of comparing the same retinas over a period of time (progressors' group, i.e. subjects with diabetes that progressed to DR and whose retrospective data are available).

Unfortunately, when it comes to reviewing, comparing and summarising studies like the previously presented ones, some problems arise that make the interpretations problematic. Most of the studies do not use the same techniques for measurements and they do not follow the same protocol for selecting participants. The results cannot be directly compared when the cohorts have different demographic background and/or different medical records (other diseases etc.). On the other hand, these results are enough to give a good indication of what is happening in the retina, from a hemodynamical and a geometrical perspective. This helps in planning future studies, which can be based on the previous findings, trying at the same time to include and consider more factors that are taken into account, in order to enhance the robustness and reliability of the analyses and the results.

2.12 Connection of Previous Studies with this Thesis

It is crucial to emphasise that to the best of my knowledge, all of the above studies and others in literature involve non-progressed cohorts (subjects with diabetes vs different subjects with advanced stages of diabetic retinopathy) or progressed ones with wide and unstandardised periods of time, in which they try to identify or exclude any associations between the disease and the vascular changes. The Wisconsin study was held over a large period of time, but it did not focus on the changes that occur during the last stages of the diabetic eye and the immediate onset of DR, but on the long-term changes. Besides, in such a long period of time many factors might have changed that could affect the associations.

Therefore the main novel study of this thesis, the progressors' study (the three year period pre-DR and the onset of DR, i.e. four years of investigations within the same subjects), and the comparisons with the non-progressors' one (four consecutive years of the same subjects with diabetes but no DR), aim to bridge the gap that exists in literature, quantifying the vascular changes, by taking into account many different candidate biomarkers, that some of them, in one way or another, have not ever been evaluated before. General impressions and knowledge exists about what happens to the retinal vasculature during the transition from normal to the diabetic eye and/or to the severe non-proliferative or proliferative diabetic retinopathy eye. However it still remains unknown, how the vasculature behaves during the time period that the work of this thesis investigates, and also whether the biomarkers follow the general trends described previously.

In addition, a novel framework and extensive evaluation of the investigated features is presented, by designing and conducting various classification combinations, within the progressors' and non-progressors' groups, and also between them. In this way, the purpose is to evaluate both the discriminative power of the investigated features, as well as how this can be used in predicting different stages of the disease. Tak-

ing into account that all of the studies that will be presented in the upcoming chapters are focusing on the late stages of the diabetic eye, as well as the early and still manageable stages of DR, the importance of the outcome of these studies is vastly amplified. Any possible important findings can be used for further investigations and/or clinical applications, since focusing on the advanced or proliferative retinopathy, where the changes are already extensive and more difficult to reverse or manage, is not an objective or part of this thesis's studies. It has to be emphasised that it is important to understand how the progression of diabetes affects the retinal vasculature at these crucial last stages of the diabetic eye, but always inside a controlled and accurately specified period. Duration of diabetes, as commonly reported, is a factor that cannot be accurately defined, because the date that a subject was diagnosed with diabetes can be significantly different from the actual date that diabetes occurred. Undoubtedly, duration of diabetes still remains a very important indicator of the progression when reliably estimated, however it cannot give an accurate indication in which stage of diabetes a patient is.

Finally any comparisons with the results of other studies in literature shall therefore be made with caution, given that this thesis includes studies of a specific period of time and with data extracted from the same subjects (follow-up) during the transitional years from diabetic- to diabetic retinopathy- eye. More reliable comparisons can be made between the non-progressors' group and the progressors' DR group, which include independent subjects with diabetes and DR. This last combination is more similar to the approaches of the other studies in literature. However, as it is going to be stressed more extensively later, more accurate and representative results can be obtained when one compares the same retinas over a period of time, and also when, instead of conducting a cross-sectional experiment, one includes follow-up examinations, in order to define a more accurate and meaningful trend.

Chapter 3

Proposed Framework and Biomarkers

3.1 Introduction

Retinal image analysis remains a very challenging and versatile field of study. Different regions of the retina, starting from the optic nerve head (large vessels) until the precapillaries (small vessels), include vessels of dissimilar size, having different properties (Gariano *et al.*, 1996). As already shown, most of the previous studies in literature that investigate the retinal pathologies, in regards to the changes in the retinal vasculature, present a few deficiencies; a) they are only focusing on limited number of features, b) the measurements within each retina are averaged prior to the analysis (no mixed effects model then), c) the analysis is made across independent groups, letting other factors that cannot be standardised to contribute to the errors, d) ignore the more localised changes in different areas of the retina and e) they do not account for the diversity of the retina, which includes many small and large vessels. This latter situation leads to collecting measurements that might show differences that cannot be safely attributed to the investigated pathologies, but rather to the inconsistency in selecting vessels of the same order, hence creating inflated unjustified differences.

The purpose of this chapter is to present a novel framework and methodology for

studying the retinal vasculature, for identifying the progression from diabetes to diabetic retinopathy, that also addresses the issues that were previously described. In addition, multiple geometric and haemodynamic features are introduced and described, which will be investigated and evaluated in the next chapters, in a manner that no other previous study has ever attempted.

3.1.1 Proposed Approach

Despite the many studies in literature, which try to make associations of mainly geometric alterations with a number of different pathologies, still no comprehensive framework and/or guidelines exist that combine a series of robust image processing approaches, statistical analysis tailored to the specific requirements and also a comprehensive evaluation of the discriminative power of the candidate biomarkers.

For the first time, in this thesis, such a framework and methodology are proposed, in order to shed some light on the vascular changes that occur during the progression and proliferation of diabetes. To achieve that, given the complexity of the retinal functionality and structure, a series of algorithms that either have been proposed in literature or were implemented as part of this thesis, are used. As a novel outcome, a total number of 101 geometric and haemodynamic features are measured or derived, analysed and evaluated as candidate biomarkers of progression to DR, extracted from the retina as a whole, as well as from some predefined areas of interest (Hove *et al.*, 2004).

The first major novel contribution is that a series of haemodynamic features are estimated, following the segmentation and connection of vascular trees, using both haemodynamic principles and the retinal geometry. These features are also analysed and evaluated as candidate biomarkers of progression from diabetes to diabetic retinopathy, by studying the same vascular trees in a four year period of time that includes the last three years of the diabetic eye and the onset of diabetic retinopathy

(progression study). In addition to the above the same exact procedure and analysis are followed for the non progression study, as well as for the various geometric features that come from different areas of the retina.

The second major contribution is that a novel framework for studying the retinal vasculature is presented, following a new methodology that aims to minimise the errors of the analysis, by building and comparing suitable and robust statistical and classification models. In addition, a tool is implemented that automatically matches the bifurcations across the different images of the same patient, handling the data in a way that lets us study the progression and the effects on specific segments over a period of time and on a large scale basis. The retina is also partitioned into predefined areas of interest, in order to be separately evaluated.

3.1.2 Chapter Overview

In the studies that follow, many algorithms are utilised for accommodating the needs of this thesis, e.g. for the segmentation and registration of the retinal images, for extracting features, etc., but it is also totally out of the scope of this thesis to implement any such algorithms. There are numerous algorithms in literature doing the above or purely focusing on the technical aspects, but no studies, to the best of my knowledge, exist that estimate haemodynamic features, combined with various image-level, bifurcation-level and localised (areas of interests) geometric ones, for studying the progression of the disease.

The **first part** of this chapter will give details about the data collection process, fulfilled in two phases. The **second part** will be devoted on describing the proposed framework, which can be used to conduct complete and comprehensive studies for identifying the retinal vascular alterations during the progression of a disease. At the **third part**, an introduction of all the geometric and haemodynamic features will be made alongside the processes that were involved for measuring and extracting them.

Finally the **last part** will conclude the chapter.

3.2 Data Collection and Studies

Prior to commencing any of the studies, an ethical approval was obtained by the ethical committee of the University of Lincoln in the United Kingdom. The most important part of all the studies presented in this thesis is the data collection process, which is driven by the planned experiments. In order to investigate and study the progression of the disease over a few year period, eliminating any errors coming from factors that cannot be taken care of in advance, like age, gender, lifestyle, other unknown diseases or the diversity of different retinas, a retrospective study was planned. To accomplish that, data coming from patients that progressed from diabetes to DR were needed, in order to be able to study the exact same segments over a period of time, identifying vascular changes that one can be confident enough that they can be attributed to the effects that diabetes has to the retinal vasculature. Therefore, in all of the studies of this thesis, only retinal fundus images and the age of each subject were used. No other information was extracted or utilised whatsoever.

The data were collected in two phases. At the initial phase, during which different methods and techniques were investigated and explored, ninety two macula-centred (temporal) and ninety two optic disc centred (nasal) digitised colour fundus images were utilised, coming from twenty three patients with Type 2 diabetes and no other known diseases, over a period of time that includes the last three years of diabetic eye and the onset of DR. This set of images came from a UK diabetic screening database, kindly provided to me by the Health Intelligence company in the UK.

The second phase, during which the proposed framework was developed and all the utilised methods were finalised (focus of this thesis), was conducted after obtaining additional data and hence extending the available images that were provided dur-

ing the first, exploratory phase. This was achieved by myself visiting the diabetic eye screening service of the Pilgrim hospital in Boston, UK, in order to carefully extract additional data. The selection of the images was very careful and thorough, reviewing every single of them, in order to make sure that the quality is high enough and consistent, for measuring all the features in an equally reliable manner within the same patient and across the whole image set. The images are graded (R0: Subject with diabetes but no DR, R1: Subject with diabetes and background retinopathy) by the DR graders of the hospital that the screening is taking place, according to the standards and protocols defined by the UK national health system (RCOphth, 2012). In addition, more than 5.000 patients' data were reviewed by myself, in order to finally extract the images from 104 of them, which had a full history of four consecutive years and the quality of the images was decent enough with minimal to no artefacts, so that all the vessels appear clearly (figure 3.1). This was also made possible by the fact that in the UK, diabetic patients are urged to monitor their retinas annually, albeit in practice many of them neglect it. This last fact made it very difficult and time consuming to personally extract all the images and go through so many patients' data to identify the suitable ones that follow the strict requirements. In fact, the vast majority of these data were inadequate, because either some images were missing, or the quality was not high enough or due to other factors, such as suffering also from other diseases (e.g. hypertension, glaucoma, etc.) or having cataract that blurs part of the image, which made them ineligible. Figure 3.2 shows two examples of how bad quality some of the retinal images can be, hence the careful selection of them. On the other hand, figure 3.1 shows an example of two qualified images, coming from the dataset that was used in the studies of this thesis.

In addition, a control group of 27 patients was also extracted, that had not progressed to diabetic retinopathy at the time of the data collection. For that group also the last four consecutive years of the disease were chosen and eight images for each

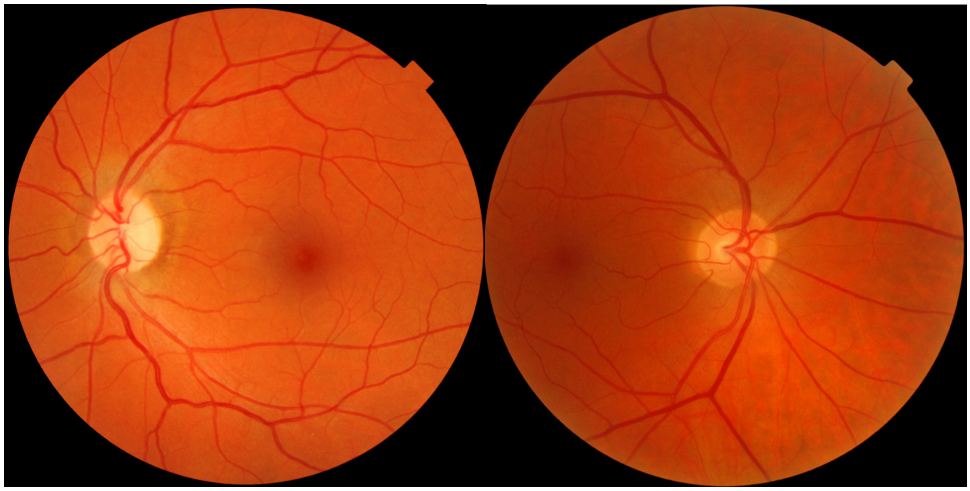

Figure 3.1: Two good quality and gradeable retinal images can be seen. In both images, all the areas of interest inside the retina can be seen clearly, without artefacts or blurriness whatsoever.

patient were extracted, just like with the other group. This group was used in order to make some additional analyses, beyond the progression study, both within this non-progressors' group and also in conjunction with the progressors' group, for comparison purposes.

For all of the above data, both the temporal (macula centred) and nasal (optic disc centred) images were extracted, with the latter category to be used for estimating the AVR, and the former for all the other studies. Moreover, the average group age for both the non-progressors and progressors groups was very close (mean age 54.5 (± 7.8) and 53.4 (± 8.4) for progressors and non-progressors, respectively). All the studies and results presented in this thesis were conducted using the final unified dataset with retinal images, which includes all the images that were collected during the two separate phases.

In addition, for implementing the proposed method of summarising the calibre of the parent vessel trunk (healthy group, chapter 5), another set of images, coming from an experiment organised by the School of Computer Science at the University of Lincoln, UK, was used. This set includes 25 images only from healthy subjects and was exclusively used for extracting bifurcations, utilising them for implementing the

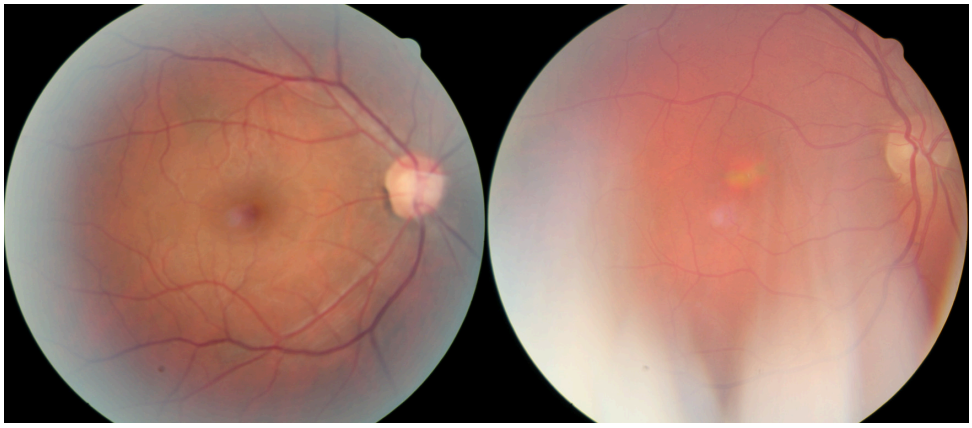

Figure 3.2: Two very low quality and ungradeable retinal images can be seen. The image on the left is a quite common situation, where the macula area is of acceptable quality, but the periphery is blurry. On the other side, the image on the right is a very bad one that cannot be used for DR grading or any other study whatsoever.

Table 3.1: Overview of the retinal image datasets used in the studies of this thesis

Datasets collection	Source	Number of Images (Subjects)	Resolution (FOV)	Nasal or Temporal	Studies
Initial Phase	Health Intelligence, Cheshire, UK	184 (23)	2352X1568 (45°)	Both	All, excluding non-progressors
Second Phase	Pilgrim hospital, Boston, UK	1048 (131)	3216X2316 (45°)	Both	All
Healthy Group	University of Lincoln, UK	25 (12)	3888X2592 (45°)	Temporal	Summarising calibre of parent vessel

healthy group part of the proposed method in chapter 5. A summary of the datasets can be found in Table 3.1.

All of the above data were used to create the six categories of experiments, as seen in Table 3.2, that will be analysed later, in order to identify any vascular changes during different stages of both diabetes and DR, evaluating at the same time all the candidate biomarkers.

Table 3.2: Summary of the studies of this thesis

Studies	Experimental Design	Description	Particulars
Category 1	Progressors' group - four year period of progression with matched segments (four groups - main study)	• Four years of progression, starting from the three years pre-DR until the onset of DR • Exact same segments and subjects are included over the whole four year period	Truly identify and associate any vascular changes with the progression of the disease, given that the same exact segments are investigated
Category 2	Progressors' group - four year period of progression with non-matched segments (four groups)	• Similar to category 1, but the bifurcations are not matched, so no measurements are discarded • Imbalanced groups	Direct comparison with category 1 to figure out whether matching the bifurcations affects the results
Category 3	Non-Progressors' group - four year period of progression with matched segments without DR (four groups)	• Same subjects and segments across a four year period, which does not include DR • Subjects with diabetes but have not progressed DR	Evaluate to what extent diabetes affects the retinal vasculature in a four year period and compare with the progression to DR, as outlined in category 1
Category 4	Non-Progressors' group - four year period of progression with non-matched segments without DR (four groups)	• Similar to category 3, but the bifurcations are not matched, so no measurements are discarded • Imbalanced groups	Direct comparison with category 3 to figure out whether matching the bifurcations affects the analysis and the results whatsoever
Category 5	Averaged non-progressors' group (one group) versus averaged progressors' group (one group) versus DR-only group (one group)	• Does not include the same subjects, but compares the three different independent groups of the disease's stage (milestones of progression)	The purpose is indeed to find out, if any significant vascular changes occur, during the transition from the non-progressed level to the progressed and finally to DR
Category 6	Progressors' group - four year period of progression with independent groups (different individuals in each group)	• Similarly to category 2, this category intends to show whether there is any notable reason to prefer the more demanding and sensitive analysis of category 1 • Different subjects and segments across the different groups	The stages of the disease are similar to category 1, but the subjects are independent and non-matched across the four year period. Direct comparisons with the results of category 1 will be made, given that the data reflect the same period of progression, but within independent groups

3.3 Proposed Framework

One of the main objectives of this thesis is to provide a complete framework for the analysis of the retinal vasculature, in order to show how more robust the statistical inferences can be that way. Some of the processes involved are automated and other could possibly become automated in the future, when some unsolved problems in retinal imaging are addressed, e.g. vessel connectivity and vessel labelling (Arteries vs Veins). However, it is worth pointing out, that the purpose of this thesis is not to implement the methods that automate the framework (e.g. image registration, image segmentation, etc.), rather to take all the necessary steps in order to conclude with more accurate, precise and meaningful statistical inferences and classification models.

The pipeline of the framework that is proposed includes the following general steps, which are going to be elaborated next.

1. **Image registration**
2. **Image segmentation and junction measurements**
3. **Extraction of the areas of interest**
4. **Automated matching of bifurcations**
5. **Extraction of candidate biomarkers**
6. **Statistical analysis and inferences**
7. **Feature selection and classification**

Mathworks Matlab software (Matlab, 2014) was utilised for all the algorithms that were used or developed for steps 1-5, and will be addressed in this chapter. Regarding steps 6 and 7, the open source language R (R Core Team, 2013) and Rstudio (RStudio Team, 2015) with some core packages were used, and they will be discussed in details during the next chapter.

It is worth noticing, that at the **initial** exploratory phase, until the proposed framework was conceived and developed, no framework existed and none of the above steps was included in any process whatsoever. The features were manually extracted, using a tool developed by (Al-Diri *et al.*, 2010) that will be described in details in the next sections, and only statistically evaluated. However, during the **main** (second) phase, in which this thesis and contributions are based on, the proposed framework was developed. In this framework, which is applied in practice to all the studies of the progression of the diseases (diabetes and DR, chapter 6), all 7 steps are automated, with the only manual intervention being the labelling of the vessels as arteries or veins. Detailed information and technical aspects about all these processes are given in the next sections.

3.3.1 Image Registration

3.3.1.1 General Approach

When studying the progression of a disease, it is paramount to be able to compare the same segments over a period of time, primarily for two reasons. Firstly, for eliminating, as far as possible, the errors due to other conditions that affect the measurements, e.g. age, gender, lifestyle, other unknown conditions etc., and secondly for being confident enough that any measured differences can be attributed to the effects of the disease, with all of these being conditional to a large enough amount of observations and subjects. Doing the opposite, i.e. comparing independent groups and unmatched measurements, requires a much larger sample, and that is because one can inconsistently include and measure large and small vessels across the whole cohort, leading to inflated unreal differences. An efficient and practical way to resolve some of these issues is to register the bifurcations that come from the same subjects, in order to automatically identify and include only the matched bifurcations across the images of the

same patient.

3.3.1.2 Registration Technique Adopted in this Thesis

For the purpose of my studies, the open source library VLFeat (Vedaldi & Fulkerson, 2008) was used, implemented in Matlab, created by Vedaldi and Fulkerson, which offers a complete registration algorithm, based on the Scale-Invariant Features Transform (SIFT) and homography estimation with Random sample consensus (Ransac). Following the above processes, the images can be registered prior to segmenting them, in order the bifurcations to approximately be aligned, letting us automatically match the segmented bifurcations, based on their individual coordinates.

In brief, SIFT combines together a feature detector and a feature descriptor (Lowe, 1999). The detector extracts from the image a number of attributed regions in a way that is consistent with some variations of the illumination, the viewpoint and other viewing conditions. The descriptor associates to these regions a signature, which identifies their appearance in a compact and robust way. Then a regular Euclidean distance is used, i.e. the distance between descriptor i from image A and descriptor j from image B (Vedaldi & Fulkerson, 2008). For each descriptor in image A, the algorithm calculates the distance to all the descriptors in image B, matching it with the closest one from B; otherwise the descriptor in A is not matched at all. Following that, a homography is estimated that describes the transformation between the two images (equation 3.1) (Kriegman, 2007).

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \Leftrightarrow x_2 = Hx_1 \tag{3.1}$$

Since the homography is defined up to scale, it has 8 degrees of freedom, hence requiring four point correspondences to estimate (Dubrofsky, 2009). At this point, since

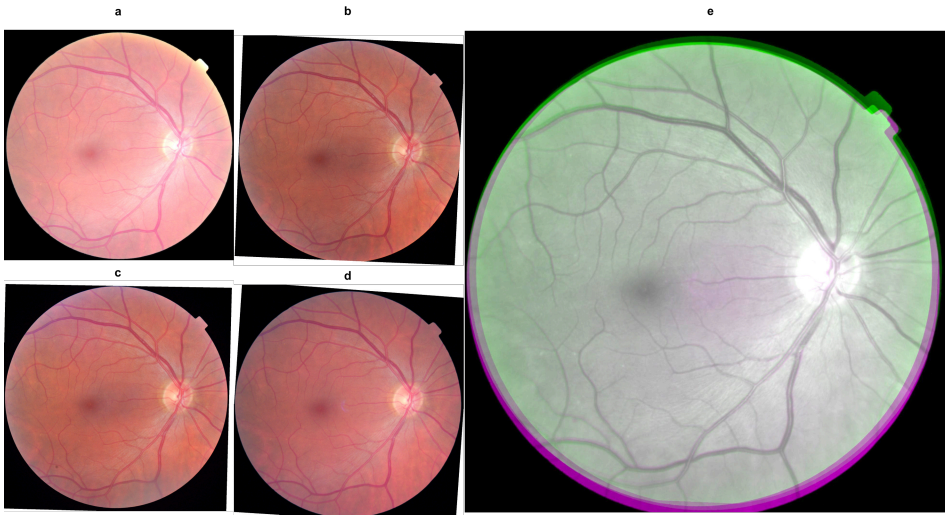

Figure 3.3: Four registered images and the fused one. a: Reference image. b,c,d: Subsequent registered images. e: Fused image. The bifurcations are now adequately aligned, in order to be reliably matched.

not all correspondences are correct, Ransac is used to find a set of good matches (inliers). This is achieved by randomly choosing 4 matches, estimating the homography by using these matched pairs, and counting how many other matches agree. A match between the point of image A = (Xa,Ya) and the point of image B = (Xb,Yb) agrees with the transformation, if the transformed point H(A)=(X'a, Y'b) is close to point B (Fischler & Bolles, 1981).

The VLfeat algorithm (Vedaldi & Fulkerson, 2008) adjusts the resolutions of the images, for creating a mosaic. In my case, since the four images of the four consecutive years need to be aligned altogether, I parametrised the way that the algorithm creates the vectors for each image, in order to preserve the same resolution, and also each of them to be aligned with the unaltered reference image. In this way, the same coordinate system is applied in all the images and the resolutions remain the same across all of them. In figure 3.3 an example can be seen that shows four registered images both separately and fused together.

3.3.2 Image Segmentation and Junction Measurements

3.3.2.1 General Approach

The main step for studying every retinal disease is the image segmentation. This process must follow the image registration, otherwise matching the bifurcations and studying the progression of any disease within the same patients will be impossible. It is an indispensable part that is needed in order to obtain the vascular trees and bifurcations, which will allow any type of vascular measurements. Every candidate biomarker, either geometric or haemodynamic, requires an accurate segmentation with profiles of the vessel segments, before making any estimations. Any segmentation algorithm that can provide binary trees, profiles of the vessel segments and geometric measurements can be used in the pipeline of this framework.

3.3.2.2 Image Segmentation Techniques Used in this Thesis

Following the image registration, the registered images are then segmented, using a segmentation algorithm, implemented by Bashir *et al.*, whose accuracy, reliability and robustness has been evaluated in literature (Al-Diri *et al.*, 2009). This algorithm works by growing a ”Ribbon of Twins” active contour model, which uses two pairs of contours to capture each vessel’s edge, while maintaining width consistency. The algorithm starts using a generalised morphological order filter to identify approximate centrelines. After identifying the vessel segments, the network topology is also determined, using an implicit neural cost function, defined by using self-organising feature maps, in order to resolve junction configurations. The output of these neural networks is a “novelty signal”, which indicates whether the input vector is similar to the vectors presented during the training stage; they are therefore ideal to provide an implicit cost function, characterising whether a configuration is consistent with normal expectations. The local junction measurements, such as vessel widths, angles and derivatives

of those, are also estimated by this algorithm. In addition the lengths of the vascular trees are also returned, which is one of the parameters, together with the vessel widths, that are used for the estimation of the haemodynamic features.

A second segmentation algorithm (Hunter *et al.*, 2005), which was used for the fractal dimension calculation and tortuosity measurements, is the one proposed by Hunter *et al.* It is based on a non linear tram-line filter, which is applied at a number of orientations, rotating the kernel by 22.5° and selecting the strongest response at each pixel. This process calculates the contrast between a central line and satellite tram-lines (two short parallel segments). Order statistics filtering was also used, where the response values within the inner and outer filters are sorted. Finally the differences between the third darkest and third lightest values are obtained, and the resulting strength map is thresholded in order to receive the segmentation. The noise and the false positive segments that occur are sorted out a) by doing some morphological thinning to reduce vascular segments to a single-pixel ones, b) by removing small segments and c) by a clean-up of the remaining segments, including hole-filling and pruning of small side-spurs.

The segmentation algorithm in (Hunter *et al.*, 2005) was preferred over the one in (Al-Diri *et al.*, 2009) for the above estimations (fractal dimension and tortuosity), because it provides a) binary vessel segmentations needed for the fractal dimension estimations, and b) the profiles of the vessel segments are returned in a single file (in contrast to the other segmentation algorithm), as required by the algorithm that calculates the tortuosity metrics. Figure 3.4 shows an example of the segmentation of two images used in this thesis, as obtained from the two algorithms.

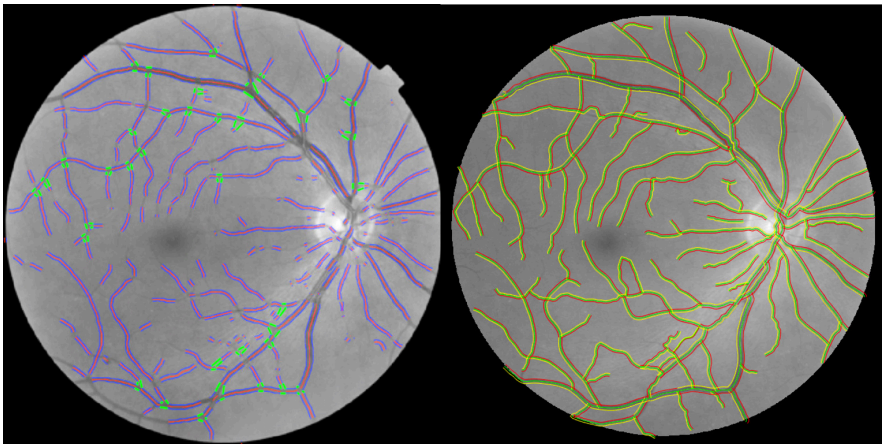

Figure 3.4: On the left, a segmentation from the algorithm proposed in (Al-Diri *et al.*, 2009) can be seen. On the right is an example of the segmentation based on the method proposed in (Hunter *et al.*, 2005).

3.3.3 Extraction of the Areas of Interest

3.3.3.1 General Approach

An important part and novelty of the studies presented in this thesis is the examination and exploitation of the possible importance of different areas of the retina, in addition to the retina as a whole. Different regions inside the retina include different sizes of vessels, which can be larger in the periphery (E3) and near the optic nerve head, and finer around the macular area (E1 and E2). Therefore, it is likely the disease to be affecting the vasculature differently, depending on the size of the vessels, as was explained in chapter 2. To make sure that all these possible scenarios are taken into account, it is recommended to include in every study of any retinal disease the areas that are described next, as an additional step in order to identify possible localised alterations.

3.3.3.2 Recommended Approach for the Areas of Interest

To do so, the areas that Hove *et al.* proposed were adopted in this study, which were defined based on previously published knowledge of the development of retinopathy.

This included information about the early lesions that develop temporally from the fovea (Dobree, 1970; Taylor & Dobree, 1970), the correlation between blood pressure and the location of the lesions around the arcades (Bek & Helgesen, 2001), about how the occurrence of lesions in the periphery is associated with the retinopathy grade (Group *et al.*, 1991), and also how the severity of retinopathy is connected to the distance of the lesions from the fovea area (Group *et al.*, 1987). To come up with these areas, they also studied 11 Type 2 patients with diabetes, who had developed visual loss secondary to DR. They took into account the total number and area of all the haemorrhages and exudates located inside the three ellipses (E1, E2 and E3) and circles (C1 and C2), as seen in figure 3.5 (Hove *et al.*, 2004). The two circles, with their centre in the fovea area, have a diameter of one and two diameters of the optic disc. The ellipse E1 includes the retinal area temporal from the fovea, the ellipse E2 includes the macular area and the ellipse E3 includes the upper and lower vascular arcades.

Following up the patients (Hove *et al.*, 2004), they also observed an increase in the total number and area of the haemorrhages inside E1 and outside E3, only an increase in the area of the haemorrhages between E1 and E2 and an increase in the number of the haemorrhages between E1 and E2; no significant changes were observed around the arcades whatsoever. The number and area of haemorrhages and exudates were found to be significant to the later development of vision-threatening diabetic maculopathy.

All of these observations they made (Hove *et al.*, 2004), support the evidence that retinal hyperperfusion co-occurs with diabetic retinopathy, due to malfunctions in tone regulation of smaller arterioles that supply the microcirculation (Kohner *et al.*, 1995; Grunwald *et al.*, 1996; Bek, 1999). On the other hand, increased arterial blood pressured and impaired autoregulation facilitate the apprearance of lesions around the larger arterioles near the arcades (Bek & Helgesen, 2001; Rassam *et al.*, 1995).

Finally, in definining these areas, no information about the vascular geometry was taken into account. Therefore, our purpose is to find out whether the vessel widths,

angles and tortuosity are affected in these areas, given that the location of the lesions seems to be connected to the later development of vision-threatening disorders. The

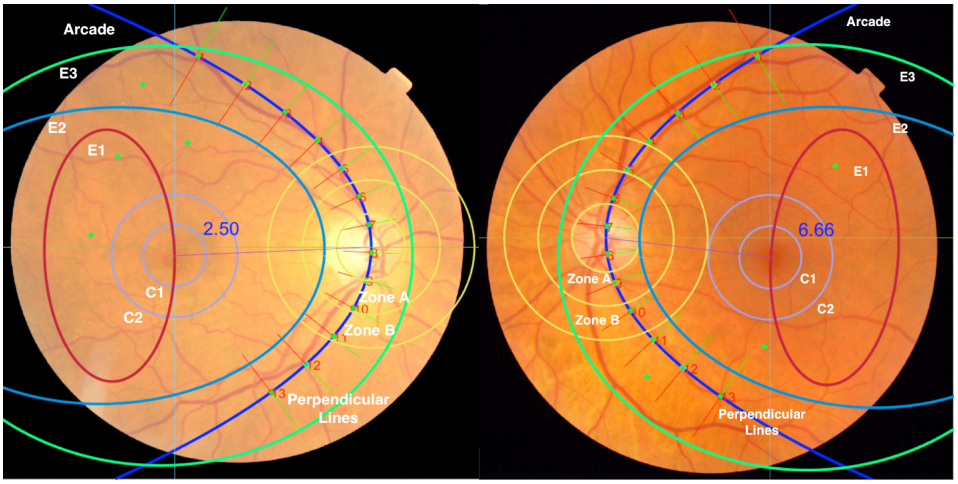

Figure 3.5: Two examples of how the areas of interest are defined. E1, E2 and E3 are the ellipses, C1 and C2 are the two circles around the fovea, and Zone A and Zone B are the excluded and included areas for the AVR estimation respectively. The arcade is a parabola with its vertex at the optic disc centre. The perpendicular lines, which cross the arcade in various points, help to standardise the selection of arteries and veins for the estimation of the local arteriovenous ratio. All these areas, across all the images, are either semi-automatically or automatically generated.

ellipses were generated based on the general ellipse equations and according to the methodology in (Hove *et al.*, 2004), adjusting however the distance and dimension parameters to account for the difference in the resolution of the images. Similarly, the circles C1 and C2 are generated according to the general circle equation. In our studies, the values of the two radii of each ellipse, **a** and **b**, are 227 and 450, 530 and 750, and 750 and 880 pixels, for E1, E2 and E3, respectively.

In addition to the above areas, in figure 3.5, some additional areas can be seen. Zone A and Zone B are the two areas that were proposed in (Knudtson *et al.*, 2003; Hubbard *et al.*, 1999), in order to estimate the central retinal artery (CRAE) and vein (CRVE) equivalent, as well as their quotient the arteriovenous ratio (AVR). The region of interest (Zone B) is defined as shown in figure 3.5, and includes the region where the edges of the vessels course through at 0.5 to 1.0 disc diameters from the optic disc margin. The region between this area and the optic disc is excluded, as not having the

vessels attained their status inside the retina yet. Within this area, the six largest veins and the six largest arteries are measured, following an iterative procedure of pairing up the largest vessels with the smallest ones, until a final single number is obtained. More information about this is provided in the next sections and even more details about the methodology is given in chapter 5.

An important step that was corrected for this thesis's studies, which improves the accuracy of the estimation of the areas, is the correction of the rotation. In figure 3.5, especially in the image on the right, it can be clearly seen, that the centre of OD and the centre of fovea are not in the same horizontal line, having a small offset of 6.66 degrees. This is estimated, by calculating the angle that is created between the line parallel to the x axis that crosses the centre of the OD, and the line that connects the centres of OD and fovea. For visualisation purposes, the lines and the angles appear on both images of figure 3.5. Following that, all three ellipses are then rotated according to this offset, in order to finally include the correct areas, which are consistent across the consecutive years. Finally, the algorithm is adjusted so as to account for both left and right eyes, since some of the equations for the estimations of the areas need a change of sign (minus to plus and vice versa). That is achieved by comparing the estimated coordinates of the OD with the x axis, i.e. whether the coordinates are in the first half, e.g. between 0 and $x/2$, or in the other half, between $x/2$ and x (x is the horizontal dimension of the image in pixels).

In addition, the arcades/parabolas that can be seen in the two images in figure 3.5 follow approximately the main vein and artery that exit the optic disc and head towards the retinal periphery. Their vertex point is on the centre of OD and they are symmetric in respect to the x axis. In various points along the arcades, the tangents and then the slopes are calculated, in order to draw the perpendicular lines that cross this point, at both sides of the arcade. The arcade has its vertex at the centre of the OD and its curvature is empirically selected so as to approximately follow the arteries and veins

that span towards the periphery on the side of the macula area.

Optic Disc and Fovea Localisation

As can be inferred, the important part for defining all these areas is the optic disc and fovea localisation. There are two steps involved in this process. At the first step, both fovea and optic disc are automatically estimated, based on an implementation provided by Andrew Hunter, University of Lincoln (Lowell *et al.*, 2004a), for which more details will follow. Given that here the priority is not the accurate and automated localisation of the fovea and OD, but the areas of interest and the arcade, the process involves an extra step. After the algorithms estimate the coordinates of the centres of the OD and the fovea across the whole image set, and before any areas are defined, the user is prompted to confirm that both the fovea and OD are correctly estimated. If so, the process continues; if not, then the user manually clicks and draws the margins of the OD and fovea, overpassing the automated process. In that way the goal is achieved, which is to estimate as accurately as possible the areas, in order to reliably extract the features needed for the progression and non-progression studies.

As aforementioned, in order to automate the process of defining the areas of interest, OD and fovea localisation algorithms were incorporated, which precede the execution of the algorithm that defines the areas. The method, developed by (Lowell *et al.*, 2004a), provides us with the initial coordinates needed to proceed with the whole estimations. As far as the OD localisation is concerned is based on the correlation with a specialised filter, which matches key elements of the structure of the optic disc. The correlation peak gives the approximate centre of the optic disc. The template consists of a Laplacian of Gaussian with a vertical channel in the middle to correspond to a major vessel band (Lowell *et al.*, 2004a). This template is correlated with the intensity component of the retinal image. The different values of the optic disc radius that the authors tried are based on estimates of the pixel to micrometer ratio, using an average value of $1850\mu m$ for the OD diameter (Bertelsen *et al.*, 2013; Patton *et al.*, 2006b).

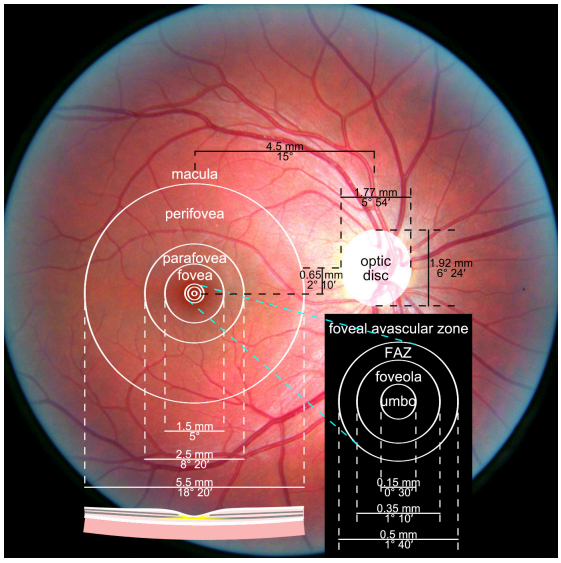

Figure 3.6: A very illustrative example, showing the dimensions and the exact areas that correspond to the macula and OD inside the retina. Source: Danny Hope from Brighton & Hove, UK - https://en.wikipedia.org/wiki/Macula_of_retina/media/File:Macula.svg.

Regarding the fovea localisation, this relies on the OD one. As can be seen in figure 3.6, the positions of the macula lutea and the OD inside the retina are anatomically related. Based on that fact, the fovea localisation starts with a given starting position for the OD centre. Similarly to the OD, a Laplacian of Gaussian filter is used to match the fovea, assuming macula-centred images. Given an initial offset from the OD, which is trimmed according to the maximum and minimum OD diameters, a correlation search is conducted for finding the best foveal match on a fixed granularity, based on the OD centre. The sampling is made on a wide spaced grid around the estimated position that is more or less anatomically defined. If the best correlation is below a global threshold, the expected position is used instead, followed by firstly resampling on a finer grid around the peak of the first search, and then on a single pixel granularity grid, which gives us the final position (Lowell *et al.*, 2004a).

It is worth pointing out that all of the specific techniques that are adopted in this thesis to implement the framework (SIFT and RANSAC for image registration, active contours for image segmentation, etc.) could be replaced with alternative ones, as long

as they can accommodate the requirements and the purpose of this framework.

3.3.4 Matching of the Bifurcations

Just like the optic disc and fovea, bifurcations are also anatomical landmarks inside the retina. The position of each bifurcation is more or less fixed, with only negligible changes occurring, just like with the retinal vessels, mainly due to the dynamic nature of the retina. Given that the vast majority of the candidate biomarkers in this thesis are extracted in reference to the bifurcations, then a clever way to make sure that the automated system will select and compare the same segments during the investigated periods of time, is matching the bifurcations. In that way, we make sure that the same exact vessel segments are compared, minimising the chances of making wrong comparisons. Therefore, although the bifurcations themselves, or more precisely the bifurcation points, do not have any particular value as candidate biomarkers, their almost fixed spatial location and the vessels that form the bifurcation are highly important and indispensable part of the studies of this thesis.

The purpose of the previously presented areas is to finally be able, in conjunction with suitable segmentation and vascular measurements' algorithms, to extract each feature, matched across the whole period of time, and at the same time group them according to the area they are coming from. The algorithm is also checking whether a bifurcation or a given segment belongs to this area, i.e. if it is inside, on, or outside of it. Therefore, each point's coordinates are checked against the equations of each area (ellipses, circles, arcade and perpendicular lines). At the end, all the bifurcations can be organised in accordance with the area they belong to, store the measurements and create the relevant feature. The same exact thing occurs and in the case of matching the junctions across the whole investigated period. The condition defines to firstly identify the same junction in all four images of the progression, and if successful, to check in which area they have to be assigned to.

To clarify the process further, the area of ellipse 1 includes everything that is inside or on the ellipse 1; the area of ellipse 2 includes everything that lies between ellipse 1 and ellipse 2; and likewise, the area of ellipse 3 includes the enclosed area that lies between ellipse 2 and ellipse 3, including Zones A and B. It is also worth pointing out that Zones A and B, the arcade and the perpendicular lines are not exclusive, as the ellipses are among themselves, because they serve a different purpose, i.e. to calculate the CRVE, CRAE, AVR and the local arteriovenous ratio. For instance, a point can both belong to Zone B and ellipse 2 (or ellipse 3), given the different purpose that these two areas serve. In particular, the perpendicular lines help both to visually identify the vein and artery that serve the same area of the retina and also, alongside the segmentation algorithm, to return the widths of the two points; the one on the vein and the one on the artery.After that, the local arteriovenous ratio is calculated, near the OD, and after one and two branchings of both vein and artery.

Regarding the process of the **local arteriovenous ratio**, the first measurement (near OD) is used as a reference, which does not have any importance, since the value can be affected by the possible different order of branching of the vein and the artery before they extrude the OD. However the second and third measurements can be controlled, and are indeed selected after the vein and artery branch once and twice, towards the same direction. Therefore, it will be later evaluated, whether the local arteriovenous ratios' differences (point 2 minus reference, point 3 minus reference and point 3 minus point 2), present any significant differences across the years of progression.

Regarding the final iterative process for matching all the segments across the four year period, this is based on multiple conditionings.

1. Firstly, a segment is manually labelled as artery or vein. All the labellings were fulfilled by two raters, ending up with a 97,5% Kappa inter-rater agreement, hence keeping only the mutually agreed segments and discarding the rest.

2. Secondly, the images that come from the same patient are identified, by search-

ing and matching their unique identification numbers.

3. Following that, only a vein or artery (A or V in our data) is selected at a time, and thus restricting the search to the respective type of segment, until all of them have been identified.

4. The process initially commences with the x and y coordinates of e.g. arteries, and continues with the selection of the first pair of coordinates of the first image of the first subject. The search is only conducted around the same coordinates of the other 3 images of the same patient, in a 5x5 window, which is adequate in order not to miss a not perfectly registered bifurcation, but at the same time not mistakenly include another different nearby bifurcation (unlikely to have another bifurcation of the same vessel type, A/V, that close though. In any case the algorithm will prompt the user to choose one of them). If both x and y coordinates are within the margins of the defined windows in all the four images, then the geometric measurements (widths, angles and derivatives of them), that correspond to the four matched bifurcations are saved in the respective matrix, including both the type of feature and the period they are coming from (e.g. for the progression study: three years pre-DR, two years pre-DR, one year pre-DR or onset of DR).

5. Finally, the identified entries are replaced with zero values and the previous steps are repeated. After the process has finished, the remaining, non-zero values are returned, which correspond to the non-matched bifurcations, for possible further investigations. The matched bifurcations are now correctly stored. The same process is also repeated for the veins.

3.3.5 Candidate Biomarkers

As was previously stressed, this is the first time that such a comprehensive study is conducted, investigating such a great deal of geometric and haemodynamic features. This section will give information regarding the investigated features and the tools that were used to extract them, prior to describing the experiments and the findings.

In the previous chapters, it was shown how complicated the retinal functionality is and how previous studies have investigated the effects of hyperglycaemia and oxidative stress to the retinal tissue, with the consequences being observed in the geometry and the heamodynamics.

The retinal vascular network is non-uniform, including both symmetrical and assymetrical bifurcations. Although retina is curved in space, its vascular network mostly lies on a two-dimensional surface, therefore the angle δ, which the parent makes with the plane, is either zero or near zero (Zamir, 1978; Al-Diri, 2008).

The vascular branching is a main characteristic of the cardiovascular system. The simplest and most common form of branching is that in which a single stream of blood is divided into two distinct streams. These two streams may differ significantly, in terms of size and direction (Al-Diri, 2008).

In the beginning of the 20th century the first attempts to understand the branching angles are found (Thompson, 1917), with the comprehensive and extensive studies arriving in 1926 (Murray, 1926b,a). Since that, by the mid of 70s, many more studies were conducted, dealing with the measuring of the properties of the vascular network and its patterns (Kamiya & Togawa, 1972; Kassab & Fung, 1995; Schreiner & Buxbaum, 1993; Zamir, 1976, 1999).

Murray's Law

The two main early papers of Murray (Murray, 1926b,a) focused on defining the principles of minimum work, for predicting the relationship between radii and blood flow, applying them to predict the branching angles. The vascular system with a uni-

formly larger vessel radii will require a larger volume of blood and a higher metabolic rate, while requiring less pumping power for circulating the blood. These facts led to the hypothesis that the radii of vessels in the cardiovascular system are the result of a compromise between the pumping power H_p (equation 3.2) and the cost of maintaining the blood volume *u* (equation 3.3).

$$H_p = \frac{8\rho l Q^2}{\pi r^4}, \tag{3.2}$$

where ρ is the viscosity of the blood, Q is the flow of the blood in the vascular system and *r* is the segment radius.

$$u = K\pi r^2 l, \tag{3.3}$$

where *K* is a constant.

3.3.5.1 Geometric Features

In this chapter, much of the focus goes to the bifurcations and the vascular measurements that can be extracted from them. In figure 3.7 (Al-Diri *et al.*, 2010), the main configuration of a bifurcation can be seen.

As shown in chapter 2, there are numerous studies in literature, having investigated the importance of the retinal geometry in studying different pathologies, however some of the underlying biological processes that might trigger these alterations, are still not well-known (Stitt & Curtis, 2005; Stitt *et al.*, 2013; Singh *et al.*, 2014).

This section will provide details about the general concepts of the candidate biomarkers that appear in Table 3.3 and the methods that were used to measure and/or extract them.

Widths & Angles

The vessel widths (or interchangeably calibres, or diameters) and angles (1-24,

Table 3.3: Investigated features as used in the statistical analysis

Features (Short form)	
1. Alpha value of arteries (**AlphaA**)	50.Parent's arterial width in ellipse 3(**Ellipse3parA**)
2. Branching coefficient of arteries (**BetaA**)	51.Small child's arterial width in ellipse 3(**Ellipse3ch1A**)
3. Bifurcation index of arteries (**LambdaA**)	52.Large child's arterial width in ellipse 3(**Ellipse3ch2A**)
4. Diameter ratio λ_1 of arteries(**Lambda1A**)	53.Arterial branching angle in ellipse3(**Ellipse3ThetaA**)
5. Diameter ratio λ_2 of arteries(**Lambda2A**)	54.Central Retinal Artery Equivalent(**CRAE_LEON**)
6. Arterial branching angle(**ThetaA**)	55.Central Retinal Vein Equivalent(**CRVE_LEON**)
7. Adjusted Gamma ratio of arteries (**Adj.GammaA**)	56.Arteriovenous ratio(**AVR_LEON**)
8. Parent's arterial width(**Width_parentA**)	57.Central Retinal Artery Equivalent(**CRAE_KNUD**)
9. Large child's arterial width(**Width_child1A**)	58.Central Retinal Vein Equivalent(**CRVE_KNUD**)
10.Small child's arterial width(**Width_child2A**)	59.Arteriovenous ratio Knudtson(**AVR_KNUD**)
11.Average junction's arterial widths(**Width_allA**)	60.Average tortuosity (Chandrinos *et al.*, 1998) (**mean_phi**)
12.Arterial Angle to Branching coefficient ratio(**Angle.BC_A**)	61.Median tortuosity(**median_phi**) 62.Third quartile tortuosity(**75th_phi**)
13.Alpha value of veins (**AlphaV**)	63.Standard deviation of tortuosity(**sd_phi**)
14.Branching coefficient of veins (**BetaV**)	64.Average tortuosity (Hart *et al.*, 1999)(**mean_tau**)
15.Bifurcation index of veins (**LambdaV**)	65.Median tortuosity(**median_tau**)
16.Diameter ratio λ_1 of veins(**Lambda1V**)	66.Third quartile of tortuosity(**75th_tau**)
17.Diameter ratio λ_2 of veins(**Lambda2V**)	67.Standard deviation of tortuosity(**sd_tau**)
18.Branching angle of veins (**ThetaV**)	68.Average tortuosity(Grisan *et al.*, 2008)(**mean_psi**)
19.Adjusted Gamma ratio of veins (**Adj.GammaV**)	69.Median tortuosity(**median_psi**)
20.Parent's venular width(**Width_parentV**)	70.Third quartile tortuosity(**75th_psi**)
21.Large child's venular width(**Width_child1V**)	71.Standard deviation tortuosity(**sd_psi**)
22.Small child's venular width(**Width_child2V**)	72.Combined tortuosity(Poletti *et al.*, 2011)(**Tort_blend**)
23.Average junction's widths of veins(**Width_allV**)	73.Combined tortuosity in ellipse 1(**Tort_ellipse1**)
24.Venular Angle-to-Branching coefficient ratio(**Angle.BC_V**)	74.Combined tortuosity in ellipse 2(**Tort_ellipse2**) 75.Combined tortuosity in ellipse 3(**Tort_ellipse3**)
25.Fractal dimension(**Fractal**)	76.Parent's arterial wall shear stress(**wssparentA**)
26.Lacunarity(**Lacunarity**)	77.Arterial pressure in the whole junction(**pressureA**)
27.Local arteriovenous ratios' difference between the second and the reference values(**AVmid_AVbeg**)	78.Parent's arterial velocity(**vparentA**) 79.Parent's arterial Reynolds number(**ReparentA**)
28.Local arteriovenous ratios' difference between the third and the reference values(**AVmid_AVbeg**)	80.Large child's arterial wall shear stress(**wsschild1A**) 81.Large child's arterial flow rate(**qchild1A**)
29.Local arteriovenous ratios' difference between the third and the second values(**AVmid_AVbeg**)	82.Large child's arterial flow velocity(**vchild1A**) 83.Large child's arterial Reynolds number(**Rechild1A**)
30.Parent's venular width in ellipse 1(**Ellipse1parV**)	84.Small child's arterial wall shear stress(**wsschild2A**)
31.Small child's venular width in ellipse 1(**Ellipse1ch1V**)	85.Small child's arterial flow rate(**qchild2A**)
32.Large child's venular width in ellipse 1(**Ellipse1ch2V**)	86.Small child's arterial flow velocity(**vchild2A**)
33.Venular branching angle in ellipse1(**Ellipse1ThetaV**)	87.Small child's arterial Reynolds number(**Rechild2A**)
34.Parent's arterial width in ellipse 1(**Ellipse1parA**)	88.Arterial pressure according to PinQout(**PinQoutA**)
35.Small child's arterial width in ellipse 1(**Ellipse1ch1A**)	89.Parent's venular wall shear stress(**wssparentV**)
36.Large child's arterial width in ellipse 1(**Ellipse1ch2A**)	90.Venular pressure in the whole junction(**pressureV**)
37.Arterial branching angle in ellipse1(**Ellipse1ThetaA**)	91.Parent's venular flow velocity(**vparentV**)
38.Parent's venular width in ellipse 2(**Ellipse2parV**)	92.Parent's venular Reynolds number(**ReparentV**)
39.Small child's venular width in ellipse 2(**Ellipse2ch1V**)	93.Large child's venular wall shear stress(**wsschild1V**)
40.Large child's venular width in ellipse 2(**Ellipse2ch2V**)	94.Large child's venular flow rate(**qchild1V**)
41.Venular branching angle in ellipse2(**Ellipse2ThetaV**)	95.Large child's venular flow velocity(**vchild1V**)
42.Parent's arterial width in ellipse 2(**Ellipse2parA**)	96.Large child's venular Reynolds number(**Rechild1V**)
43.Small Child's arterial width in ellipse 2(**Ellipse2ch1A**)	97.Small child's venular wall shear stress(**wsschild2V**)
44.Large child's arterial width in ellipse 2(**Ellipse2ch2A**)	98.Small child's venular flow rate(**qchild2V**)
45.Arterial branching angle in ellipse2(**Ellipse2ThetaA**)	99.Small child's venular flow velocity(**vchild2V**)
46.Parent's venular width in ellipse 3(**Ellipse3parV**)	100.Small child's venular Reynolds number(**Rechild2V**)
47.Small child's venular width in ellipse 3(**Ellipse3ch1V**)	101.Venular pressure according to PinQout(**PinQoutV**)
48.Large child's venular width in ellipse 3(**Ellipse3ch2V**)	
49.Venular branching angle in ellipse3(**Ellipse3ThetaV**)	

Table 3.3) were estimated using two methods. At the **initial phase** of my studies, the semi-automatic computerised method was used (Al-Diri *et al.*, 2010). This algorithm, proposed by Bashir *et al.*, creates a graphical user interface that gives the observer the

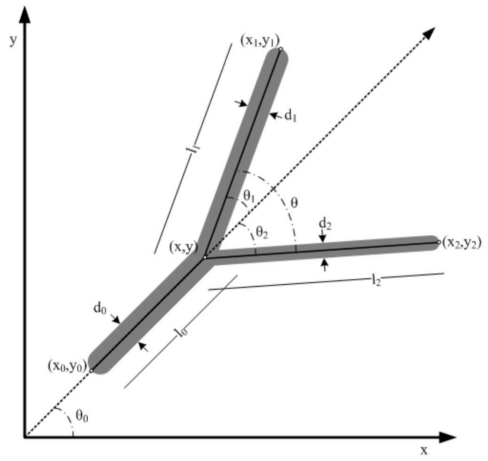

Figure 3.7: A schematic drawing of the configuration of a bifurcation. the symbols represent the diameters (d), the lengths (l) an the angles of the bifurcations (θ, θ_1 and θ_2). The subscripts 0, 1 and 2 are used throughout this thesis to denote the parent vessel, the larger branch and the smaller branch, respectively. The points (x_0,y_0), (x_1, y_1) and (x_2, y_2) are the coordinates of the end points of each segment, whereas the point (x, y) is the position of the bifurcation. Diagram taken from (Al-Diri, 2008).

ability to label specific junctions inside the retina and then mark them manually. The observer firstly selects two points along the approximate centreline and then a third one on one of the segment edges. Finally, an aligned rectangle is drawn, whose width and direction can easily be adapted, until the observer is confident that the rectangle is properly aligned. This method is effective as the width is defined along a part of the vessel, without needing segment profiles. This is assisted by the fact that the tool presents the image junction zoomed up, allowing taking measurements with sub-pixel accuracy. The angle in turn, is calculated by firstly clicking to place an intersection point, followed by clicking to select the ends of the three vessel centrelines. The final drawn lines can be adjusted until the observer is satisfied with the result.

At the **second phase**, for the main study, a fully automatic tool was used (Al-Diri *et al.*, 2009), which is a subpart of the segmentation algorithm (section 3.3.2). This active contour algorithm provides precise widths and angles measurements with sub-pixel average width errors.

For every image that was used in these studies, the diameter of the optic nerve head

was also estimated with the semi-automatic tool. Although the images are rescaled to have the same resolution (1700X1700 pixels), in order to compensate for possible magnification errors, each width measurement was standardised, dividing by the OD diameter of the respective image. However, the measurements of the widths across all the images, did not suggest the existence of important outliers, which would highlight possible measurement or calibration errors. In fact, the standardisation (versus non-standardised values) did not affect the direction of the results.

Another comparison was also made, between a small sample (210 widths, 70 angles) of pre-registered and post-registered semi-automatically measured bifurcations. The purpose was to find out whether the two-dimensional registration algorithm affects or not the estimation of the branching angles and/or vessel widths. A paired t-test did not suggest any significant differences (p-values equal to 0.35 and 0.48 for widths and angles respectively).

In addition to the above direct measurements, some additional relative ones have also been proposed in literature and are adopted in theses studies.

- The *bifurcation index* λ, or otherwise degree of asymmetry (Zamir, 1978). The value of λ is defined by diving the diameter of the smaller child d_2, by the diameter of the larger one, d_1. Therefore the value of λ is greater than 0 and less or equal to 1.

$$\lambda = \frac{d_2}{d_1} \tag{3.4}$$

- The *asymmetry ratio* α, which is the cross-sectional area of the smaller branch divided by the larger one. The final value of α is greater than 0 and less or equal to 1 (Zamir, 1978).

$$\alpha = \frac{d_2^2}{d_1^2} \tag{3.5}$$

  Both λ and α are indexes of the degree of asymmetry. The bifurcation is symmetrical when the values of λ or α equal unity, meaning that the two branches

have the same diameter. As the values converge towards zero, the bifurcation is considered to be asymmetric or non symmetric.

- The *diameter ratios*, λ_1 and λ_2, are the diameters of the children branches divided by the diameter of the parent vessel (Karch *et al.*, 2000).

$$\lambda_1 = \frac{d_1}{d_0}, \lambda_2 = \frac{d_2}{d_0} \tag{3.6}$$

- The *area ratio* or *branching coefficient* β, is the sum of the cross-sectional areas of the two children branches divided by the one of the parent vessel (Zamir, 1978; Knudtson *et al.*, 2003).

$$\beta = \frac{d_1^2 + d_2^2}{d_0^2} \tag{3.7}$$

- Branching angle θ (defined by the points (x_0,y_0), (x_1, y_1) and (x_2, y_2) in figure 3.7) is the one I will be measuring and evaluating throughout this thesis.

$$cos(\theta_0) = \frac{x - x_0}{((x - x_0)^2 + (y - y_0)^2)^{1/2}} \tag{3.8}$$

Fractal Dimension & Lacunarity

Fractal dimension (FD) and lacunarity (25-26, Table 3.3) are another two important features that are included in these studies. The former can give us a measure of complexity of a structure, as long as it can be considered a fractal, just like the retinal network (Mainster, 1990). The latter is a measure of heterogeneity of a fractal structure (Tolle *et al.*, 2008).

Fractality

Fractals present various degrees of self-similarity of different scales. Human retina has been found to almost be a self-similar structure, thus being possible to be analysed

as such, giving us a measure of complexity, and letting us also investigate, whether it changes at different periods of time (Family & Vicsek, 1991). Its discriminatory power was also evaluated within the classification system, in conjunction with the other features. Higher values of FD indicate more complex structure.

For the calculation of FD, the well established method of box-counting algorithm (Minkowski - Bouligand dimension) was used (Mandelbrot, 1983; Foroutan-Pour *et al.*, 1999; Mandelbrot, 1977), based on equation 3.9 (Li *et al.*, 2009). For this purpose, all the images were segmented using the segmentation algorithm described previously (Hunter *et al.*, 2005), obtaining the binary vascular trees, in order to apply the box-counting and gliding box methods.

$$FractalDim. = \lim_{r \to 0} \frac{LogN(r)}{Log1/r}, \tag{3.9}$$

in which N(r) refers to the number of boxes of side length *r* that has to be used to cover a given area in the Euclidean *n*-space, by using a sequential number of descending size boxes. This occurs in multiple orientations. The final dimension in the 2D space is between 1 and 2 ($1 \leq D \leq 2$) (Li *et al.*, 2009).

Lacunarity

Complementary to the FD, lacunarity was also evaluated, which is a counterpart of FD, describing the gappiness between the structures, or alternatively how the fractals fill the space.

Lacunarity was estimated using the gliding-box algorithm, for different grid orientations (Tolle *et al.*, 2008). A unit box of size *r* is randomly chosen and the number of set points *p* are counted i.e. the mass. The procedure is repeated with the box centred consecutively for each point within the set, creating a distribution of masses B(p, r). Finally, we get the probability, by converting the distribution into probability distribution $Q(p, r)$, dividing by the total number of boxes (*B*) of size *r*. Finally, the

gliding box equation can be written in terms of the accumulated sum of the mean and the second moments of all boxes (equation 3.10).

$$L_{GB}(r) = \frac{B_{(r)} \sum_{i=1}^{B_{(r)}} p(i,r)^2}{[\sum_{i=1}^{B_{(r)}} p(i,r)]^2}, \tag{3.10}$$

where the denominator is the square of the total number of elements in the data set (Tolle *et al.*, 2008).

Arteriovenous ratio

Chapter 5 of this thesis describes in details the current methodology for estimating the AVR, including both the widely-used approach in literature, as proposed in (Knudtson *et al.*, 2003) and an alternative method proposed as part of this thesis (Leontidis *et al.*, 2016b). The area of interest is defined as shown in figure 3.8, and includes the region where the edges of the vessels course through at 0.5 to 1.0 disc diameters from the optic disc margin. The vessels are labelled and measured using the semi-automatic tool described previously. The region between this area and the optic disc is excluded,

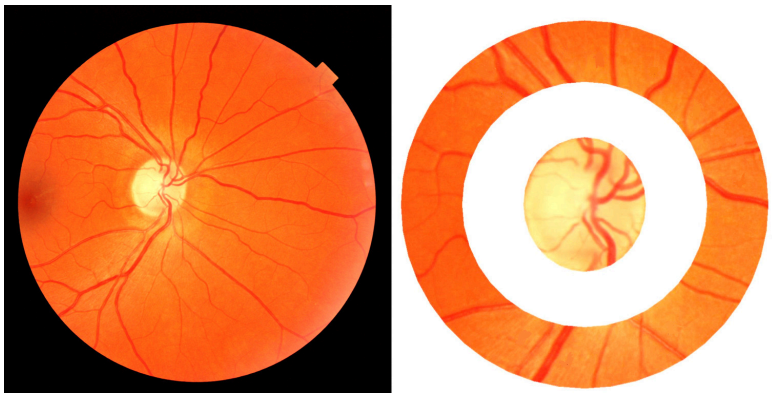

Figure 3.8: On the left, we can see the original, macula centred image. On the right, the processed image is shown, which includes only the area, where the vessels are labelled and measured; optic disc is only shown for reference purposes.

as not having the vessels attained their status inside the retina yet. Within this area, the six largest veins and the six largest arteries are measured, following an iterative procedure of pairing up the largest vessels with the smallest ones, until a final single

number is obtained.

The final value for the vein is termed central retinal vein equivalent and the respective final value for the artery is termed central retinal artery equivalent. Their quotient is known as arteriovenous ratio (54-59, Table 3.3).

Local arteriovenous ratio

The local arteriovenous ratio (27-29, Table 3.3) was estimated in three points along the arcades, starting from near the OD until two branchings apart. An example of two successive points can be seen in figure 3.9. The included images were carefully selected in order the branchings to be clear. The first point works as a reference point, therefore its absolute value does not provide any information, given that one cannot know the actual branching order, i.e. the number of times a vessel splits into two smaller ones within a vascular tree, as they come off the OD.

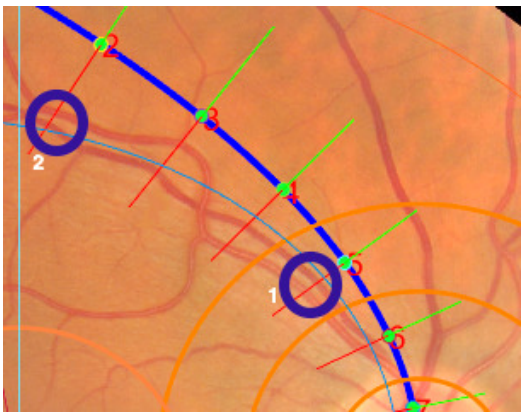

Figure 3.9: The two blue circles show the first two points, where the local arteriovenous ratio is taken. The first one is near the OD, whereas the second one is after one branching that both the artery and vein have made.

Tortuosity

Tortuosity (60-72, Table 3.3) is another very important candidate biomarker, considered to be among the first indicators of the alterations in the retinal vascular network to appear in retinal pathologies (Nguyen *et al.*, 2007). Tortuosity measures lack of an official clinical definition; nevertheless, in practice, it refers to the state or quality of being tortuous, i.e. twisted, having many turns.

As Grisan *et al.* describes it (Grisan *et al.*, 2008), a vessel can be considered as a curve

$$s(l) = [x(l), y(l)] : D \subset R \rightarrow R^2, \tag{3.11}$$

where $s \in C^1(R)$ and l being the curvilinear coordinates defined on an interval D of R. The cord length L_x is defined as the distance between the curve end points

$$L_x = \|s(max(D)) - s(min(D))\| \tag{3.12}$$

and the curve length L_c as:

$$L_c = \int_{min(D)}^{max(D)} \left\| \frac{\partial s}{\partial l} \right\| dl \tag{3.13}$$

Finally, the curvature $\kappa(l)$ of *s(l)* is defined as follows:

$$\kappa(l) = \frac{\frac{dx}{dl}\frac{d^2y}{dl^2} - \frac{d^2x}{dl^2}\frac{dy}{dl}}{\left(\left(\frac{dx}{dl}\right)^2 + \left(\frac{dy}{dl}\right)^2\right)^{\frac{3}{2}}} \tag{3.14}$$

In the present studies, specifically for the image-level (whole retina) tortuosity - three main different methods were utilised, for calculating the local tortuosity. The image level tortuosity was expressed in terms of mean, median, standard deviation and third quartile values. In addition, a fourth measure proposed by Poletti *et al.* was also used, which is a supervised approach that provides a tortuosity index capable to reproduce the clinical experts' assessment (Poletti *et al.*, 2011). This index is a linear, weighted combination of the other three.

The expertise and algorithms were provided by a research group I am collaborating with at the Department of Information Engineering, at the University of Padova in Italy.

The ***first method*** was proposed and elaborated by Grisan *et al.* (Grisan *et al.*, 2008). It starts with defining the curvature of the curve as in equation 3.15, and then a

subsegment s_i of a vessel segment s is also defined as a turn curve if:

$$[\kappa(l) \leq 0, \forall l \in D_i] \vee [\kappa(l) \geq 0, \forall l \in D_i] \tag{3.15}$$

with D_i a compact subset of D. In order to account for the possible issues arising from straight segments, which affects the tortuosity measures, the authors split the straight segments into two halves, assigning one to the previous and one to the following turn curve. After diving the curve *s(l)* into *n* turn curves,

$$s_i : s = s_1 \oplus s_2 \oplus ... \oplus s_n \tag{3.16}$$

they propose a vessel tortuosity measure as follows:

$$\tau(s) = \frac{n-1}{n} \frac{1}{L_c} \sum_{i=1}^{n} \left[\frac{L_{c_{s_i}}}{L_{x_{s_i}}} - 1 \right] \tag{3.17}$$

The ***second method***, is a curve-based measure proposed by Hart *et al.* (Hart *et al.*, 1999). This method suggests the use of the ratio between the integral of the absolute curvature (tc) (or the squared absolute curvature (tsc)) and the chord length L_x (or the vessel length L_c). The rationale behind this is that integral measures of κ along the domain D give a measure of the variability of vessel direction.

$$tc = \int_{min(D)}^{max(D)} |\kappa(l)| dl, \tag{3.18}$$

$$tsc = \int_{min(D)}^{max(D)} |\kappa(l)|^2 dl \tag{3.19}$$

The ***third method*** is an angle-based method proposed by Chandrinos *et al.* (Chandrinos *et al.*, 1998). Tortuosity is estimated by averaging the change of angle at discrete points along the vessel's length, which is independent of scale. For each pixel X that

belongs to a specific track, they consider two additional centrelines, creating two vectors $\mathbf{u}_{i+n}$ and $\mathbf{u}_{i-n}$. At the final step, they take the inverse cosine of the dot product that form these two vectors.

$$\theta(i) = \arccos(\mathbf{u}_{i+n} \cdot \mathbf{u}_{i-n}) \tag{3.20}$$

The mean tortuosity is then given by:

$$T = \frac{1}{N - 2n} \sum_{i=1}^{n} \theta(i) \tag{3.21}$$

The ***fourth method*** is a linear supervised combination of the previous three methods, aiming to provide a tortuosity index capable of reproducing the clinical experts assessment (Poletti *et al.*, 2011).

Because of the few segments that could be chosen out of the three ellipses (E1, E2, E3), the employed tortuosity in this case is the simplest and widely used *Arc Length over Chord Length ratio*. This vessel tortuosity measure is simply the ratio between its length L_c and the length of the underlying chord L_x (Lotmar *et al.*, 1979; Heneghan *et al.*, 2002). In addition to this tortuosity measure, the widths and angles of two bifurcations were extracted out of each of the previously described three areas (30-53 & 73-75, Table 3.3).

A detailed list of all the 101 investigated features is given in Table 3.3.

3.3.5.2 Haemodynamic Features

The haemodynamic features (76-101, Table 3.3) were estimated in cooperation with a colleague at the University of Lincoln, Francesco Caliva, based on some of his current work on modelling of the haemodynamics of the retinal vasculature. In order to avoid any overlap and conflict with his future doctoral thesis, I will succintly define my contribution on this specific part of the thesis.

In order to conduct this part of the study, the segmented images were carefully selected by myself, in order to make sure that the same vascular trees can be selected in all four successive images of each patient. This was fulfilled for both the progressors' and non-progressors' studies. Using a semi-automatic tool, developed by Caliva *et al.* (Calivá *et al.*, 2015), each bifurcation that was not properly connected, due to errors during the segmentation process, was corrected in order to form a proper vascular tree (figure 3.10). This process, which gives us all the vascular trees across the whole image set, is then followed by the estimation of the haemodynamic features that are described in details in the next section. Similar to the geometric features, all of the analyses and evaluations of these features were conducted as part of the contributions of this thesis.

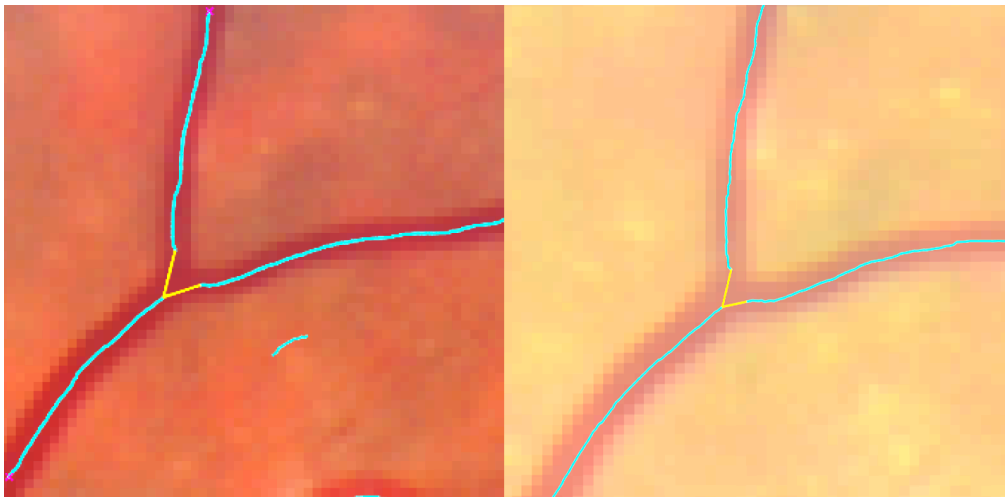

Figure 3.10: Two connected bifurcations, in which the segments were initially disconnected. The cyan segments are the properly segmented ones; the yellow "Y" junction is the one we manually draw with the tool in order to define the connectivity and therefore make sure that the vessels are properly connected.

The retinal network is a hydraulic system, which can be studied by applying Kirchoff's circuit laws in lumped element models (Rashid, 2010). According to Ohm's law, current I is given by the difference in voltage ΔV divided by the electric resistance R (Hayes & Horowitz, 1989). Relating Ohm's law to fluid dynamics, the electric current I is the electric analogue of blood flow rate Q, the voltage difference ΔV is

the analogue of pressure difference ΔP and the electrical resistance R is the analogue of blood resistance R to flow (Hubbert *et al.*, 1956; Sherman, 1981; Sutera & Skalak, 1993). A schematic representation of the above can be seen in figure 3.11.

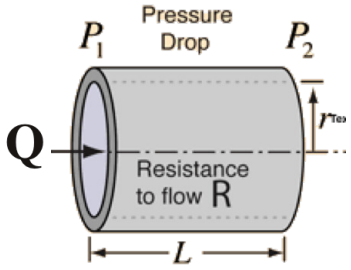

Figure 3.11: Schematic representation of the pressure drop in a vessel. Q is the inlet flow, L is the length, r is the diameter, and $P1$ and $P2$ are the pressures in the inlet and outlet of the vessel, respectively.

The features were estimated based on haemodynamic principles and boundary conditions, separately for arterial and venular trees. As we are dealing with vessels with a diameter less than 200 μm, then microcirculation principles are applicable (Wong *et al.*, 2003; Pournaras *et al.*, 2008). In such small vessels, Fåhræus-Lindqvist effects and plasma skimming are paramount in determining the distribution of the hematocrit and the blood flow velocity profile (Pries *et al.*, 1992, 1989). The purpose of this study is to see how the variations in the geometry of the vessels affect the haemodynamics without deepening into the rheology or haematocrit's distribution within the network, hence assuming a Newtonian fluid (Aletti *et al.*, 2015). To achieve that, a zero dimensional model (0-D) is designed to estimate the fluid dynamics within the network.

In the present study, all the vascular trees include one bifurcation, with the inlet (root of the tree) in the parent vessel and the outlets in the two children ones. At the branching points, the fluid dynamic conditions were estimated under the assumptions of Hagen-Poiseuille's law: stiff, straight and uniform tube; Newtonian fluid; laminar, steady, non-pulsatile and fully developed flow with null velocity at the walls. Blood flow velocity (V), blood flow rate (Q), Reynolds number (Re), pressure (P) and wall shear stress (WSS) were estimated, for both arteries and veins.

Each vascular segment (S) consists of two end-sides (E1 and E2); at each bifurcation (B), the parent vessel (F) splits into two children (C1, C2) vessels, and the blood flow (Q) is driven from the F to the C1 and C2. Blood viscosity ($\mu = 0.04poise$), tube's length (L) and diameter (D), were used to compute the fluid resistance to flow through each vessel.

$$R = \frac{128\mu L}{\pi D^4} \tag{3.22}$$

where μ is the blood viscosity, L is the length and D is the diameter of the vessel (Kirby, 2010). The flow rate (Q), following the conservation of mass ($Q_{parent} = Q_{ch1} + Q_{ch2}$) was calculated by applying Hagen- Poiseuille's law $Q = (P1 - P2)/R$, where P1 and P2 refer to the values of the pressure at E1 and E2 (Zamir, 2016). In literature different values of the pressure (P) at the central retinal artery and vein (CRA, CRV) are reported that range from 10 to 90 $mmHg$ (Pemp *et al.*, 2013; Kim *et al.*, 2015). In other studies they measured (in-vivo) blood flow velocity with colour doppler imaging, getting a wide range of values (Keyser *et al.*, 1994; Pemp *et al.*, 2013). For this study, the adopted values are V_{CRA} = 6.4 cm/S (Fuchsjäger-Mayrl *et al.*, 2001) and V_{CRV} = 3.6 cm/S, selected within the values reported at (Kaiser *et al.*, 1997; Gobel & Lieb, 1995). Accordingly, the diameters of the CRA and CRV are D_{CRA} = 166μm (Pemp *et al.*, 2013) and D_{CRV} = 210μm, based on a value of AVR = 0.79 (average values in healthy subjects (Triantafyllou *et al.*, 2014; Heitmar *et al.*, 2010)).

In this thesis the aim is to estimate how fluid dynamics vary at the branching points, and not to simulate the behaviour of the retinal vasculature. Wall shear stress at the vessel wall (WSS) was computed as

$$WSS = \frac{32\mu Q}{D^3} \tag{3.23}$$

where μ is the blood viscosity, Q is the Blood flow and D is the diameter of the vessel (Papaioannou & Stefanadis, 2005).

Reynolds number was calculated as

$$Re = \frac{VD\rho}{\mu} \tag{3.24}$$

where $\rho = 1.06gr/mL$ is the value of blood density adopted, V is the velocity and D is the width of the vessel (Reynolds, 1883).

Boundary conditions

Some necessary boundary conditions were adopted to solve the haemodynamic problem.

The pressure in CRA = pressure in CRV = 45$mmHg$; blood flow (Q_a) in CRA and blood flow (Q_v) in CRV were derived from the velocity in the central retinal artery V_{CRA} and the diameter D_{CRA}, and the velocity in the central retinal vein V_{CRV} and the diameter d_{CRV}, respectively, by using the formula $Q = VA$, where V is the flow velocity and A is the cross-sectional area of the blood vessel. Two different boundary conditions were imposed in this network. In the **first one (Q_{in} - Flow boundary condition)**, the pressure is directly estimated from the blood flow (Q_{in}) of the system at the bifurcation point, according to $P = QR$, where R is the hydraulic resistance of the tube according to the Hagen-Poiseuille's law. The inlet flow rate Q_{in} equals to 0.0014 cm^3/s and 0.0012 cm^3/s for arteries and veins respectively, based on equation 3.25 (Riva *et al.*, 1985) and the values of blood flow velocity reported previously. Across the parent vessels this value does not change, but it does in the children vessels.

$$Q_{in} = \frac{\pi D^2}{4} V \tag{3.25}$$

where D and V are the diameter and velocity of the CRV and CRA, respectively. The boundary conditions at the outlet were enforced applying Murray's law (Murray, 1926b), according to which the blood flow distribution at the bifurcation depends only on the vessels' calibre. In the **second one (P_{in} - pressure boundary condition)**, a spe-

cific inlet pressure is forced into the system, through the parent vessel (P=45$mmHg$ for both arteries and veins), which was chosen as an average of the values mentioned previously. This estimation reflects the hydrostatic and frictional pressure losses from the aorta to the Central Retinal Artery (CRA). At the outlet of the network, a pressure of 15 mmHg is considered, which is equivalent to the intraocular pressure under normal conditions (Causin *et al.*, 2016). After that the pressure is estimated according to the pressure drop at the bifurcation point (figure 3.11). Finally all the selected trees were chosen to be as near to the optic disc as possible.

All of the above estimations are based on the boundary conditions that were adopted in this thesis, with whatever limitations these impose. However the conditions are adequate to provide an estimate of the relative haemodynamic changes that occur during different stages of the disease.

Some of the core work of Francesco Caliva will elaborate on the above methodologies, therefore, in order to avoid any overlap, only essential information was provided, given that this thesis is focusing on the evaluation of the above features as candidate biomarkers. In figure 3.12, three examples of the estimation of the haemodynamics can be seen.

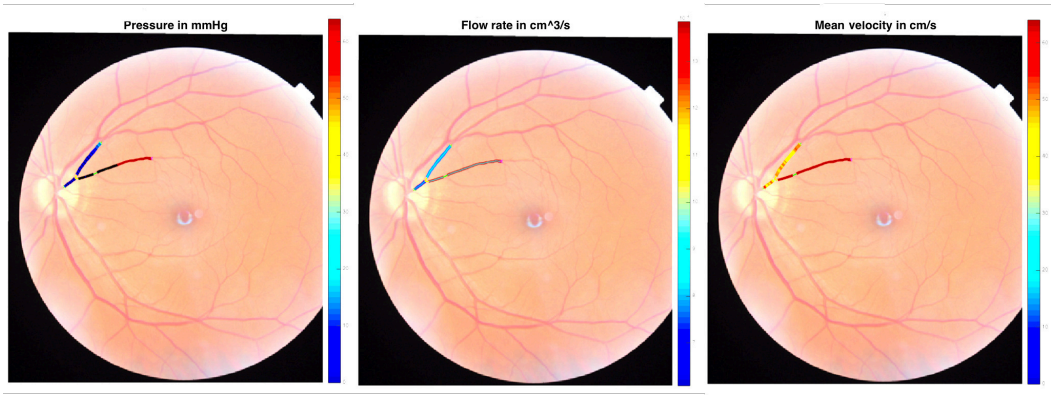

Figure 3.12: Three examples of how the heamodynamic features in a given bifurcation are estimated. All images come from the same subject and include the same connected bifurcation. On the left, the image refers to the pressure; on the centre refers to blood flow rate; on the right refers to blood flow velocity.

3.4 Conclusion

This chapter presents all the necessary information about the data, the proposed framework and the investigated biomarkers. A detailed introduction was made about the areas of interest, including how the junctions are being matched, and also the way the geometric and haemodynamic features are measured and estimated. All of these are going to be used in conjunction with the techniques that follow in chapter 4 and 5, in order to conclude with the results of all the studies in chapter 6.

Chapter 4

Statistics, Feature Selection and Classification

4.1 Introduction

After providing the description of all the features that are under investigation, the crucial part remains the proper evaluation, in order the results to be as comprehensive, meaningful and conclusive as possible. In statistics and machine learning it is important to identify the optimum feature subset that gives the best performance. In addition, a simplified and more interpretable model is preferred to a more complicated one.

The purpose of this chapter is to give a detailed description of the evaluation methods used in this thesis, including the statistical analysis, the feature selection process, the classification and the validation process. The first aim is to find out whether the retinal vascular geometry and haemodynamics can provide us with robust biomarkers of progression to DR, and also quantify their discriminative power in multiple different occasions. The second aim is to find the minimal feature set that can differentiate between multiple combinations of binary classifications.

The chapter is organised as follows. The **first section** includes the background

information and the type of statistics that were used in the present studies, followed by the **second section**, which will be dedicated to the feature selection and classification process. **The third section** will describe the different experiments that were conducted and the corresponding models that were designed for making the inferences for each of them. Finally, **the fourth section** will be the conclusion and discussion of this chapter.

4.2 Statistical Analysis

In every statistical analysis process it is paramount to decide the correct way of analysing the collected data at the moment of designing the experiments and/or studies. In addition, the selection of the proper methods is subject to the hypothesis that one is willing to test. The same data can be analysed in various ways, depending on what inferences one wants to make out of them. In the studies presented in this thesis, different types of experimental designs are used, which will be described below.

In general, two main approaches were used in this thesis. In the initial experiments, mixed-effects models were designed, which were analysed using the Analysis of Variance (ANOVA) parametric test (Fisher, 1936). In the second and most complicated phase of the experiments, the models were designed and evaluated using Linear Mixed Models (LMM) (McCulloch & Neuhaus, 2001; Bates *et al.*, 2014b). Before talking about their similarities and differences, it is worth clarifying what a mixed model stands for.

Mixed models or better mixed-effects models refer to a variety of models that have as main feature both ***fixed*** and ***random*** effects. However, the distinction between fixed and random effects is not always clear. Fixed effects are those where the possible values of the variable are fixed. On the other side, random effects refer to variables in which the possible set of values can change (McLean *et al.*, 1991). A clear way of thinking about the distinction between fixed and random effects is at the observation

level. Fixed effects assume independency of scores or observations, whereas random effects assume some type of relationship between some scores or observations. For instance, gender is a fixed effect variable because one knows all the values of that variable, and at the same time they are mutually exclusive (independent). A good example for a random effects variable is a group of patients with diabetes at a hospital department; one can only sample some of the departments that exist. Besides, patients can and many of them do move into and out of those departments from time to time indeed, depending on the treatment and the disease.

There are many types of random effects; repeated measures of the same individuals is one of them. This is the main random effect that the progression study of this thesis includes. The scores at each time constitute samples from the same individual for ”indefinite” number of times. Another example, which is also part of the models here, is a nested design, e.g. for the previous example, the multiple retinal screenings of patients that are nested within departments and these departments nested within hospitals. A nested structure assumes a relationship among groups such that measurements that belong to a department (individuals) are considered to be similar to the other scores of the same department, in such a way as to be able to distinguish them from measurements of other departments (Bates *et al.*, 2014a). Some of the candidate features that are being investigated in this thesis include multiple vascular measurements, which are nested within individuals. This is also a random effect, allowing to deal with the non-independence. Individuals and measurements are the random effects factors; the disease stages (e.g. three years pre-DR, two years pre-DR, one year pre-DR and onset of DR), which is defined and cannot be changed, is the fixed effects factor. These factors represent the independent variables, since they stand alone and are not changing by other variables. On the other hand, the different features under investigation are the dependent variables, because their values depend on the other factors, such as the stage of the disease or a specific individual. The general design can be seen in figure

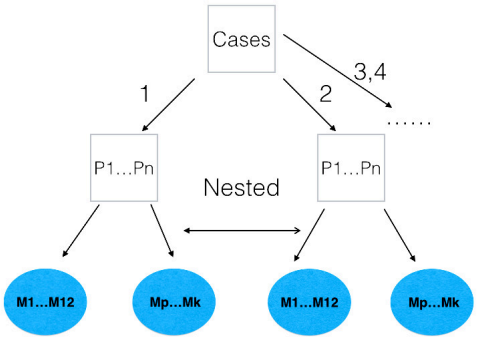

Figure 4.1: A mixed model design for the progression study. The same patients are used in all four groups and the multiple measurements within each of them represent the nested factor. In this way we account for the fact that the measurements within each individual are not independent, hence being nested.

4.1.

4.2.1 Repeated Measures ANOVA

In the initial stage of the analysis, where the data were more limited and balanced, a repeated measures ANOVA was preferred. ANOVA is a statistical method used to test general rather than specific differences between two or more means. Although the technique is called ”Analysis of Variance” rather than ”Analysis of Means, the name is indeed appropriate because inferences about means are made by analysing variance.

In order to use the ANOVA parametric test, instead of a non-parametric one, one has to make sure that the dependent variable is approximately normally distributed, even though it is robust in slight violations of normality. Another important assumption for the repeated measures is the sphericity, which refers to the fact that the variance of the population difference scores for any two conditions should be the same as the variance of the population difference scores for any other two conditions. A possible violation causes the test to become unstable (i.e. leads to an increase in the Type I

error; that is, the likelihood of detecting a statistically significant result when there is not one) (Girden, 1992).

The Shapiro-Wilk tests were used for testing for the normality of each feature (Shapiro & Francia, 1972). Regarding sphericity, Mauchly's test was used to test for this assumption (Mauchly, 1940).

4.2.2 Linear Mixed Models

Linear mixed effects models (LMM) model the fixed and random effects as having a linear form. The outcome variable is contributed by additive fixed and random effects and of course the error term. Using the hierarchical notation of (Laird & Ware, 1982), one can express the LMM as

$$Y_i = \underbrace{X_i\beta}_{fixed} + \underbrace{Z_i b_i}_{random} + \underbrace{\epsilon_i}_{error} \tag{4.1}$$

where

$Y : (n \times 1)$ response vector

$X : (n \times \mathrm{p})$ design matrix for fixed effects

$\beta : (p \times 1)$ regression coefficients for fixed effects

$Z : (n \times \mathrm{q})$ design matrix for random effects

$b : (q \times 1)$ random effects

$\epsilon : (n \times 1)$ error vector

Assumptions for the distributions: b_i and ϵ_i are independent, with prior distributions:

$$\mathrm{b} \sim N\,(\mathrm{O}, \sigma^2 D(\theta))$$

$$\epsilon \sim N\,(0,\, \sigma^2\, I)$$

where D is a symmetric, positive semi definite matrix, which can be parametrised by a variance component vector θ, I is an *n*-by-*n* identity matrix and σ^2 is the error variance (Fox, 2002).

In the above described model, the parameters one needs to estimate are the fixed-effects coefficients β and the variance components θ and σ^2. In LMMs, the two main methods to estimate the parameters are the maximum likelihood (ML) and restricted maximum likelihood (REML) (Bates *et al.*, 2014b). In the comparisons of models one needs to fit the models with both methods, and the explanation for that follows.

Maximum Likelihood (ML) estimates both the regression coefficients and the variance components, i.e. both fixed-effects and random-effects terms in the likelihood function (Swaminathan & Rogers, 2008). For the previously defined linear mixed model, the conditional response of the response variable y given β, b, θ and σ^2 is

$$y|b, \beta, \theta, \sigma^2 \sim N(X\beta + Zb, \sigma^2 I_n) \tag{4.2}$$

So the likelihood of y given β, θ and σ^2 is

$$P(y|\beta, \theta, \sigma^2) = \int P(y|b, \beta, \theta, \sigma^2) P(b|\theta, \sigma^2) db, \tag{4.3}$$

The problem with ML is that it treats β as fixed, but unknown quantities, when it estimates the variance components, but does not take into account the degrees of freedom lost by estimating the fixed effects. This ends up with ML estimates to be biased with smaller variances.

Restricted Maximum Likelihood Estimation (REML) includes only the variance components, i.e. those parameters that parametrise the random-effects terms in the linear mixed-effects model. The parameter β is estimated in a second stage (Bates, 2014). Provided a uniform improper prior distribution for β and integrating the like-

lihood ($P(y|\beta, \theta, \sigma^2)$) with respect to β, one ends up with the restricted likelihood:

$$P(y|\theta, \sigma^2) = \int P(y|\beta, \theta, \sigma^2)P(\beta)d\beta = \int P(y|\beta, \theta, \sigma^2)d\beta \tag{4.4}$$

Following that, β is estimated with respect to the posterior distribution

$$P(\beta|y, \hat{\theta}_R, \hat{\sigma}^2_R) \tag{4.5}$$

The process starts by profiling out $\hat{\sigma}^2_R$ and maximising the remaining objective function with respect to θ, in order to find $\hat{\theta}_R$. The restricted likelihood is maximised with respect to σ^2 to find $\hat{\sigma}^2_R$, and then find the expected value of β.

Given that REML accounts for the degrees of freedom lost for estimating the fixed effects, makes a less biased estimation of random effects variances. Most importantly, compared to the ML estimates, the estimates of θ and σ^2 are invariant to the value of β as well as less sensitive to possible outliers in the data (Cheung, 2013). Therefore all the estimates in LMMs are made by fitting a REML. However, when it comes to comparing two models with a likelihood ratio test, then one needs to refit them to compare the fixed- and random-effects terms. Comparisons with REML can only be made between models that are nested in their random-effects terms, thus having the same fixed-effects term. In the results of the next chapters, whenever is needed to compare the full model (the model with the fixed effect, i.e. the disease stage, and the random-effects part) with the null (or restricted) model (the model without the fixed effects term), then the models need to be refitted with the ML (Bates *et al.*, 2014a).

When designing a LMM, there are two main design issues that one needs to resolve regarding the random effects. The first is the **random effects intercept** and the second is whether one needs to include a **random slope** or not. Every individual is different and so is his/her retina. The slightly different retinal vascular geometry and functionality is going to be an idiosyncratic factor that affects all responses from the

same individual, rendering these multiple responses inter-dependent rather than independent. Random effects let us deal with this non-independence by assuming a different baseline response value for each individual (Bates *et al.*, 2014a). For instance, individual 1 in group 1 (e.g. DR group) may have a mean vessel angle of $85{,}4^o$ across measurements and individual 2 in group 1 may have a mean vessel angle of $88{,}3^o$.

These individual differences can be modelled by assuming different random intercepts for each subject, i.e. each subject is assigned a different intercept value, and the mixed model estimates these intercepts for us (Winter, 2013). In mixed models, adding one or more random effects to our fixed effects, gives structure to the error term ϵ. In the case of the studies here, the random effect for individuals characterises idiosyncratic variation that is due to individual differences. The error term ϵ is necessary because even if one accounts for the individual by-subject variations, there are still going to be random differences between different utterances of the same subject. Another example follows that does not apply to the studies of this thesis, but is mentioned in order to elaborate even further. If one has the stress levels of every individual every time they had their retina screened, then this would be an additional random effect in the model, which could help explain some more variation and also account for it, given that hypothetically higher levels of stress might possibly lead to the dilation of the arteries, for instance. It is now clear why avoiding averaging the within-subjects' values leads to a more robust and accurate model, offering more flexibility and also allowing to take into account the full data (Locker *et al.*, 2007; Baayen *et al.*, 2008).

Going a step further, until now, all the described models account for the baseline differences among individuals, but for that the assumption is that whatever the effect that the progression of diabetes has to the retina is going to approximately be the same for all individuals (Winter, 2013). But this is an assumption one also needs to check, because some individuals might be more or less affected by the progression of diabetes than others. LMM can account for that by adding a **random slope** as well,

where subjects are not only allowed to have different baselines, but are also allowed to have different slopes for the effect of the progression of diabetes. As can be easily understood, adding more parameters to be estimated by the LMM, increases the complexity and interpretability of the model, thus creating the need to think of the trade-off. Luckily, there are metrics and comparisons one can make to decide whether a more complicated model makes any difference or not. The next section will elaborate on this part. All the LMM models implemented in this thesis were fitted by using the R package lmer and its dependent functions (Bates *et al.*, 2014a).

4.2.3 ANOVA versus Linear Mixed Models

The main reason to prefer LMMs over ANOVA is that the former are considerably more general, e.g. they work well with both **balanced** and **imbalanced designs** and can also be extended in multilevel modelling (Bates, 2007; Krueger & Tian, 2004). In addition, linear mixed models deal very well with **missing data**, something that ANOVA does not. If, for example, an individual is missing one time point, it needs to be dropped off the entire analysis. On the contrary in LMMs, only this point will be discarded, retaining all the remaining data (Little & Rubin, 2014; Krueger & Tian, 2004). Moreover, because of the way that the sums of squares are calculated in repeated measures nested ANOVA, **post hoc tests** are not appropriate (Howell, 2006) or need some adjustments (Nolan & Heinzen, 2010); but they are using LMMs. It is worth highlighting that the results from an ANOVA and a LMM are almost identical when the design is balanced. In addition, an advantage of ANOVA is that it can be communicated easier, reporting the F-statistics, degrees of freedom and p-values. In LMMs especially in imbalanced designs, the p-values need to be approximated, due to the not direct way of calculating the degrees of freedom (Schluchter & Elashoff, 1990; Bates *et al.*, 2014b). Therefore other metrics (which are supported by the purpose of the development of the LMMs, (Bates *et al.*, 2014a; Bates, 2014)), like the Akaike

(Akaike, 1998) and Bayesian Information criterion (AIC and BIC) (Posada & Buckley, 2004) and likelihood ratio test (Vuong, 1989) are more indicative metrics of the performance of the model, or in simpler words, for making inferences out of our data. These are the main differences that in the second phase of the experiments the LMMs were preferred over ANOVA. Further details for each of the two approaches are given in the next sections.

4.2.4 Evaluation Metrics

Traditionally, before conducting any statistical test, one has to form some kind of null hypothesis to be tested. Hereafter, every single feature that is tested is forming its own hypothesis. That is the fixed effect (disease) has a significant effect on the measured feature, which is tested against the null hypothesis. Null hypothesis asserts that the samples being compared or contrasted are drawn from the same population with regards to the outcome variable. This means that a) any observed differences in the dependent variable (features) must be due to sampling error (chance) or b) the independent (predictor) variable (disease) does not make a difference (Blackwelder, 1982).

In literature, to test the above, the most widely utilised metric is the p-value, or more correctly, the calculated probability (Fisher, 1925). P-value refers to the probability of finding the observed result when the null hypothesis of the study in question is true. This metric has undergone a great deal of controversy in literature about its importance in communicating results (Anderson *et al.*, 2000; Wagenmakers, 2007).

Other metrics are also useful in deciding whether a predictor has an effect on the outcome variable. Therefore for each feature that is being investigated in this thesis, a series of metrics will be reported in order to give an overall performance indication, hence avoiding to only follow a specific evaluation method. The p-values for the LMMs cannot be directly extracted, because the estimation of the denominator degrees of freedom is not straightforward (Bates *et al.*, 2016; Kuznetsova *et al.*, 2015).

Since the parameters are calculated by the REML, they are not based on observed and expected mean squares or on error strata, thus being able to handle imbalanced designs with multiple nested, fully or partially crossed grouping factors for the random effects. Therefore, there are a few approximations for the degrees of freedom leading to approximated p-values, which will be described next.

For each feature in chapter 6, the following metrics will be reported:

1. **Likelihood ratio test X^2 (chi-square)** is a goodness of fit between two models, one being the null model and the other being the full one (Vuong, 1989). Removing variables from the model will most probably reduce the model fit, but the purpose is to evaluate whether this possible reduction is statistically significant. The likelihood ratio test (LRT) does this by comparing the log likelihoods of the two models; supposing that this difference is statistically significant, then the less restrictive model (the one with the most variables) is said to fit the data significantly better than the null one. The formula for the LRT is:

$$LR = -2 \ln \frac{L(nullmodel)}{L(fullmodel)} \tag{4.6}$$

   where $L(nullmodel)$ and $L(fullmodel)$ are the likelihoods of the respective models. The resulting probability distribution of the test statistic is approximately chi-squared distributed, with degrees of freedom equal to the number of parameters that are constrained (i.e. the number of variables removed from the full model).

2. **Akaike Information criterion (AIC)** is another penalised measure of the relative quality of a statistical model (Akaike, 1998). AIC is used to compare models relative to each other, hence making a model selection. AIC is not a hypothesis test, rather a relative estimate of the information lost, when a specific model is used to represent the process that generates the data. In other words, it pro-

vides us with a trade-off between the goodness of fit and the complexity of the model. AIC can provide no information about the quality of the model in absolute sense, meaning that if a model fits the data poorly, this cannot be monitored by the value of AIC. The value of AIC is given by

$$AIC = 2\kappa - 2\ln(\hat{L}) \tag{4.7}$$

where $\hat{L}$ is the maximum value of the likelihood function for the model and κ is the number of estimated parameters in the model.

3. **Bayesian Information Criterion (BIC)** is a similar to AIC penalised measure for the comparison between two models. Its value is given by

$$BIC = -2\ln\hat{L} + \kappa\ln(n) \tag{4.8}$$

where $\hat{L}$ is the maximised value of the likelihood function of the model, *n* is the number of data points and κ is the number of free parameters to be estimated. BIC provides us with an estimate of a function of the posterior probability of a model being true, and just like AIC, it is based on various assumptions and asymptotic approximations, but it penalises model complexity more heavily. Lower BIC denotes a better fit. Both BIC and AIC aim to avoid overfitting (Posada & Buckley, 2004). Empirically, in both AIC and BIC, a difference between two models greater than or equal to 2 is considered positive evidence that one model is better than the other.

4. **Welch–Satterthwaite approximation** is a method for obtaining approximated p-values. In statistics, degrees of freedom is the number of values in the final calculations that are free to vary. As aforementioned, for various reasons, in LMMs calculating the effective degrees of freedom is not straightforward, hence the

need for approximations. Based on uncertainty analysis, the aim is to combine multiple variances characterised by various distributions. The linear combination of the sample variances gives a probability distribution, which cannot be expressed analytically. However, it can be approximated by a chi-squared distribution, giving the pooled degrees of freedom, needed for the inference tests, thus coming up with an approximated p-value (Satterthwaite, 1946).

5. **Kenward-Roger's approximation** proposes a scaled Wald statistic, involving an approximate covariance matrix, which accounts for the variability introduced by the estimation of the variance components. The advantage of this method is that it performs well under small sample conditions; they showed that its small sample distribution can be approximated by an F-distribution, obtaining the denominator degrees of freedom, using similar to Satterthwaite's approximations (Kenward & Roger, 1997).

6. **Intraclass correlation** is mainly used for two purposes. Firstly as a measure of reliability or agreement between two or more raters, or for the evaluation of methods on the same set of subjects. Secondly, it allows to assess whether the random effect is present in a dataset or not. To do so, one first creates a null model, which only includes the fixed- and random- effects intercepts and the random effect for the highest level variable, in case of nested model (Snijders, 2011). After fitting the null model the ICC is calculated by dividing the random effect's variance by the total variance estimate (random effect's variance and the residual's variance). The final value ranges from 0 to 1 and gives an indication of how much of the variance in the response variable can be explained by the random effect (West *et al.*, 2014).

7. **Explained variation** refers to the variation that can be explained by the covariates under the linear mixed model. In ordinary linear regression the R^2 is a mea-

sure of explained variation and is very popular in the applied fields of statistics (O'Grady, 1982). When promptly defined, R^2 consistently estimates the proportion of explained variation in the population, since it is not affected by the sample sizes, as happens with p-values. Another advantage of R^2 is the assessment of the importance of a covariate, overcoming the limitation of the regression coefficient, which depends on the scale; in multiple regression it summarises the effects of all the covariates. For the LMMs, (Xu, 2003) proposed a generalised measure similar to R^2, the Ω_0^2, which also provides information about the effect size. It can be estimated by

$$\Omega_0^2 = 1 - \frac{\sigma^2}{\sigma_0^2} \tag{4.9}$$

where σ^2 is the residual variance of the full model, i.e. the amount of variation in the response Y that is not explained by the covariates, and σ_0^2 is the residual variance of an intercept-only null model (for fixed- and also possibly random-effects; depends on one's purpose). It is worth pointing out that goodness of fit and/or coefficient of determination metrics are still an on-going field of research in LMMs (Xu, 2003).

In case one wants to simply find the amount of variance of a specific grouping random effect, another way is to divide the variation attributed to this effect by the total variation of the full model and based on the outcome decide whether this effect is meaningful or not. Another way to address this is to compare the two created models, by including or excluding this effect; if the outcome is significant then one can keep it, otherwise the simplest model is preferred (Starkweather, 2010).

8. **Bootstrap confidence intervals** for the estimates is a way of obtaining a range of values (intervals) that act as good estimates of the unknown population pa-

rameters and are calculated from a given set of sample data (DiCiccio & Efron, 1996). The level of a confidence interval gives the probability that the estimated interval includes the true value of a parameter (e.g. variance estimates or area under the curve (AUC)). The most common confidence level value is 0.95 (95%). In non-Bootstrapped approaches the estimation begins with a sampling distribution of a statistic, e.g the sample mean. From the sampling distribution, based on a 95% confidence level, one argues that 95% of the time the sample mean $\bar{X}$ will fall between $\mu \pm 1.96$ SE, where SE is the standard error of $\bar{X}$ as the estimator of the population mean μ. Practically, the underlying distribution of the data is most of the times unknown and all one has are the data. So the basic idea of bootstrap is to take (re-)samples $x_1^*, ..., x_n^*$, taken with replacement from the original sample $x_1, ...x_n$. To get the boostrapped confidence intervals, a variety of different concepts have been proposed (Efron & Tibshirani, 1997; Davison & Hinkley, 1997). However it is common in the LMMs to use either a parametric or nonparametric approach. Without getting into much details, given the vast information on that subject, one can say that the nonparametric bootstrap makes no assumptions about the distribution of the data. The data are assumed to be a vector y_i of *n* independent observations and one is interested in a confidence interval $\hat{\theta}_{(y_i)}$. On the other side, a parametric bootstrap assumes that a parametric model for the data, $F_Y(y; .)$, is known up to the unknown parameter vector, θ. Then, the bootstrap data are sampled from $F_Y(y; \hat{\theta})$, where θ actually is the maximum likelihood estimate of the original data (DiCiccio & Efron, 1996).

9. **Post-hoc tests** is an advantage that one has with the LMMs, even when the data are imbalanced or have nested factors. As the term implies (latin phrase for "after this"), one is looking at the data after the experiments have concluded, looking for patterns not specified a priori (Jaccard *et al.*, 1984). With these tests one can reach to further conclusions about the data, usually focusing on the pos-

sible pairwise comparisons. It is very important to clarify the difference between ”proper” post-hoc tests and multiple comparisons. Starting with the latter, it occurs when one considers multiple statistical inferences simultaneously. Failure to compensate for that by adjusting the p-values will most likely lead to TYPE I errors, i.e. falsely rejecting the null hypothesis. The reason is simple: If one performs a test at 5% level, there is only a 5% chance of incorrectly rejecting the null hypothesis, if the hypothesis is true. On the other hand, the former accounts for the above issue and in literature there are many methods. In this thesis the Tukey post-hoc test (Tukey, 1949) and the Bonferroni correction post-hoc test (Weisstein, 2004) are utilised. Both methods correct for family-wise error rate, but following a different approach. In **Bonferroni** the adjustment is simple and it can be used as a post-hoc test or with planned contrasts, being more suitable when one has a small set of planned comparisons. If $\alpha = 0.05$ is the level and $\kappa = 10$ is the number of pairwise comparisons, the adjusted p-values will be $\alpha/\kappa = 0.005$. **Tukey's** approach follows upon a significant main test and it can tell us, which subgroups actually differ. Tukey's formula is very similar to the one of the t-test, except that it corrects for the family-wise error rate. At first, the difference between the largest and the smallest of the two means is divided by the standard error of the data. The obtained value q_t is then compared to a q value from the studentised range distribution. If q_t is larger than the q, the two means are said to be significantly different.

4.3 Classification & Feature Selection

In order to test the discriminative power of the proposed features, two different approaches were followed, given their popularity in literature. Firstly, a regularised random forests classifier was used, which is an adjusted for the feature selection process

random forests classifier, as proposed in (Deng & Runger, 2012, 2013). Secondly, an elastic-net logistic regression process was also used, which includes a feature selection step based on the Least Absolute Shrinkage and Selection Operator (Lasso) and Ridge regression, creating a hybrid penalty for the coefficients of the features (L1- and L2-norms) (Friedman *et al.*, 2010; Zou & Hastie, 2005) .

4.3.1 Regularised Random Forests

Random forests is a well-established supervised classifier and very popular in machine learning. It was proposed by Breiman as an improvement to the decision trees bagging method (Breiman, 2001). It consists of multiple decision trees, each of which is grown on a bootstrap sample, taken from the original training data. The Gini impurity index (Gini(u)) at node u, is defined as

$$Gini(u) = \sum_{c=1}^{c} p_c^u (1 - p_c^u) \tag{4.10}$$

where ${p_c}^u$,is the proportion of class-c observation at node u. Subsequently, the Gini information gain of Xi for splitting node u, is the difference between the impurity at node u and the weighted average of impurities at each child node of u. This can be seen in equation 4.11 (Deng & Runger, 2013).

$$Gain(X_i, u) = Gini(X_i, u) - w_L Gini(X_i, u^L) - w_R Gini(X_i, u^R) \tag{4.11}$$

where u^L and u^R are the left and right children nodes of u respectively. Similarly w_L and w_R are the proportions of instances assigned to the left and right children nodes. Gini index is used in random forests as a measure of variable importance. An important part of random forests is the mtry parameter (i.e. the square root of the number of predictor variables (rounded down)), which defines the random set of

features out of P that is evaluated each time (Cutler *et al.*, 2007). The feature with the highest Gain(X_i, u) is used for splitting the node *u*. The importance score for variable X_i is then calculated as,

$$Importance_i = \frac{1}{ntree} \sum_{u \in S_{Xi}} Gain(X_i, u) \tag{4.12}$$

where S_{Xi} refers to the set of nodes split by X_i in random forests with ntree number of tree. In short, the regularised version of random forests (RRF) can select a compact feature subset, by including an additional penalty coefficient, creating a regularized information gain (equation 4.13) (Deng & Runger, 2013),

$$Gain_R(X_i, u) = \begin{cases} \lambda \cdot Gain(X_i, u) & i \notin F \\ Gain(X_i, u) & i \in F \end{cases} \tag{4.13}$$

in which F refers to the set of indices of features used for splitting in the previous nodes. The parameter $\lambda \in (0,1]$ is the penalty coefficient. When i$\notin$ F the coefficient penalises the *ith* feature for splitting node *u*. Smaller λ leads to a larger penalty. Regularised random forests uses $Gain_R(X_i, u)$ at each node, and adds the index of a new feature to F. For instance a RRF with $\lambda = 1$, has the minimum regularisation, however a new feature has to be more informative at a given node than the features that have already been included to the feature subset. The feature subset selected by RRF($\lambda = 1$) is termed the least regularised subset, as it offers minimum regularisation. Apart from the feature selection process, the rest of the algorithm is exactly the same as the initially proposed random forests classifier (Deng & Runger, 2013).

For the evaluation of the performance of RRF, in addition to the other metrics, the Out of Bag error (OOB) is also used, which is the internal way of validating the performance of random forests classifier (Breiman, 2001).

4.3.2 Logistic Regression with Elastic-Net

In all the instances of this study, in addition to the RRF, a regularised logistic regression model is also used, as has been proposed in (Friedman *et al.*, 2010; Zou & Hastie, 2005). The difference with the ordinary logistic regression has to do with the penalty parameter applied to the coefficients. In the case of ridge regression, the coefficients of correlated predictors are shrunk towards each other, letting all of them work together. From a Bayesian perspective, the ridge regression performs better, if there are many predictors and all have non-zero coefficients.

On the other side the least absolute shrinkage selector operator (Lasso) is to some extent indifferent to very correlated predictors, tending to pick one and discard the rest (Tibshirani, 1996). The Lasso penalty corresponds to a Laplace prior, which expects many coefficients to be zero or close to zero and also a small subset of non-zero coefficients. In the middle of this, elastic net with the mixing parameter (i.e. the balance between Lasso and ridge regression) α = 1 - ϵ for small $\epsilon > 0$ (parameter defines the amount by which one moves away from Lasso-only penalty,i.e. α=1), performs similar to Lasso, dealing with any extreme behaviour caused by highly correlated predictors. The general formula *Pa* of elastic net, as seen in equation 4.15, introduces a trade-off between ridge and Lasso. As α increases from 0 to 1 for a specific value of parameter λ, the sparsity of the solution in equation 4.17 (referring to the coefficients equal to zero), increases monotonically from 0 to the sparsity of the Lasso solution. More specifically, assuming that the response variable Y = {1, 2}, then the logistic regression model represents the class-conditional probabilities, through a linear function of the predictors, which in the logarithmic form is given by equation 4.14 (Friedman *et al.*, 2010).

$$log\frac{Pr(Y=1|x)}{Pr(Y=2|x)} = \beta_0 + x^T\beta \tag{4.14}$$

Where in this case the model is fit by regularised maximum binomial likelihood.

$$P_\alpha(\beta) = \sum_{j=1}^{p} \left[\frac{1}{2}(1-\alpha)\beta_j^2 + \alpha|\beta_j|\right] \tag{4.15}$$

Let $p(x_i) = P(Y = 1|x_i)$ be the probability according to equation 4.16.

$$P(Y = 1|x_i) = \frac{1}{1 + e^{-(\beta_0 + x_i^T\beta)]}} \tag{4.16}$$

For an observation *i* at specific values for the parameters (β_0, β), the penalised log likelihood is maximised.

$$\max_{(\beta_0,\beta)\in\mathbf{R}^{(p+1)}} \left[\frac{1}{N}\sum_{i=1}^{N}\left\{I(g_i = 1)logp(x_i) + I(g_i = 2)log(1 - p(x_i))\right\} - \lambda P_\alpha(\beta)\right] \tag{4.17}$$

Replacing , the log-likelihood part of equation 4.17 takes the form,

$$v(\beta_0, \beta) = \frac{1}{N}\sum_{i=1}^{N} y_i \cdot (\beta_0 + x_i^T\beta) - log(1 + e^{(\beta_0 + x_i^T\beta)}) \tag{4.18}$$

a concave function of the parameters. In this approach, for every value of λ, an outer loop is created for the computation of the quadratic approximation v_Q in equation 4.19 about the current parameters $(\tilde{\beta}_0, \tilde{\beta})$.

$$v_Q(\beta_0,\beta) = -\frac{1}{2N}\sum_{i=1}^{N} w_i(z_i - \beta_0 - x_i^T\beta)^2 + C(\tilde{\beta}_0,\tilde{\beta})^2 \tag{4.19}$$

where

$$z_i = \tilde{\beta}_0 + x_i^T\tilde{\beta} + \frac{y_i - \tilde{p}(x_i)}{\tilde{p}(x_i)(1 - \tilde{p}(x_i))} \tag{4.20}$$

$$w_i = \tilde{p}(x_i)(1 - \tilde{p}(x_i)), (weights) \tag{4.21}$$

Finally, the penalised weighted least-squares problem can be solved by equation 4.22, using the coordinate descent approach (Friedman *et al.*, 2010).

$$\min_{(\beta_0,\beta)\in\mathbf{R}^{(p+1)}} \Big[- v_Q(\beta_0, \beta) + \lambda P_\alpha(\beta) \Big]. \tag{4.22}$$

A number of sequential nested loops are created :

• **Outer loop:** Decrement λ.

• **Middle loop:** New quadratic approximation v_Q for the current parameters $(\tilde{\beta}_0, \tilde{\beta})$.

• **Inner loop:** Execute the coordinate descent algorithm on the penalised weighted least-squares problem (equation 4.22).

Further information of the above method is given by Friedman *et al.* (Friedman *et al.*, 2010; Zou & Hastie, 2005; Tibshirani, 1996).

4.3.3 All-Relevant Feature Selection (Boruta)

The minimal-optimal problem, which refers to the selection of a small feature set that gives the best possible classification result has been intensively studied in literature (Nilsson *et al.*, 2007; Kursa *et al.*, 2010). This is quite practical when one is mainly interested in building a classifier with high performance. The all-relevant features approach aims to accomplish something complementary to that, by identifying all attributes that are somehow relevant for the classification. As the creators of this approach claim, finding all relevant attributes, instead of only the non-redundant ones, is necessary when one is also interested in understanding the underlying mechanisms related to the topic of interest, instead of simply creating a black box with high predictive power.

This algorithm is a wrapper around random forests, because the important part is to have an efficient and relatively quick classifier that can provide a numerical feature ranking, just like random forests classifier does (Kursa *et al.*, 2010). As importance measure the authors use the Z-score, since it takes into account the fluctuations of the mean accuracy loss among the trees in the random forests. However, the Z-score is not directly related to statistical significance of the feature importance returned by the random forests. So the algorithm proceeds by creating a corresponding ”shadow” attribute, the values of which are obtained by shuffling the values of the original attribute across objects.

After that, all the attributes, both the shadowed and the original ones, are used to perform a classification and compute their importance (Z-scores). Then the maximum score among the shadowed attributed (MSA) is found (in the corresponding figures in chapters 4 and 6 this is referred as ”shadowmax”), which is compared with the corresponding scores of the other attributes. The attributes that are found to score statistically significantly higher than the MSA ones are marked as important. If there are attributes with undetermined importance, then a two-sided test of equality is performed between the attribute under question and the MSA, to determine the outcome. Similarly, the attributes that have importance significantly lower than the MSA are marked as unimportant and are permanently removed from the system. Finally the shadow attributes are removed and the process is repeated again until the importance is assigned for all the attributes, or the algorithm has reached the maximum number of runs or when it cannot converge (Kursa *et al.*, 2010).

In the beginning this algorithm is preceded by three start-up rounds, which at this stage they have less restrictive importance criteria. The purpose of the start-up rounds is to deal with the high fluctuations of the Z scores, when the number of features is large at the beginning of the process. All of the features are compared respectively with the fifth, third and second best shadow attribute. The rejection test is performed

at the end of each initial round, whereas the test for confirmation is not performed at all (Kursa *et al.*, 2010).

4.3.4 Validation & Metrics

In order to evaluate the performance and validate the classification models, a number of approaches were adopted. As was previously described, three feature selection processes were utilised. In addition, the evaluation process was conducted by using a repeated 10-fold cross validation (Devijver & Kittler, 1982) and 0.632 bootstrap (Kohavi *et al.*, 1995). The different models were compared according to their performance in terms of the average area under the receiving operators curve (AUROC or AUC), but additionally, other metrics are also utilised and/or reported, such as kappa, specificity and sensitivity (both of these are only used for building the ROC curve) and out of bag error (not for the model comparison, but as complementary to AUC metric; only for the random forests).

Cross validation (Geisser, 1993) is possibly the most widely used model validation technique to assess the performance of a classifier. Initially, the original sample is partitioned into k equally sized subsamples. Of these k subsamples, a single subsample is kept out as the validation sample for testing the model, with the remaining k - 1 subsamples used as training data. This process is then repeated k times, but making sure that each fold is used only once as the validation data. In our studies, given the relative small amount of features and observations, the above process is repeated 1000 times, getting 1000 AUC values and hence reporting the average AUC with the confidence intervals. This approach was used for the final, post-feature selection classification process.

0.632 Bootstrap has been proposed as a way of reducing the bias arising from the ordinary bootstrap method (Kohavi *et al.*, 1995). In general, bootstrapping refers to the process of using the data of a sample as a surrogate population, in order to

approximate the sampling distribution of a statistic, i.e. to resample with replacement from the sample data, known as bootstrap samples. Finally, the sample summary is estimated on each of the bootstrap samples. The 0.632 bootstap aims to reduce the bias arising from the ordinary bootstrap, which is due to non-distinct observations in the bootstrap samples that result from sampling with replacement. To resolve this they proposed the 0.632 estimator,

$$\hat{Error}_{0.632} = 0.368\bar{error} + 0.632\hat{Error}^{(1)} \tag{4.23}$$

in order to correct the upward bias in $\hat{Error}^{(1)}$ (leave-one-out bootstrap error estimate) , by averaging it with the downwardly biased estimate $\hat{Error}^{(0.632)}$. $\bar{Error}$ is the apparent error rate, or resubstitution rate. The two coefficients $0.368 = e^{-1}$ and 0.632 were suggested based on the fact that the bootstrap samples are supported on roughly 0.632*n* of the original data points (Kohavi *et al.*, 1995; Efron & Tibshirani, 1997).

Receiver Operating characteristic (ROC) curve (figure 4.2) is a graphical plot that provides us with important information about the performance of a binary classifier, as its discrimination threshold is varied. The curve in its *x* axis includes the false positive rate (FPR), also known as fall-out, and can be calculated as 1-*specificity* (Hanley & McNeil, 1982). This term gives us the proportion of positive samples that are erroneously identified as such. In its *y* axis it includes the true positive rate (TPR), also known as sensitivity or recall. ROC analysis provides a useful tool to select optimal models irrespectively from the cost context or the class distribution. The TPR (sensitivity) is a rate telling us the proportion of positive samples that are correctly classified as such (e.g. the percentage of DR individuals who are correctly identified as having the disease). A similar metric, specificity, also called the true negative rate (TNR), measures the proportion of negative samples that are correctly identified as such (e.g. diabetic subjects that are correctly identified as not having DR; convention-

ally one of the two classes represents the negative outcome without implying anything further). Therefore sensitivity, additionally quantifies the avoiding of false negatives in the same way that specificity does for false positives. Lastly the false negative rate (FNR) refers to the proportion of negative samples that are mistakenly classified as being negative. From a statistical point of view, false positives and false negatives can be thought of as concepts analogous to Type I and Type II errors, where a positive outcome corresponds to rejecting the null hypothesis, whereas a negative outcome corresponds to not rejecting the null hypothesis. From these four metrics one can create a 2-by-2 contingency table or else confusion matrix, to additionally visualise the results (Bradley, 1997).

An important additional metric that we get out of the ROC curve is the Area under the ROC curve (AUROC, or from now on AUC), or $c - statistic$, which together with the ROC curve gives an overall estimate of the performance of a binary classifier (Hanley & McNeil, 1982). AUC is equal to the probability that the classifier will rank a randomly chosen positive sample higher than a randomly chosen negative sample, i.e. $P(sample(x^+) > sample(x^-))$. When the classifier cannot distinguish between the two classes, i.e. there is no difference between the two distributions, the AUC will be equal to 0.5. In the case of perfect separation of the classes the value of AUC will be equal to 1. One important characteristic of the AUC, which depends on integral calculus, is that it makes no assumptions about the form of the x^+ and x^- distributions. The statistical properties of this important metric have also been stressed in literature, as the AUC is intimately connected with the Wilcoxon or Mann-Whitney statistical tests (Bamber, 1975; Hanley & McNeil, 1982).

Out-of-Bag error (OOB), or else out-of-bag estimate is a method of estimating the prediction error in random forests classifier, in decision trees but also in other machine learning models (Breiman, 2001). It utilises bootstrap aggregating to sub-sample the training data. This estimate is reliable and efficient in that it can be computed during

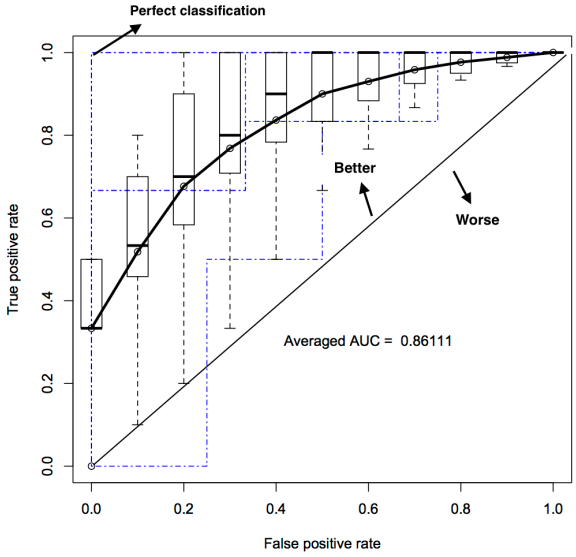

Figure 4.2: An example of the ROC curve that is going to be part of the results. On the x and y axes the false positive rate and the true positive rate can be found respectively. As can be seen the more we move towards the left upper corner the performance of the classifier improves. This plot is the averaged ROC curve over all the cross-validation runs, including the variation and the spread estimates around the average curve, by using box plots. The spread estimated are shown at 11 equally spaced positions along the curve.

the same run that constructs the bagged predictor. OOB is a random number, since it is based on random resamples of the data. Given a training set *P* with an output vector *y* and an input *x*, a predictor $f(x, P)$ is constructed and a given loss function $M(y, f)$ measures the error in predicting y by f. Next the bootstrap training sets $P_{(\kappa,B)}$ and the predictors $f(x, T_{(\kappa,B)})$ are formed, aggregating them in a way to form the bagged predictor $f_B(x)$. For each y, x in the training set, the predictors are only aggregated over those κ for which $P_{(\kappa,B)}$ does not contain y, x. If we denote these OOB predictors by f^{OOB}, then the OOB estimate for the generalisation error is the average of $M(y, f^{OOB}(x))$ over all the samples in the training set.

Accuracy is another common metric given by

$$ACC = \frac{TruePositive + TrueNegative}{Positive + Negative} \tag{4.24}$$

Table 4.1: Summary of the classifiers, feature selection process and evaluation metrics

Techniques	Overview	Process
Logistic regression	Linear classifier	Classification
Random forests	Non-linear classifier	Classification
Elastic-net	Regularisation parameter & mixing parameter for LASSO and Ridge regression	Feature Selection
Regularised random forests	Regularised information gain using penalty parameter	Feature selection
AUC	Represents a probability, the closer to 1 the better	Evaluation metric
ROC	Graphical plot for the performance of the classifiers	Evaluation metric
kappa	Adjusted accuracy based on the balance of the classes	Evaluation metric
OOB	Prediction error using bootstrap aggregating	Evaluation metric

and it merely refers to the corrected classified instances of the classifier in a specific cut-off point (denominator refers to the total sample size). Accuracy itself is an unreliable and misleading metric if used alone, especially when comparing the performance of different classifiers (Provost *et al.*, 1998). Therefore, it has to be accompanied by the ”no information rate”, which is the largest class percentage. That is because a 75% of accuracy in a balanced design (50-50) is not the same to an imbalanced design, where for instance the largest class represents the 70% of the whole sample. Therefore, there is another metric known as ”kappa”, which compares the observed accuracy with the expected accuracy, yielding a metric that can be used for evaluating both the performance of a single classifier and also many classifiers among themselves (Ben-David, 2008).

$$kappa = \frac{ObservedAccuracy - ExpectedAccuracy}{1 - ExpectedAccuracy} \tag{4.25}$$

where expected accuracy is defined as the accuracy that any random classifier would be expected to achieve based on the confusion matrix. A summary of the processes can be seen in Table 4.1

4.3.5 Classification Models

The classification process was conducted for the following combinations:

1. Year three versus onset of DR (progressors' group)
2. Year two versus onset of DR (progressors' group)
3. Year one versus onset of DR (progressors' group)
4. Mean diabetics versus Onset of DR (progressors' group)
5. Mean diabetics (non-progressors' group) versus onset of DR
6. Mean diabetics (non-progressors' group) versus mean diabetics (progressors' group)
7. Year 3 minus year 2 versus onset of DR minus year 1 (progressors' group)
8. Mean diabetics (non-progressors' group) versus mean diabetics (progressors' group) versus onset of DR (One versus All method)

The primary objective of this work is to evaluate the discriminative power of the entire feature set (candidate biomarkers), as an additional, complementary to the statistical analysis step. The secondary objective is to find the optimum model for each of the above combinations, in order to assess the appropriateness of these classifiers in having a possible clinical interest. To accommodate these purposes, the above combinations include balanced classes. The number of features are 112 in total and the relevant list is provided in Tables 3.3 and 6.15.

The purpose of the various evaluated combinations is to find out whether different periods of time and different cohorts can actually be efficiently classified, using exclusively geometric and haemodynamic features. This will give us a very important indication of how the feature sets perform in practise, and also whether they change

according to the combination. **Combinations 1, 2 and 3** will show how extensive the alterations are as we move towards the onset of DR, always in reference to the onset of DR. **Combination 4** includes the average of the three year period before the onset of DR, which offers a more realistic and representative sample for a wider period of time. **Combination 5** is similar to **combination 4**, but during this time, the group contains a non-progressors' diabetic cohort, for which the last four consecutive years of diabetes are combined and averaged, in order to form this group that will be tested against the DR group. **Combination 6** aims to actually find out, whether two diabetic groups that on first sight belong to the same category, in reality they come from predefined periods, for which the prior information that one of them will progress to DR shortly, is available. However in a prospective study this information will not be available and therefore subjects from both categories are mixed up, adding much unnecessary variation in the sample. The purpose is therefore to find out if indeed two such diabetic cohorts are starting early to be affected by DR to a level that defines the progression. **Combination 7** intends to show whether the vascular changes from the three years- to two years- pre-DR are more or less extensive or even non-existent at all, compared to the changes when progressing from one year pre-DR to the onset of DR. Finally, **combination 8**, which is an extension of combination 5, 6 and 7, will provide us with an overview of whether the three main groups, non-progressors diabetics, progressors diabetics and DR subjects, can actually be efficiently differentiated among each other, but this time altogether. As it will be stressed more in details in chapter 6, if this combination yields a good performance, it can possibly have a direct clinical application, since the first two categories, i.e. non-progressors diabetics and progressors diabetics, can denote the progression towards DR, which can be an indication for the clinicians to run more tests or try to adjust the patient's treatment plan.

4.3.5.1 Feature Selection Process

The feature selection process includes various steps. For the combinations 1-7 the steps are exactly the same, but for the combination 8 the process is slightly different. Everything was fulfilled using Rstudio and a number of packages (RStudio Team, 2015; Kuhn, 2016; Kursa & Rudnicki, 2010; Friedman *et al.*, 2010; Deng & Runger, 2013; Sing *et al.*, 2005).

As was previously described , each of the feature selection algorithms follows a specific process for coming up with a feature subset. Given the randomness behind all the processes involved, the feature subsets slightly change every time the algorithms are executed. Initially the data are randomly partitioned into 10 balanced folds. Leaving one fold aside at a time, the feature selection process is executed 10 times in total, ending up with 30 different models (10 from each selection process). Each of these models will inevitably probably include a slightly different feature subset. In order to arrive at this point, the whole feature set is penalised by the algorithms, trying to find the optimum subset. Given the complexity of the feature selection processes, in order to select the optimum regularisation parameter for each fold that gives the final subset, a 0.632 bootstrap process is run, for 1000 iterations in total.

Regarding the elastic-net, the parameters of α and λ were bootstrapped within the range of 0 and 1, with a step of 0.02. Similarly, the regularised random forests process was bootstrapped for the parameters mtry (4 and 5; 6 and onwards were giving inferior results), i.e. the square root of the number of predictors sampled for splitting at each node rounded down , the penalty λ and the importance coefficient κ within the range of 0 and 1, with a step of 0.02. At the end of this process the optimum parameters that give the best classification result (based on AUC) is estimated, which in turn gives the final feature subset for the respective fold. Figures 4.3, 4.4 and 4.5 show an example of the bootstrap process for each of the three feature selection techniques. As previously described, this process defines the final regularisation parameters, and consequently a

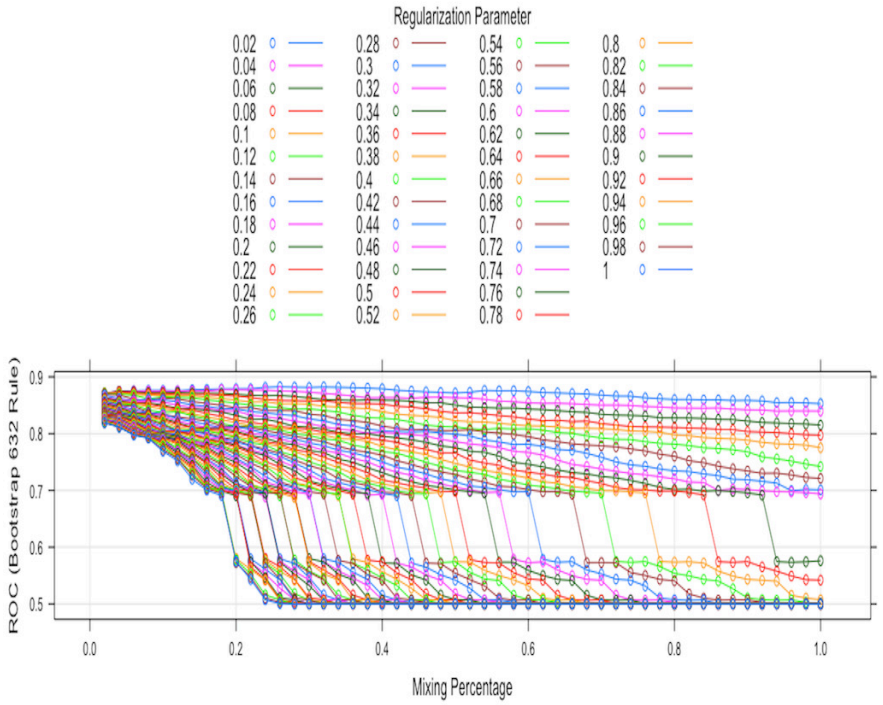

Figure 4.3: The plot shows the bootstrap process for ending up with the optimum regularisation parameters. The x axis shows the mixing parameter, $alpha$, which is the trade-off between Lasso and Ridge regression. The y axis includes the 0.632 Bootstrapped AUC. Each line inside the plot represents the penalty parameter λ.

feature subset, that yield the best performance for a given fold (each of the 10 folds). Figure 4.3 is presented here for purely visualisation purposes, in order to show how the combinations of the different values of the regularisation (λ) and the mixing parameters (α) in the elastic net yield varying performances. The combination of values that scores higher in the y axis (ROC) is selected, providing us with the corresponding feature subset.

4.3.5.2 Validation Process

Having created the 30 models in total from the elastic-net, the regularised random forests and the all-relevant features selection process (boruta) for each combination (apart from the combination 8), it is now time to find out, which feature subset performs best in the whole dataset. A 1000 repeated 10-fold cross validation process is followed,

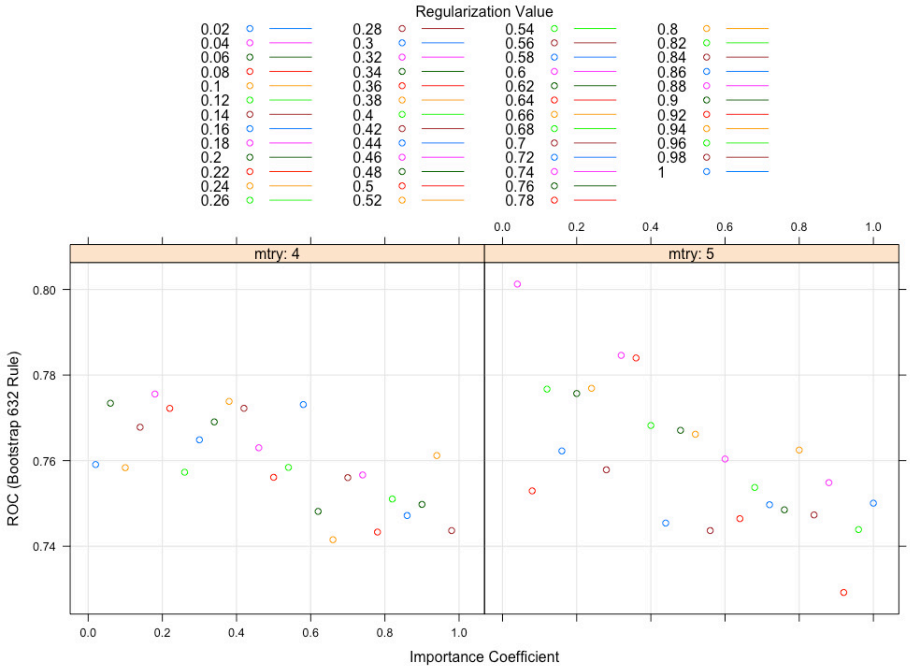

Figure 4.4: This plot is split into two parts, one for each of the two mtry values. It includes the importance coefficient on x axis and the 0.632 bootstrapped AUC on y axis. The dots are the different regularisation values, which in turn define the corresponding feature subsets.

which gives an averaged ROC curve with box plots in different points on the curve, and also the average AUC with bootstrapped confidence intervals. And all that for each of the 60 classifications (30 with the random forests and 30 with the logistic regression classifiers). A plot with a bootstrapped AUC will also be reported for every combination, mostly for visualising how each bootstrap sample performs.

The five best models out of this process for each of the seven combinations will be presented in chapter 6. The final models provide a solid representation and a final verdict of which features are effective predictors and good biomarkers for the respective combination that is under investigation each time. A ranking of the twenty best features, based on the times that they were selected in all of the 420 models will also be shown.

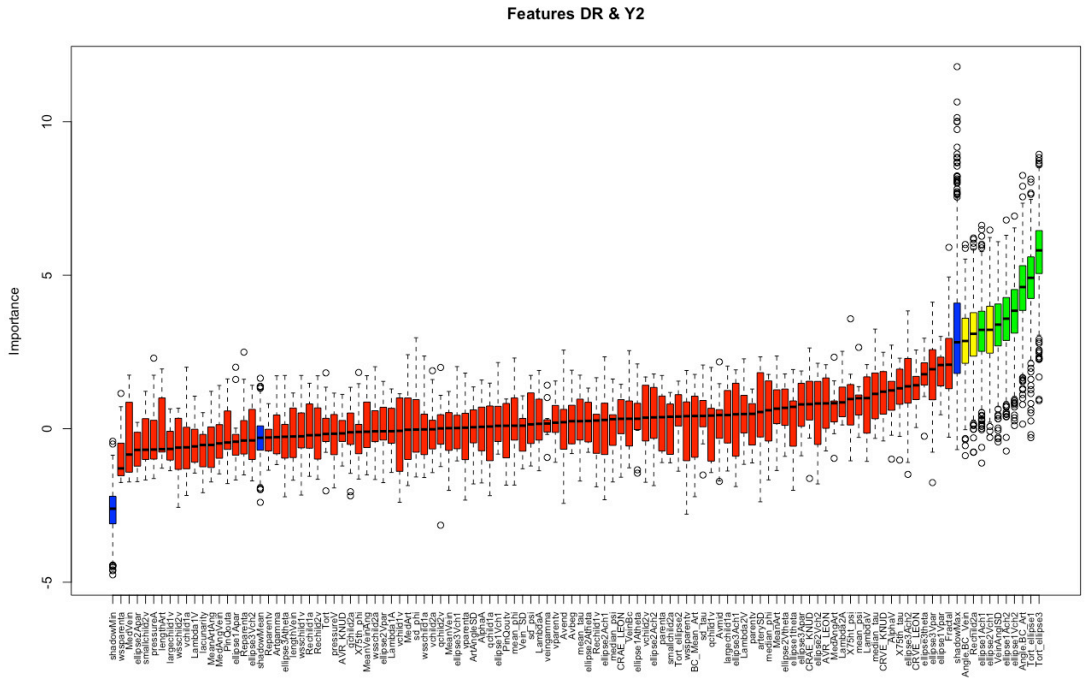

Figure 4.5: Example of the selection process with the all-relevant features algorithm. On x axis the group of features with the relevant box plots can be seen, starting from the less important (left) to the more important (right). Red coloured features represent the non-selected ones, whereas the green coloured are the important ones. The blue coloured boxes show the minimum, mean and maximum importance of the shadow attributes, in order to have an idea where the candidate features stand in comparison with them. Values that are statistically significantly higher than the shadowmax are declared as confirmed features.

4.3.5.3 One Versus All Approach

In contrast to the other combinations, in combination 8 the aim is to figure out whether a classification model that combines the three main scenarios that are of our interest, when building a classification system for diagnosing the disease and the progression, is plausible. These scenarios refer a) to the non-progressors's group, b) to the progressors' group (years of diabetes but no DR) and c) to the group of patients with DR. This combination is of high clinical importance because the progressors' group state is the pre-DR situation, which we would ideally like to know. Given that DR is just around the corner, possible treatment interventions to slow down the proliferation might be feasible. Therefore, this classification model, which includes these three scenarios, could be used to identify not only the transition between the diabetic state and the

pre-DR one (or alternatively progressors's state), but also DR.

The one versus all method (Rifkin & Klautau, 2004) reduces the problem from multi-class to multiple binary classification problems, and was preferred due to the complexity and the few available observations that are not enough for building a meaningful multi-class classifier.

The process starts by training a single classifier per class, with the samples of this class as positive and the rest as negatives. To achieve that, for each combination of classifiers (in this case three), the probabilities must be returned instead of the predicted class labels. The final prediction is then performed by executing these binary classifiers and finally choosing the prediction instance with the highest confidence score (probability). Similarly to the other combinations, the evaluation of this ensemble of individual models is decided by a 1000 repeated 10 fold cross-validation, getting one ROC curve for each of the three cases. Three feature subsets were used to test this ensemble, which are the best feature subsets of the combinations 4, 5 and 6. Given that the best performance in these combinations was achieved by the random forests, this classifier was used for the ensemble, without any regularisation.

4.4 Conclusion

This chapter gives some important information regarding the different ways that the biomarkers are going to be evaluated and tested in practise. The methods of the statistical analysis together with the metrics that are adopted for the evaluation and validation processes were described, focusing on the linear mixed models, which are the basis of the whole statistical analysis. Furthermore, the feature selection and classification processes were also elaborated, giving detailed information about the different combinations, based on which the multiple classification models are going to be built and evaluated. In chapter 6, the results and findings of the all the studies will be presented,

after applying the methods and techniques described in the present chapter as well as in chapter 3.

Chapter 5

Summarising the Retinal Vascular Calibres

5.1 Introduction

5.1.1 Background

As previously described, the human retina is an important and non-invasive window for monitoring the blood vessels, being part of the brain's vascular system (Dowling, 1987). In numerous studies in literature, alterations of the retinal vessels have been associated with diseases, like diabetic retinopathy (DR) (Leontidis *et al.*, 2015a), hypertension (Wigdahl *et al.*, 2015) and glaucoma (Arend *et al.*, 2002). In diabetes and DR especially, changes have been identified both in the blood flow and the vessel wall structure, leading to geometric alterations, which can be measured in fundus images (Bursell *et al.*, 1996; Gardner *et al.*, 2002; Kohner *et al.*, 1995; Barber *et al.*, 2011). DR, at the advanced stages, has already caused serious damages and can lead to blindness. Early detection of these alterations, before the onset of DR, could possibly assist the clinicians to decide a better series of treatment interventions.

The retinal vascular geometry offers a unique opportunity to study a number of

different diseases. However, to do so, reliable and robust methods are needed in order to extract all of the relevant information for running the statistical analysis. Accurately measuring and summarising the calibre (width) of the parent vessel trunks in the most efficient way, no matter how subtle the improvement will be, can contribute towards defining biomarkers of the progression of a disease. Additionally, taking into account that during the proliferation of retinal diseases, as defined by the anatomical structure of the retinal vessels, any alterations are very fine, any possible improvement, regardless of how negligible it might be, is undoubtedly important.

The main purpose of this chapter's study is to provide an alternative method for quantifying the retinal vessel calibres. Even the slightest improvement of this quantification can have a direct effect on studies of retinal pathologies, especially if we take into account that in different diseases, like diabetes and DR, the vascular changes are quite subtle (even less than 5%), especially during the consecutive years of the progression of the disease (Leontidis *et al.*, 2015a). In addition, these changes are difficult to quantify in digitised images, which thus may warrant and justify the innovation of methods to quantify those small changes of the vessel calibres. The proposed method aims to improve and for the first time to specify the quantification for healthy, diabetic and DR eyes.

5.1.2 Related Work

As it was briefly described in chapter 2, the central retinal artery (CRA) and vein (CRV) are the two main vessels of the retina. The CRV leaves the optic nerve head, draining the blood from the capillaries directly into either the superior ophthalmic vein or to the cavernous sinus. Meanwhile, the CRA branches off the ophthalmic artery, passing under the optic nerve head within its dural sheath to the eyeball (Rodieck, 1973). Initially Parr *et al.* (Parr & Spears, 1974a,b) and then Hubbard *et al.* (Hubbard *et al.*, 1999) completed some important work in the quantification of the retinal

vessel calibres, which was then extended and improved by Knudtson *et al.* (Knudtson *et al.*, 2003). In their methods, the central retinal vein and artery equivalents (CRVE and CRAE) were proposed. In both cases, the calibre measurements around the optic nerve head were combined into one index, giving us the CRVE and CRAE. All of the measurements are collected around the optic nerve head, in an area of interest that lies within 0.5 to 1.0 optic disc diameters from the optic disc margin. To improve the estimation, Knudtson *et al.* revised the formulas used in the previous studies, in order to be independent of the units of measurement (e.g. pixels or microns) and without constant terms. The derived formulas for arteries and veins can be seen in equations 5.1 and 5.2 respectively, which are based on branching coefficient values of 1.28 for arteries and 1.11 for veins.

$$Arterioles : \hat{w} = 0.88 * \sqrt{(w_1^2 + w_2^2)} \tag{5.1}$$

$$Veins : \hat{w} = 0.95 * \sqrt{(w_1^2 + w_2^2)} \tag{5.2}$$

where $\hat{w}$ is the estimate of the parent trunk arteriole or venule and w_1, w_2 are the two branches (children).

It is worth mentioning that the data used to derive these formulas included only normotensive subjects. However, as has been shown in previous studies (Leontidis *et al.*, 2015a, 2016a), applying these formulas to diabetic and DR subjects give us different estimates of the BC. This could lead to errors in the measurements, in the event that they are about to be used in summarising the vessel calibres in diabetic and/or DR patients.

Following a similar logic, Patton *et al.* attempted to improve the above formulas for calculating the BC, recruiting healthy subjects to measure the retinal vessels calibres (Patton *et al.*, 2006a). Following that, they tested the performance on a different

group of people that they also recruited. Their main method included the design of a least-squares linear regression model for the arteries only, since the main independent variable, the asymmetry index, was found to be related only to the arteries and not the veins. Their developed model can be seen in equation 5.3.

$$BC\_arteries\_model = 0.78 + 0.63 * Asymmetry\,index \tag{5.3}$$

5.1.3 Proposed Approach

The relationship between parent and daughter vessels at vascular junctions has been expressed by using the junction exponent (x) (Murray, 1926b). It has been shown that deviations from the optimal value, as predicted by Murray's cubic law (x=3), can be an indication of a vascular disease. Witt *et al.* (Witt *et al.*, 2010) proposed an alternative parameter, which they called the optimality ratio, that is simpler to calculate and more robust in the presence of noise than Murray's junction exponent.

In addition, in another study (Takahashi *et al.*, 2009), they assessed the arteriovenous distribution of haemodynamic parameters throughout a microvascular network inside the human retina. They suggested an alternative value for the exponent, based on the fractal dimension and a branch exponent, with the latter derived from data on cerebral vessels, mainly because they share similar structural and functional characteristics as the retina (Kamiya & Takahashi, 2007). The reasoning behind this is that Murray's theoretical value of 3, actually ranges between 2.7 and 3.0 in various vascular beds (Sherman, 1981; Suwa & Takahashi, 1971). Fractal dimension has been vastly studied in literature and is primarily used to quantify complex vascular networks like the retina (Mandelbrot, 1983). The branch exponent, which is also consistent with a fractal recursive rule, can also define the relation between the length and radius of a branch segment as $L(r) = 7.4r^a$ (Suwa & Takahashi, 1971). In the study of Takahashi *et al.* they assume a value of 1.7 for fractal dimension and 1.15 for the branch exponent

(Takahashi *et al.*, 2009).

In the proposed method, the mean diameter ratio is combined with a revised exponent, which has been derived from data of healthy, diabetic and DR subjects. Therefore, the final exponent, although part of it utilises the standard branch exponent of 1.15, as suggested in (Takahashi *et al.*, 2009; Kamiya & Takahashi, 2007), it is combined with the fractal dimension, as was experimentally calculated in these three groups. The newly derived formulas were compared to the BC of Knudtson *et al.* and the linear regression model of Patton *et al.* for all the above cases, using the mean absolute percentage error (MAPE).

The following study is organised into four main sections. In **Section 1**, a thorough description of the methods and techniques is given, alongside the tools and the data that were used. In **Section 2**, more details will be given about the study of the evaluation of the AVR and the arteries and veins individually, as biomarkers of progression to DR. **Section 3** will be devoted to presenting the results in two parts. The first part will present the comparisons between the three techniques for summarising the retinal vessel calibres. The second part will present the results and the statistical analysis for the study of progression to DR, in the four different groups, using a repeated measures mixed model design. In the **last section**, the implications and importance of this study will be stressed, followed by the conclusion.

5.2 Methods and Tools

5.2.1 Data Collection

As previously mentioned , numerous studies have shown that the retinal vasculature changes during the progression of diabetes and DR (Barber *et al.*, 2011; Wong *et al.*, 2006; Antonetti *et al.*, 2006; Klein *et al.*, 1998; Yang *et al.*, 2016).Therefore, in order to improve the accuracy of the calibre's estimation of the parent vessel trunk for both

arteries and veins, different data for healthy, diabetics with no-DR and DR subjects were utilised. All of the individuals in the diabetic and DR groups had no history of hypertension or any other cardiovascular disease. The healthy group included 25 images, coming from a previous unpublished study, conducted at the School of Computer Science at the University of Lincoln, UK (3888X2592 pixels). The diabetic group included data from 25 diabetic subjects taken annually, over a three-year period, before the onset of DR (3 years pre-DR, 2 years pre-DR and 1 year pre-DR). The same exact junctions were used in all three cases, which were then averaged and used to derive the formulas, constituting one group (age range 35-60 years, mean 47.5 (± 7.48)). The rationale behind this is that in order for the formulas to be useful in estimating the vessel trunk in diabetic subjects, who can be at different stages of diabetes, more than one year is needed to have a more accurate representation of the junctions.

Regarding the DR group, 25 DR patients (first year that the first lesions appeared). The data of both these groups come from a diabetic/DR screening database in the UK (3216X2316 pixels) (age range 37-70, mean 53.2 (± 9.63)). The image resolution in the above groups does not affect the implementation of the method, because the final estimated value comes from the ratio between the calibres of the two children vessels (nominator) and the calibre of the parent vessel. Moreover, in all of the above cases, either the left or the right eye was used, chosen at random before the beginning of the study.

5.2.2 Tools

The junctions were labelled and measured using a semi-automated tool described in (Al-Diri *et al.*, 2010) and in chapter 3. In brief, using this tool, each junction is firstly manually labelled as being a vein or an artery and then in a second stage the calibre is determined by fitting a rectangle between the wall edges of each vessel (in pixels). The entire process is conducted in a semi-automated manner, going through each junction.

The rectangle is adjusted by the user until she/he is confident that the measurement is correct. In that way, if a junction cannot be accurately measured, then the user can discard and select another one, making sure that the included measurements are accurate. The process was repeated twice for every junction, taking the average of the two measurements, after calculating the intra-rater reliability (95.6% (95%CI=94.3, 96.8) of agreement). The performance of the expert user of the tool (myself) has been validated against a fully automated tool as proposed in (Al-Diri *et al.*, 2009), with the average agreement being 96.1% (95%CI=95.1, 97.9).

The fractal dimension estimation was calculated using the well-established box-counting algorithm, which was described in chapter 3 (Mandelbrot, 1983; Li *et al.*, 2009). Initially, the fundus images are segmented using a tram-line algorithm (chapter 3) as proposed in (Hunter *et al.*, 2005) (figure 5.1). In the second stage, the box-counting method is applied in the binary segmented trees, getting the fractal dimension.

5.2.3 Γ Ratio

It is well known that the relationship between the diameters of the parent and daughter vessels at a junction is described by the junction exponent(x) defined by equation 5.4.

$$d_0^x = d_1^x + d_2^x, \tag{5.4}$$

where d_0 is the diameter of the parent vessel and d_1, d_2 are the daughter vessels' diameters (Murray, 1926b). Murray's law predicts that this exponent is equal to 3 under conditions of optimum power loss in the bifurcation. A few studies have shown that healthy veins and arteries follow Murray's law with the exception of very large vessels (Sherman, 1981; Chapman *et al.*, 2002).

Witt *et al.* suggested an alternative parameter to characterise the optimality of the relationship between the vessel's diameter at a bifurcation, which they called the

optimality ratio, aiming to reduce the bias and variability in the presence of noise in the measurements of individual vascular diameters, from which the junction exponent calculation suffers. This optimality ratio is based on the non-dimensional variants ζ_1, ζ_2 of the daughter diameters at a bifurcation (equation 5.5).

$$\zeta_1 = \frac{d_1}{d_0}, \zeta_2 = \frac{d_2}{d_0} \tag{5.5}$$

After a few transformations, the mean diameter ratio Γ is calculated by the equation 5.6.

$$\Gamma = \frac{d_1 + d_2}{2d_0} = \frac{\zeta_1 + \zeta_2}{2} \tag{5.6}$$

Finally, after some algebraic manipulations of the mean diameter ratio and the asymmetry factor, they define the optimality ratio for a bifurcation obeying Murray's cubic law as

$$\Gamma_{ratio} = \left(\frac{d_1^3 + d_2^3}{2d_0^3} \right)^{\frac{1}{3}} \tag{5.7}$$

More information for the above derivation of the formula is given in Witt *et al.* (2010).

5.2.4 Experimental Exponent

As previously mentioned , Murray's cubic law of x=3 has actually been found to range between 2.7 and 3.0 for various vessels. Following the mathematical proofs in (Takahashi *et al.*, 2009), who showed that this exponent can be more accurately approximated, by combining the fractal dimension (estimated in this study) and a branch exponent (taken by (Kamiya & Takahashi, 2007; Takahashi *et al.*, 2009), various exponents were derived for this study, from data of healthy, diabetic and DR eyes.

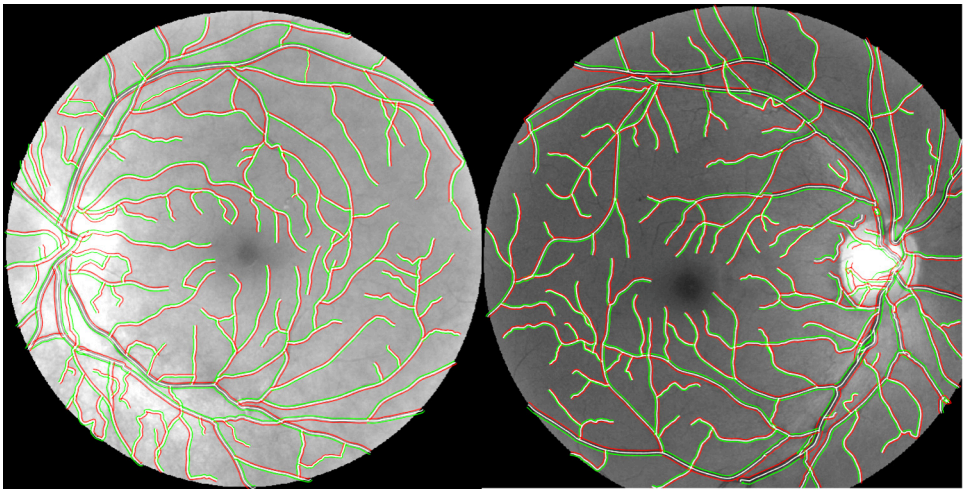

Figure 5.1: Two segmented images, as used in this study.

5.2.5 Estimation of Fractal Dimension

All of the 125 images used for this specific study (25 for healthy, 75 for diabetic and 25 for DR eyes) were segmented and the binary trees were used to calculate the fractal dimension (figure 5.1). In the diabetic group, the fractal dimensions of the same retinas were averaged over the three-year period, in order to get a more accurate representation of the multiple and versatile changes that occur during the progression of diabetes. The final 25 measurements for each of the three groups (healthy, diabetic and DR) were summarised using both the mean and median values. Therefore, all of the experiments were run using both the mean and median fractal dimension to derive the exponent.

5.2.6 Adjusted Γ Ratio

The general form of the adjusted Γ ratio, using the experimentally measured fractal dimension, combined with the branch exponent of 1.15 is summarised in equation 5.8,

$$Adj.\Gamma_{ratio} = \left(\frac{d_1^m + d_2^m}{2d_0^m}\right)^{\frac{1}{m}}, \tag{5.8}$$

where m is taking six different values, for the three groups and for the mean and median fractal dimensions. By solving equation 5.8 with respect to the parent vessel d_0, the generalised equation 5.9 is obtained, which, after adjusting for the different exponents, will be used in all of the following experiments and will be compared with the other two methods in the literature (Knudtson *et al.*, 2003; Patton *et al.*, 2006a).

$$trunkvessel\_d_0 = \sqrt[m]{\frac{d_1^m + d_2^m}{2 * \sqrt[m]{adj\Gamma_{ratio}}}} = \sqrt[m]{\frac{d_1^m + d_2^m}{2 * (adj\Gamma_{ratio})^m}} \tag{5.9}$$

5.2.6.1 Group Specific Γ Ratio

In Table 5.1, the experimental results for the three groups can be found. Depending on the value of the exponent, the adjusted $\Gamma ratio$ has been calculated based on equation 5.8. For each exponent, the corresponding equation was derived, in order to calculate the performance within the different groups, for both veins and arteries, using the mean and median FD. In total, twelve equations were obtained, to be used and compared with the other two methods (Knudtson *et al.*, 2003; Patton *et al.*, 2006a).

$$Vein : Healthy\_d_0\_mean = \sqrt[2.8]{\frac{d_1^{2.8} + d_2^{2.8}}{0.8894}} \tag{5.10}$$

$$Vein : Healthy\_d_0\_median = \sqrt[2.78]{\frac{d_1^{2.78} + d_2^{2.78}}{0.8936}} \tag{5.11}$$

$$Vein : Diabetic\_d_0\_mean = \sqrt[2.76]{\frac{d_1^{2.76} + d_2^{2.76}}{0.9377}} \tag{5.12}$$

$$Vein : Diabetic\_d_0\_median = \sqrt[2.79]{\frac{d_1^{2.79} + d_2^{2.79}}{0.9317}} \tag{5.13}$$

$$Vein : DR\_d_0\_mean = \sqrt[2.74]{\frac{d_1^{2.74} + d_2^{2.74}}{0.9314}} \tag{5.14}$$

$$Vein : DR\_d_0\_median = \sqrt[2.77]{\frac{d_1^{2.77} + d_2^{2.77}}{0.9253}} \tag{5.15}$$

$$Arteries : Healthy\_d_0\_mean = \sqrt[2.8]{\frac{d_1^{2.8} + d_2^{2.8}}{1.079}} \tag{5.16}$$

$$Arteries : Healthy\_d_0\_median = \sqrt[2.78]{\frac{d_1^{2.78} + d_2^{2.78}}{1.0826}} \tag{5.17}$$

$$Arteries : Diabetic\_d_0\_mean = \sqrt[2.76]{\frac{d_1^{2.76} + d_2^{2.76}}{1.118}} \tag{5.18}$$

$$Arteries : Diabetic\_d_0\_median = \sqrt[2.79]{\frac{d_1^{2.79} + d_2^{2.79}}{1.1129}} \tag{5.19}$$

$$Arteries : DR\_d_0\_mean = \sqrt[2.74]{\frac{d_1^{2.74} + d_2^{2.74}}{1.152}} \tag{5.20}$$

$$Arteries : DR\_d_0\_median = \sqrt[2.77]{\frac{d_1^{2.77} + d_2^{2.77}}{1.1465}} \tag{5.21}$$

Table 5.1: Group specific measurements

Group	Number of Junctions (Images)	Mean/Median FD (SD)	Adjusted Γ_{median} (CI 95%)	Adjusted Γ_{mean} (CI 95%)	Exponent m (mean/median)
Healthy Veins	150(25)	1.646/1.633 (0.03)	0.7484 (0.736,0.7564)	0.7487 (0.7363,0.7568)	2.8/2.78
Diabetic Veins	450(75)	1.613/1.638 (0.035)	0.7605 (0.7514,0.77)	0.76 (0.751,0.7696)	2.76/2.79
DR Veins	150(25)	1.591/1.621 (0.028)	0.7571 (0.7519,0.7714)	0.7566 (0.7418,0.771)	2.74/2.77
Healthy Arteries	150(25)	1.646/1.633 (0.03)	0.8019 (0.7886,0.8084)	0.8022 (0.7888,0.8086)	2.8/2.78
Diabetic Arteries	450(25)	1.613/1.638 (0.035)	0.8105 (0.789,0.8185)	0.81 (0.787,0.862)	2.76/2.79
DR Arteries	150(25)	1.591/1.621 (0.028)	0.818 (0.801,0.837)	0.8176 (0.799,0.833)	2.74/2.77

As can be seen in Table 5.1, the final values for the exponent range between 2.74 and 2.8. These values are within the previous suggestions found in (Sherman, 1981; Suwa & Takahashi, 1971), and also different from the rough estimation of 2.85 in (Takahashi *et al.*, 2009).

5.2.7 Methodology of the Estimation of the AVR

Knudtson *et al.* suggested using a region around the optic nerve head in order to estimate the CRVE, the CRAE and the AVR. They also observed that the number of included vessels does not significantly affect the final estimation (Knudtson *et al.*, 2003). This methodology was adopted here to make the AVR estimation. Therefore the area used can be seen in figure 5.2.

However, although the number of vessels might not significantly affect the final estimation, the branching order might play an important role. In an example that will be presented later on (cases 1, 2 and 4 versus case 3), the final estimation is affected when vessels of different branching order are combined. Nevertheless, just looking at the fundus images, one cannot be certain about the branching order of the vessel that is measuring, since the vessels might have branched before they exit the optic disc. This also makes it practically impossible to identify in which subtree a specific segment originally belongs to .

Having said this, in many retinas, the excluded area, which is 0.0-0.5 optic disc diameters from the optic nerve head margin, might include the branching of some segments. Therefore zone B (figure 5.2) might include vessels of different branching order. The same problem however exists with the area inside the optic disc, where the branchings cannot be identified as well.

Taking into consideration the above reasons and that the branching order cannot be accurately defined, despite the limitations of this methodology, the established zone B is adopted here, where the vessels are larger and more accurately measured, with their status already attained.

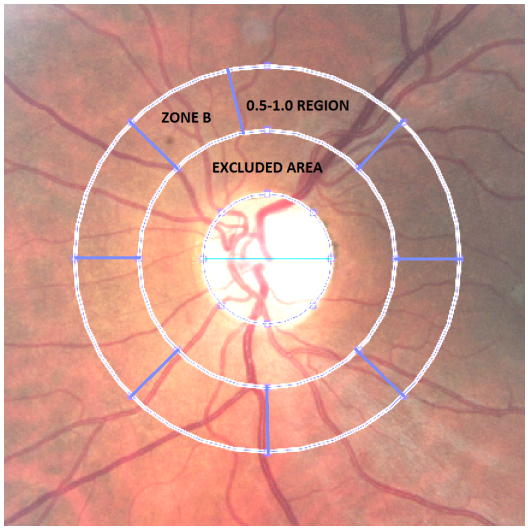

Figure 5.2: Schematic representation of the region used for calculating the CRVE, CRAE and AVR (Knudtson *et al.*, 2003).

Another issue is that in (Knudtson *et al.*, 2003), the six vessels are summarised using an iterative process, where the largest vessel is paired with the smallest one and then the second largest with the second smallest and so on. However, experimentally testing that, by summarising the calibres of four different parent trunk vessels (4 cases), using the six subsequent vessels (children) of the same tree, showed that both this method and simply taking the vessels in an ascending order, leads to mixed results (two versus two), when compared with the actual measured trunks (Table 5.3). Therefore, according to these estimations, there is no evidence that one method is consistently more accurate than the other. This suggests that either can be used to obtain the AVR, as long as it remains consistent.

5.2.8 Comparison of the Proposed Method with the State of the Art

In contrast to the method in (Knudtson *et al.*, 2003), the method in (Patton *et al.*, 2006a) includes only the arteries, hence the comparison with this one will be limited to the arteries. For the other method a full comparison will be made. In addition, despite the fact that in previous studies (Leontidis *et al.*, 2015a, 2016a), updated branching coefficients for both diabetic and DR groups have been reported, the purpose of this study is not to revise the established methods rather to propose a new one. Therefore these methods are utilised exactly as they have been described in literature and used by other researchers.

5.2.8.1 Statistical Analysis

In order to compare the performance and accuracy of the estimations of the proposed method against the other two, the mean absolute percentage error (MAPE) is used,

which is defined as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - P_i}{A_i}\right| * 100\%, \tag{5.22}$$

where A_i is the actual measurement and P_i is the predicted one.

The whole sample of junctions is initially partitioned into 10 equal parts, in order to run a repeated 10-fold cross validation. For instance, for the healthy group, each time the formulas are derived using 135 junctions and tested with the other 15; this process is repeated for every combinations, for a total 1000 times. This means that all of the above equations, as well as the reported errors, are the average of all the iterations. In this way it is avoided the bias that arises from using the same data for training and test purposes.

In addition, in order to justify taking multiple junction measurements from the same image, the within each image variation was compared with the between images variation, using the intraclass correlation coefficient (Snijders, 2011; Fisher, 1925).

For the progression study, using the individually measured matched vessels of each image, a mixed-effects model was used, analysed by a repeated measures analysis of variance (ANOVA) (Fisher, 1936), as described in (Leontidis *et al.*, 2015a), as well as a linear mixed model (Bates, 2007; McCulloch & Neuhaus, 2001) with a random intercept for every subject. Both model types account for the multiple measurements taken within the same fundus images, and in balanced designs, they produce similar results. Adding random effects in a model allows us to make inferences about a population, which is the desired situation in this study. Dealing with a non-repeatable covariate, one aims to characterise the variation induced in the response by the different levels of this covariate, thus adding it as a random effects term. As described in details in chapter 4, the random intercept lets us model the fact that each individuals' baseline can be different, hence each subject is assigned a different intercept value

(differences can be seen in two examples in figure 5.3). An even more complicated model could include a random slope as well, which can model varying correlations within a subject. However, in this case, comparisons between the more complicated random intercept/random slope models and the random intercept only models, yielded no significant difference for any case, therefore the simplest model approach (only random intercept) is preferred.

Taking the opportunity of measuring the vessels around the OD, these will also be evaluated, in order to test whether they follow the same trend like the junctions of the main study (category 1 of chapter 6). However they are going to be elaborated in this chapter, because they are not part of the main studies (no junctions).

Normality, using the Shapiro-Wilk test (Shapiro & Wilk, 1965), and sphericity, using the Mauchlys test (Mauchly, 1940), were run prior to the analysis, to make sure that the data can be analysed using the ANOVA. Similarly, the residuals in each model were also evaluated in order to make sure that they are normally distributed (figure 5.4). Accordingly, the same assumption stands for the random effects of each model.

Post-hoc comparisons were also made for the significant results, adjusted according to the Bonferroni correction (Bland & Altman, 1995), for retaining the family-wise error rate, in order to specify which group means actually differ. All of the analyses were fulfilled using a balanced design. It is worth pointing out, that the dependent variables are the veins, arteries, CRVE, CRAE and AVR, and the independent variables are the different stages of the disease (fixed effect) and the factor ”subjects” with random intercept (random effect). In the case of the analysis of the multiple vessel measurements, the ones used to calculate the CRVE and CRAE, an additional random effects factor is used (the within each subject’s vessel measurements), which is a nested within the subjects factor .

In the case of the linear mixed model, the (AIC) (Akaike, 1998) is used to evaluate each model, together with the corresponding p-value of the ANOVA (the smaller the

AIC the better; the relative value does not imply anything). To do so, the full model (the one with the fixed effect) is compared with the restricted model (without the fixed effect), in order to see whether this effect leads to a better model, i.e. having a significant effect to the dependent variable. As was described in a previous chapter, due to the bias of maximum likelihood (ML) for the estimation of variance components, it is suggested that linear mixed models are fitted using the restricted maximum likelihood (REML) (Corbeil & Searle, 1976). However, in order to compare the two models (full and restricted), we need to refit them using the maximum likelihood (ML) method. The reason is that ML estimates the variances as if the fixed parameters are known, so it does not account for the degrees of freedom lost during the estimation. On the other hand, the REML adjusts for the uncertainty of the fixed parameters. Therefore, we generally cannot use REML to compare models, because whatever difference in the fixed effects part invalidates the comparison.

Finally, Table 5.5 is complementary to Table 5.4, showing for each feature, the corresponding variance and standard deviation values of the random effects (intercept and residuals), accompanied by the parametric bootstrapped confidence intervals (after 1000 simulations) (DiCiccio & Efron, 1996). In addition, the ICC is also reported (derived by the restricted model), which gives us the proportion of the total variance that is explained by the respective blocking/grouping factor (random effect), telling us if it is worth fitting a mixed model instead of a fixed effects only (West *et al.*, 2014). The ICC (Starkweather, 2010) values in Table 5.5 definitely suggest that the proper and more meaningful way to fit the model is using a mixed effects model.

5.2.8.2 Summary of the Trunk Vessel

In figure 5.5 two example images out of the four can be seen that were used to estimate the trunk veins. In each of these cases, the trunk vein is estimated from the subsequent daughter veins of different orders. The number 1 is the trunk vein, and 2, 3, 4, 5, 6, 7

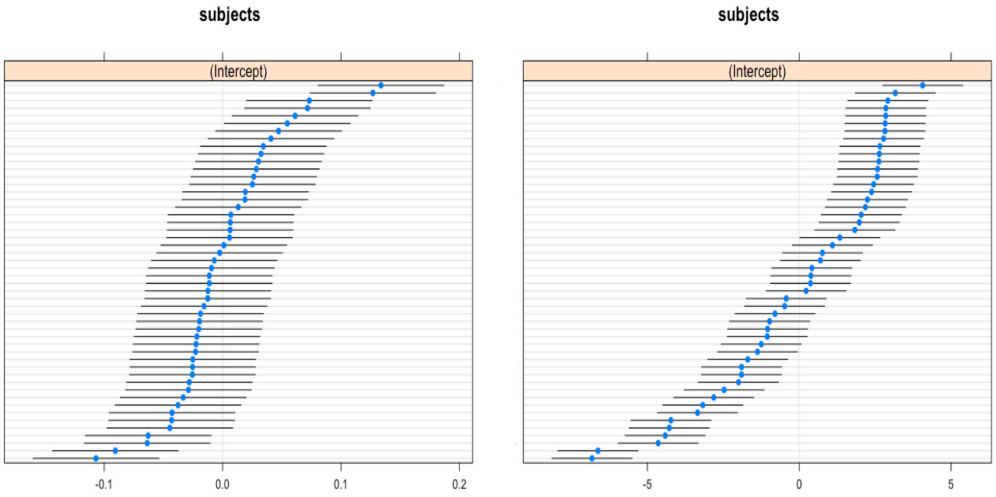

Figure 5.3: AVR (Left plot) and CRAE (Right plot) according to the proposed method. Both plots show how different the random intercepts for each level of the random factor ”subjects” are, including also the prediction intervals. The *x* axis refer to the intercept value whereas y axis to the different levels of the factor.

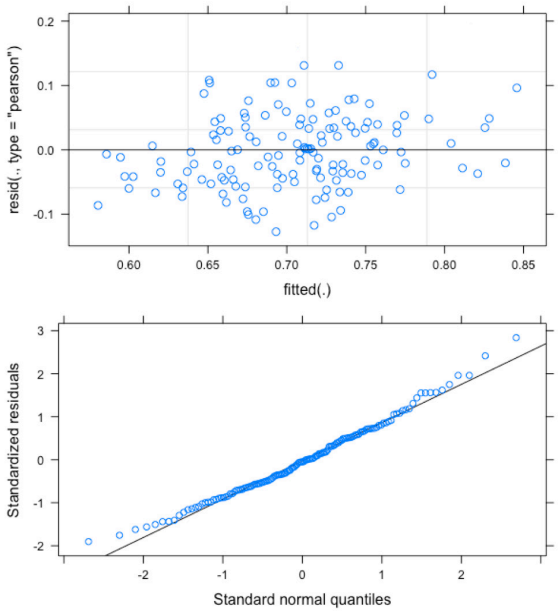

Figure 5.4: Both of the above plots are used to check the normality assumption for the errors. Up: Pearson residuals versus the fitted values for the linear mixed model of the AVR (the estimated according to the proposed method). Down: a Q-Q plot of the standardised residuals versus the standard normal quantiles. Both plots suggest normally distributed residuals.

are the six veins that are used to estimate the trunk. The same four images were used to compare the ascending order iteration with the largest-smallest.

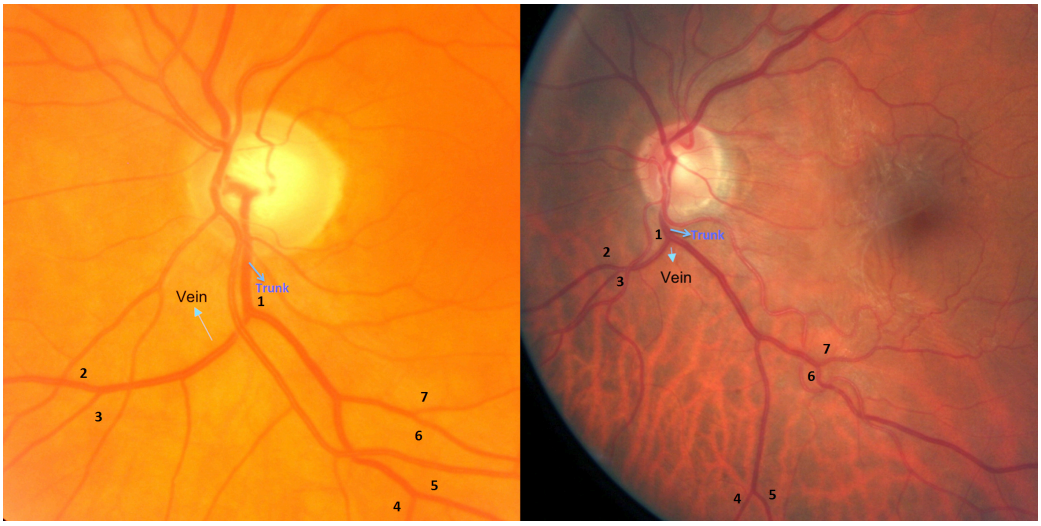

Figure 5.5: Two example images that were used for quantifying the vein trunk and also compare the two methods against each other, using both the established methodology and an ascending iteration process. On the left the branching order is equal, whereas on the right it is uneven. Such case is common when estimating the CRVE/CRAE, where some daughter segments might belong to a lower order.

5.2.9 Evaluation of the Estimated Features

All of the candidate biomarkers are going to be evaluated in chapter 6. However the purpose of this chapter is primarily to assess the proposed method of quantifying the calibres of the vessel trunks and subsequently apply them in a short diabetic/DR progression study, in order to figure out whether the results of the analysis change according to the utilised method. Therefore, it has to become clear and avoid any confusions that in this chapter, some of the results that are needed in order to elaborate and validate the proposed method will be presented, with the full details of them (CRVE, CRAE and AVR) to be presented in the next chapter with all the other studies.

Supplementary to this, since the six veins and six arteries are measured to estimate the CRVE, the CRAE and the AVR, the same ones are used for evaluating them over the four consecutive years of progression (last three years of diabetic eye and the first

year of DR), but they will only be included in this chapter. In both these studies, 200 fundus images from 50 diabetic patients who progressed to DR were used, extracting and measuring 2,400 vessels in total (1,200 veins and 1,200 arteries).

5.3 Results of the Evaluation and the Comparisons

Firstly, the intraclass correlation coefficient was 0.12 (95%CI=0.11-0.15) for the arteries and 0.14 for the veins (95%CI=0.12-0.17). In this way, we are confident that each fundus image has no more variation than the variation between the images. Additionally, the proposed method and the method in Knudtson *et al.* (2003), are positively correlated, with the correlation coefficients ranging from 0.975 to 0.989.

5.3.1 Analysis and Evaluation of the Methods

In Table 5.2 the estimates of the mean absolute percentage errors for all the methods are included. The proposed method indeed improves the quantification of the vessel calibres for every group under consideration. In addition, in Table 5.3 the results of the quantification of four different vein trunks from four retinas of different subjects can be found. In absence of data of the actual measurements of the central retinal arteries and veins, this is the closest alternative one can do to simulate a real case scenario, and present the performance and comparison of both methods as well. Therefore, as can be seen in Table 5.3, the column ”Measured values” refers to the reference ground truth, against which the methods are compared. This is an intuitive way of simulating and hence evaluating the performance of the methods, and it is as close to the real scenario of estimating the CRV as possible. That means that the trunk vessels, in the simulations of this study, ”represent” (play the role of) the CRVE (or CRV/CRA, when referring to the actual central vessels).

The first two cases refer to the trunk veins of two healthy eyes, whereas the third

Table 5.2: Comparison between the methods for the vessel calibre quantification (MAPE)

Group	Proposed method (Mean/Median)(CI95%)	Knudtson et al. (CI95%)	Pratton et al. (CI95%)
Healthy arteries	7.83%(0.0705,0.086)/ 7.81%(0.0707,0.0858)	8.05% (0.0727,0.883)	8.01% (0.0730,0.0881)
Diabetic arteries	9.21%(0.0819,0.1023)/ 9.21%(0.0825,0.1012)	9.45% (0.084,0.1087)	9.41% (0.087,0.1101)
DR arteries	7.1%(0.0575,0.0845) 7.09%(0.0583,0.0853)	7.58% (0.0623,0.0887)	7.56% (0.0634,0.0895)
Healthy veins	6.8%(0.0611,0.0774)/ 6.83%(0.0615,0.0777)	7.22% (0.0642,0.0803)	-
Diabetic veins	6.79%(0.0609,0.0750)/ 6.81%(0.0611,0.0752)	7.11% (0.0637,0.0780)	-
DR veins	6.74%(0.0577,0.0741)/ 6.75%(0.0579,0.0744)	7.15% (0.06,0.0794)	-

and fourth refer to the trunk veins of two diabetic eyes. The results of the estimation of the trunk veins for these cases can be seen in Table 5.3.

1. 10.30, 6.7, 5.9, 7.9, 9.4, 8.9
2. 10.60, 6.9, 6.4, 7.6, 8.9, 8.2
3. 13.1, 9.1, 8.5, 10.2, 9.7, 5.8
4. 6.3, 10.8, 10.7, 13.1, 9, 5

As it was anticipated, the results are quite close, however a notable difference does exist indeed. The more accurate the estimation of the calibre of a vessel trunk is, the more constructive the statistical analysis will be as well (reference to chapter 6 as well). And despite the close mean absolute percentage errors between the methods, when used in real case examples (Table 5.3), the differences are more extensive and

Table 5.3: Quantification of the Calibres of the Vein Trunks

Veins	Proposed method	Knudtson *et al.*	Ascending grouping (Knudtson/Proposed)	Measured value
case1	17.88	17.8	17.91/17.74	19.54
case2	17.66	17.55	17.69/17.49	19.2
case3	20.03	20.60	20.79/19.91	19.65
case4	20.02	20.46	20.64/19.91	18.33

can make an actual difference, especially when studying the progression of a disease (clearer in chapter 6).

5.3.2 Classification

In order to illustrate the proposed improvement in an additional manner, the elastic-net logistic regression classifier was utilised. Although there is no actual feature selection process intended, the elastic-net was preferred over the ordinary logistic regression, in order to find out whether all three features are actually needed. The purpose behind this classification process is not to build an optimum classification system at this stage (objective of chapter 6), but rather show an example of how this newly proposed method of quantifying the retinal trunk vessels could improve the performance of such a system, in conjunction with other features. To do so, an elastic-net regularised logistic regression classifier was trained, according to the method proposed in (Friedman *et al.*, 2010). As described in details in chapter 4, this algorithm linearly combines the L1 (least absolute shrinkage and selection operator) and L2 (ridge regression) penalties, using a cyclic coordinate descent, computed along a regularisation path, in order to improve prediction accuracy. After running a 1000 bootstrap , for both parameters α and λ, the optimum result with the highest AUC and minimum mean squared error was achieved with α=0.07 and λ=0.108 for both cases (proposed and method in (Knudtson *et al.*, 2003)), penalising but keeping at the same time all three features (reminder: α = balance between L1 and L2, and λ = magnitude of regularisation/penalty).

AVR, CRVE and CRAE were the three features used inside the binary classification system. Given the amount of data and the use of only three features, this system was trained and applied between two groups: the group of subjects with diabetes and the DR group. The ROC curves of the two methods under comparison, the proposed one (Leontidis *et al.*, 2016b) and the one in (Knudtson *et al.*, 2003), can be seen in figure 5.6, which reflect the performance of the logistic regression classifier, after being trained using the three features under investigation. Both of them show the average classification performance and the values of the AUC, after following a 10-fold cross validation process for both methods.

The average AUC (maximum possible value equals to 1) is 0.594 for the proposed method and 0.581 for the method of Knudtson *et al.* The improvement in this example case by 0.013 (or 1.3%) in terms of the AUC is particularly important. First, because it clearly shows that the proposed method can provide better estimates of these three features, leading to biomarkers with greater discrimination potential in real world applications than the similar ones in (Knudtson *et al.*, 2003) (Table 6.23 clearly depicts that).

As initially indicated by the decrease of the MAPE using the proposed method, the outcomes of the classifications also suggest that getting more accurate estimates of the CRVE, CRAE and AVR can in turn lead to a better performance of a multi-feature classification system.

5.3.3 Inferences about the AVR

Firstly, no violation of normality and sphericity was observed in either of the investigated cases (p-values ranging from 0.27 to 0.46 and from 0.18 to 0.26 respectively). In Tables 5.4 and 5.5 the results of the analysis for the AVR data can be seen, as well as for the CRVE and CRAE separately, using both the proposed method and the method in (Knudtson *et al.*, 2003), since the method in (Patton *et al.*, 2006a) refers only to the

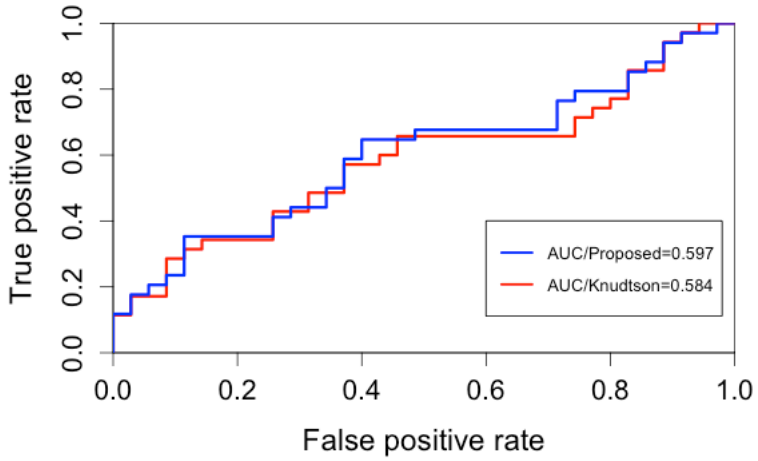

Figure 5.6: ROC curves showing the performance of the two methods. The blue line shows the proposed method (Leontidis *et al.*, 2016b), whereas the red line represents the method in (Knudtson *et al.*, 2003).

arteries.

In this experiment, the semi-automated tool described in (Al-Diri *et al.*, 2010) was used to manually label the arteries and veins. The methodology in (Knudtson *et al.*, 2003) was followed, using the six largest veins and six largest arteries to estimate the CRVE and CRAE.

In figure 5.7 the within subjects confidence intervals for both methods can also be seen. Adjusting the *p*-values according to the Bonferroni correction, pairwise comparisons using the method in (Knudtson *et al.*, 2003) gave significant results for both the CRVE and CRAE for the combinations three years pre-DR/onset of DR, two years pre-DR/onset of DR and one year pre-DR/onset of DR, but not for any other combination (*p*-values ranged from 0.0001 to 0.007 for CRVE and from 0.0006 to 0.008 for CRAE). No significant result was observed for the AVR whatsoever. Similarly, using the proposed method, significant results were found for the same combinations as above, apart from the group two years pre-DR/onset of DR of the CRAE (*p*-values ranged from 0.0009 to 0.005 and from 0.003 to 0.008 for CRVE and CRAE respectively). Likewise, no significant results were observed for any combination of the

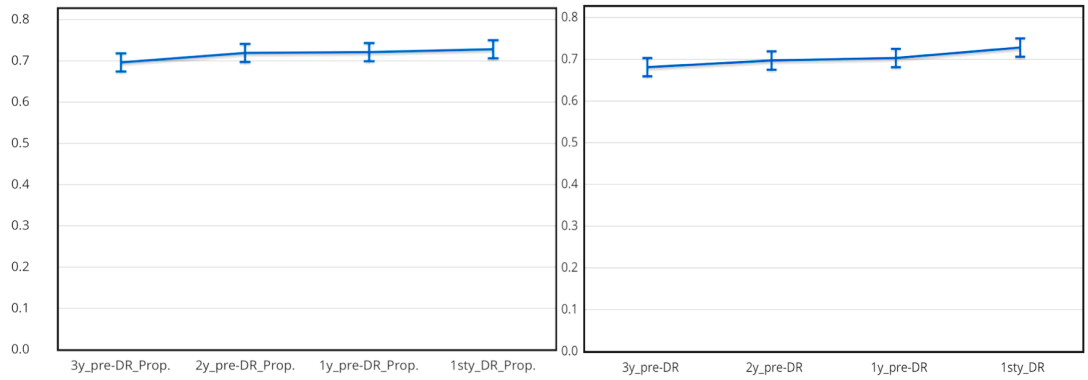

Figure 5.7: On the left: Error bars for the within subjects confidence intervals of the AVR, using the proposed method; on the right a similar plot using the method in (Knudtson *et al.*, 2003).

AVR.

Table 5.4: Results of the Progression study

Scenario	*p*-value (a=0.05)	F-value (dfn,dfe)[a]	AIC (p-value) (full/restricted)[b]	Group means (SD) (3y, 2y, 1y, DR)
CRVE(Knud.)	0.002	5.49 (3,147)	744.3/753.4 (<0.00)	29.56(4.93),27.50(4.45), 28.19(5.03),27.14(4.89)
CRAE(Knud.)	0.009	4.11 (3,147)	589.7/595.8 (<0.00)	20.1(2.99),19.33(2.86), 19.56(3.07),19.87(2.82)
AVR(Knud.)	0.10	2.04 (3,147)	302.1/301.01 (0.19)	0.688(0.082),0.704(0.082), 0.702(0.094),0.723(0.093)
CRVE(Leon.)	0.003	5.16 (3,147)	736.3/745.4 (<0.00)	28,88(4.77),26.88(4.40), 27.28(5.03),27.04(4.73)
CRAE(Leon.)	0.02	3.02 (3,147)	608.3/611.8 (0.01)	20.12(3.04),19.37(2.93), 19.52(3.03),19.28(2.75)
AVR(Leon.)	0.31	1.33 (3,147)	294.3/289.8 (0.45)	0.691(0.081),0.712(0.084), 0.718(0.091),0.723(0.092)
Veins	<0.00	9.1 (3,897)	250.3/261.4 (<0.00)	13.32(4.36),12.38(4.03) 12.72(4.2),12.51(4.21)
Arteries	<0.00	6.01 (3,897)	311.1/320.5 (<0.00)	11.25(2.99),10.89(2.83) 10.95(2.9),11.01(2.96)

[a]: F-Value for the ANOVA, [b]: AIC for the Linear Mixed Model

5.3.4 Analysis of the Calibres of the Vessels

Similarly, no violation of either normality or sphericity was observed for the individual analysis of the calibres of the vessels (p-values = 0.28 and 0.35 for veins and arteries for normality and p-values = 0.3 and 0.37 for veins and arteries for sphericity, respec-

Table 5.5: Random effects summary of each model

Scenario	Subjects Variance (CI95%)	Residuals Variance (CI95%)	Subjects Std. Dev. (CI95%)	Residuals Std. Dev. (CI95%)	ICC
CRVE(Knud.)	18.65 (10.93, 30.84)	5.84 (4.12, 7.72)	4.31 (3.32, 5.54)	2.41 (2.32, 2.93)	0.737
CRAE(Knud.)	7.12 (4.33, 11.71)	1.87 (1.46, 2.49)	2.66 (2.15, 3.44)	1.44 (1.13, 1.57)	0.776
AVR(Knud.)	0.0034 (0.0018,0.0057)	(0.0039) (0.0027,0.0059)	0.062 (0.043,0.081)	0.073 (0.057,0.075)	0.442
CRVE(Leon.)	18.09 (10.89, 30.09)	5.48 (4.11, 7.06)	4.25 (3.30, 5.48)	2.34 (2.02, 2.65)	0.745
CRAE(Leon.)	7.03 (4.23, 11.72)	2.22 (1.66, 2.87)	2.65 (2.05, 3.42)	1.49 (1.29, 1.69)	0.748
AVR(Leon.)	0.0038 (0.0019,0.0070)	(0.0048) (0.0036,0.0063)	0.062 (0.044,0.083)	0.069 (0.06,0.079)	0.445
Veins	0.41 (0.24,0.62) nested:0.23 (0.16,0.34)	0.19 (0.17,0.22)	0.63 (0.49,0.78) nested:0.48 (0.40,0.58)	0.44 (0.42,0.47)	0.624
Arteries	0.22 (0.11,0.35) nested:0.043 (0.016,0.103)	0.047 (0.038,0.054)	0.46 (0.34,0.59) nested:0.19 (0.11,0.32)	0.22 (0.19,0.24)	0.813

tively). Tables 5.4 and 5.5 present the analysis of the progression study in details. As can be seen, the differences are significant for both veins and arteries. Studying the progression in a repeated measures way, at the exact same segments over a period of few years, yields more accurate results, and as has been suggested, increases the power of the study at the same time (Guo *et al.*, 2013).

Finally, significant results were also observed for all of the combinations of arterial and venular calibres, apart from the group three year pre-DR/two years pre-DR, when analysed in a two-group design (p-values ranged from 0.0003-0.002 and 0.0009-0.006 for the veins and the arteries accordingly). These findings are also in line with the respective results presented in Table 6.1 of chapter 6 that follows.

5.4 Discussion

As has been shown in previous studies (Grunwald *et al.*, 1996; Sasongko *et al.*, 2010; Leontidis *et al.*, 2015a) and in the other chapters of this thesis, both geometric and haemodynamic features can be studied for different diseases of the human retina. However, due to factors such as the different modalities used, the low image quality of real-life data and the difficulty of implementing tools that will account for all these properties consistently and accurately, the study of diseases is a difficult process, leading to ambiguous and contradicting results. For this reason, this study focuses on trying to bridge this gap and introduce some new and more specific formulas for the vessel quantification for healthy, diabetic and DR subjects.

A crucial point when calculating the CRVE and CRAE around the optic disc, is that the exact branching order remains uncertain, and it is also unclear whether the vessels have branched before they come out of the optic disc or not. As can be seen in figure 5.5, this exact scenario was simulated, by quantifying the vessel trunk in uneven branching order daughter vessels. Even in this case, the quantification proved to be more accurate (cases 1, 2 and 4 in Table 5.3), which covers this important possible situation. The more accurate the quantification is, the more robust and meaningful the statistical inferences will be as well, because the opposite scenario implies approximate estimates that it is uncertain whether they actually reflect the calibre of the vessel trunk that one wants to estimate. In this study, an alternative, more generalised and improved method was presented in comparison to the state of the art in the literature (Knudtson *et al.*, 2003), for quantifying the retinal vessel calibres, adding up to the previous works found in (Parr & Spears, 1974b; Hubbard *et al.*, 1999; Knudtson *et al.*, 2003; Patton *et al.*, 2006a).

5.5 Conclusion

A new approach for summarising the calibre of the retinal vessel trunk has been proposed. The improvement in terms of the mean absolute percentage error, ranged from 0.24% to 0.49%, whereas the improvement in the classification capability of the proposed method against the method in (Knudtson *et al.*, 2003) between the diabetic and DR group was 0.013 of the AUC. Using the newly derived formulas one can estimate and summarise more accurately the retinal vessel trunks and use them in studies of any retinal disease. In addition, if we take into account that during the iteration process, six vessel calibres are summarised before the final vessel trunk value is estimated, then the actual improvement is more extensive than what the mean absolute percentage error suggests, since the errors are accumulated during the process (Table 5.3). In addition, many studies rely only on estimating the CRVE, CRAE and AVR as biomarkers that can be standardised across different cohorts, the improvement in their estimation becomes even more important.

The progression study that was presented shows a difference between the two compared methods, both in the results of the estimation of the CRVE, CRAE and AVR and in the statistical analyses. Therefore it is apparent from the results of the statistical analysis and the values of the AUC that the improved estimations lead to even more accurate analysis of the data.

Chapter 6

Evaluation of the Candidate Biomarkers

6.1 Introduction

This chapter is focusing on presenting the results of the evaluation of all the geometric and haemodynamic features that were described in the previous chapters. It also examines the main hypothesis of this thesis that the retinal vascular geometry and haemodynamics can give us important biomarkers, that not only significantly differ among the different periods of the progression of the disease, but can also have a certain predictive and discriminative power. The results that follow shed some light on some not previously investigated situations, like the progression of DR within a progressors' group across a period of time, and the simultaneous examination of a vast collection of geometric and haemodynamic features that provide an important base for future studies.

The Tables that will be presented in this chapter, include all the necessary information, in order this study to become a reference point for future studies and clearly convey the results. For that purpose, the results of every significant feature will be

documented, allowing one to make additional inferences, even beyond the scope of this thesis. Each Table contains different features, starting with the general ones (vessel widths and angles, fractal dimension, etc.), followed by the individual areas, the tortuosity metrics and the haemodynamic features. Before the presentation of the results, a short description of the data that were used for each study will be presented. In addition, at the end of each category, another Table is also reported, which is the post-hoc analysis, wherever applicable. All of the non-significant results are documented in appendix A.

The **first section** of this chapter includes the introduction of the chapter, followed by **sections 2 and 3** that present the results of the statistical analysis and the classification, respectively. Finally, **section 4** will include the conclusion and discussion, providing a summary of the results and highlighting the importance of the findings.

6.1.1 List of Candidate Biomarkers

This chapter contains several plots and Tables, in which all the investigated features appear. The main goal is every feature to be self-explainable, including information that will help others make comparisons and/or additional inferences. In order to make the visualisation possible, the features will appear in a short form, details of which can be found in Table 3.3.

6.2 Statistical Inferences

Following the methodology and approaches that were described in the previous chapters, the results of each of the six categories that will follow, include the following distinct analyses:

1. Progressors' group - four year period of progression with matched segments (four groups).

2. Progressors' group - four year period of progression with non-matched segments (four groups).

3. Non-progressors' group - four year period of progression with matched segments without DR (four groups).

4. Non-progressors' group - four year period of progression with non-matched segments without DR (four groups).

5. Averaged non-progressors' group (one group) versus averaged progressors' group (one group) versus DR-only group (one group). The averaging is made on the matched segments.

6. Progressors' group - four year period of progression with independent groups (different individuals in each group)

6.2.1 Data Included in Each Category

The data that were extracted for conducting the above six studies can be found as follows:

Category 1: Out of the 127 subjects and 508 temporal images, 852 venular junctions were successfully matched for the main features across the four year period , giving us 2556 venular width measurements. For the arteries, the total amount was 388 junctions, hence successfully measuring 1164 arterial widths. Tortuosity, fractal dimension and lacunarity were estimated for the whole set of images. Regarding the individual areas, for each of the three ellipses, 240 junctions (same number for the local tortuosity measurements as well) for each vessel type were measured in total (720 venular and 720 arterial widths). CRVE, CRAE and AVR features were extracted out of 200 nasal images. Finally, the haemodynamic features were estimated out of each of the 240 vascular trees

(120 arterial and 120 venular), based on the boundary conditions that were defined in chapter 3. For categories 1, 3, 5 and 6, the two different conditions of pressure are applied to the whole tree (therefore, only one pressure per tree and per boundary condition exists) - also the inlet blood flow rate through the parent vessel is always the same, hence changing only after the bifurcation point (no statistical evaluation for the blood flow rate of the parent vessels then).

Category 2: 904 arterial and 2534 venular junctions (same number for vessel angles) were included in this non-matched progressors study (2712 and 7602 arterial and venular vessel widths, respectively).

Category 3: The non-progressors' group contains 27 subjects across a four year period, including 108 temporal images. 136 arterial junctions were successfully matched, measuring 408 widths in total. Accordingly, 220 venular junctions were also matched, giving a total number of 660 width measurements. Tortuosity, fractal dimension and lacunarity were estimated for the whole set of images. CRVE, CRAE and AVR were extracted out of 80 nasal images. Regarding the individual areas, for each of the three ellipses, 108 junctions (same number for the local tortuosity measurements as well) for each vessel type were measured in total (324 venular and 324 arterial widths). Finally, the haemodynamic features were estimated out of 160 vascular trees (80 arterial and 80 venular).

Category 4: 379 arterial and 594 venular junctions were included in the non-matched non-progressors' study (1137 arterial and 1782 venular widths, respectively).

Category 5: For the arteries, 228 junctions were included, and similarly 471 venular junctions. Tortuosity, fractal dimension and lacunarity were estimated on a total 282 temporal images. CRVE, CRAE and AVR features were extracted out of 100 nasal images. As far as the areas are concerned, each of the three

ellipses included 147 junctions (same number of segments for the local tortuosity). 70 venular and 70 arterial trees were used for all the haemodynamic features' estimation.

Category 6: This final category included only the main features. The 127 subjects were randomly assigned to one of the four groups, discarding at the same time three of the subjects, in order to have equal number of subjects in each group. Therefore, each group contained 31 subjects, which contributed with their non-matched measurements for the analysis. The total number of junctions were 219 for arteries (657 vessel widths) and 456 for veins (1368 vessel widths). Accordingly, for the FD and lacunarity a total number of 124 images were used (31 for each of the four groups).

All of the investigated features and categories conform with the necessary assumptions of the LMMs. An example of the residual plots of the arterial widths and random intercepts of category 1 can be seen in figure 6.1. The blue-coloured features (main analyses; post-hoc analyses appear in black) reported in this chapter are considered to be significantly changing throughout the investigated periods, with the non-significant ones appearing in appendix A. This is decided by the *p*-value, with values lower than or equal to 0.05 to be considered significant. Regarding the other reported metrics (their importance already described in details in chapter 4), AIC and BIC provide an additional way of quantifying how better one model is compared to the other, with a higher difference denoting a much better model. ICC and Ω^2 give us an overview of how extensive is the random effect in our model (restricted one) and also the effect size of it, respectively. As far as F^c is concerned, it refers to the ratio of the variances (random effects and nested factor of the full model, wherever applicable) to the variance of the residuals. In addition to all these, the values of the variances of each model are reported, alongside the group mean values and standard deviations.

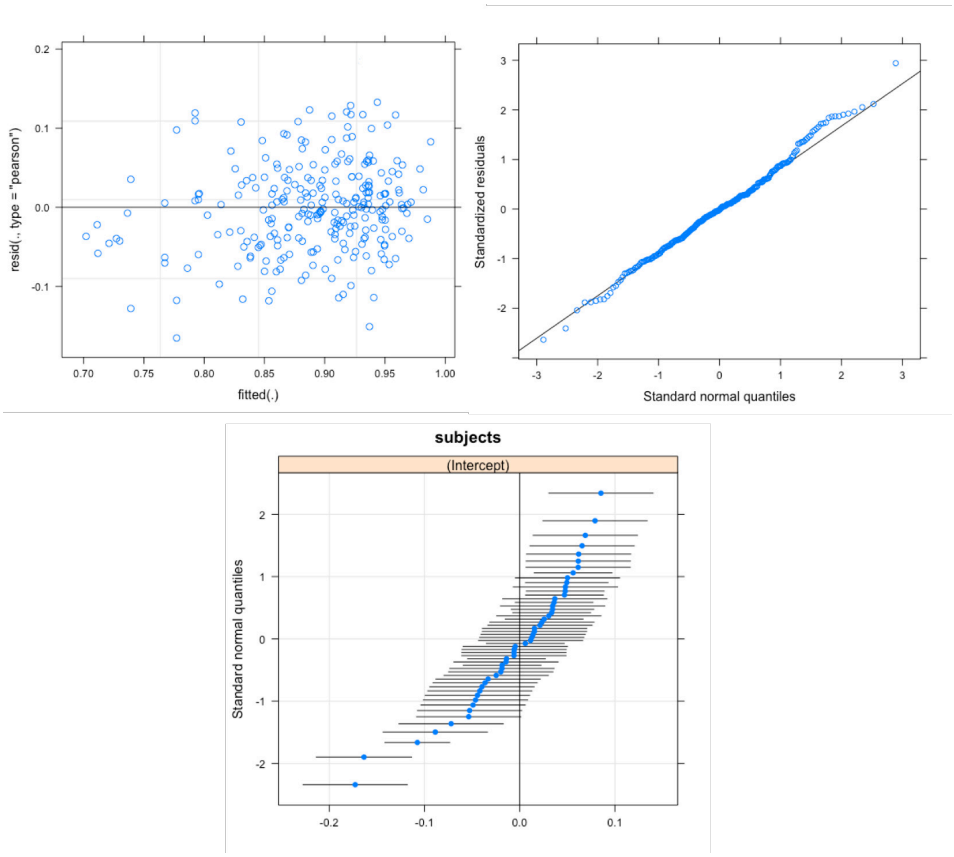

Figure 6.1: Up: Both plots are used to check the normality assumption for the errors. Down: The plot shows how different the random intercepts for each level of the random factor ”subjects” are, including also the prediction intervals for each level.

6.2.2 Four Year Matched Study of the Progressors

This group represents the main study of this thesis, around which all the other categories and analyses will be compared against. The main hypothesis suggests that the alterations in the retinal vascular geometry and haemodynamics during the last stages of the diabetic eye are distinct; but given the diversity of the retina, a proper design is needed, accompanied by a suitable analysis, in order to make robust and solid inferences. In this category, the same subjects are used in all four groups, and the exact same segments are included in all of the four year period. Therefore the whole study is absolutely balanced. The purpose is to show how different the results can be depending on the design of an experiment that aims to study the effects of the disease in the

retinal vasculature - and also make the statistical inferences about the progression to DR.

Table 6.1 shows the results of the main features, followed by Tables 6.2, 6.3 and 6.4, which include the results of the different areas, the tortuosity metrics and the haemodynamic features, respectively. Finally Table 6.5 includes the post-hoc results.

Table 6.1: Analysis of the main features for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
AlphaA	**306.5/**	**283.9/**	**0.039/**	**0.46**	**0.605**	**0.378**	**0.0089/**	**0.77(0.16),0.82(0.15)**
	308.8	**292.4**	**0.042**			**nested:0.126**	**0.0116**	**0.81(0.12),0.79(0.16)**
LambdaA	**600.1/**	**578.6/**	**0.023/**	**0.484**	**0.622**	**0.416**	**0.0032/**	**0.87(0.1),0.90(0.09)**
	603.5	**585.9**	**0.025**			**nested:0.108**	**0.0036**	**0.901(0.07),0.885(0.095)**
Lambda2A	**595.9/**	**575.1/**	**0.017/**	**0.554**	**0.726**	**0.389**	**0.0024/**	**0.807(0.062),0.837(0.061)**
	600.1	**581.7**	**0.019**			**nested:0.259**	**0.0033**	**0.826(0.068),0.827(0.072)**
Width_parentA	**276.3/**	**294.9/**	**0.015/**	**0.786**	**0.882**	**0.592**	**0.294/**	**3.93(0.71),3.81(0.69)**
	280.7	**301.2**	**0.017**			**nested:0.254**	**0.076**	**3.86(0.73),3.79(0.65)**
Width_child1A	**208.3/**	**226.2/**	**0.05/**	**0.81**	**0.863**	**0.762**	**0.287/**	**3.65(0.61),3.57(0.62)**
	210.9	**228.2**	**0.05**			**nested:0.064**	**0.065**	**3.56(0.6),3.57(0.6)**
Width_allA	**144.5/**	**157.9/**	**0.043/**	**0.645**	**0.875**	**0.722**	**0.21/**	**3.585(0.528),3.524(0.559)**
	146.2	**159.8**	**0.038**			**nested:0.121**	**0.047**	**3.544(0.565), 3.493(0.51)**
AlphaV	**427.4/**	**404.7/**	**0.001/**	**0.33**	**0.577**	**0.177**	**0.007/**	**0.68(0.19),0.71(0.19)**
	436.6	**409.2**	**0.002**			**nested:0.298**	**0.021**	**0.737(0.21),0.725(0.19)**
BetaV	**203.5/**	**174.8/**	**0.026/**	**0.344**	**0.62**	**0.168**	**0.01/**	**1.11(0.18),1.15(0.19)**
	206.7	**185.5**	**0.027**			**nested:0.356**	**0.028**	**1.14(0.19),1.16(0.21)**
LambdaV	**1123/**	**1100/**	**0.002/**	**0.338**	**0.593**	**0.181**	**0.002/**	**0.815(0.12),0.832(0.12)**
	1133	**1105**	**0.002**			**nested:0.313**	**0.007**	**0.85(0.13),0.843(0.12)**
Lambda2V	**1371/**	**1349/**	**<0.00/**	**0.362**	**0.636**	**0.178**	**0.002/**	**0.733(0.087),0.754(0.082)**
	1381	**1352**	**<0.00**			**nested:0.363**	**0.005**	**0.759(0.081),0.761(0.091)**
Width_parentV	**1282/**	**1312/**	**0.045/**	**0.624**	**0.812**	**0.483**	**0.41/**	**4.38(0.9),4.29(0.91)**
	1285	**1316**	**0.047**			**nested:0.28**	**0.199**	**4.33(0.92),4.21(0.89)**
Width_child1V	**1078/**	**1117/**	**0.046/**	**0.62**	**0.805**	**0.502**	**0.311/**	**4(0.83),3.935(0.75)**
	1080	**1119**	**0.049**			**nested:0.25**	**0.153**	**3.931(0.792),3.84(0.746)**
Width_allV	**840.5/**	**856.8/**	**0.047/**	**0.333**	**0.823**	**0.518**	**0.244/**	**3.861(0.694),3.815(0.667)**
	843.6	**858.9**	**0.048**			**nested:0.258**	**0.105**	**3.848(0.698),3.792(0.653)**
Angle.BC_V	**6494/**	**6510/**	**0.05/**	**0.501**	**0.611**	**0.26**	**229.3/**	**124.53(38.32),121.61(38.37)**
	6498	**6515**	**0.048**			**nested:0.263**	**420.4**	**122.96(42.74),119.91(38.15)**
Fractal	**531.1/**	**515.5/**	**0.047/**	**0.796**	**0.844**	**0.802**	**0.0031/**	**1.662(0.06),1.648(0.05)**
	533.6	**521.3**	**0.048**				**0.0007**	**1.647(0.05),1.632(0.06)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.2: Analysis of the features of the different areas for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
ellipse1parV	**234.8/**	**247.4/**	**0.007/**	**0.56**	**0.705**	**0.601**	**0.924/**	**6.1(1.2),5.32(1.03)**
	240.3	**249.1**	**0.011**				**0.613**	**5.98(1.16),5.93(1.5)**
ellipse2ch1V	**250.2/**	**263.1/**	**0.008/**	**0.36**	**0.558**	**0.407**	**0.599/**	**6.14(1.22),5.4(0.99)**
	256	**264.5**	**0.01**				**0.87**	**6.08(1.4),5.37(1.19)**
ellipse2ch2A	**279.1/**	**291.3/**	**0.016/**	**0.431**	**0.602**	**0.469**	**1.065/**	**5.87(1.88),4.91(1.22)**
	284.8	**294.1**	**0.013**				**1.201**	**5.87(1.5),5.37(1.32)**
CRAE_LEON	**608.3/**	**626.5/**	**0.01/**	**0.75**	**0.818**	**0.759**	**7.038/**	**20.12(3.04),19.37(2.93)**
	611.8	**629.3**	**0.02**				**2.228**	**19.52(3.03),19.28(2.75)**
CRVE_LEON	**736.3/**	**753.5/**	**0.001/**	**0.745**	**0.826**	**0.767**	**18.09/**	**28.88(4.77),26.88(4.4)**
	745.4	**756.8**	**<0.000**				**5.48**	**27.28(5.03),27.04(4.73)**

CRAE_KNUD	**589.7/**	**604.5/**	**0.007/**	**0.776**	**0.843**	**0.791**	**7.059/**	**20.10(2.99),19.33(2.86)**
	595.8	**607.4**	**0.009**				**1.857**	**19.56(3.07),19.87(2.82)**
CRVE_KNUD	**744.3/**	**761.1/**	**0.001/**	**0.737**	**0.821**	**0.761**	**18.63/**	**29.56(4.93),27.50(4.45)**
	753.4	**763.1**	**0.002**				**5.84**	**28.19(5.03),27.14(4.89)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.3: Analysis of tortuosity features for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
sd_tau	**2452/**	**2422/**	**0.05/**	**0.295**	**0.406**	**0.296**	**1.7x10^{-4}/**	**(7.99(2.4),8.3(2.3)**
	2455	**2425**	**0.051**				**4.1x10^{-4}**	**8.1(2.5),8.4(2.3))x10^{-2}**
Tort_ellipse3	**271.8/**	**345.2/**	**0.026/**	**0.242**	**0.419**	**0.269**	**0.001/**	**0.037(0.069),0.039(0.086)**
	275.1	**348.3**	**0.024**				**0.0027**	**0.037(0.046),0.047(0.026)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.4: Analysis of the haemodynamic features for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
qchild1A	**1349/**	**1339/**	**0.008/**	**0.398**	**0.586**	**0.438**	**8.4x10^{-9}/**	**(8.4(1.1),7.8(1.3)**
	1354	**1342**	**0.01**				**1.1x10^{-8}**	**7.4(1.7),8(1.2))x(10^{-4})**
qchild2A	**1348/**	**1337/**	**0.011/**	**0.399**	**0.587**	**0.439**	**8.5x10^{-9}/**	**(5.3(1.1),6(1.3)**
	1355	**1340**	**0.008**				**1.1x10^{-8}**	**6.4(1.7),5.8(1.3))x(10^{-4})**
Rechild2A	**800/**	**812/**	**0.015/**	**0.622**	**0.752**	**0.653**	**421.9/**	**119.5(26.3),131.1(26.6)**
	804	**815**	**0.017**				**223.7**	**132(26.57),125.4(23.6)**
wssparentV	**550.5/**	**561.3/**	**0.021/**	**0.942**	**0.936**	**0.951**	**355.48/**	**24.01(15.16),25.62(11.59)**
	554.2	**564.7**	**0.024**				**18.22**	**22.46(11.87),25.43(17.23)**
vparentaV	**335.4/**	**346.3/**	**0.021/**	**0.932**	**0.933**	**0.942**	**4.468/**	**9.29(3.75),9.82(3.1)**
	339.2	**349.6**	**0.024**				**1.102**	**8.96(3.12),9.59(4.19)**
ReparentV	**693.4/**	**703.9/**	**0.024/**	**0.922**	**0.931**	**0.933**	**41.85/**	**187.8(37),194.1(32.2)**
	696.8	**707.5**	**0.027**				**11.19**	**185.1(32.4),190.2(40.6)**
wsschild1V	**472.1/**	**486.3/**	**<0.000/**	**0.917**	**0.952**	**0.939**	**110.64/**	**19.52(10.52),21.59(10.14)**
	484.8	**491.9**	**<0.000**				**7.11**	**19.59(9.27),22.7(10.82)**
qchild1V	**1279/**	**1267/**	**0.034/**	**0.955**	**0.973**	**0.959**	**2.5x10^{-8}/**	**(6.21(1.72),6.3(1.63)**
	1281	**1272**	**0.031**				**1.1x10^{-9}**	**6.4(1.7),6.1(1.6))x(10^{-4})**
vchild1V	**231.8/**	**246/**	**<0.000/**	**0.896**	**0.941**	**0.927**	**4.43/ 0.34**	**6.25(2.03),6.85(2)**
	247.2	**254.3**	**<0.000**					**6.39(1.94),7(2.14)**
Rechild1V	**593.9/**	**608/**	**<0.000/**	**0.892**	**0.946**	**0.925**	**449.92/**	**107.14(20.13),114.57(21.7)**
	610.3	**617.4**	**<0.000**				**36.12**	**109.98(22.59),113.8(22.53)**
wsschild2V	**490.6/**	**504.8/**	**<0.000/**	**0.901**	**0.927**	**0.924**	**117.7/**	**18.77(7.87),21.8(9.67)**
	501.2	**508.3**	**<0.000**				**9.68**	**21.04(10.19),22.25(10.48)**
qchild2V	**1269/**	**642.8/**	**0.035/**	**0.955**	**0.973**	**0.959**	**2.56x10^{-8}/**	**(6.23(1.81),6.11(1.62)**
	1271	**646.8**	**0.031**				**1.1x10^{-9}**	**6.14(1.72),6.3(1.6))x(10^{-4})**
vchild2V	**247.5/**	**261.6/**	**0.002/**	**0.915**	**0.973**	**0.932**	**5.71/ 0.41**	**6.35(1.97),6.88(2.06)**
	256	**263.9**	**0.003**					**6.73(2.27),6.73(2.25)**
Rechild2V	**608.2/**	**618.1/**	**0.034/**	**0.944**	**0.96**	**0.95**	**734.35/**	**110.7(27.1),113.3(24.4)**
	610.9	**622.4**	**0.04**				**38.41**	**112.4(27.4),116.5(25.6)**
PinQoutV	**444/**	**455.2/**	**0.017/**	**0.833**	**0.894**	**0.854**	**38.35/**	**31.63(5.62),29.34(8.31)**
	448.2	**458.1**	**0.02**				**6.54**	**32.12(4.85),30.5(7.1)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.5: Post-Hoc analysis for category 1

Significant LMM features	Significant Linear Hypotheses	P-values
AlphaA	y3-y2	0.045
Lambda	y3-y2	0.04
Lambda2A	y3-y2	<0.000
Width_allA	y3-DR	0.032
	y2-DR	0.043
Width_parentA	y3-DR	0.032
	y3-y2	0.045
Width_child1A	y3-DR	0.048
AlphaV	y3-DR	0.019
	y3-y1	0.001
BetaV	y3-DR	0.035
Lambda	y3-DR	0.017
	y3-y1	0.002
Lambda2	y3-DR	0.001
	y3-y1	0.003
	y3-y2	0.03
Width_allV	y3-DR	0.012
	y2-DR	0.048
	y3-y1	0.029
Width_parentV	y3-DR	0.049
Width_child1V	y3-DR	0.048
Angle.BC_V	y3-DR	0.032
Fractal	y3-DR	0.026
ellipse1parV	y2-y1	0.036
	y3-y2	0.008
ellipse2ch1V	y3-DR	0.042
ellipse2ch2A	y3-DR	0.024
qchild1A	y3-y1	0.0058
qchild2A	y3-y1	0.0052
Rechild2A	y3-y2	0.035
	y3-y1	0.024
wssparentV	y3-y2	0.015
vparentaV	y3-y2	0.0147
ReparentV	y3-y2	0.0178
wsschild1V	y3-DR	<0.000
	y3-y2	0.002
	y3-y1	0.019
qchild1V	y2-DR	0.048
	y3-y2	0.049
vchild1V	y3-DR	0.001
	y3-y2	0.001
	y3-y1	0.008
Rechild1V	y3-DR	<0.000
	y3-y2	<0.000
	y3-y1	0.009
wsschild2V	y3-DR	0.003
	y3-y2	<0.000
	y3-y1	0.009
qchild2V	y3-DR	0.049
	y3-y2	0.049
Rechild2V	y3-DR	0.017
vchild2V	y3-DR	0.002
	y3-y2	0.02
	y3-y1	0.005
PinQoutV	y3-DR	0.024
	y2-y1	0.013
sd_tau	y3-DR	0.035
Tort_ellipse3	y2-DR	0.032
CRAE_LEON	y3-DR	0.017
	y1-DR	0.049
CRVE_LEON	y3-DR	0.034
	y2-DR	0.05
	y1-DR	0.012

CRAE_KNUD	y3-DR	0.023
	y2-DR	0.026
	y1-DR	0.039
CRVE_KNUD	y3-DR	0.036
	y2-DR	0.021
	y1-DR	0.029

Additionally, multivariate comparisons combining arteries, veins and angles altogether yielded no significant results. On the contrary, combining all the haemodynamic features for veins and arteries, gave significant results only for the children of veins (p-value= 0.006).

6.2.2.1 Comparison with the Studies of the Initial Phase

All of the above results are coming from the final stage of all the experiments, where all the features were calculated and/or estimated on the extended amount of data. It is worth pointing out that these results actually confirm most of the initial results. In (Leontidis *et al.*, 2015a), the studies followed a similar logic to the above ones, where the analysis was conducted over the same four year period. In this study, the angles and widths of the vessels, fractal dimension, lacunarity and some derivative features were measured concluding to the same results as above.

In some other early experiments, we focused on the paired comparisons between the three year pre-DR and first year of DR periods, as well as the one year pre-DR and first year of DR periods. For the former, the investigated features were the fractal dimension, lacunarity, tortuosity, AVR, CRVE, CRAE and vessels' widths and angles. Significant results were observed for the tortuosity (p-value=0.021), arteries widths (p-value=0.01) and angles (p-value=0.022), fractal dimension (p-value=0.024) and widths of the veins (p-value=<0.000) (Leontidis *et al.*, 2015a,b, 2016a). Regarding the latter combination, the investigations were limited to the widths and angles of the vessels, where the only significant result was observed for the angles of the veins (p-value=0.034). It is important to take into account that planning a pairwise experiment

between two groups is a totally separate hypothesis and shall not be confused with the results of an ANOVA or of a LMM. In addition, as was mentioned earlier, the post-hoc results refer to follow-up examinations after running the main tests, so any comparisons between a planned pairwise experiment and post-hoc outcome shall be carefully made. Besides, the interest here is on the progression as this can be identified in an extended period that includes a period before DR and the onset of DR; so investigating these stages altogether, minimises the chances of committing Type I error, with the opposite being the case, if multiple paired comparisons without any adjustments were conducted instead.

6.2.3 Four year Non-Matched Study of the Progressors

This part includes the same data as category 1 does, but the main features are not matched, so all the measurements are used without any selection. The purpose is to find out whether there will be any differences in the results in comparison to the similar category 1, to justify or not the use of matched segments. Table 6.6 shows the results of the main features, followed by the post-hoc analysis in Table 6.7.

Table 6.6: Analysis of the main features for category 2

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
Width_child2A	**1049/**	**1072/**	**0.02/**	**0.234**	**0.504**	**0.323**	**0.079/**	**3.049(0.475),3.178(0.504)**
	1052	**1082**	**0.021**			**nested:0.097**	**0.142**	**3.137(0.501),3.112(0.481)**
Width_child2V	**2981/**	**3022/**	**<0.000/**	**0.123**	**0.407**	**0.34**	**0.091/**	**3.031(0.494),3.108(0.475)**
	3003	**3027**	**<0.000**			**nested:0.048**	**0.159**	**3.149(0.521),3.108(0.466)**
Angle.BC_V	**24684/**	**24705/**	**0.045/**	**0.3**	**0.154**	**0.074**	**76.42/**	**117.49(37.93),117.38(36.67)**
	24687	**24711**	**0.046**			**nested:0.031**	**927.47**	**115.74(36.81),114.17(35.21)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.7: Post-Hoc analysis for category 2

Significant LMM features	Significant Linear Hypotheses	P-values
Width_child2A	y3-y1	0.042
	y3-y2	0.033
Width_child2V	y3-DR	0.003
	y3-y1	<0.000
	y3-y2	0.004
Angle.BC_V	y3-DR	0.039

6.2.3.1 Short Discussion for Categories 1 and 2

Comparing the Tables 6.1 ans 6.6, is apparent why not matching the segments can lead to controversial results. A series of important biomarkers that define the vascular changes, e.g. arterial and venular widths, branching coefficient etc., were missed out in the non-matched design. The same exact segments compared over a period of time can provide more accurate representation of the underlying changes, regardless of the direction of the effect. A non significant result can be turned into a significant with the presence of a few outliers in the sample and similarly a significant result can be missed inside noisy data. Unless the data are of adequate magnitude to averagely compensate for these noisy data, by more and more converging to a true random error, then studies with small and different cohorts will be prone to leading to contradicting results.

6.2.4 Four Year Period Matched Study of the Non-Progressors

This category investigates a group of non-progressors, i.e. diabetic subjects that at the moment of the data collection had not progressed to DR. They are followed over a four consecutive year period of time, and the aim is to figure out whether in this period, any significant alterations due to diabetes occur. Table 6.8 shows the results of the main features, followed by Table 6.9, which includes the post-hoc results.

Table 6.8: Analysis of the tortuosity features for category 3

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, Y0)
median_psi	**748.19/ 750.84**	**735.7/ 740.6**	**0.034/ 0.039**	**0.44**	**0.592**	**0.466**	**(9.21/ 10.53)x10^{-6}**	**0.0158(0.0034),0.0152(0.0038), 0.0178(0.0053),0.0168(0.0048)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.9: Post-Hoc analysis for category 3

Significant LMM features	Significant Linear Hypotheses	P-values
median_psi	y2-y1	0.0325

6.2.5 Non-Progressors Four Year Period Non-Matched Study

Similar to category 2, this particular study is the non-matched study of category 3, using the same data but without any matching of the bifurcations whatsoever. Accordingly, the purpose is to figure out whether any of the differences of category 3, are also observed here and vice versa. However no significant results were found, suggesting, just like with category 2, that using the same bifurcations and comparing them over a period of time, can identify vascular alterations that could be missed otherwise. A list of the analysis for all those non-significant features can be found in appendix A.

6.2.6 Progressors Versus Non-Progressors Versus DR

This particular analysis is very important, because it combines all the years within the progressors' and non-progressors' groups, creating one category for each of them. Since the matched segments are used in both of them, the final values are the average of the same segments across the whole period of time. For instance, in progressors' group, the three years of measurements are averaged segment-wise. Finally, in the two groups that are created, a third also independent group is added, which represents a DR group. Therefore, the analysis is run including three independent groups in total. The results

of this novel study will show us two things. Firstly, whether either of the progressors' or non-progressors' groups differ compared to the DR group, and secondly whether they also differ between each other. Given that when selecting a cohort with current data is unlikely to have any information about the future progression, it is likely that many of the selected subjects will soon progress to DR, i.e. belonging to a progressors' group. Therefore with this study the aim is to find out if a random non-progressors' diabetic group, also differs from a carefully selected progressors' diabetic group, and therefore intonate the importance of more carefully selecting the cohorts.

Table 6.10 shows the results of the main features, followed by Tables 6.11, 6.12 and 6.13, which include the results of the different areas, the tortuosity measurements and the haemodynamic features, respectively. Finally Table 6.14 includes the post-hoc analysis.

Table 6.10: Analysis of the main features for category 5

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Prog, Non-Prog, DR)
Width_parentA	**1471/ 1479**	**1501/ 1504**	**0.002/ 0.003**	**0.341**	**0.957**	**0.294 nested:<0.525**	**0.225/ 0.138**	**4.413(1.071),4.062(0.671) 3.937(0.856)**
Width_child1A	**1357/ 1362**	**1376/ 1379**	**0.008/ 0.009**	**0.373**	**0.409**	**0.338 nested:<0.000**	**0.221/ 0.433**	**3.932(0.914),3.663(0.648) 3.675(0.795)**
Width_child2A	**908.11/ 915.12**	**928.39/ 930.24**	**0.004/ 0.005**	**0.271**	**0.778**	**0.234 nested:<0.331**	**0.067/ 0.124**	**3.461(0.641),3.167(0.453) 3.181(0.51)**
Width_allA	**1155/ 1162**	**1171/ 1174**	**0.003/ 0.004**	**0.373**	**0.402**	**0.326 nested:<0.031**	**0.154/ 0.3139**	**4.002(0.839),3.564(0.524) 3.597(0.669)**
Width_parentV	**2031/ 2034**	**2047/ 2053**	**0.03/ 0.029**	**0.344**	**0.943**	**0.295 nested:<0.521**	**0.223/ 0.139**	**4.83(1.215),4.423(0.896) 4.476(1.154)**
Width_child1V	**1925/ 1928**	**1941/ 1948**	**0.035/ 0.043**	**0.276**	**0.3**	**0.257 nested:<0.000**	**0.2862/ 0.8262**	**4.458(1.099),4.116(0.893) 4.147(1.061)**
Width_child2V	**1171/ 1175**	**1188/ 1194**	**0.021/ 0.028**	**0.276**	**0.778**	**0.237 nested:<0.333**	**0.063/ 0.12**	**3.482(0.691),3.173(0.431) 3.294(0.585)**
Width_allV	**1618/ 1622**	**1636/ 1641**	**0.021/ 0.027**	**0.346**	**0.345**	**0.315 nested:<0.021**	**0.244/ 0.519**	**4.256(0.925),3.904(0.652) 3.972(0.861)**
Angle.BC_V	**7845/ 7865**	**7912/ 7918**	**0.001/ <0.000**	**0.321**	**0.506**	**0.126 nested:<0.000**	**660.3/ 4575**	**119.56(37.87),121.45(38.56) 117.62(36.08)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.11: Analysis of the features of the different areas for category 5

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Prog, Non-Prog, DR)
CRAE_LEON	**1025/ 1036**	**1042/ 1046**	**<0.000/ <0.000**	**0.506**	**0.593**	**0.412**	**4.793/ 6.828**	**19.57(3.13),16.29(4.08) 19.51(2.75)**
CRVE_LEON	**1163/ 1165**	**1175/ 1177**	**0.05/ 0.053**	**0.515**	**0.607**	**0.483**	**12.07/ 12.87**	**27.71(4.92),24.78(5.14) 27.69(4.74)**

CRAE_KNUD	**955.9/ 964.2**	**972.3/ 974.8**	**0.002/ 0.001**	**0.595**	**0.677**	**0.535**	**5.041/ 4.374**	**19.58(3.03),17.01(3.18) 19.98(2.82)**
CRVE_KNUD	**1245/ 1248**	**1255/ 1257**	**0.016/ 0.017**	**0.413**	**0.511**	**0.365**	**11.95/ 20.76**	**28.38(5.03),24.61(7.09) 27.99(4.78)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.12: Analysis of the tortuosity features for category 5

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Prog, Non-Prog, DR)
sd_phi	**1062/ 1065**	**1048/ 1052**	**0.035/ 0.029**	**0.144**	**0.198**	**0.096**	**(2.87/ 27.26)x10^{-5}**	**0.246(0.017),0.237(0.016) 0.249(0.018)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.13: Analysis of the haemodynamic features for category 5

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Prog, Non-Prog, DR)
qchild1A	**1521/ 1528**	**1513/ 1517**	**0.012/ 0.018**	**0.354**	**0.515**	**0.375**	**(7.11/ 11.86)x10^{-9}**	**(7.19(1.48),7.94(1.87), 8.04(1.25))x10^{-4}**
qchild2A	**1516/ 1523**	**1523/ 1529**	**0.015/ 0.018**	**0.356**	**0.516**	**0.377**	**(7.12/ 11.89)x10^{-9}**	**(5.93(1.49),5.94(1.87), 5.81(1.26))x10^{-4}**
qchild1V	**1407/ 1413**	**1409/ 1412**	**0.03/ 0.034**	**0.952**	**0.888**	**0.975**	**(2.91/ 0.14)x10^{-8}**	**(6.28(1.7),6.49(2.17), 6.15(1.66))x10^{-4}**
vchild1V	**320.98/ 324.72**	**340.48/ 342.22**	**0.018/ 0.017**	**0.848**	**0.906**	**0.853**	**3.663/ 0.63**	**6.49(1.98),6.32(1.45), 6.99(2.14)**
qchild2V	**1410/ 1414**	**1400/ 1403**	**0.012/ 0.023**	**0.952**	**0.975**	**0.954**	**(2.9/ 0.133)x10^{-8}**	**(6.18(1.71),5.97(2.18), 6.31(1.68))x10^{-4}**
vchild2V	**322.71/ 326.1**	**342.21/ 346.6**	**0.024/ 0.035**	**0.876**	**0.923**	**0.881**	**4.27/ 0.57**	**6.66(2.09),6.45(1.99), 7.06(2.25)**

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table 6.14: Post-Hoc analysis for category 5

Significant LMM features	Significant Linear Hypotheses	P-values
Width_parentA	prog-nonprog	0.001
	prog-DR	0.014
Width_child1A	prog-nonprog	0.011
	prog-DR	0.021
Width_child2A	prog-nonprog	0.006
	prog-DR	0.013
Width_allA	prog-nonprog	0.003
	prog-DR	0.011
Width_parentV	prog-DR	0.042
Width_child1V	prog-DR	0.046
Width_child2V	prog-DR	0.043
Width_allV	prog-DR	0.03
Angle.BC_V	prog-nonprog	<0.000
	nonprog-DR	<0.000
CRAE_LEON	prog-nonprog	<0.000
	nonprog-DR	<0.000
CRVE_LEON	prog-nonprog	0.04

CRAE_KNUD	prog-nonprog	<0.002
	nonprog-DR	<0.001
CRVE_KNUD	prog-nonprog	0.01
sd_phi	prog-nonprog	0.021
qchild1A	prog-DR	0.012
	prog-nonprog	0.023
qchild2A	prog-DR	0.015
	prog-nonprog	0.028
qchild1V	prog-DR	0.042
	prog-nonprog	0.039
vchild1V	prog-DR	0.011
qchild2V	prog-DR	0.022
	prog-nonprog	0.019
vchild2V	prog-DR	0.033

6.2.6.1 Short Discussion for Category 5

It comes as no surprise that the vascular geometry both in a junction- and in image-level and the haemodynamic features present differences across this three group category. It also confirms the initial hypothesis, that the functional impairment and the vascular adjustments, do not only occur when approaching DR, but also within the diabetic eye period. It is clear that the changes start much earlier, concluding with some strong and solid indications that DR starts affecting the retinal vessels quite some time before the first lesions appear.

6.2.7 Progressors Four Year Study - Independent Groups

This section is focusing on presenting the results in a similar to category 1 way, with just a difference on the design. The subjects in each group are not the same across all groups, but different ones. Nonetheless, it is known that they are coming from the same cohort as in category 1, so they can be directly compared with each other and make inferences about how the results are affected when the subjects are different and when the sample is not as large as well. Having to include different subjects in each group, the available data in each of them are by three quarters less, but still empirically adequate to define the trends. Due to the limited amount of data, the analysis is only made for the main features, where, due to the grouping factor, multiple mea-

surements are included within each subject. In case that the matching of the segments and the follow-up design of category 1 does not make any difference, then we should be able to observe the same or quite similar results with this type of design. However, as can be seen in Table A.15 in appendix, this is not the case. No wonder, why in the above analysis no statistically significant results were observed. However, this category shows something that should be stressed further. It is easy to conclude and make decisions based on misleading results. The same features, with the same dataset but slightly differently analysed, change the whole interpretation and direction of the results. In our case, which is also a general issue with random effects, the variability among our cohort, does not let us identify the subtle vascular changes, suggesting that an alternative way (category 1) is needed.

Given the higher variability that can also be seen across the features, in comparison to the results in category 1, it is much more difficult to identify a difference. To do so, the sample has to be a lot larger and the outliers have to be quite limited. Taking into account that the above groups are coming from the same cohort as the subjects in category one, albeit not the same ones across all four groups, then it is even more convincing that committing Type I and Type II is easier. Needless to mention again that the retina by definition can drive us to more easily committing these errors, bringing again to the surface the importance of carefully selecting representative amount of large and fine vessels, if matching the same segments is not an option.

6.2.8 Discussion About the Results of the Statistical Analysis

The above results come from the most comprehensive and complete study yet to appear in literature. Seeing the problem from multiple different angles can give us a more thorough representation of what is happening in the retina during the progression of a disease, either inside a longitudinal follow-up design, or an independent one, or even a non-matched one etc. It was made clear at the previous chapters that the

exact mechanisms that affect the retina during the proliferation of diabetes are not accurately defined, although the general concepts have been clarified, i.e. pericyte loss, capillary occlusion, vascular permeability etc (Hammes *et al.*, 2002; Joussen *et al.*, 2004). Besides, there are factors that inevitably cannot be taken into account or included in the model when studying a disease. In addition it needs to be clear, which exact periods one is investigating and not try to vaguely generalise. A healthy group, a diabetic group and a DR group have already enough variability themselves to be able to come up with accurate conclusions about them. Besides, a diabetic group can include subjects in their early stages or at a stage which is near to progressing to a different one, e.g. to the early stages of DR. To account for this variability one either needs to include a really large representative sample, or instead target a specific period of time, just like it was done in this retrospective study, given that obtaining a large dataset is rather difficult.

In addition, as can be seen in category 5, these results indicate something extremely important. Two groups, that both include diabetic patients, and representative measurements averaged over a consecutive three- and four- year period accordingly, are statistically significantly different across many features (CRVE, CRAE, venular and arterial widths, blood flow etc.). That tells us something that has to be taken into account in future studies. The selection of a diabetic cohort has to have stricter requirements, since the alterations in the retinal vasculature are shown to change even during the progression of diabetes, before the onset of DR. Although it is also observed that between the late stages of the diabetic eye and the onset of DR quite pronounced alterations exist, they still remain less distinct than when compared with a diabetic group, in which one does not know the exact period that the images were taken from.

Regarding the above results, it is important to point out that, although the non-significant ones are also documented, the exact numbers of the variances and Ω^2 should not be used to make sound inferences, but rather for reference purposes and indica-

tions. The reason is that one cannot be confident at the defined significance level , that the observed results can be generalised beyond the random results that have been observed. Therefore, fluctuations in the variances and/or explained variances can be considered as random and more like point estimates of the specific sample, which can actually take any value. On the other hand, for the significant results, these metrics can provide us with important information. Intraclass correlation shows whether the random effect is actually present in our data justifying the use of LMMs. In all of the above results, wherever applicable, the use of both the nested effect and the random effects is justified, hence making sense conducting this type of analysis, producing more accurate estimates, explaining more variance, and finally reducing the residual variance. Moreover, the Ω^2 can give us an additional magnitude of the explained variance and the effect of the random covariate and the additional grouping. The interpretation of Ω^2 in LMMs is not that straightforward, however larger values generally indicate more important and higher effect of the covariates, with lower values indicating the opposite. Most of the significantly different biomarkers in the above categories are observed to have fair values of Ω^2, which indicates that the effect is high and not just significant.

In addition, no significant influence was observed when the factor age was added to the models of category 1 and 2, meaning that the comparison between a model with the fixed effect of age and a model with just the fixed effect of the disease gave no significant difference, which can be attributed either to the fact that the age groups of our data were relatively close, with many subjects around the same age, or that the number of subjects is not so great to offer enough information about this factor. An example of the non-existent correlation between the age and the differences in veins and arteries between the three year period pre-DR and the onset of DR can be seen in figure 6.2, which comes from the early phase's results. However, it needs to be highlighted, that for accurate conclusions about such factors, like age or other

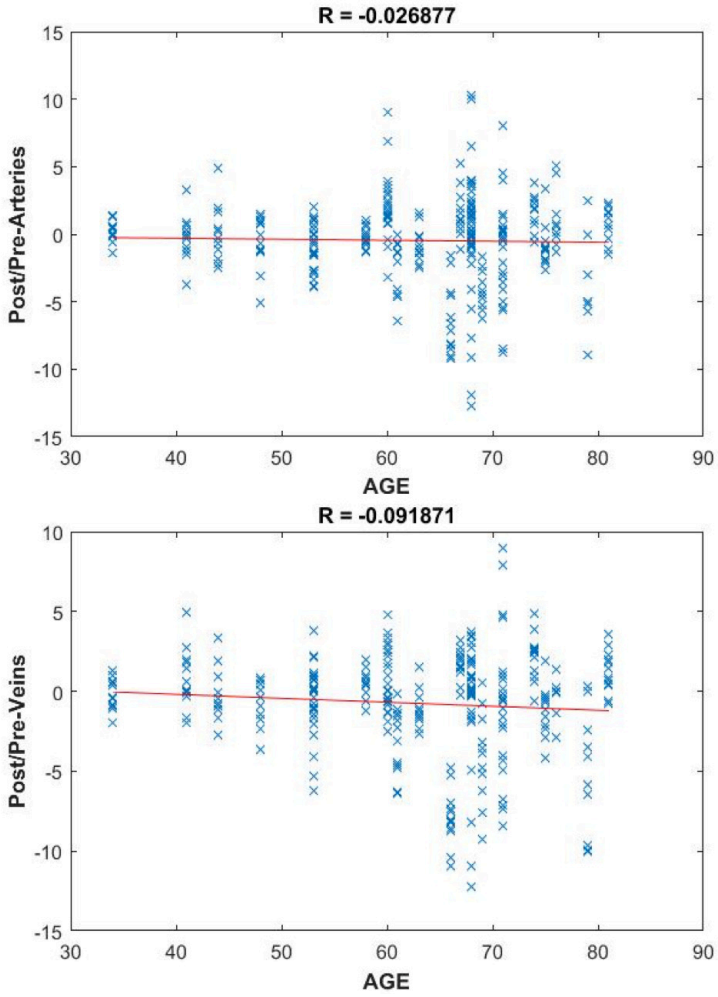

Figure 6.2: Up:Plot shows the difference between the three years pre-DR and the onset of DR for the arteries in relation to the age. Down: Similar plot but for the veins.

demographic information, a large sample is required, with enough representation of different levels of those factors. In addition, this study does not focus on demographic data, rather to the geometry and haemodynamics of the retina, so this information is only offered for reference purposes.

Another interesting fact to point out is that in all the combinations above, a random slope model was not justified over a simpler random intercept-only one. All the comparisons between them yielded no significant differences, strongly suggesting that the simpler model is adequate to model the changes. In terms of the studies presented here, this implies that it makes sense to assume that different subjects start from a dif-

ferent base, hence including a random intercept model. However, making a random slope no difference, suggests that the rate that the disease affects each subject is not significantly different among subjects.

To sum up, the statistical analyses of this novel study, gave us some very important findings.

1. Two different diabetic cohorts (analysis of category 5) can give totally different results, which can lead to contradicting findings, by committing either TYPE I or TYPE II errors. This enhances the importance of studying the same subjects over a period of time, and as far as possible, the same exact segments/junctions over a period of time. Planning such an experiment, and recruiting diabetic subjects, one cannot know at the moment of recruitment whether some of these subjects will progress to DR very soon after. This suggests that the variability within the cohort might be high enough to be possible to identify subtle vascular changes, when compared to a DR group. This difference could also suggest one additional thing. The diabetic cohort that is about to progress to DR in the near future, presents some noticeable differences compared to another diabetic group, which actually supports the hypothesis, that the retinal vasculature is affected and adapts to the underlying functional alterations that occur right before the onset of DR. Examples of that can be seen in Table 6.14, where many of these features, such as CRAE, arteries widths, tortuosity, blood flow, etc., differ significantly between the two diabetic cohorts.

2. In the investigated areas, apart from the widths of the parent veins in ellipse 1, the width of the large arterial child in ellipse 2 and the tortuosity in ellipse 3, no other significant result was observed, at least from a statistical point of view. Given that the same features, among others, were also found to significantly differ in the retina as a whole, shows no actual meaningful importance in specifically focusing on these areas, instead of studying the retina as a whole. This

observation also includes the local arteriovenous ratio, which does not seem to be changing during the progression of the disease, just like occurs with the AVR.

3. The haemodynamic features definitely add to the inferences, and makes sense to include them in any analysis of the disease's progression, especially the haemodynamics of the venular network, which as it was observed, it seems to be affected more than the arterial network.

   It needs to be kept in mind that micro-aneurysms are formed due to the loss of pericytes, whereas vascular occlusion is a consequence of basement membrane thickening and endothelial cell proliferation (King *et al.*, 1994). These processes, which are initially subtle, affect the retinal haemodynamics. In the normal retina, blood flow remains constant by the autoregulated vascular responses over a range of systemic blood pressures and intra-ocular pressures (Riva *et al.*, 1981). Local factors, which target smooth muscle cells in arterioles and capillary pericytes in patients with diabetes are controlling the vessels (De La Rubia *et al.*, 1992). In addition, there are changes occurring in the local vasoactive factors and in the response of pericytes due to these local factors (Burgansky-Eliash *et al.*, 2010).

4. In the progressors' group, prior to the onset of DR, all of the significantly altered features (CRVE, CRAE, fractal dimension, venular and arterial widths) are following a trend of decrease. On the other hand, the generalised tortuosity is increased, suggesting that the progression of diabetes to DR is linked to more tortuous vessels.

6.2.9 Conclusion

Undoubtedly, the retinal vascular geometry and haemodynamic functionality is shown to be changing during the progression of diabetes, which can be possibly attributed to

processes that include, among others, the neurodegeneration, the subclinical inflammation, the oxidative stress and the dysfunction of the endothelial cells that hyperglycaemia also causes (Leontidis *et al.*, 2014). In addition, the dysfunction of the endothelial cells affects the production of nitric oxide in response to shear stress, which can lead to the development of atherosclerosis, i.e. the hardening and thickening of the arterial wall, thus the observed arterial alterations.

As it was hypothesised, matching the vessel junctions and studying them over a period of time, one can identify changes that otherwise might not be identified or even misidentified. All of the above results strongly suggest that it is very easy to be misled and come up with conclusions that are not representative of what is being studied. Nonetheless, our studies showed that the geometry and the underlying haemodynamics are indeed affected during the progression of diabetes, which are more or less in line with our early results (Leontidis *et al.*, 2015a, 2016a, 2015b). The vessel widths, FD, venular angle to BC ratio, CRVE, CRAE, the blood flow rate of both arteries and veins, the pressure and velocity in the veins, etc. (Tables 6.5 and 6.14), are all biomarkers of progression to DR. CRVE and CRAE observations are in line with the findings of Klein and , when looking at Table 6.14, and partly with (Yang *et al.*, 2016; Tsai *et al.*, 2011; Islam *et al.*, 2009; Kifley *et al.*, 2007). These observations are linked possibly to the endothelial dysfunction, inflammatory changes, and hyperglycemia, all of which are factors involved in the pathogenesis of DR (Schmidt *et al.*, 1999; Klein *et al.*, 2000, 2012, 2004). However, on the exact same table, it can be seen that not only there is an increase to both these biomarkers (CRVE and CRAE) between non-progressors' diabetics and DRs, but also more importantly, the same occurs between the non-progressors and progressors. This extends the observations of Klein *et al.*, suggesting that the changes are not necessarily occurring during or because of DR, but rather during the last few years just before the onset of DR, when the retina is adjusting to the upcoming vascular and functional alterations. Fractal dimension is

also an important biomarker, with significant differences between diabetics and DR, which comes in contrast to the findings in (Kunicki *et al.*, 2009).

Finally, the most important and rigorous observation is that by not following a matched and repeated measures design, it is more likely to not observe an otherwise significant difference than the opposite, i.e. observe a significant one when there is not one. As a final remark, regarding the non-significant features (Appendix A), it is clear that the alterations attributed to the progression of the disease cannot be illustrated, as expected, in every single investigated feature. This is totally in line with the purpose of this thesis, which aims to identify important features, given the lack of similar pre-existing studies. This exploratory process yielded some important biomarkers, as mentioned previously in details, and this is something that was made possible by including all these candidate biomarkers in the analyses.

6.3 Classification Results

Following the statistical evaluation of the features in each category, which gives us a specific type of information, this section aims to check these features in practice and also decide whether a subset of them can be useful in discriminating between groups that correspond to different type of cohorts. The novelty of this approach includes two parts. The first is that the proposed approaches and experiments have never been attempted in the past, and also include different combinations that come from different periods of time. The second one is that this is done exclusively by using information directly from the vascular geometry and haemodynamics, with no information or features extracted from the images per se, whatsoever.

In this section, for each of the eight combinations that were described in chapter 4, the results of the feature selection process and the classification models will be presented. For each combination and each feature selection process, the final five best

models and feature subsets will be presented. Additionally, one plot for each feature selection method will be included, which will correspond to the best model and to two other models that were selected by a different selection process.

All the features were standardised before any of the processes is executed, by subtracting, feature-wise, each value from the corresponding feature's group mean value and diving by the standard deviation. All of the combinations contain balanced classes, 112 features with thirty observations per class for those that come from the progressors' data, and twenty obervations for those that come from the non-progressors' data. The combinations that include classes between non-progressors and progressors are also balanced, selecting randomly twenty out of the thirty observations that are available from the progressors' group. An exemption is the one vs all process, in which the three classes are still balanced, but the binary combinations merge two classes each time, creating a new one with forty observations, and therefore performing the classification versus the third class. Nevertheless, the folds of the cross validation process were kept balanced, in order not to have overpopulated from one class folds.

In comparison to the features in Table 3.3, which were used for the statistical analysis, for the classification the majority of them is the same, except for some differences that exist, which are summarised in Table 6.15. The differences mostly refer to the widths and angles of the vessels of the main features, which are summarised based on their descriptive statistics, in order to have equal observations for each feature in the classification, and also one observation per image.

6.3.1 Year Three Versus Onset of DR (Progressors' Group)

Starting with the classification of the three year period pre-DR versus the onset of DR, the aim is to actually visualise and evaluate how well the features can discriminate between these two different periods, which are still apart by three years. All the classification and feature selection processes previously described are applied in this

Table 6.15: Additional features for the classification

Features (Short form)	
1.Arterial parent's width of haemodynamics (**parenta**)	12.Arterial average length of junctions (**lengthArt**)
2.Arterial small child's width of haemodynamics (**smallchild2a**)	13.Arterial average branching coefficient(**BC_Mean_Art**)
3.Arterial large child's width of haemodynamics (**largechild1a**)	14.Standard deviation of arterial widths (**arterySD**)
4.Venular parent's width of haemodynamics (**parentv**)	15.Venular average length of junctions (**lengthVein**)
5.Venular small child's width of haemodynamics (**smallchild2v**)	16.Average Branching coefficient of veins (**VeinBc**)
6.Venular large child's width of haemodynamics (**largechild1v**)	17.Standard deviation of venular widths (**Vein_SD**)
7.Average arterial widths (**MeanArt**)	18.Average venular widths (**MeanVein**)
8.Median arterial widths (**MedArt**)	19.Median venular widths (**MedVein**)
9.Standard deviation of arterial angles (**ArtAngleSD**)	20.Standard deviation of venular angles (**VeinAnglSD**)
10.Average arterial angles (**MeanAngArt**)	21.Average venular angles (**MeanAngVein**)
11.Median arterial angles (**MedAngArt**)	22.Median venular angles (**MedAngVein**)

combination and also to all the subsequent ones, apart from combination 8.

In Table 6.16 the summary of the results of the whole process can be seen. In this combination the best classification performance and feature subset selection was achieved by the regularised random forests classifier, yielding an average AUC = 0.878 (Table 6.16), which outperforms the previous reported results (Leontidis *et al.*, 2015a, 2016a).

Table 6.16: Summary of the classification performance for Year three versus Onset of DR (Combination 1)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests AUC (CI95%)	**0.878 (0.732-0.921)**	0.842 (0.694-0.935)	0.835 (0.671-0.936)	0.813 (0.659-0.924)	0.74 (0.582-0.852)
Logistic Regression AUC (CI95%)	0.754 (0.643-0.842)	0.76 (0.648-0.856)	0.73 (0.629-0.816)	0.793 (0.703-0.89)	0.811 (0.694-0.902)
Random Forests kappa	0.632	0.614	0.609	0.602	0.498
Logistic Regression kappa	0.483	0.492	0.476	0.534	0.619
Random forests OOB	0.191	0.213	0.228	0.253	0.298
Feature selection process	RF	RF	BO	EN	EN

As described in chapter 4, for each of the seven combinations of the classifications, sixty different models are created, based on the feature selection process (according to the penalisation parameters as provided by the bootstrapping process), classifier (logistic regression or random forests) and the different folds of the cross validation. However, instead of only presenting the one with the best performance, the five best

of them that yielded the highest performance are described, alongside their results and their feature subsets. The features for the above models, ordered according to their importance, are the following:

Best model: ArtAngleSD, lengthVein, BC_Mean_Art, Angle.BC_Art, MeanArtAng, pressureA, MedAngArt, Tort_ellipse3.

Model 2: LengthVein, Angle.BC_Art, BC_Mean_Art, Angle.BC_Vein, ArtAngleSD, Tort_ellipse3.

Model 3: ArtAnglesSD, BC_Mean_Art, Angle.BC_Art, MedAngArt, lengthVein.

Model 4: Qchild1v, Angle.BC_Art, Vein_SD, Tort, Tort_ellipse3, lengthVein.

Model 5: Angle.BC_Art, Vein_SD, BC_Mean_Art, Tort, Tort_ellipse3, lengthVein.

In the following figures 6.3, 6.4, 6.5 and 6.6 the examples of the feature selection process of the three techniques can be seen, including the ROC curves of the two best models. Figure 6.7 also shows the bootstrapped AUC.

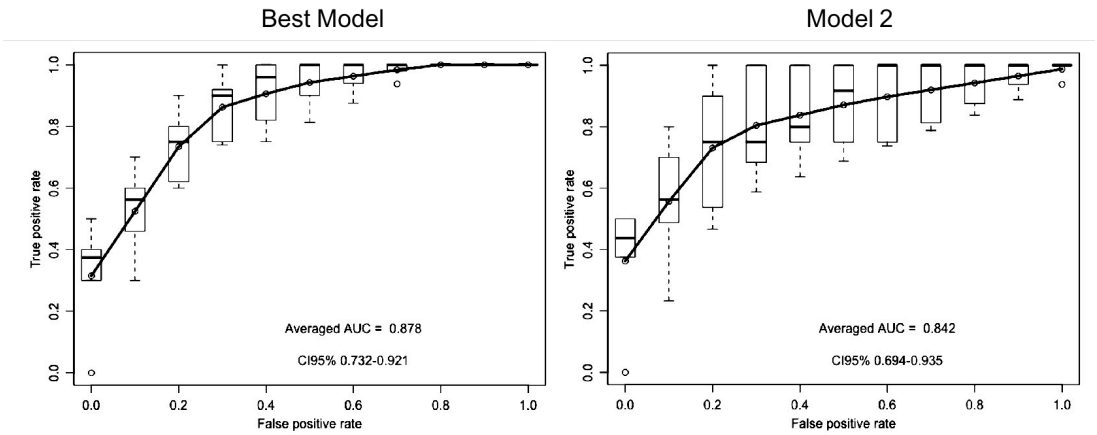

Figure 6.3: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 1; on the right similar information is given for the second best model.

The above plots give a proper visualisation of the whole process of both feature selection and classification performance. The bootstrapped AUC and the ROC plots suggest that the classifier has a good performance in differentiating between the observations that come from the three year period pre-DR and those that come from the

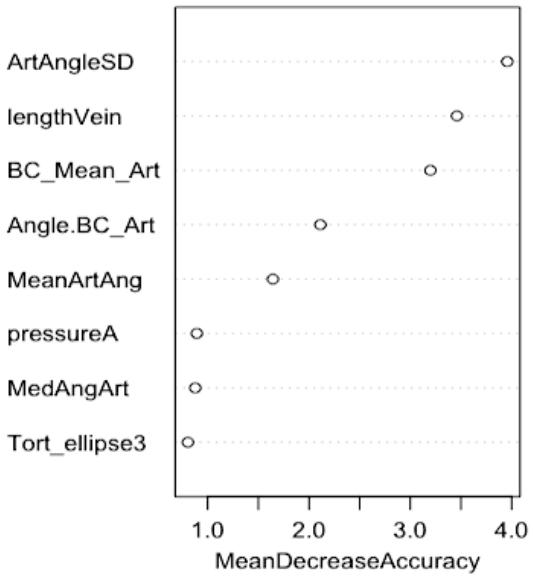

Figure 6.4: Feature ranking of the regularised random forests classifier for the best model of combination 1, according to the mean decrease accuracy, which shows how the performance of the classifier is affected if one of these features is removed.

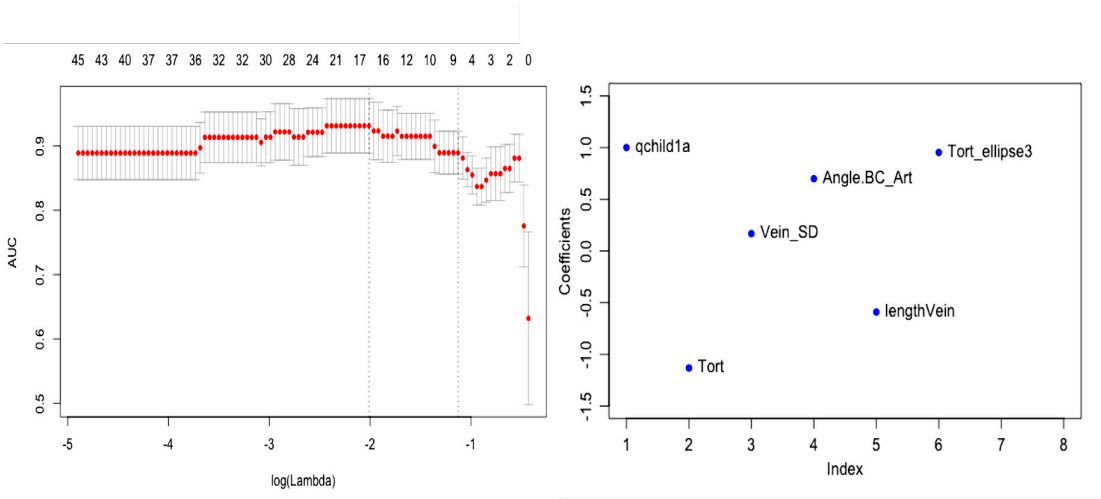

Figure 6.5: This plot refers to the elastic-net feature selection process of the model 4 of the combination 1. On the left: The regularisation path of the elastic-net can been seen, with the x axis including the penalty parameter λ in a logarithmic scale, the y axis being the performance according to the AUC and on top of the plot the number of features. In addition the upper and lower standard deviation curves along the red cross-validation curve are included; on the right the coefficients and the final selected features can be seen, where x axis is the index of the number of features, with the y axis including the values of the coefficients of the predictors.

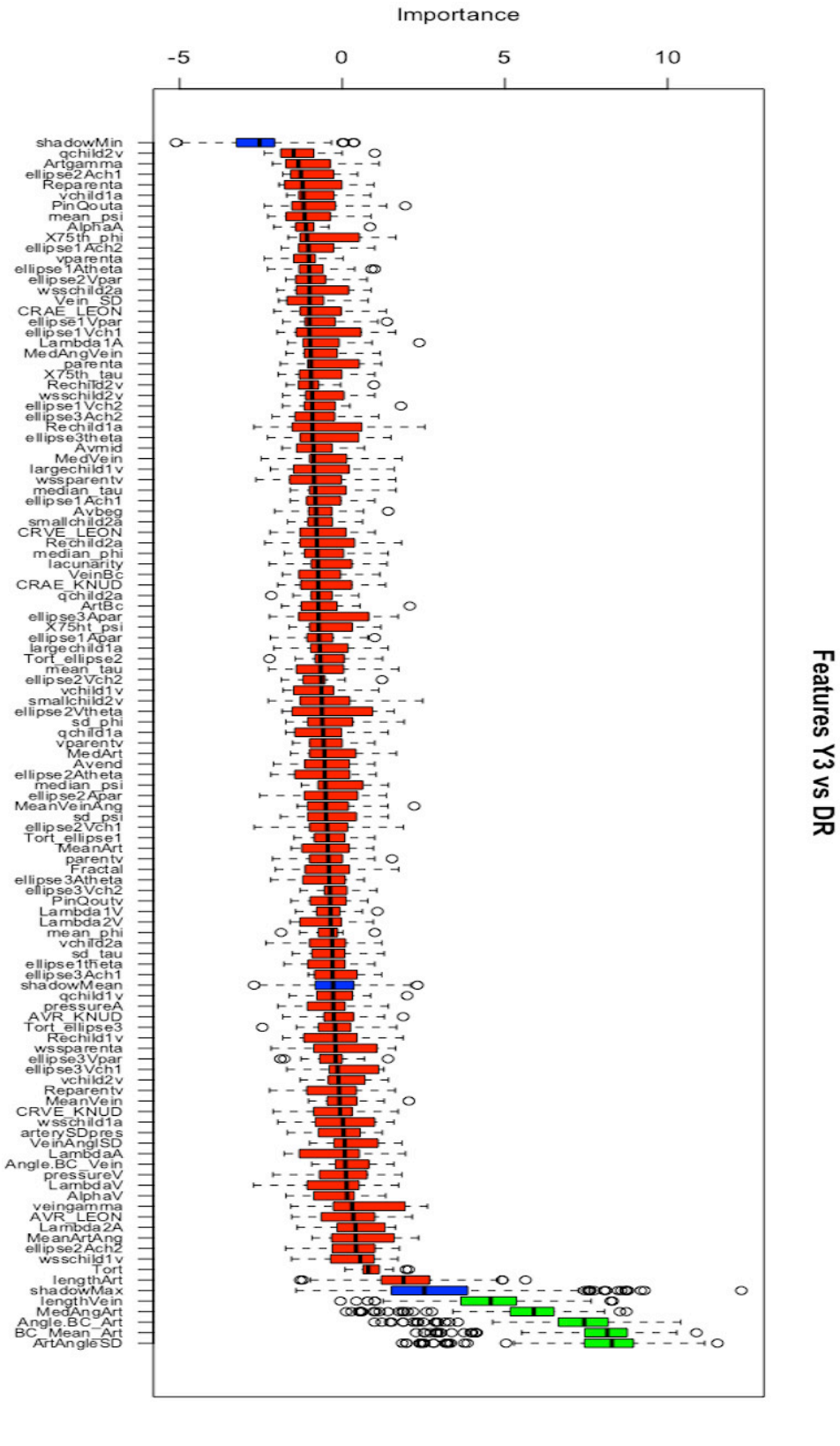

Figure 6.6: Feature selection process according to boruta method for model 3 of combination 1; y axis shows the relative importance of each feature, whereas x axis includes the features. Green coloured features are the finally selected ones, whereas the red are the rejected ones.

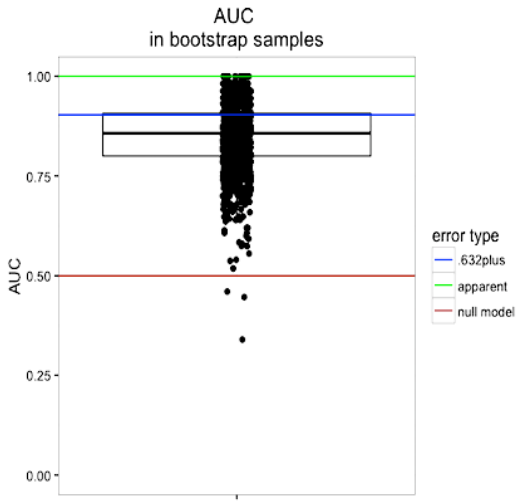

Figure 6.7: Bootstrapped AUC for the best model of the combination 1. It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested using the same training set and the random model, respectively.

onset of DR, which follows the same trend as the statistical analysis. The final model includes a combination of the derivatives of arterial angles, the lengths of the veins, the tortuosity of ellipse 3, the arterial pressure and the ratio of the arterial to the the branching coefficient.

6.3.2 Year Two Versus Onset of DR (Progressor's Group))

Just like the previous section's process, the best performance according to the AUC as achieved by the best model equals to 0.745 (Table 6.17). This performance was achieved by the logistic regression classifier, using a feature subset that was selected by the elastic-net (α = 0.13 and λ = 0.23). All the details of the performance of the five best models can be seen in Table 6.17.

The features for the above models, ordered according to their importance, are the following:

Best model: PressureV, Angle.BC_vein, Angle.BC_Art, ellipse1vpar, ellipse1Vch1, ellipse2Ach2, ellipse3Vtheta, Tort_ellipse3.

Model 2: Angle.BC_Vein, Angle.BC_Art, Lambda1V, MedAngArt, ellipse1Vpar, ellipse3Vtheta,

Table 6.17: Summary of the classification performance for Year two versus Onset of DR (Combination 2)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests AUC (CI95%)	0.668 (0.541-0.752)	0.643 (0.512-0.732)	0.658 (0.529-0.754)	0.698 (0.514-0.813)	0.710 (0.523-0.802)
Logistic Regression AUC (CI95%)	**0.745 (0.576-0.843)**	0.724 (0.587-0.845)	0.714 (0.554-0.848)	0.713 (0.532-0.822)	0.619 (0.483-0.723)
Random Forests kappa	0.293	0.282	0.274	0.307	0.312
Logistic Regression kappa	0.478	0.483	0.466	0.473	0.244
Random forests OOB	0.391	0.403	0.382	0.343	0.324
Feature selection process	EN	EN	BO	BO	RF

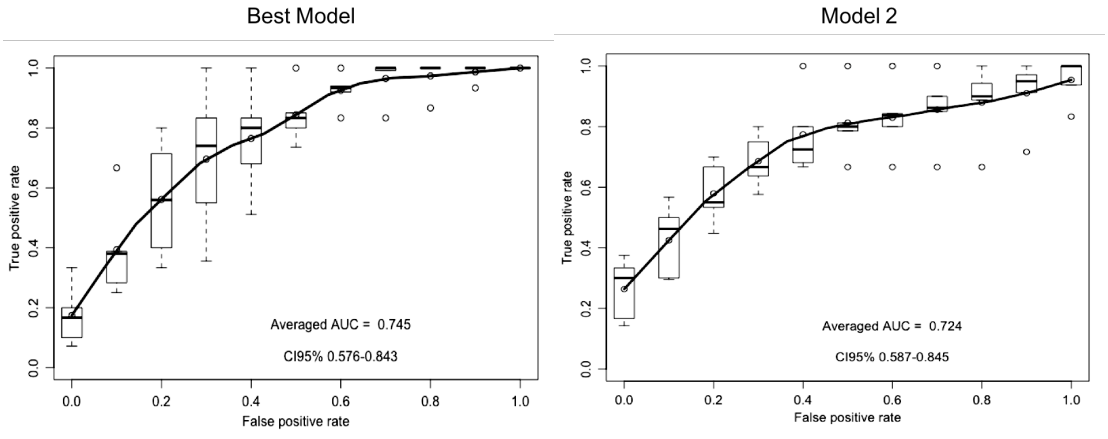

Figure 6.8: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 2; on the right similar information is given for the second best model.

Tort_ellipse2.

Model 3: Median_tau, Angle.BC_Vein, ellipse3thera, Fractal.

Model 4: CRAE_LEON, Tort_ellipse3, ellipse1Vpar, smallchild2a, CRVE_LEON, CRVE_KNUD, mean_psi.

Model 5: Tort_ellipse3, veingamma, Fractal, CRVE_LEON, Tort_ellipse1, ellipse3ch2, arterySD, ellipse3VTheta, Lambda2A, ellipse1Vch1.

In the figures that follow, i.e. figures 6.8, 6.9, 6.10, 6.11, 6.12, all three feature selection processes can be seen, the ROC curves of the two best models and the bootstrapped AUC as well.

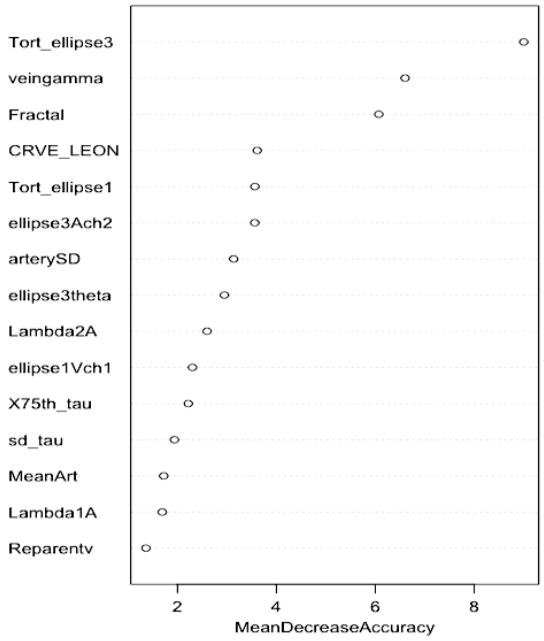

Figure 6.9: Feature ranking of the regularised random forests classifier for the model 5 of combination 2, according to the mean decrease accuracy, which shows how the performance of the classifier is affected if one of these features is removed.

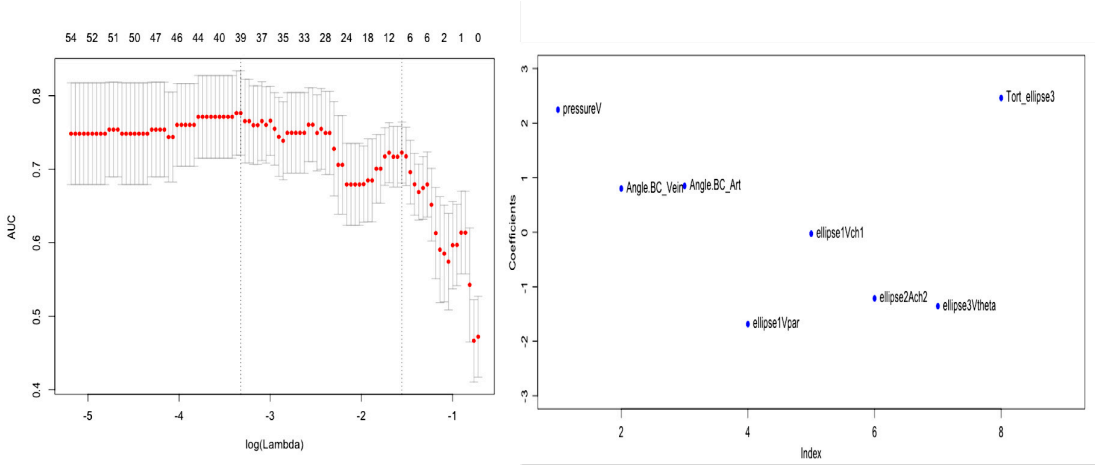

Figure 6.10: This plot refers to the elastic-net feature selection of the best model of the combination 2. On the left: The regularisation path of the elastic-net can been seen, with the x axis including the penalty parameter λ in a logarithmic scale, the y axis being the performance according to the AUC and on top of the plot the number of features. In addition the upper and lower standard deviation curves along the red cross-validation curve are included; on the right the coefficients and the final selected features can be seen, where x axis is the index of the number of features, with the y axis including the values of the coefficients of the predictors.

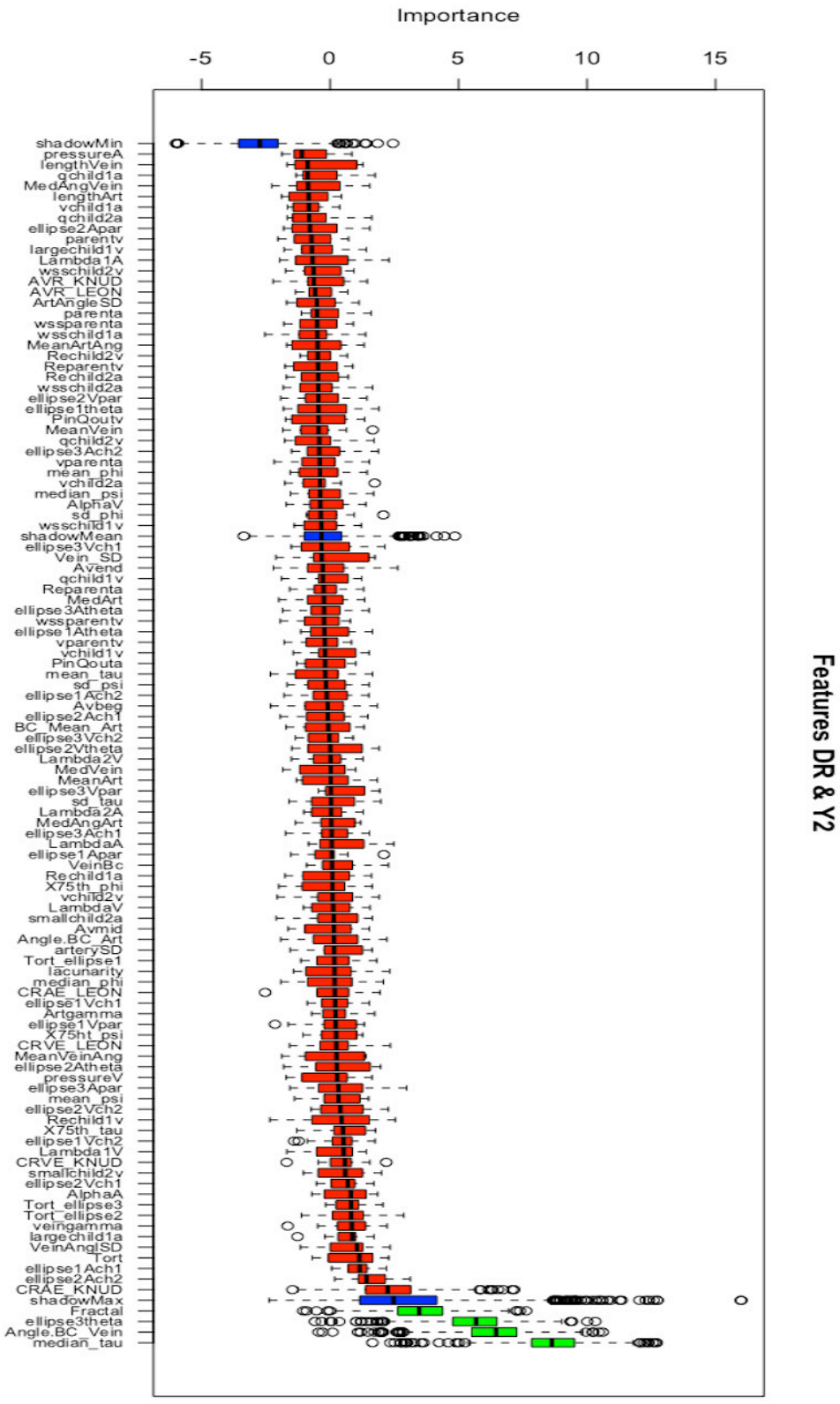

Figure 6.11: Feature selection process according to boruta for model 3 of combination 2, with the y axis showing the relative importance of each feature. The x axis includes the features. Green coloured features are the finally selected ones, whereas the red are the rejected ones.

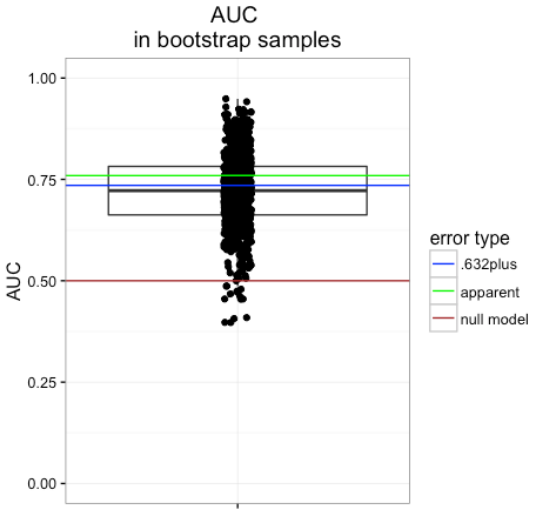

Figure 6.12: Bootstrapped AUC of the best model of the combination 2. It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested using the same training set and the random model, respectively.

In contrast to the results of the combination 1, passing into the two year period pre-DR, makes the changes less distinct. Therefore predicting the two classes is more difficult and the performance is fair but not enough to provide a reliable way of doing so. Nevertheless, the performance was achieved by a combination of venular pressure, venular angle to BC ratio, arterial angle to BC ratio, and venular widths and angles coming from the areas.

6.3.3 Year One Versus Onset of DR (Progressors' Group)

This combination includes two groups that are very close in time. The changes at this level between the two groups are possibly more intense, since the retinal vasculature undergoes a series of functional alterations that precede DR. The best performance based on the AUC equals to 0.776 (Table 6.18), achieved by the random forest's classifier, using a feature subset as selected by the boruta feature selection process.

The features for the above models, ordered according to their importance, are the following:

Table 6.18: Summary of the classification performance for Year one versus Onset of DR (Combination 3)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests	**0.776**	0.728	0.637	0.678	0.725
AUC (CI95%)	**(0.633-0.836)**	(0.587-0.843)	(0.483-0.758)	(0.512-0.792)	(0.546-0.836)
Logistic Regression	0.653	0.765	0.745	0.732	0.689
AUC (CI95%)	(0.521-0.744)	(0.598-0.873)	(0.568-0.854)	(0.567-0.765)	(0.538-0.853)
Random Forests Kappa	0.508	0.489	0.243	0.278	0.458
Logistic Regression kappa	0.311	0.512	0.498	0.475	0.378
Random forests OOB	0.273	0.324	0.415	0.397	0.314
Feature selection process	BO	EN	EN	RF	RF

Best model:Ellipse1Vch2, PinQoutV, X75th_psi, pressureV.

Model 2:Qchild1a, qchild2a, pressureV, wsschild1V, ellipse3Vch1.

Model 3: Qchild1a, pressure, wsschild1V, PinQoutV, Lambda1A, AlphaV, MeanArtAng, MedAngart, ellipse2Ach2, ellipse3Vch1, ellipse3Vch2.

Model 4: Lambda1A, pressureV, wsschild1v, qchild2a, PinQoutV, Angle.BC_art, median_phi, ellipse3Atheta, x75th_psi, CRVE_LEON, ellipse3Ach1, Lambda1V, MeanArt.

Model 5: Ellipse2Ah2, pressureV, Lambda1A, Rechild2A, median_phi, ellipse1Vch2, ellipse2Ach1, largechild1V, ellipse3Ach2, Tort_ellipse1, AvmidAvbeg, Tort_ellipse2.

Figures 6.13-6.17 present a series of plots that visualise the performance of the selection processes and classifications for combination 3.

This combination achieved the best performance by using geometric and haemodynamic features from the venular network. The performance also suggests that there are still underlying changes occurring at this transitional year, despite the closeness to progressing to DR. This can be better justified by the fact that the changes on the exact same segments are evaluated. A combination of the venular pressures, tortuosity and venular widths in ellipse 1 provided the best classification performance.

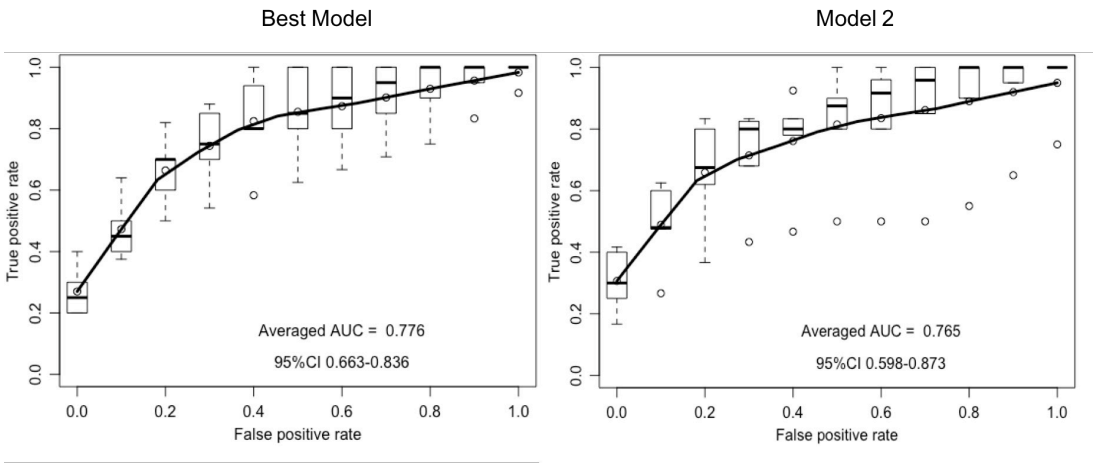

Figure 6.13: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 3; on the right similar information is given for the second best model.

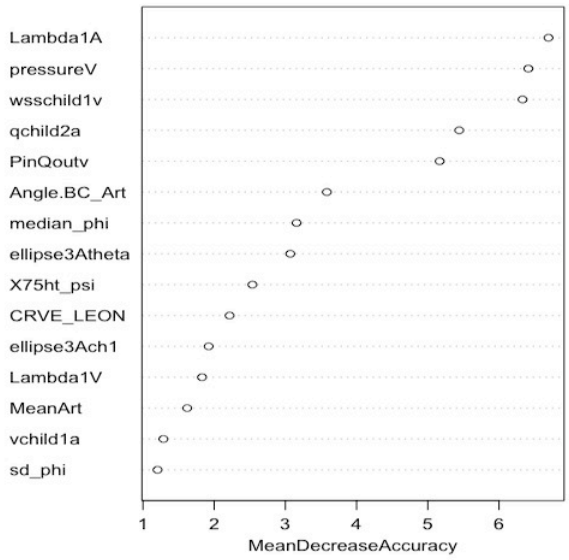

Figure 6.14: Feature ranking of the regularised random forests classifier for model 4 of combination 3, according to the mean decrease accuracy, which shows the effect that removing a feature has to the performance of the classifier.

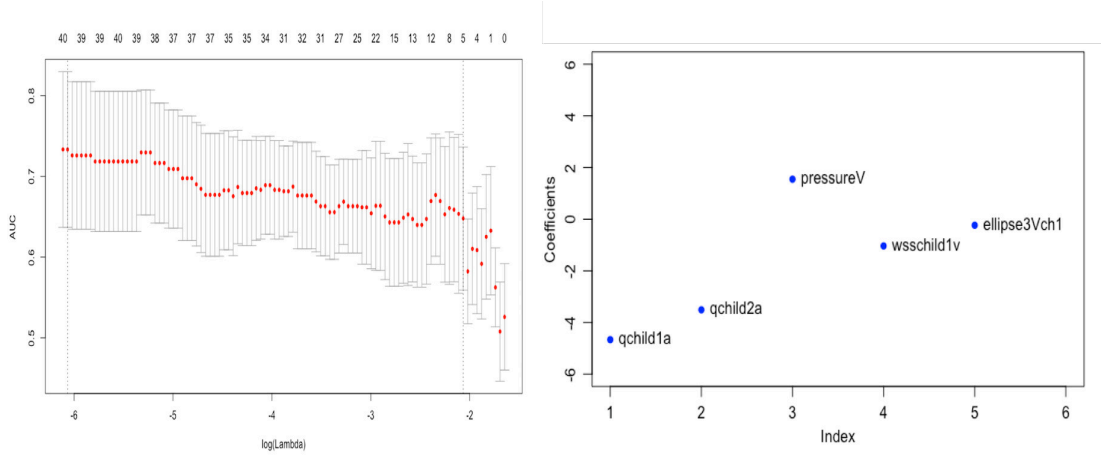

Figure 6.15: This plot refers to the elastic-net feature selection of model 2 of the combination 3. On the left: The regularisation path of the elastic-net can been seen, with the x axis including the penalty parameter λ in a logarithmic scale, the y axis being the performance according to the AUC and on top of the plot the number of features. In addition the upper and lower standard deviation curves along the red cross-validation curve are included; on the right the coefficients and the final selected features can be seen, where x axis is the index of the number of features, with the y axis including the values of the coefficients of the predictors.

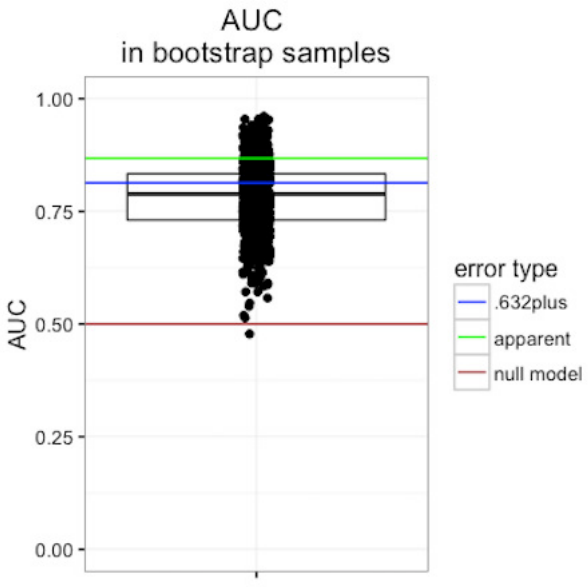

Figure 6.16: Bootstrapped AUC for the best model of the combination 3. It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested using the same training set and the random model, respectively.

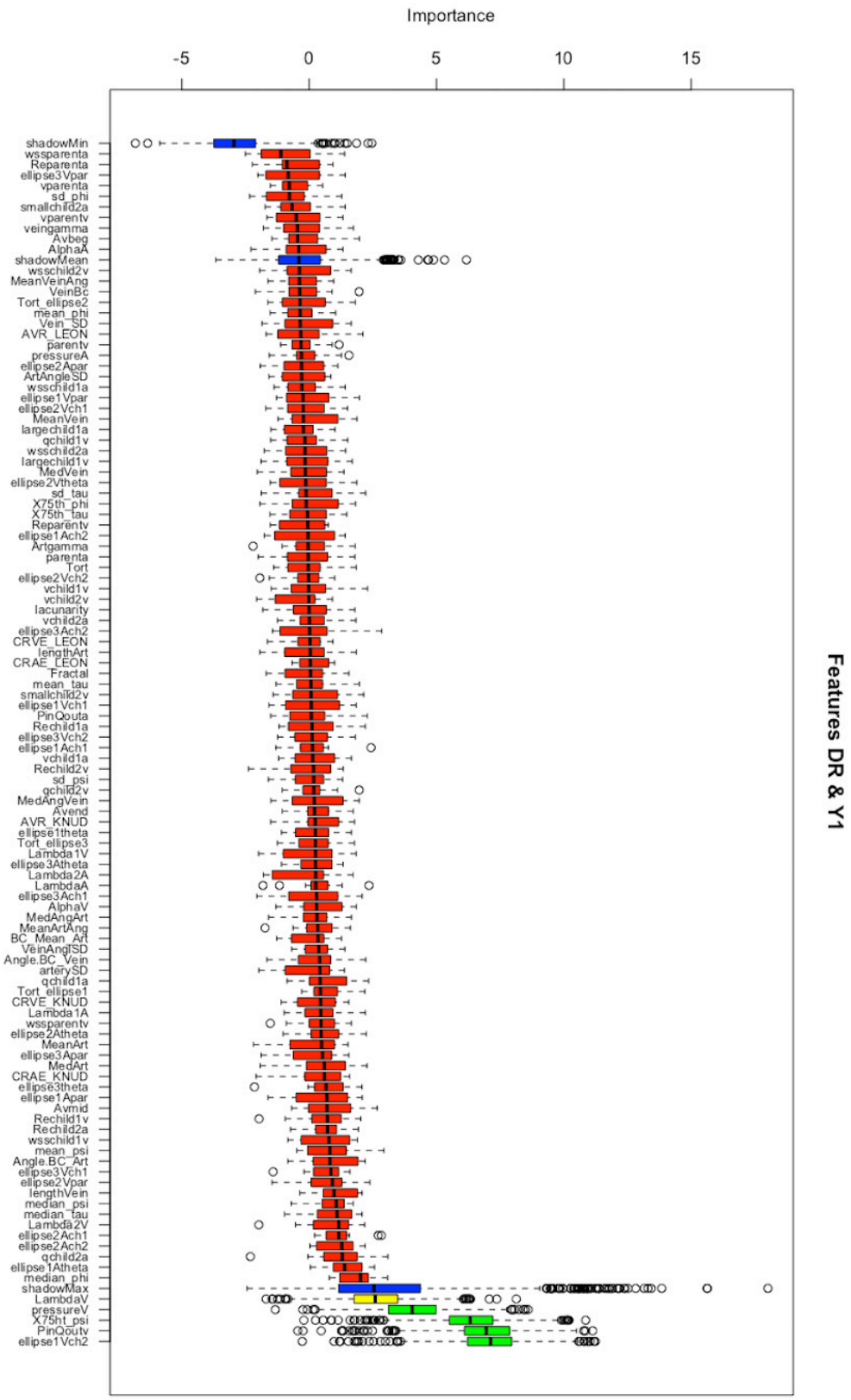

Figure 6.17: Feature subset according to boruta selection process. The plot refers to the best model of combination 3; y axis shows the relative importance of each feature, whereas x axis includes the features. Green coloured features are the finally selected ones, whereas the red are the rejected ones.

6.3.4 Mean Diabetics (Progressor's Group) Versus Onset of DR

This combination includes two classes; one being the DR group, just like in the previous combinations, and the other one being the average of the entire three year period pre-DR, i.e the average of the diabetic classes of combinations 1, 2 and 3. The previous combinations aim to give us an indication of the discriminative power of the features as the disease progresses, whereas combination 4 is a more realistic situation where a group that takes information from a few year period is included, in order to create a more representative group of the progessors of the subjects with diabetes. Still the purpose of this combination remains to evaluate the definite changes that have occurred to the retinal vasculature between the last stages of the diabetic eye and the onset of DR. The best performance in this combination was achieved by the feature subset and classification of the regularised random forest's classifier, yielding an AUC of 0.841 (Table 6.19).

Table 6.19: Summary of the classification performance for meandiab vs DR (Combination 4)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests	**0.841**	0.819	0.813	0.647	0.672
AUC (CI95%)	**(0.698-0.932)**	(0.645-0.912)	(0.654-0.921)	(0.489-0.777)	(0.543-0.798)
Logistic Regression	0.735	0.675	0.754	0.763	0.724
AUC (CI95%)	(0.534-0.853)	(0.521-0.765)	(0.567-0.843)	(0.574-0.853)	(0.531-0.834)
Random Forests kappa	0.542	0.514	0.511	0.245	0.269
Logistic Regression kappa	0.468	0.343	0.503	0.509	0.462
Random forests OOB	0.223	0.254	0.245	0.367	0.389
Feature selection process	RF	BO	RF	EN	EN

The features for the above models, ordered according to their importance, are the following:

Best model Ellipse1Vch2, Tort_ellipse1, Angle.BC_Art, Tort_ellipse3, Fractal.

Model 2: Tort_ellipse1, ellipse1Vch2, ellipse3Vtheta, pressureV, pressureA.

Model 3: Ellipse1Vch2, Tort_ellipse1, Fractal, Tort_ellipse3, pressureV.

Model 4: Fractal, Angle.BC_Vein, Angle.BC_Art, ellipse3theta, Tort_ellipse3.

Model 5: PressureV, Fractal, Angle.BC_Vein, Angle.BC_Art, Artgamma, MedAngArt, ArtAngleSD, ellipse1Vch2, ellipse2Vch1, ellipse3Vtheta, Tort_ellipse3.

Just like previously, the figures in this section show the performance of the feature selection processes and the classification.

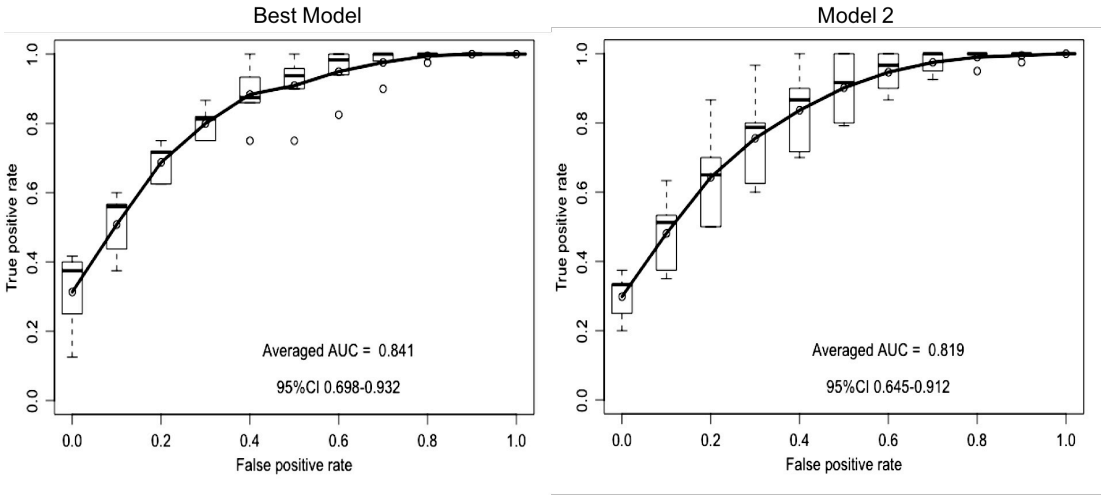

Figure 6.18: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 4; on the right similar information is given for the second best model.

The features that yielded that performance include the ones coming from the different areas, the ratio of the arterial angles and BC, and also the fractal dimension. The performance suggests that still distinct changes are occurring that could differentiate between diabetic and DR cohorts. Such classifier could have an application in automatic diagnosis of DR, where instead of looking only at the lesions, information from the vascular geometry can also be included for an initial sorting of the pre- to post-DR cases. In addition, the good performance of the classifier implies that it can be directly used for classifying a progressors' cohort versus a DR one, since the three year period of available information is possibly enough to create a representative sample.

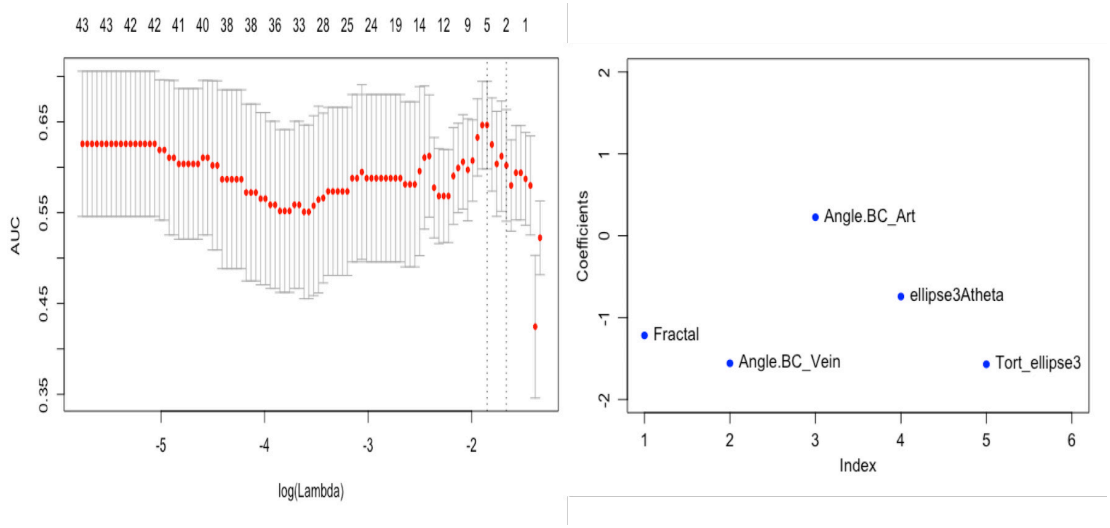

Figure 6.19: This plot refers to the elastic-net feature selection of model 4 of the combination 4. On the left: The regularisation path of the elastic-net can been seen, with the x axis including the penalty parameter λ in a logarithmic scale, the y axis being the performance according to the AUC and on top of the plot the number of features. In addition the upper and lower standard deviation curves along the red cross-validation curve are included; on the right the coefficients and the final selected features can be seen, where x axis is the index of the number of features, with the y axis including the values of the coefficients of the predictors.

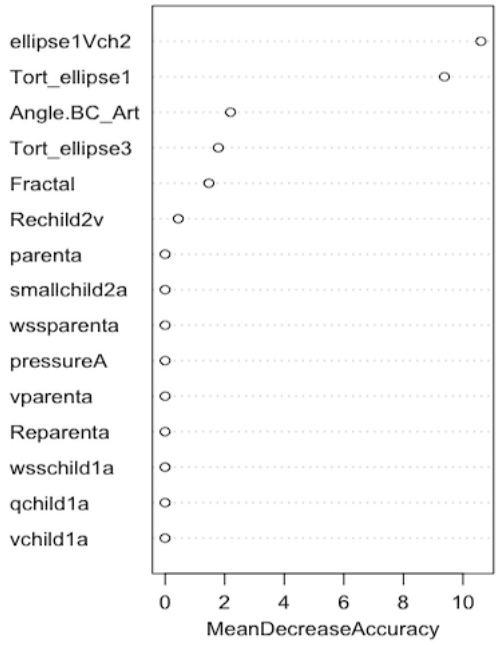

Figure 6.20: Feature ranking of the regularised random forests classifier for the best model of combination 4, according to the mean decrease accuracy, showing the effect that removing a feature has to the classifier's performance

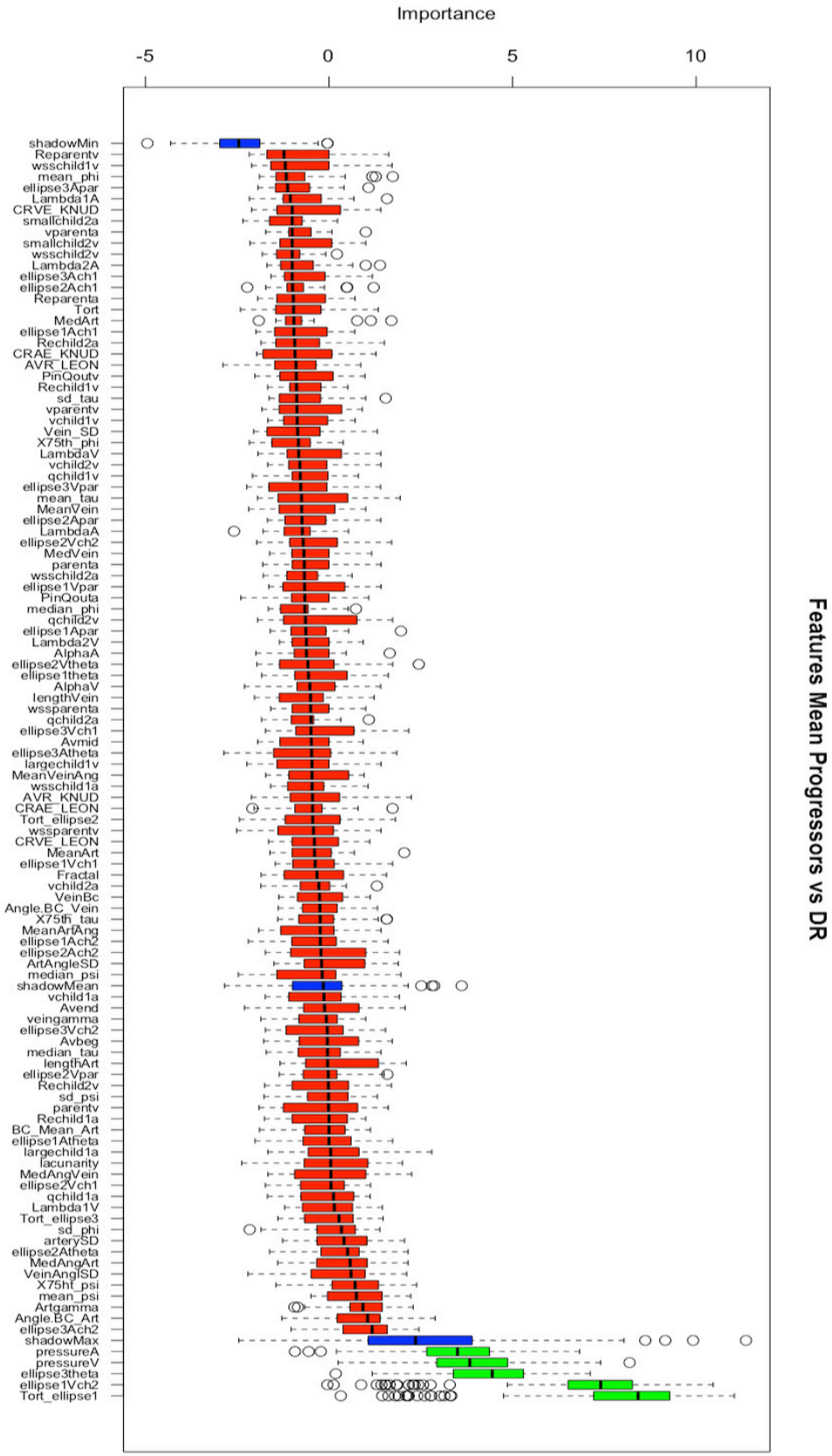

Figure 6.21: Feature subset according to boruta selection process. The plot refers to model 2 of combination 4; y axis shows the relative importance of each feature, whereas x axis includes the features. Green coloured features are the finally selected ones, whereas the red are the rejected ones.

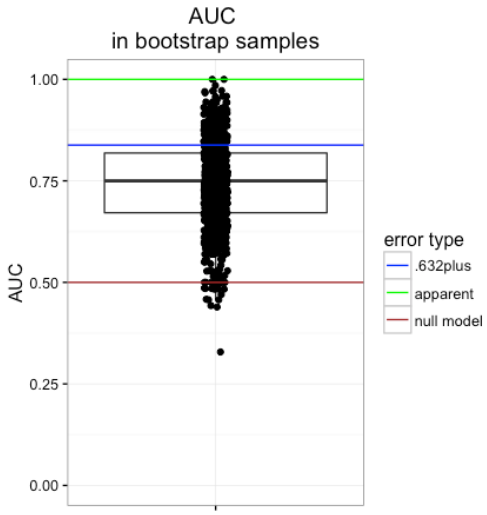

Figure 6.22: Bootstrapped AUC for the best model of the combination 4. It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested using the same training set and the random model, respectively.

6.3.5 Mean Diabetics (Non-Progressor's Group) Versus Onset of DR)

In contrast to the purpose of the previous four combinations, where the objective is to evaluate the extent of the vascular changes in an alternative to the statistical analysis manner, and also how these can be used inside a classification system, in this combination the focus is to try to see how a classifier can possibly perform in a pragmatic and realistic situation. If in the future, a high performance and reliable classification model is to be proposed that could classify diabetic and DR subjects, it needs to include independent information, given that the future instances that will need to be classified will come from independent subjects. To test this scenario, the classifier was trained by including information from the non-progressors' group and the DR group.

The maximum achieved performance was made possible by the logistic regression's classifier from a feature subset coming from the elastic-net process (α = 0.08 and λ = 0.17), giving an AUC of 0.968 (Table 6.20).

The features for the above models, ordered according to their importance, are the

Table 6.20: Summary of the classification performance for non progressors vs DR (Combination 5)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests AUC (CI95%)	0.945 (0.817-0.991)	0.961 (0.869-0.992)	0.943 (0.835-0.987)	0.942 (0.844-0.975)	0.931 (0.812-0.972)
Logistic Regression AUC (CI95%)	**0.968 (0.923-0.999)**	0.823 (0.656-0.912)	0.786 (0.643-0.889)	0.794 (0.587-0.887)	0.903 (0.789-0.981)
Random Forests kappa	0.887	0.907	0.875	0.867	0.854
Logistic Regression kappa	0.891	0.654	0.567	0.587	0.803
Random forests OOB	0.113	0.124	0.133	0.122	0.127
Feature selection process	EN	BO	BO	RF	EN

following:

Best model: Vparenta, wsschild1v, PinQoutV, Vein_SD, lengthVein, MedArt, MedVein, MeanVeinAng, VeinAnglSD, CRAE_KNUD, CRAE_LEON, sd_phi, mean_tau, Tort_ellipse3.

Model 2: CRAE_LEON, CRAE_KNUD, CRVE_KNUD, arterySD, Artgamma, VeinAnglSD, CRVE_KNUD, lengthVein, wsschild2a, Rechild2v, Medart, wsschild1v, Rechild1a, VeinBc, Reparenta, vparenta, ArtAnglSD, BC_Mean_Art.

Model 3: CRAE_LEON, CRAE_KNUD, CRVE_LEON, Artgamma, arterySD, wsschild2a, sd_phi, BC_Mean_Art, MedAnglVein, ArtAngleSD, Medvein, MedArt, wsschild1v, Rechild1a, CRVE_KNUD, Rechild2v, largechild1v, veingamma.

Model 4: CRAE_LEON, MedAngVein, lengthart, lengthvein, Rechild2v, wsschild1v, arterySD, wssparenta, Rechild1a.

Model 5: Vparenta, Reparenta, PinQoutA, lengthVein, MedVein, VeinAngSd, CRAE_KNUD, CRAE_LEON, CRVE_KNUD, CRVE_LEON, sd_phi, Tort, Tort_ellipse3.

The following figures 6.23-6.27, show the examples of the feature selection process of the three processes, the ROC curves of the two best models and the bootstrapped AUC.

The features that gave this performance include mostly venular haemodynamic and geometric features but also a few arterial geometric ones. The interesting part is that these features are different from the ones in combination 4. This suggests that there are

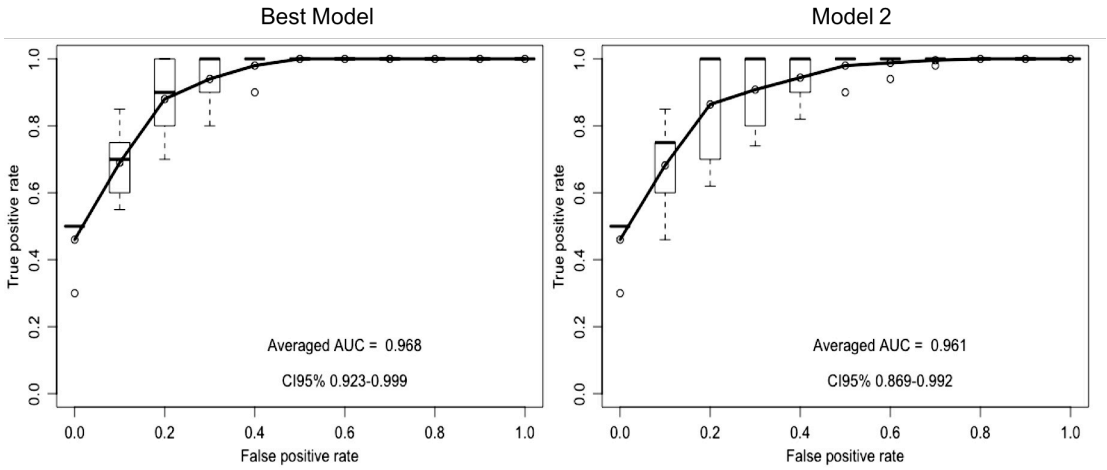

Figure 6.23: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 5; on the right similar information is given for the second best model.

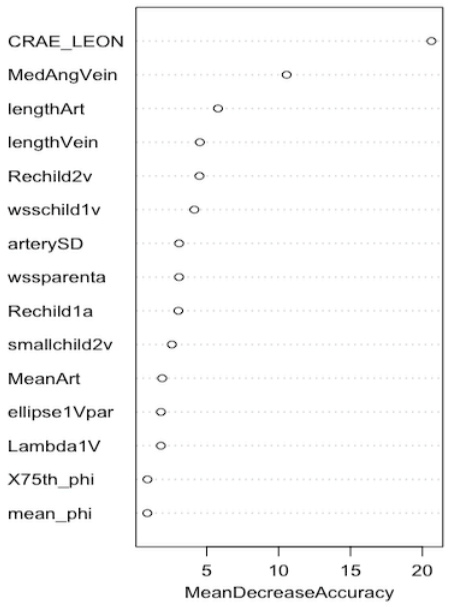

Figure 6.24: Feature ranking of the regularised random forests classifier for model 4 of combination 5 based on the mean decrease accuracy, which shows the effect that removing a feature has to the performance of the classifier.

different attributes that differentiate the two different diabetic cohorts from DR, implicating that the vascular changes are not the same across all the periods of time that the disease proliferates. This is something that is also supported by the statistical analysis. Therefore it is worth investing on collecting representative data from both diabetic

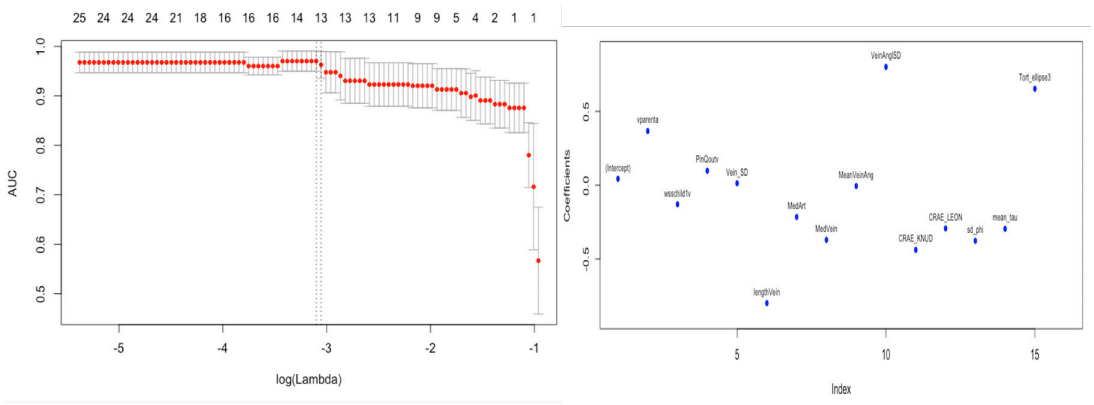

Figure 6.25: This plot refers to the elastic-net feature selection of the best model of the combination 5. On the left: The regularisation path of the elastic-net can been seen, with the x axis including the penalty parameter λ in a logarithmic scale, the y axis being the performance according to the AUC and on top of the plot the number of features. In addition the upper and lower standard deviation curves along the red cross-validation curve are included; on the right the coefficients and the final selected features can be seen, where x axis is the index of the number of features, with the y axis including the values of the coefficients of the predictors.

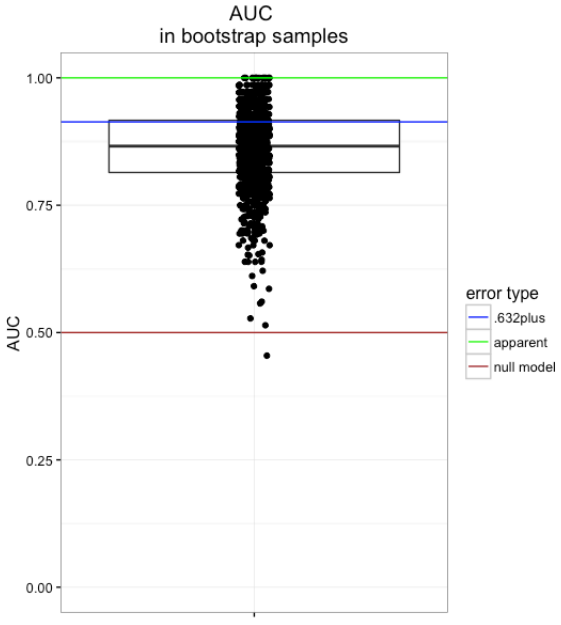

Figure 6.26: Bootstrapped AUC for the best model of the combination 5. It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested using the same training set and the random model, respectively.

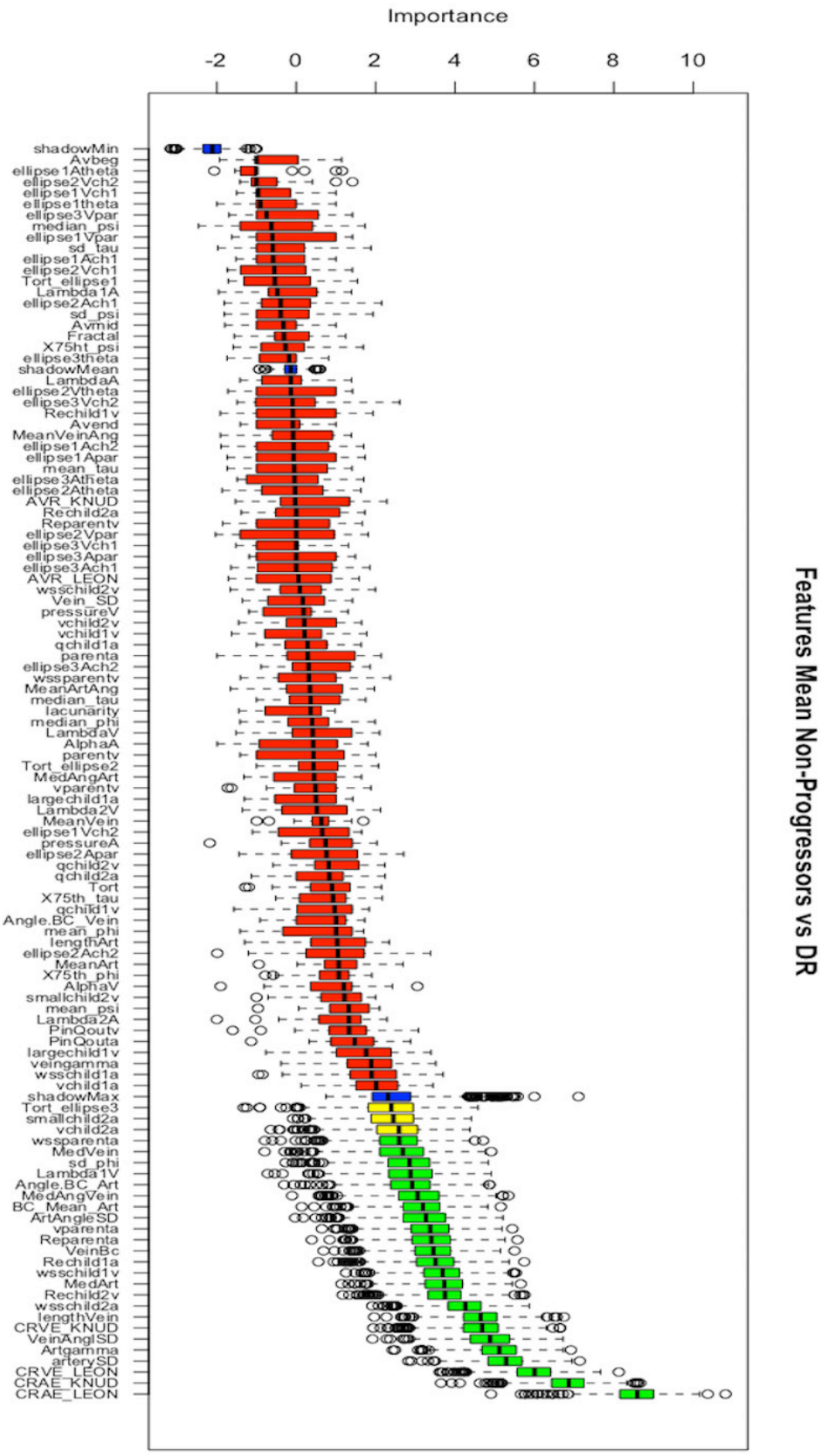

Figure 6.27: Feature selection process according to Boruta for the second best model of combination 5. Green labeled features are the finally selected ones, whereas the red are the rejected ones. The yellow represent the tentative features, which can either be rejected or accepted, subject to repeating some of the steps of the algorithm.

cohorts and also make all the efforts to be considered separately in the future studies. The remarkable performance of this classification model, with the tight confidence intervals as well, gives evidence that such a model could be used for differentiating between subjects with diabetes and subjects with DR.

6.3.6 Mean Diabetics (Non-Progressor's Group) Versus Mean Diabetics (Progressor's Group))

Similar to the objective of category's 5 statistical analysis, this combination aims to actually try to differentiate between a group of subjects with diabetes that are not progressors and a group that they are about to progress to DR soon. This combination has high clinical importance, given that being able to accurately differentiate between these two groups, can give precious indications that one is about to progress to DR in the near future. The random forests classifier gave the best performance of 0.963 (Table 6.21) according to the AUC, using a feature subset that was selected by the boruta feature selection process.

Table 6.21: Summary of the classification performance for groups Meandiab vs meanNonProg (Combination 6)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests	**0.963**	0.957	0.957	0.949	0.937
AUC (CI95%)	**(0.898-0.998)**	(0.887-0.991)	(0.893-0.997)	(0.896-0.987)	(0.867-0.977)
Logistic Regression	0.923	0.865	0.902	0.845	0.856
AUC (CI95%)	(0.854-0.995)	(0.667-0.934)	(0.756-0.983)	(0.673-0.921)	(0.667-0.946)
Random Forests kappa	0.903	0.896	0.856	0.843	0.823
Logistic Regression kappa	0.834	0.723	0.805	0.709	0.715
Random forests OOB	0.092	0.098	0.107	0.111	0.126
Feature selection process	BO	RF	RF	RF	RF

The features for the above models, ordered according to their importance, are the following:

Best model: CRAE_LEON, CRVE_LEON, ArtAngSD, CRVE_Knud, CRAE_KNUD, Tort_ellipse2, arterySD, pressureV, BC_Mean_Art, artgamma, vparenta, Reparenta, MeanArt, wsspareanta,

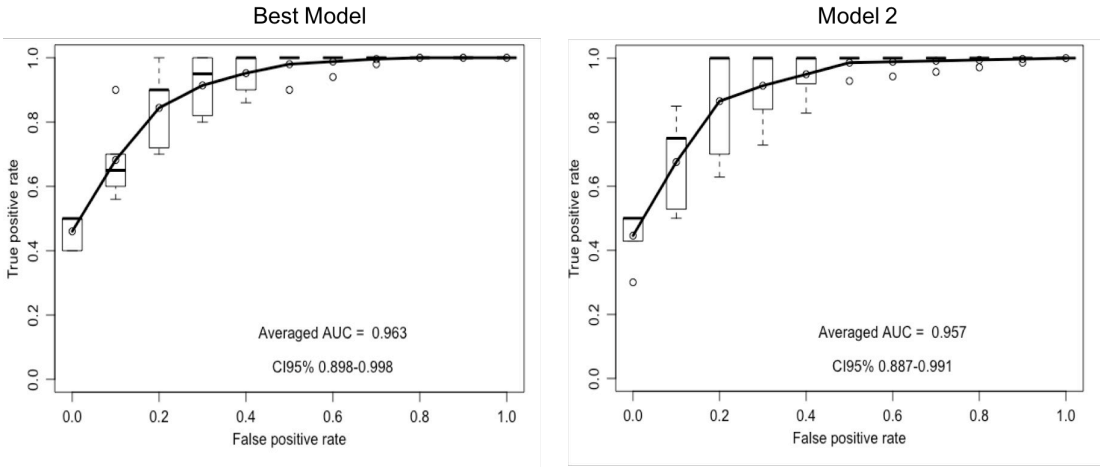

Figure 6.28: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 6; on the right similar information is given for the second best model.

lengthart, MedArt, wsschild1v, Angle.BC_Art, wsschild1a, Rechild2v, MedAngVein, vchild1a.

Model 2: CRVE_LEON, CRVE_KNUD, CRAE_LEON, ArtAngSD, Veingamma, Medvein, Rechild2, wssparenta.

Model 3: CRVE_LEON, ArtAngSD, MeanVeinAng, mean_tau.

Model 4: CRVE_LEON, BC_Mean_Art, CRAE_LEON, ArtAngleSD, Reparenta, VeinAnglSD, MeanArtAng, sd_phi, ellipse3theta.

Model 5: CRAE_LEON, artAngSD, CRVE_LEON, Angle.BC_Art, CRVE_KNUD, wsschild1v, PinQoutv, wsschild2v.

The figures that follow include the feature selection process according to boruta and regularised random forests, the ROC curves and the bootstrapped AUC.

Just as was indicated by the statistical analysis, it comes as no surprise that features like CRAE, CRVE, tortuosity, arterial and venular widths and angles, as well as a series of haemodynamics ones, can provide a very good predictive performance of the disease's progression. The performance is decent enough to suggest that such a classifier could be used in the near future for identifying and predicting the progression of a not yet progressed subject with diabetes. This is a very strong observation, which

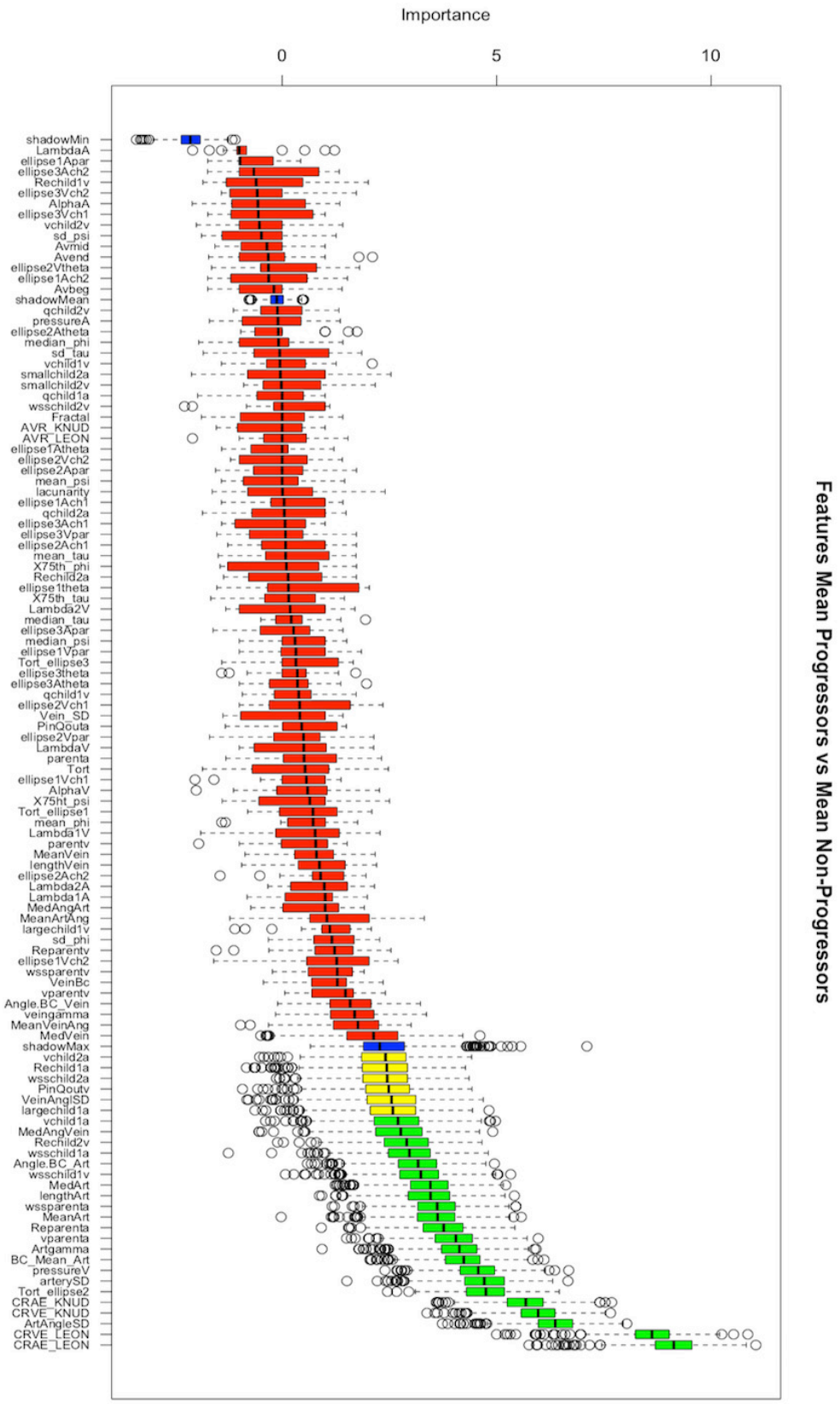

Figure 6.29: Feature selection process according to Boruta, which also gave the best performance. Green labelled features are the finally selected ones, whereas the red are the rejected ones. The plot refers to the combination 6.

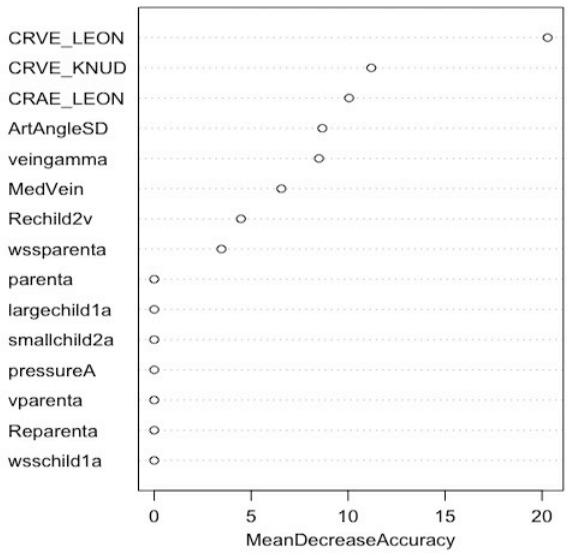

Figure 6.30: Feature ranking of the regularised random forests classifier for model 2 of combination 6.

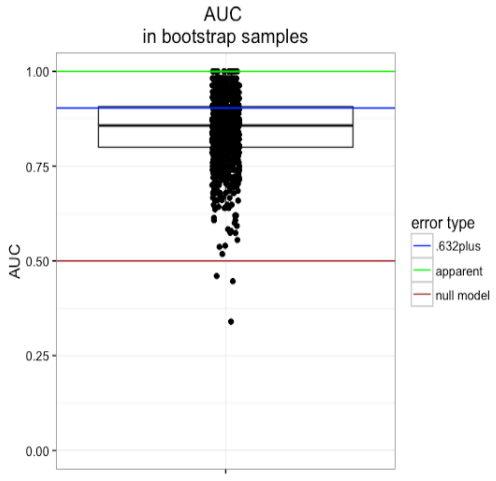

Figure 6.31: Bootstrapped area under the curve of the best model for combination 6. It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested with the same training set and the random model, respectively.

can be catalytic for managing to have in the future a clinical system that can assist with the assessment of the disease's progression, prior to arriving in an irreversible stage. It goes without saying that the amount of data needed to do so are much more than the ones that were available for these studies, but the results are promising enough to suggest the possibility of introducing such a classifier with definite clinical importance.

6.3.7 Year 3 Minus Year 2 Versus Onset of DR Minus Year 1 (Progressors' Group)

This specific combination is adopted mostly for reference purposes in order to figure out whether there is any difference to the extent of the vascular alterations between the transitional phase from diabetic eye to DR and the transition from three years pre-DR to two years pre-DR. The best performance was achieved with the the feature subset selected by the boruta selection process and the random forests classifier. The final AUC is 0.842 (Table 6.22).

Table 6.22: Summary of the classification performance for the differences Y3-Y2 vs DR-Y1 (Combination 7)

Summary	Best model	Model 2	Model 3	Model 4	Model 5
Random Forests	**0.842**	0.834	0.831	0.828	0.821
AUC (CI95%)	**(0.698-0.931)**	(0.678-0.921)	(0.667-0.921)	(0.687-0.931)	(0.658-0.921)
Logistic Regression	0.798	0.789	0.783	0.718	0.743
AUC (CI95%)	(0.612-0.904)	(0.598-0.887)	(0.586-0.889)	(0.545-0.853)	(0.578-0.868)
Random Forests kappa	0.688	0.675	0.665	0.671	0.658
Logistic Regression kappa	0.589	0.592	0.579	0.423	0.465
Random forests OOB	0.231	0.256	0.248	0.267	0.272
Feature selection process	BO	BO	RF	RF	RF

The features for the above models, ordered according to their importance, are the following:

Best model: Wsschild2v, arterySd, artAngleSD.

Model 2: ArtAngSd, wsschild2V, arterySD, lacunarity.

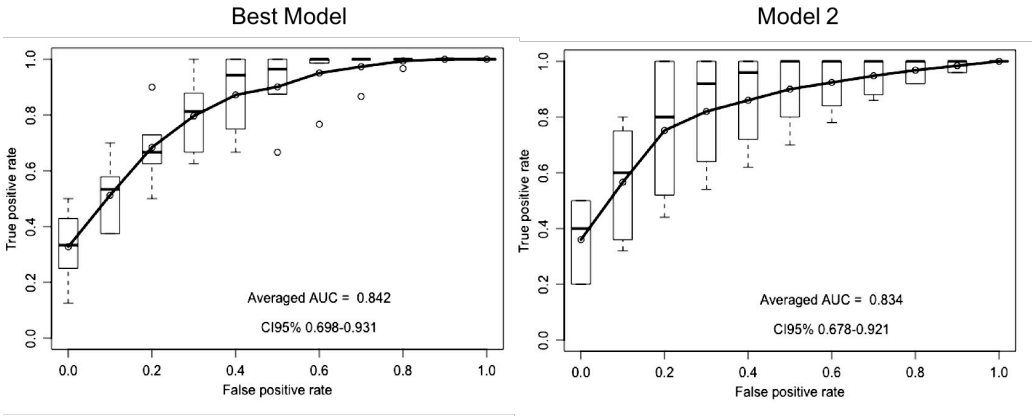

Figure 6.32: On the left: ROC curve and AUC of the best model with the confidence intervals for combination 7; on the right similar information is given for the second best model.

Model 3: Wsschild2v, arterySD, artAngleSD, lacunarity, veinAnglSD, smallchild2a, vein_SD, pressureV, ellipse2Vch1.

Model 4: ArterySD, wsschild2v, VeinBc, VeinAngSD, ellipse2Vch1, ArtAngleSD, ellipse1Vpar, Vein_SD.

Model 5: Wsschild2v, arterySD, ArtAngleSD, pressureA, lacunarity, VeinAnglSD, VeinBc, Tort_ellipse1.

The figures that come next include the feature selection process of boruta (best model) and regularised random forests (third model) and also the ROC curves and the bootstrapped AUC.

The performance of this combination suggests something that was presumed; the direction of the changes based on the values of the two classes show that the alterations during the transitional year, i.e. from the last year of diabetic eye to DR, are more pronounced than the three years- to two years- pre-DR. The fact that all five models of this combination are close in performance suggests that actually a series of haemodynamic and geometric features of both the venular and arterial network are supporting the above statement.

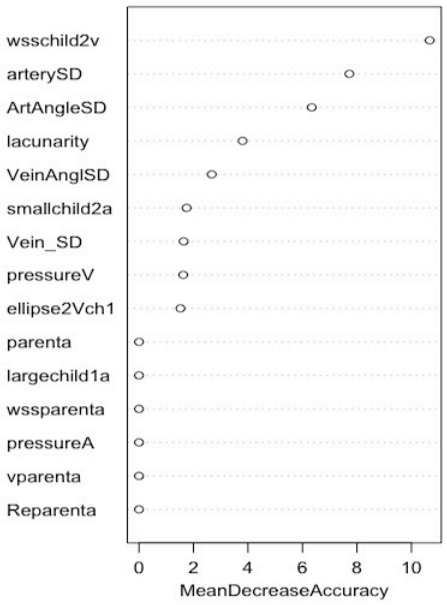

Figure 6.33: Feature ranking of the regularised random forests classifier for the model 3 of combination 7. The importance of each feature is expressed in terms of the mean decrease accuracy, that is how the classifier is affected by removing a specific feature.

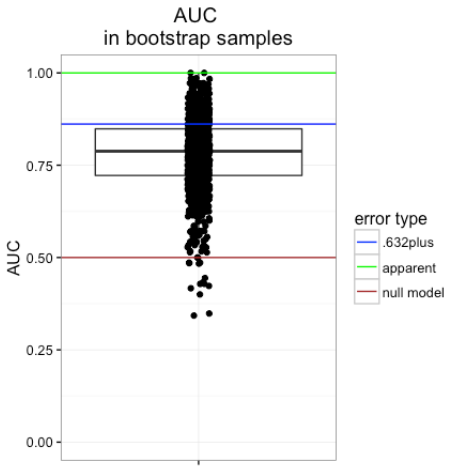

Figure 6.34: Bootstrapped area under the curve of the best model for combination 7.It shows all the 1000 iterations of the bootstrap, including the box plot, and the lines of the apparent and null error models, i.e. the model tested with the same training set and the random model, respectively.

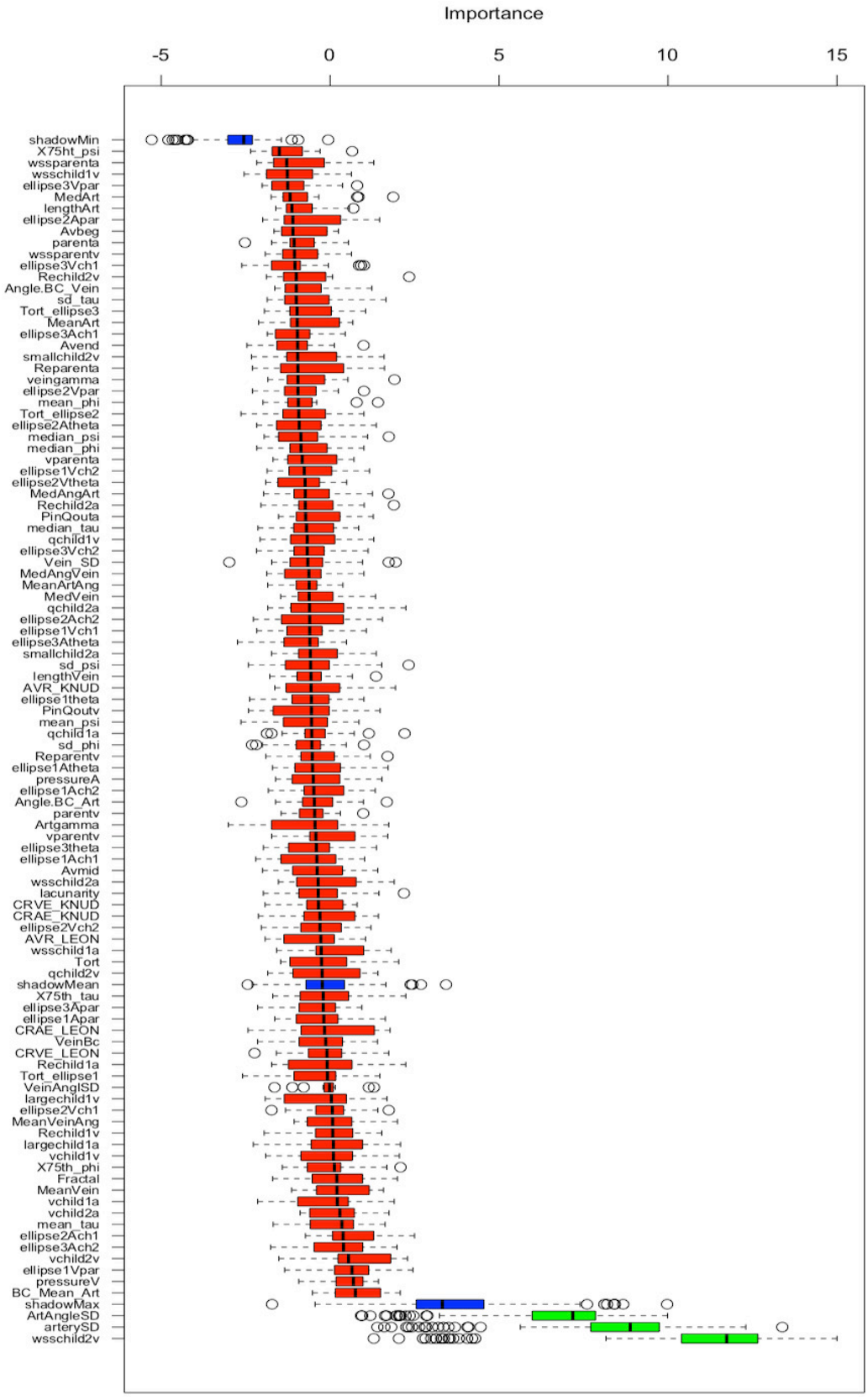

Figure 6.35: Feature selection process according to Boruta of the best model for combination 7. Green labelled features are the finally selected ones, whereas the red are the rejected ones. The plot refers to the combination 7.

6.3.8 Mean Diabetics (Non-Progressors' Group) Versus Mean Diabetics (Progressors' Group) Versus Onset of DR (One Versus All Method))

This final combination presents a pilot classification model, in order to figure out whether a complete classifier that can differentiate the whole progression of the disease and the stages could be feasible. Such a model can have a direct application in identifying the stage that a subject is and/or the probability that it belongs to a specific stage. Just like in combination 5 and 6, it is paramount to be able to predict the stage of the disease and assist the clinicians with adopting a treatment plan and/or conducting further tests.

As mentioned in the previous chapter, in this specific comparison, the best feature subsets from the combinations 4, 5, 6 were used to train the random forests classifier. The best result was obtained by the feature subset of combination 6 (AUC = 0.821). In figure 6.36 the ROC curves of all the combinations can be seen, as well as the individual ROC curves for each binary classifier.

The performance is decent enough to suggest that it is possible to identify these important stages of the disease. It is also promising that such a performance was achieved with limited observations and only with information of the retinal geometry and haemodynamics. The application and use of such a classifier in a clinical study could be possible and definitely useful.

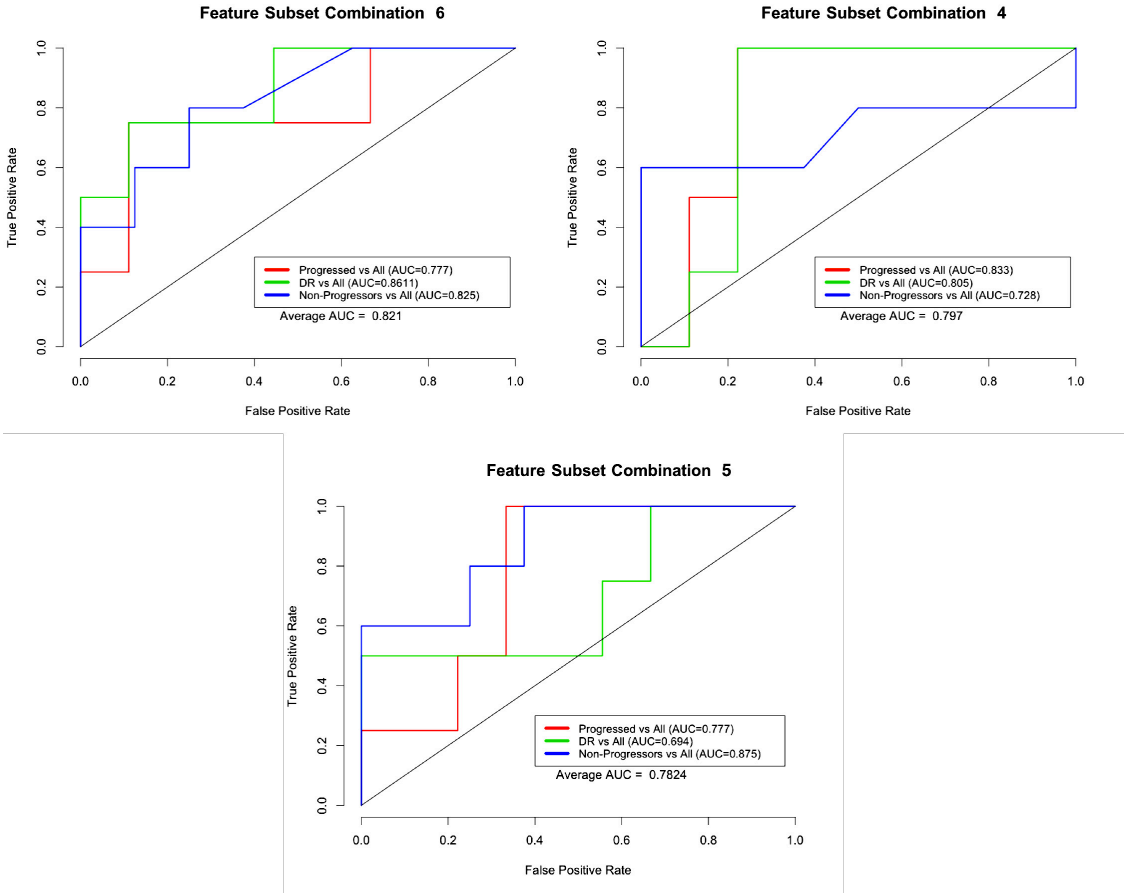

Figure 6.36: ROC curves for each of the three combinations used in building the one vs all classification models. For each feature subset, the plot includes three ROC curves that correspond to the three different possible classifications: progressors versus the rest, non-progressors versus the rest and DR versus the rest. The upper left plot, coming from combination 6, presents the best results, followed by the one on the upper right, which belongs to combination 4, and finishing with the one down, coming from combination 5.

6.3.9 Feature Ranking

The top 22 features as they appeared in all of the feature selection processes (210 in total) across all the different combinations can be seen in Table 6.23. This table provides us with very important information, showing that the retinal vascular geometry, and mainly the vessel widths, vessel angles and tortuosity are strong biomarkers, with extensive discrimination potential, across all the combinations that were investigated. This is even further highlighted by the fact that they appear in the models of all the seven combinations. The 20 out of the 22 features are coming directly from the retinal

Table 6.23: Ranking of the 22 most selected features

Features	Appearances
ArtAngleSD	78
CRAE_LEON	72
CRVE_LEON	70
Angle.BC_art	64
PressureV	62
CRVE_KNUD	62
Tort_ellipse3	62
ArterySD	58
CRAE_KNUD	54
VeinAnglSD	54
lengthVein	54
BC_Mean_Art	48
Artgamma	40
Ellipse1Vch2	40
Fractal	40
Reparenta	40
wsschild1v	38
Angle.BC_Vein	38
Tort_ellipse1	34
MedVein	34
Ellipse3Vtheta	34
Sd_phi	34

vascular geometry, with the remaining two representing the haemodynamic features. Pressure of the veins appears as the fifth best feature, contributing to the higher performance of the classification models. As was observed both within the statistical analysis and the classifications, the features of Table 6.23 shall be taken very seriously into account when building a diagnostic system, because they are among the first ones to be affected from the proliferation of the disease (e.g. oxidative stress, local metabolic stimuli and pericyte loss).

6.3.10 Summary

For the first time in literature, multiple classification models were trained and tested, in order to investigate and achieve two things. Firstly to evaluate the discriminative power of the investigated biomarkers and secondly to figure out whether a classification system based exclusively on information of the retinal geometry and haemodynamics can have a performance that could warrant a possible clinical application. It is shown in an explicit way, based on robust feature selection processes and by using

repeated cross validation and bootstraping to validate the performance of every model, that many candidate biomarkers (Table 6.23) can discriminate between different stages of the progression of the disease, tested in various combinations. The achieved performance in all eight combinations ranged from 0.745 to 0.968 in terms of the AUC.

Finally the results suggest that a model built solely with information of the retinal geometry and heamodynamics can possibly provide reliable classifiers that could be tested and incorporated in a real clinical application. These features accommodate the purpose of this thesis, which is the early detection and diagnosis of DR. Therefore these novel findings have a clear clinical importance, which is originally reported in this thesis.

6.4 General Conclusion and Discussion

This chapter presents the main results of the evaluation of all the candidate biomarkers. The first part is devoted to the complete statistical analysis for every category that is of our interest. Taking into account that the main goal of this thesis is to identify and early detect the progression to DR, many different categories were evaluated that cover every possible analysis, accomplishing this goal. A large enough progressors' diabetic group was included alongside a non-progressors' diabetic one, in order to emphasise and show that in order to make proper inferences, a carefully designed experiment is required. Significant results were observed for a number of biomarkers, both geometric and haemodynamic, including venular and arterial widths, tortuosity, fractal dimension, CRVE, CRAE, blood flow, venular wall shear stress, venular velocity and others (Tables 6.5 and 6.14). Changing the novel design to a more conventional one, masked all these differences, affecting massively the analysis. This strongly suggests that in order to quantify the vascular changes, a deep investigation and a follow-up of the same segments is needed.

Another important observation is that there are extensive changes between the non-progressors' diabetic group and the progressors' one, e.g. in the CRVE, CRAE, the venular widths and tortuosity, the arterial and venular blood flow, among others. This is one of the key finding, suggesting that the changes start earlier, much before the DR, and continue until the years before the onset of DR, adopting at this stage the characteristics of a progressors' group. Following that, more differences are under way, until the first lesions finally appear in the retina. The findings for the CRAE and CRVE confirm the ones in (Klein *et al.*, 2004) and also (Islam *et al.*, 2009).

The second part of this chapter took the analysis a step further, evaluating all those biomarkers in practise. Many different models were created, which included 8 different combinations, and evaluating 210 different features subsets, as came out of the feature selection processes. 420 different classification models were created and compared in total, in order to find the optimum one for the 7 combinations. The key and rigorous finding out of this process is that it was unequivocally shown that a reliable and high performance classifier can indeed be used to differentiate between different stages of the disease's progression, trained only with geometric and haemodynamic features.

Therefore, three major and robust contributions were made. The first one is the evaluation and estimation of haemodynamic features for studying the progression of the disease, in preselected and connected vascular trees. This led to the subsequent establishment of important biomarkers of progression to DR, both heamodynamic and geometric ones, within a non-progressors' diabetic group, a progressors' diabetic group and a DR group. The second is that, for the first time, high performance classifiers are created and proposed, along with the corresponding feature subsets, solely relying on the above biomarkers, for predicting different stages of the disease's progression, not only between subjects with diabetes and DR, but also between non-progressors' diabetics and progressors' diabetics.

Lastly, the results in chapter 5 (the third major contribution), where the proposed and improved method for estimating the CRVE and CRAE is described, were confirmed on a larger scale in this chapter as well. The CRAE and CRVE features, estimated according to the proposed method, ranked second and third respectively, in the entire feature selection process across all the combinations , as can be seen in table 6.29. This explicitly points out what was previously suggested, that any improvement in the estimation of CRVE, CRAE and AVR can have a big impact in more effectively and robustly studying the progression of the disease.

Chapter 7

Conclusion

This thesis presents a novel, comprehensive and complete study of multiple geometric and haemodynamic biomarkers that have not been evaluated before. In addition it presents a new, alternative and more accurate method for summarising the calibre of the parent vessel trunk, with an application to the estimation of the CRVE, CRAE and AVR. A summary of the main contributions and achievements is given in this chapter. The previous chapters provided detailed and comprehensive information about the background of DR, previous studies, utilised and implemented methods and techniques, as well as the full results of the studies and the validation process.

7.1 Summary of Work

This thesis presents three novelties, which all define some important contributions for the early detection and screening of diabetic retinopathy. Firstly, a complete framework for the more efficient and meaningful analysis of the retinal biomarkers is presented, which was shown to produce more accurate and representative results. This was also validated with the different classification models, where indeed most of the significant or close to the significance level (0.05) features that were identified during the matched-segments progressors' group analysis, were also part of the feature subsets

in the classification process. Until now, the studies in literature were focusing on recruiting a healthy or diabetic cohort and comparing it against an advanced retinopathy one, e.g. in (Kifley *et al.*, 2007; Daxer, 1993; Yau *et al.*, 2010; Nguyen *et al.*, 2008a). However, this study concluded that the alterations are happening much earlier, during the last years of the diabetic eye, still on a stage that the interventions and treatment plans from the clinicians can possibly be adjusted for the optimum management and control of the disease.

Secondly, for the first time, a series of haemodynamic features were estimated and evaluated as candidate biomarkers of the progression of the disease, which were shown to make a difference in the statistical analysis and the classification process indeed. Despite the limitations of the boundary conditions that were adopted for their estimation, their importance was highlighted, suggesting that more appropriate, advanced and representative boundary conditions would possible make even greater difference. In addition, the performance of the classification models provided us with some strong findings with possible clinical importance. These strongly suggest that accurately and reliably predicting the progression of the disease, both from the non-progressors' to the progressors' stage and also from the progressors' stage to the onset of DR, is something that can be achieved with the proper adoption of various geometric and haemodynamic features, subject to the inclusion of large representative datasets.

Thirdly, the new method for summarising the calibre of the parent vessel trunk and the subsequent improved estimation of the CRVE, CRAE and AVR, will definitely make a difference in studying the progression of any retinal disease, not only DR. This was clearly and succinctly shown, both statistically and in practise inside the classification models. This led the estimations of the CRAE and CRVE according to the proposed method, not only to rank higher than the similar estimations that are based on the other method in literature (Knudtson *et al.*, 2003), but also to rank second and third among the entire investigated features across all the combinations (Table 6.23).

7.2 Future Work

The work presented in this thesis could be continued and extended in a variety of ways. Firstly and most importantly, given the complex, diverse and multiple effects that diabetes has to the whole body, and also the millions of people that face its consequences worldwide, additional and more representative data are needed. One of the main problems that I had to deal with in all of these studies, was to find the necessary and clean data needed for conducting all of the experiments. However, after showing in practise how all these biomarkers of progression, e.g. CRVE, CRAE, blood flow rate, vessel angles and widths, fractal dimension etc., are capable of differentiating between different stages of diabetes and the onset of DR, additional data for building more efficient classifiers are definitely needed. Moreover, despite not being the scope of this thesis, it is definitely paramount to automate some of the processes involved, such as artery/vein identification, in order to be able to apply all these analyses in a larger scale. Stressing the last thing even further, this is a crucial and sensitive step, because errors in this part pollute the available data, creating outliers that are difficult to identify, because both types of vessels can take very small and very large values. Therefore setting a threshold is not an option. That would be a fatal error that it was avoided in this thesis, by carefully, manually labelling and selecting the vessels, supporting one of the main objectives of presenting a robust, accurate and complete study. Last but not least, the range of the period that the data were obtained from could also be extended, in order to investigate the within subjects progression (progressors' group) what happens even before the three years from the onset of DR. According to what was observed in this thesis, the changes possibly start even earlier, which means that interventions and treatment plans could be even more effective.

7.3 Epilogue

The motivations and objectives of this thesis were given in chapter 1. Detecting diabetic retinopathy early, when still no lesions have appeared, can be a life-changing feat. In this thesis it was clearly shown that, subject to extending the studies to include more subjects, in order to get a more representative sample of the alterations, DR can be identified early. Clinicians could possibly be driven by that, in order to monitor the patients more closely and even try to decelerate the progression of DR.

This thesis focused on presenting a framework, evaluating candidate biomarkers and improving the estimations of some of them, in order to show that the alterations in the retinal vasculature, as a consequence of diabetes and DR, occur prior to the onset of DR, which are distinct enough to be able to detect different stages of the progression. Finally, the thesis introduces some novel fundamental ideas and a framework for being adopted by future studies, in order to lead to even more accurate and representative analyses, inferences and verdicts.

Appendix A

Results of the Non-Significant Features

Table A.1: Analysis of the non-significant main features for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
BetaA	35.4/	20.5/	0.15/	0.45	0.657	0.21	0.015/	1.25(0.17),1.32(0.15)
	36.4	21.7	0.154			nested:0.356	0.031	1.27(0.16),1.31(0.17)
Lambda1A	623.5/	612.8/	0.28/	0.372	0.604	0.137	0.0024/	0.914(0.057),0.923(0.052)
	624.8	611.2	0.29			nested:0.365	0.0033	0.908(0.057),0.926(0.052)
ThetaA	2578/	2603/	0.536/	0.487	0.648	0.13	223/	154.1(41),153.6(41.2)
	2575	2589	0.538			nested:0.43	736.9	154.9(43.1),148.2(38.4)
Adj.GammaA	689.6/	675.4/	0.137/	0.463	0.685	0.203	0.001/	0.812(0.08),0.825(0.07)
	689.2	664.3	0.137			nested:0.396	0.002	0.846(0.07),0.817(0.08)
Width_child2A	170.4/	195.3/	0.29/	0.667	0.763	0.529	0.11/	3.16(0.43),3.19(0.47)
	168.2	182.4	0.29			nested:0.171	0.062	3.197(0.46),3.125(0.43)
Angle.BC_A	2531/	2556/	0.265/	0.754	0.708	0.161	262.5/	127.4(41.7),121.1(40.5)
	2529	2543	0.271			nested:0.481	582.5	126.9(38.7), 120.2(35.5)
Lambda1V	1677/	1657/	0.35/	0.335	0.574	0.183	0.001/	0.906(0.062),0.914(0.067)
	1672	1640	0.35			nested:0.295	0.003	0.902(0.072),0.909(0.065)
ThetaV	182.4/	214.2/	0.71/	0.353	0.608	0.248	0.025/	137.1(32.9),136.3(28.8)
	177.8	196	0.71			nested:0.264	0.05	137.1(34.5),134.7(33)
Adj.GammaV	1929/	1909/	0.06/	0.351	0.611	0.171	0.0008/	0.763(0.067),0.765(0.071)
	1927	1897	0.06			nested:0.345	0.0024	0.768(0.071),0.761(0.073)
Width_child2V	734/	766/	0.125/	0.521	0.708	0.366	0.104/	3.198(0.531),3.219(0.513)
	734	765	0.127			nested:0.275	0.101	3.277(0.574),3.22(0.475)
Lacunarity	474.3/	465.3/	0.688/	0.79	0.832	0.787	0.0045/	0.16(0.082),0.17(0.073)
	469.8	451.8	0.681				0.0012	0.171(0.059),0.174(0.078)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.2: Analysis of the non-significant features of the different areas for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
AVmid_AVbeg	82.1/	75.5/	0.57/	0.514	0.612	0.51	0.013/	0.032(0.166),0.0013(0.176)
	78.6	64.7	0.58				0.013	0.038(0.127),-0.0016(0.179)
AVend_AVbeg	85.3/	78.2/	0.375/	0.627	0.713	0.628	0.019/	0.0452(0.176),-0.004(0.17)
	82.56	68.2	0.383				0.011	-0.0071(0.155),0.0062(0.193)
AVend_AVmid	67.9/	60.1/	0.413/	0.521	0.624	0.52	0.0169/	0.0133(0.167),-0.006(0.165)
	64.8	50.5	0.433				0.0155	-0.046(0.154),0.008(0.224)
ellipse1ch1V	259.4/	273.7/	0.283/	0.334	0.467	0.339	0.53/	4.33(1.12),3.9(1.31)
	257.2	264.4	0.302				1.02	4.47(1.3),4.35(1.22)
ellipse1ch2V	249.5/	263.8/	0.241/	0.387	0.519	0.394	0.567/	4.87(1.22),4.35(1.15)
	247.7	254.9	0.253				0.871	4.72(1.04),4.43(1.36)
ellipse1ThetaV	616.3/	630.6/	0.355/	0.298	0.426	0.299	6.263/	81.29(12.78),84.89(11.93)
	613.6	620.7	0.375				9.577	84.28(9.97),86.56(10.87)
ellipse1parA	285.9/	300.2/	0.678/	0.698	0.762	0.692	2.393/	6.71(1.81),6.65(1.85)
	281.5	288.6	0.693				1.065	7(1.89),6.68(1.87)
ellipse1ch1A	293.8/	308.1/	0.642/	0.676	0.745	0.67	2.445/	5.57(2.07),5.54(1.78)
	289.5	289.5	0.658				1.2	5.91(1.89),5.8(1.87)
ellipse1ch2A	292.2/	306.5/	0.35/	0.593	0.686	0.594	1.852/	5.44(1.85),5.10(1.51)
	289.5	296.6	0.37				1.264	5.71(1.9),5.26(1.75)
ellipse1ThetaA	634.1/	648.4/	0.93/	0.3	0.39	0.285	46.2/	85.99(12.64),85.77(15.22)
	628.5	635.6	0.95				115.8	87.39(12.31),86.71(10.22)
ellipse2parV	362.1/	376.1/	0.352/	0.704	0.774	0.705	6.51/	7.74(3.19),8.17(2.93)
	359.4	366.6	0.361				2.72	8.68(3),8.12(3)
ellipse2ch2V	275.9/	290.2/	0.231/	0.944	0.96	0.94	10.07/	6.64(3.2),6.42(3.31)
	274.4	281.4	0.249				0.57	6.87(3.32),6.8(3.19)
ellipse2ThetaV	601.7/	616/	0.798/	0.755	0.807	0.748	154.1/	83.31(14.38),85.17(13.93)
	596.7	603.9	0.81				51.67	83.83(14.5),85.04(14.54)
ellipse2parA	281.5/	295.8/	0.698/	0.721	0.781	0.715	2.47/	7.19(2.23),6.91(1.6)
	277	284.1	0.71				0.982	7.17(1.84),6.94(1.69)
ellipse2ch1A	258.1/	272.4/	0.53/	0.697	0.765	0.694	1.7/	5.62(1.55),5.55(1.32)
	254.3	261.5	0.542				0.75	5.93(1.74),5.7(1.61)
ellipse2ThetaA	600.2/	614.5/	0.326/	0.528	0.635	0.531	71.07/	86.43(12.21),85.24(8.95)
	597.7	604.9	0.335				62.72	83.96(11.41),81.99(13.24)
ellipse3parV	406.1/	420.3/	0.846/	0.533	0.621	0.521	6.085/	15.31(3.36),14.95(2.37)
	400.9	408.1	0.854				5.583	14.72(4.29),15.21(3.34)
ellipse3ch1V	321.6/	335.9/	0.599/	0.8	0.845	0.795	5.73/	11.89(2.47),11.42(2.73)
	317.5	324.6	0.616				1.462	11.47(2.8), 11.62(2.69)
ellipse3ch2V	351.9/	366.2/	0.291/	0.62	0.71	0.622	4.284/	11.33(2.37),10.51(2.5)
	349.6	356.8	0.3				2.599	11.02(2.31),10.53(3.20)
ellipse3ThetaV	581.9/	596.2/	0.115/	0.393	0.539	0.41	38.06/	78.4(10.61),79.82(8.1)
	581.8	594.4	0.128				54.91	81.94(8.22),83.7(11.21)
ellipse3parA	351.7/	366.1/	0.568/	0.503	0.603	0.498	2.864/	11.91(3.37),11.3(2.03)
	347.8	354.9	0.585				2.885	11.24(1.64),11.51(2.18)
ellipse3ch1A	371.1/	385.4/	0.176/	0.53	0.645	0.539	4.169/	10.16(4.08),8.97(1.67)
	370	377.2	0.193				3.551	9.13(2.26),9.21(2.49)
ellipse3ch2A	325.1/	339.3/	0.1/	0.583	0.694	0.598	2.831/	9.94(2.13),8.91(2.31)
	325/3	332.4	0.11				1.9	9.38(2.05),9.11(2.18)
ellipse3ThetaA	571.97/	586.27/	0.58/	0.472	0.575	0.466	40.51/	86.07(11.1),85.89(8.74)
	567.95	575.1	0.597				46.41	88.46(8.79),86.2(8.4)
AVR_LEON	294.5/	285.6/	0.31/	0.445	0.549	0.44	0.0038/	0.691(0.081),0.712(0.084)
	289.9	272.2	0.34				0.0049	0.718(0.091),0.723(0.092)
AVR_KNUD	302.1/	294.1/	0.1/	0.443	0.562	0.448	0.0036/	0.688(0.081),0.704(0.082)
	301.1	283.9	0.112				0.0044	0.702(0.094),0.723(0.093)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.3: Analysis of the non-significant tortuosity features for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
mean_phi	2494/ 2489	2481/ 2464	0.665/ 0.668	0.538	0.631	0.537	3.7x10^{-4}/ 3.2x10^{-4}	0.222(0.03),0.223(0.024) 0.223(0.026),0.225(0.028)
median_phi	2222/ 2218	2209/ 2193	0.488/ 0.491	0.551	0.644	0.553	6.6x10^{-4}/ 5.3x10^{-4}	0.139(0.032),0.138(0.033) 0.14(0.037),0.143(0.035)
75th_phi	1739/ 1734	1726/ 1708	0.823/ 0.812	0.491	0.589	0.489	1.3x10^{-3}/ 1.4x10^{-3}	0.347(0.052),0.35(0.04) 0.346(0.052),0.35(0.056)
sd_phi	2758/ 2754	2746/ 2728	0.72/ 0.719	0.339	0.447	0.337	1.1x10^{-4}/ 2.2x10^{-4}	0.24(0.019),0.242(0.017) 0.24(0.018),0.241(0.018)
mean_tau	3550/ 3546	3537/ 3520	0.6544/ 0.656	0.506	0.603	0.505	4.5x10^{-5}/ 4.4x10^{-5}	(3.71(0.98),3.76(0.85) 3.72(0.99),3.8(0.93))x10^{-2}
median_tau	4687/ 4684	4674/ 4658	0.371/ 0.372	0.652	0.728	0.653	8.34x10^{-6}/ 4.43x10^{-6}	(11.1(3.6),11(3.1) 10.9(3.7),11.3(3.7))x10^{-3}
75th_tau	3641/ 3637	3628/ 3611	0.593/ 0.596	0.554	0.645	0.553	4.43x10^{-5}/ 3.57x10^{-5}	(35.4(9),35.2(7.8) 35.1(9.2),36(9.4))x10^{-3}
mean_psi	3633/ 3629	3621/ 3603	0.575/ 0.573	0.539	0.632	0.538	4.29x10^{-5}/ 3.7x10^{-5}	(4.47(0.92),4.49(0.86) 4.45(0.89),4.55(0.88))x10^{-2}
median_psi	4276/ 4272	4263/ 4246	0.493/ 0.495	0.708	0.772	0.707	2.23x10^{-5}/ 9.2x10^{-6}	(16.9(5.7),16.7(5.2) 16.9(5.8),17.3(5.4))x10^{-3}
75th_psi	3254/ 3250	3241/ 3224	0.716/ 0.718	0.606	0.687	0.604	1.1x10^{-4}/ 7.16x10^{-5}	(51.4(13.6),51.9(12.7) 52.1(13.3),52.6(13.9))x10^{-3}
sd_psi	3104/ 3098	3091/ 3073	0.389/ 0.388	0.275	0.383	0.273	4.6x10^{-5}/ 12.3x10^{-5}	(7.25(1.3),7.25(1.2) 7.2(1.3),7.3(1.3))x10^{-2}
Tort.blend	989/ 986	977/ 961	0.41/ 0.42	0.576	0.664	0.575	7.4x10^{-3}/ 5.52x10^{-3}	1.009(0.114),1.016(0.103) 1.012(0.114),1.024(0.122)
Tort_ellipse1	146/ 141	138/ 125	0.86/ 0.87	0.131	0.108	0.112	0.0023/ 0.0112	0.072(0.087),0.081(0.098) 0.089(0.098),0.082(0.076)
Tort_ellipse2	74.1/ 73.9	66.5/ 58.3	0.118/ 0.129	0.191	0.336	0.203	0.0055/ 0.0216	0.061(0.129),0.145(0.249), 0.091(0.149), 0.055(0.086)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.4: Analysis of the non-significant haemodynamic features for category 1

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
wssparentA	793.5/ 787.7	795.2/ 788.4	0.865/ 0.857	0.859	0.897	0.855	923/ 156.3	54.04(30.19),56.24(29.8) 51.96(25.67),52.81(24.56)
pressureA	527.1/ 524.1	542.1/ 531.6	0.41/ 0.396	0.158	0.272	0.156	3.03/ 16.3	2.72(1.05),3.61(1.65) 2.91(1.18),3.36(1.15)
vparentaA	492.8/ 487.6	507.8/ 495.1	0.858/ 0.852	0.888	0.918	0.885	40.05/ 5.20	16.28(7.85),16.89(7.1) 16.11(6.28),16.33(6.12)
ReparentA	855.4/ 850.7	870.5/ 858.2	0.755/ 0.745	0.91	0.934	0.907	51.78/ 16.58	259.94(61.13),265.82(57.58) 260.2(53.03),262.62(50.17)
wsschild1A	703.3/ 700.4	718.3/ 707.9	0.374/ 0.392	0.663	0.752	0.662	12.18/ 8.7	25.83(11.34),29.47(14.56) 25.54(13.05),28.75(13.14)
vchild1A	434.5/ 431.33	449.5/ 438.8	0.418/ 0.436	0.689	0.771	0.687	8.173/ 3.714	8.56(3.26),9.12(3.91) 8.17(3.22),9.02(3.57)
Rechild1A	862.1/ 860.7	877.1/ 868.2	0.21/ 0.226	0.634	0.735	0.638	28.29/ 21.28	147.8(32.9),147.1(38.7) 135.7(37.8),147.9(33.8)
wsschild2A	681/ 677.3	876.9/ 872.2	0.521/ 0.538	0.764	0.828	0.761	167.8/ 52.54	32.3(13.5),35.08(18.41) 31.58(13.15),32.83(15)
vchild2A	369.8/ 367.6	384.8/ 375.1	0.282/ 0.299	0.781	0.845	0.783	5.802/ 1.604	8.65(2.63),9.38(3.19) 8.94(2.52),8.89(2.71)
PinQoutA	610/ 606	625.9/ 614.1	0.646/ 0.66	0.788	0.845	0.785	9.234/ 4.828	25.07(9.92),24.37(11.17) 23.66(11.87),24.67(9.46)
pressureV	369.9/ 368.1	384.1/ 375.2	0.241/ 0.263	0.562	0.571	0.569	2.221/ 1.931	1.79(1.41),2.36(2.02) 2.39(1.55),1.22(1.13)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.5: Analysis of the non-significant main features for category 2

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
AlphaA	505.3/ 501.6	486.7/ 468.9	0.498/ 0.499	0.388	0.324	0.144 nested:0.091	0.0052/ 0.0278	0.801(0.136),0.82(0.145) 0.808(0.125),0.807(0.127)
BetaA	192.3/ 187.2	226.4/ 206.4	0.846/ 0.815	0.567	0.247	0.067 nested:0.091	0.005/ 0.062	1.302(0.193),1.311(0.202) 1.323(0.197),1.303(0.209)
LambdaA	1509/ 1505	1490/ 1472	0.461/ 0.461	0.391	0.341	0.152 nested:0.096	0.0018/ 0.0093	0.858(0.103),0.899(0.065) 0.883(0.042),0.881(0.054)
Lambda1A	1867/ 1862	1848/ 1828	0.71/ 0.87	0.645	0.241	0.054 nested:0.098	0.0004/ 0.0064	0.922(0.068),0.92(0.071) 0.926(0.062),0.921(0.065)
Lambda2A	1570/ 1566	1551/ 1532	0.61/ 0.6	0.419	0.319	0.132 nested:0.094	0.0014/ 0.0085	0.813(0.079),0.822(0.091) 0.822(0.091),0.815(0.093)
ThetaA	1156/ 1155	1189/ 1175	0.108/ 0.109	0.687	0.369	0.081 nested:0.172	0.018/ 0.167	145.6(45.1),152.4(49.7) 149.5(48.5),153.9(46.1)
Adj.GammaA	2070/ 2065	2051/ 2031	0.865/ 0.866	0.574	0.246	0.067 nested:0.091	0.0004/ 0.0051	0.844(0.079),0.841(0.079) 0.85(0.074), 0.828(0.079)
Width_allA	1384/ 1383	1417/ 1402	0.159/ 0.161	0.236	0.609	0.399 nested:0.12	0.16/ 0.192	3.422(0.619),3.534(0.643) 3.496(0.652),3.486(0.625)
Width_parentA	1948/ 1942	1985/ 1967	0.474/ 0.475	0.284	0.559	0.334 nested:0.13	0.231/ 0.369	3.727(0.821),3.841(0.839) 3.791(0.848),3.796(0.814)
Width_child1A	1582/ 1580	1621/ 1617	0.236/ 0.237	0.229	0.628	0.413 nested:0.12	0.211/ 0.238	3.488(0.714),3.583(0.726) 3.561(0.729),3.549(0.729)
Angle.BC_A	9216/ 9213	9250/ 9232	0.444/ 0.445	0.77	0.347	0.055 nested:0.176	90.02/ 1278.9	116.40(43.81),119.1(47.39) 117.19(45.5),121.42(46.33)
AlphaV	934.2/ 932.3	910.8/ 891.9	0.198/ 0.199	0.273	0.175	0.088 nested:0.033	0.004/ 0.037	0.713(0.209),0.728(0.198) 0.735(0.205),0.734(0.207)
BetaV	413.2/ 412.8	454.1/ 436.1	0.136/ 0.137	0.497	0.215	0.07 nested:0.07	0.005/ 0.061	1.201(0.208),1.217(0.257) 1.232(0.212),1.219(0.203)
LambdaV	3421/ 3419	3397/ 3378	0.188/ 0.189	0.248	0.181	0.096 nested:0.032	0.0015/ 0.0138	0.834(0.129),0.844(0.121) 0.848(0.125),0.847(0.127)
Lambda1V	5629/ 5623	5605/ 5583	0.758/ 0.759	0.591	0.15	0.037 nested:0.054	0.0002/ 0.005	0.915(0.067),0.916(0.068) 0.919(0.061),0.915(0.063)
Lambda2V	3750/ 3749	3725/ 3709	0.075/ 0.076	0.323	0.223	0.106 nested:0.052	0.0015/ 0.0119	0.757(0.102),0.768(0.092) 0.773(0.101),0.769(0.104)
ThetaV	2328/ 2327	2369/ 2351	0.129/ 0.13	0.108	0.104	0.067 nested:0.008	0.0101/ 0.1391	137.2(39.1),138.6(38.4) 137.1(38.5),133.6(38.6)
Adj.GammaV	5990/ 5989	5967/ 5948	0.179/ 0.18	0.521	0.21	0.064 nested:0.071	0.0004/ 0.0049	0.778(0.056),0.783(0.061) 0.787(0.063),0.773(0.059)
Width_allV	4698/ 4697	4738/ 4722	0.074/ 0.075	0.125	0.421	0.341 nested:0.048	0.175/ 0.312	3.587(0.71),3.636(0.685) 3.671(0.706),3.638(0.692)
Width_parentV	6310/ 6306	6351/ 6330	0.581/ 0.582	0.149	0.392	0.299 nested:0.051	0.275/ 0.595	4.009(0.956),4.042(0.924) 4.07(0.941), 4.047(0.943)
Width_child1V	5645/ 5642	5686/ 5665	0.336/ 0.337	0.122	0.375	0.293 nested:0.04	0.205/ 0.465	3.723(0.84),3.759(0.797) 3.794(0.825),3.759(0.836)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.6: Analysis of the non-significant main features for category 3

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/ V_E)[d]	Group Means(SD) (Y3, Y2, Y1, YO)
AlphaA	60.7/ 61.1	50.7/ 43.5	0.092/ 0.102	0.2	0.145	0.247 nested:0.07	0.0078/ 0.0215	0.740(0.163),0.805(0.171), 0.82(0.161),0.842(0.142)
BetaA	5.99/ 4.76	12.81/ 4.05	0.189/ 0.205	0.04	0.341	0.07 nested:0.09	0.005/ 0.061	1.259(0.23),1.35(0.234), 1.38(0.24),1.321(0.212)

LambdaA	169.8/ 168.1	158.2/ 151.2	0.088/ 0.097	0.277	0.457	0.264 nested:0.072	0.0025/ 0.0065	0.855(0.095),0.872(0.056), 0.891(0.069),0.895(0.044)
Lambda1A	202.6/ 197.9	192.5/ 180.4	0.138/ 0.139	0.346	0.334	0.069 nested:0.170	0.0004/ 0.0049	0.846(0.079),0.856(0.082), 0.863(0.081),0.839(0.077)
Lambda2A	193/ 190.1	180.1/ 175.4	0.231/ 0.236	0.188	0.431	0.116 nested:0.173	0.0008/ 0.005	0.806(0.083),0.819(0.083), 0.815(0.061),0.821(0.055)
ThetaA	69.39/ 67.55	86.97/ 67.55	0.244/ 0.259	0.234	0.4	0.173 nested:0.118	0.022/ 0.09	143.7(34.1),143.9(33.1), 144.9(33.2), 158.6(40.6)
Adj.GammaA	236.9/ 235.1	226.9/ 217.6	0.241/ 0.258	0.05	0.334	0.05 nested:0.206	0.0005/ 0.0033	0.813(0.055),0.824(0.056), 0.835(0.053),0.819(0.051)
Width_parentA	105.1/ 104.4	121.9/ 115.1	0.085/ 0.093	0.467	0.589	0.409 nested:0.116	0.1/ 0.116	4.074(0.505),3.85(0.523), 3.838(0.457),3.912(0.402)
Width_child1A	93.58/ 92.28	111.1/ 102.3	0.195/ 0.21	0.496	0.585	0.397 nested:0.144	0.087/ 0.101	3.834(0.413),3.669(0.454), 3.689(0.474),3.669(0.421)
Width_child2A	89/ 83.71	106.6/ 93.75	0.871/ 0.877	0.373	0.442	0.348 nested:0.031	0.062/ 0.11	3.277(0.465),3.254(0.382), 3.304(0.371),3.325(0.437)
Width_allA	69.23/ 66.48	86.81/ 76.52	0.355/ 0.372	0.505	0.58	0.447 nested:0.092	0.076/ 0.079	3.73(0.415),3.593(0.403), 3.611(0.387),3.635(0.364)
Angle.BC_A	875.6/ 872.5	893.2/ 882.5	0.411/ 0.424	0.211	0.26	0.212 nested:0.0002	185.7/ 689.9	112.42(34.42),112.34(37.64), 110.94(35.01),123.99(38.99)
AlphaV	96.49/ 92.91	83.44/ 70.08	0.489/ 0.498	0.464	0.596	0.398 nested:0.144	0.021/ 0.024	0.581(0.206),0.622(0.234), 0.631(0.218),0.613(0.218)
BetaV	43.05/ 41.93	29.99/ 19.1	0.181/ 0.188	0.251	0.433	0.04 nested:0.274	0.0021/ 0.0345	1.121(0.112),1.173(0.124), 1.164(0.142),1.147(0.116)
LambdaV	267.6/ 263.9	254.6/ 241.1	0.523/ 0.531	0.436	0.57	0.376 nested:0.133	0.0077/ 0.0101	0.751(0.132),0.775(0.145), 0.782(0.137),0.771(0.141)
Lambda1V	435.3/ 430.5	422.2/ 407.7	0.745/ 0.751	0.206	0.536	0.045 nested:0.426	0.0005/ 0.018	0.816(0.057),0.831(0.053), 0.825(0.076),0.822(0.054)
Lambda2V	344.9/ 343.8	331.9/ 320.9	0.186/ 0.194	0.397	0.419	0.399 nested:0.014	0.0051/ 0.0076	0.681(0.098),0.715(0.109), 0.718(0.114),0.705(0.109)
ThetaV	9.89/ 5.56	32.73/ 18.61	0.642/ 0.648	0.499	0.623	0.484 nested:0.048	0.044/ 0.043	133.1(31.9),135.7(31.6), 136.5(31.4),131.4(30.4)
Adj.GammaV	514.6/ 513.4	501.5/ 490.3	0.211/ 0.218	0.195	0.435	0.06 nested:0.217	0.0006/ 0.0043	0.758(0.056),0.765(0.043), 0.767(0.056),0.761(0.054)
Width_parentV	394.2/ 391.2	417.8/ 404.3	0.51/ 0.518	0.58	0.611	0.55 nested:0.057	0.4245/ 0.3031	4.879(0.765),4.722(0.791), 4.844(0.853),4.818(0.827)
Width_child1V	416.9/ 411.6	439.7/ 411.6	0.854/ 0.858	0.41	0.542	0.334 nested:0.148	0.223/ 0.346	4.456(0.747),4.379(0.733), 4.467(0.861),4.435(0.818)
Width_child2V	194.4/ 191.6	217.2/ 204.6	0.36/ 0.368	0.559	0.534	0.56 nested:0.015	0.1472/ 0.1161	3.284(0.471),3.332(0.505), 3.415(0.499),3.331(0.401)
Width_allV	284.2/ 279.8	307.1/ 292.8	0.668/ 0.675	0.559	0.598	0.521 nested:0.067	0.216/ 0.171	4.206(0.571),4.144(0.588), 4.242(0.646),4.195(0.587)
Angle.BC_V	1792/ 1788	1815/ 1801	0.625/ 0.632	0.422	0.551	0.377 nested:0.095	317.5/ 444.5	123.6(36.37),122.1(32.8), 123.8(33.9),121.4(35.3)
Fractal	146.3/ 144.1	141.1/ 133.4	0.293/ 0.323	0.641	0.732	0.645	0.0019/ 0.0011	1.665(0.049),1.651(0.057), 1.64(0.059),1.643(0.055)
Lacunarity	127.5/ 126.9	122.1/ 116.2	0.139/ 0.166	0.564	0.69	0.585	0.0024/ 0.0017	0.149(0.031),0.167(0.078), 0.186(0.075),0.181(0.059)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.7: Analysis of the non-significant features of the different areas for category 3

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, YO)
AVmid_AVbeg	47.91/ 46.66/	42.91/ 36.54	0.196/ 0.231	0.453	0.599	0.469	0.0091/ 0.0102	0.052(0.155),0.0125(0.129), 0.042(0.155), 0.107(0.111)
AVend_AVbeg	36.35/ 34.09	35.28/ 34.08	0.291/ 0.331	0.185	0.342	191	0.0041/ 0.0175	0.068(0.121),0.138(0.139), 0.148(0.147), 0.061(0.174)
AVend_AVmid	61.52/ 57.03	56.45/ 47.17	0.618/ 0.651	0.29	0.404	0.275	0.0034/ 0.0091	0.005(0.122),0.003(0.116), -0.041(0.084),-0.022(0.121)
ellipse1parV	111.4/ 106.8	121.58/ 111.89	0.712/ 0.739	0.695	0.755	0.681	0.9497/ 0.4448	7.205(0.894), 6.956(1.142), 7.275(1.117), 7.147(1.491)
ellipse1ch1V	61.08/ 59.36	67.08/ 62.34	0.232/ 0.312	0.459	0.63	0.479	0.534/ 0.578	4.936(1.083),4.695(0.972) 5.604(0.470),5.24(1.152)

ellipse1ch2V	124.84/ 120.55	134.98/ 125.61	0.635/ 0.667	0.364	0.474	0.348	0.439/ 0.821	5.152(0.946),4.716(0.907), 4.807(1.387),5.070(1.181)
ellipse1ThetaV	323.65/ 323.51	333.78/ 328.58	0.118/ 0.146	0.2	0.408	0.229	38.55/ 129.34	92.15(15.30),85.84(13.77), 80.31(9.74),88.71(12.35)
ellipse1parA	160.52/ 152.01	165.21/ 162.14	0.087/ 0.094	0.113	0.51	0.231	0.532/ 1.768	6.483(1.628),7.023(1.234), 7.321(1.312),6.898(1.345)
ellipse1ch1A	142.42/ 141.72	151.85/ 147.49	0.081/ 0.105	0.73	0.819	0.754	2.686/ 0.871	6.135(2.123),5.235(1.695), 5.983(1.881),6.193(1.819)
ellipse1ch2A	142.89/ 141.25	153.02/ 146.32	0.225/ 0.261	0.241	0.413	0.241	0.4715/ 1.3835	6.022(1.316),5.375(0.791), 6.318(1.897),5.505(1.208)
ellipse1ThetaA	298.61/ 297.22	308.74/ 302.29	0.202/ 0.241	0.24	0.137	0.017	1.434/ 82.75	80.83(10.66),85.78(10.42), 88.51(7.72),82.08(7.38)
ellipse2parV	197.15/ 196.66	207.28/ 201.72	0.138/ 0.168	0.73	0.812	0.746	10.375/ 3.523	8.143(3.414),7.769(3.515), 7.785(3.288),9.460(4.556)
ellipse2ch1V	119.5/ 115.27	125.4/ 124.56	0.063/ 0.056	0.222	0.513	0.299	0.2866/ 0.6703	6.027(0.596),6.151(0.902), 6.873(1.326),6.176(0.947)
ellipse2ch2V	168.57/ 166.77	176.91/ 173.63	0.051/ 0.067	0.784	0.764	0.719	10.232/ 3.987	6.322(3.646),7.716(3.477), 6.997(3.881),6.861(3.662)
ellipse2ThetaV	338.55/ 336.5	348.68/ 341.57	0.267/ 0.305	0.254	0.416	0.262	65.29/ 183.02	101.23(13.57),100.92(15.49), 94.07(15.96),91.74(17.71)
ellipse2parA	145.5/ 143.6	155.6/ 148.7	0.247/ 0.285	0.4	0.549	0.411	0.9177/ 1.3129	6.582(1.871),7.045(1.402), 7.411(1.312),6.525(1.315)
ellipse2ch1A	130.6/ 127.5	140.8/ 132.6	0.415/ 0.455	0.213	0.351	0.210	0.2813/ 1.054	5.88(0.821),5.276(1.013), 5.261(1.437),5.687(1.255)
ellipse2ch2A	124.11/ 119.26	134.24/ 124.33	0.762/ 0.786	0.429	0.524	0.407	0.5303/ 0.7707	5.19(1.377),4.934(0.958), 5.333(1.261),5.174(0.892)
ellipse2ThetaA	296.97/ 289.41	302.05/ 299.54	0.187/ 0.198	0.410	0.683	0.521	47.82/ 43.97	90.72(8.11),87.08(8.01), 83.45(13.61),85.17(7.18)
ellipse3parV	204.1/ 201.1	214.2/ 206.08	0.398/ 0.438	0.414	0.351	0.209	1.746/ 6.61	14.81(2.02),13.86(1.51), 12.91(4.49),14.11(2.61)
ellipse3ch1V	155.8/ 151.36	161.49/ 160.82	0.087/ 0.089	0.29	0.565	0.37	0.922/ 1.566	11.58(1.68),11.91(1.58), 11.01(1.34),11.16(1.516)
ellipse3ch2V	151.57/ 148.19	158.98/ 156.63	0.092/ 0.121	0.547	0.721	0.607	1.849/ 1.196	9.905(1.839),8.783(1.281), 8.974(1.855),9.19(1.914)
ellipse3ThetaV	291.64/ 288.58	301.77/ 293.65	0.4/ 0.441	0.151	0.284	0.149	10.87/ 61.89	76.35(6.68),77.52(7.96), 81.88(10.27),79.21(8.79)
ellipse3parA	166.4/ 143.48	171.56/ 153.61	0.152/ 0.187	0.577	0.867	0.781	3.141/ 0.808	12.55(2.01),11.34(2.03), 12.08(1.57),12.23(2.32)
ellipse3ch1A	183.09/ 179.02	193.22/ 184.09	0.586/ 0.621	0.801	0.843	0.793	8.943/ 2.331	8.496(4.071),7.784(2.595), 8.522(3.219),8.569(3.379)
ellipse3ch2A	164.37/ 161.12	171.26/ 169.44	0.067/ 0.075	0.025	0.313	0.081	0.221/ 2.509	10.93(2.11),9.56(0.98), 10.23(1.73),10.05(1.71)
ellipse3ThetaA	285.75/ 284.32	295.9/ 289.39	0.207/ 0.243	0.071	0.233	0.085	5.25/ 56.40	84.51(5.47),85.84(6.81), 82.12(6.81),85.98(8.87)
CRAE_LEON	458.98/ 450.49	460.13/ 464.79	0.059/ 0.076	0.053	0.165	<0.000	<0.000/ 14.89	18.94(3.19),18.74(3.13), 18.18(3.89),18.01(4.63)
CRVE_LEON	488.93/ 483.23	497.29/ 495.52	0.098/ 0.121	0.067	0.132	<0.000	<0.000/ 22.22	25.76(5.11),26.14(3.98), 26.72(5.23),26.08(5.01)
AVR_LEON	18.94/ 15.34	20.09/ 14.56	0.668/ 0.682	<0.000	0.019	<0.000	<0.000/ 0.0571	0.735(0.321),0.716(0.234), 0.681(0.236),0.691(0.2)
CRAE_KNUD	409.26/ 407.51	421.4/ 416.4	0.052/ 0.061	<0.000	0.118	0.014	0.1295/ 8.5214	18.46(2.01),18.24(2.82), 18.02(3.28),17.89(4.32)
CRVE_KNUD	542.38/ 538.81	553.1/ 549.53	0.084/ 0.093	0.045	0.112	<0.000	<0.000/ 44.63	25.45(5.01), 26.01(3.78), 26.31(5.11),25.87(4.87)
AVR_KNUD	55.11/ 53.16	46.05/ 40.82	0.067/ 0.054	0.056	0.093	<0.000	<0.000/ 0.026	0.731(0.311),0.701(0.245), 0.684(0.245), 0.691(0.201)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.8: Analysis of the non-significant tortuosity features for category 3

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, Y0)
mean_phi	432.19/ 427.91	424.62/ 412.78	0.631/ 0.645	0.327	0.439	0.321	(1.82/ 3.93)x10^{-4}	0.218(0.021),0.216(0.022), 0.223(0.027),0.218(0.023)
median_phi	376.63/ 373.97	369.06/ 358.83	0.342/ 0.359	0.383	0.506	0.384	(4.21/ 6.74)x10^{-4}	0.133(0.029),0.133(0.03), 0.145(0.036),0.138(0.035)
75th_phi	304.87/ 300.66	297.3/ 285.53	0.615/ 0.629	0.285	0.397	0.279	(6.27/ 16.1)x10^{-4}	0.33(0.044),0.345(0.052), 0.345(0.052),0.34(0.044)
sd_phi	487.59/ 483.77	480.02/ 468.64	0.535/ 0.548	0.166	0.268	0.162	(4.68/ 24.22)x10^{-5}	0.237(0.012),0.236(0.018), 0.24(0.018),0.234(0.019)
mean_tau	617.68/ 612.13	610.11/ 597	0.929/ 0.933	0.518	0.606	0.506	(4.73/ 4.61)x10^{-5}	0.035(0.007),0.034(0.01), 0.035(0.01),0.035(0.01)
median_tau	822.32/ 821.96	814.4/ 807.19	0.095/ 0.105	0.439	0.576	0.454	(4.07/ 4.88)x10^{-6}	0.011(0.003),0.0097(0.0025), 0.0112(0.0034),0.0109(0.0033)
75th_tau	671.16/ 667.92	663.6/ 652.79	0.431/ 0.447	0.545	0.643	0.544	(2.91/ 2.43)x10^{-5}	0.0331(0.0056),0.0317(0.0065), 0.0332(0.0082),0.0340(0.0084)
sd_tau	406.01/ 400.21	398.44/ 385.07	0.978/ 0.979	0.296	0.393	0.283	(2.16/ 5.47)x10^{-4}	0.081(0.021),0.078(0.031), 0.078(0.034),0.079(0.026)
mean_psi	652.47/ 651.47	644.91/ 636.34	0.172/ 0.185	0.549	0.658	0.557	(3.63/ 2.88)x10^{-5}	0.044(0.005),0.041(0.008), 0.044(0.009),0.044(0.008)
75th_psi	616.6/ 615.69	609.04/ 600.56	0.179/ 0.165	0.585	0.687	0.593	(6.02/ 4.12)x10^{-5}	0.0516(0.008),0.0488(0.0098), 0.0528(0.0127),0.0519(0.0091)
sd_psi	552.11/ 548.14	544.54/ 533.01	0.566/ 0.582	0.407	0.517	0.402	(6.74/ 10.02)x10^{-5}	0.0724(0.0086),0.0691(0.0136), 0.0721(0.0145),0.0728(0.0141)
Tort.blend	161.55/ 157.43	153.98/ 142.4	0.612/ 0.597	0.375	0.486	0.369	0.0042/ 0.0072	0.988(0.091),0.974(0.1), 1.007(0.12),0.986(0.112)
Tort_ellipse1	78.78/ 76.97	73.72/ 66.83	0.242/ 0.282	<0.000	0.099	<0.000	<0.000/ 0.0071	0.0223(0.0193),0.0827(0.1132), 0.0286(0.0191),0.0725(0.1207)
Tort_ellipse2	31.95/ 27.56	26.89/ 17.43	0.657/ 0.689	<0.000	0.04	<0.000	<0.000/ 0.0242	0.0212(0.0181),0.0853(0.2152), 0.0882(0.2186),0.0342(0.0480)
Tort_ellipse3	71.88/ 66.49	66.78/ 56.36	0.885/ 0.899	0.307	0.396	0.279	0.0028/ 0.0072	0.0587(0.0549),0.078(0.1624), 0.0528(0.0716),0.0735(0.0749)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.9: Analysis of the non-significant haemodynamic features for category 3

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, Y0)
wssparentA	107.89/ 107.2	110.11/ 109.35	0.082/ 0.182	0.838	0.917	0.881	771.4/ 104.1	79.85(35.11),65.78(26.22), 76.52(21.26),71.39(33.6)
pressureA	68.13/ 64.36	71.04/ 65.81	0.524/ 0.654	0.347	0.483	0.295	3.096/ 7.377	5.02(2.97),3.307(2.54), 4.402(3.3),2.336(1.208)
vparentaA	68.54/ 67.75	70.66/ 69.99	0.079/ 0.136	0.832	0.915	0.877	27.99/ 3.92	21.87(6.37),18.05(5.52), 21.48(3.92),20,29(6.46)
ReparentA	117.16/ 116.46	119.37/ 118.61	0.082/ 0.123	0.814	0.905	0.861	1466.1/ 234.7	307.1(43.54),278.56(44.38), 305.6(27.47),295.35(46.71)
wsschild1A	77.32/ 76.71	80.23/ 78.16	0.145/ 0.167	0.664	0.81	0.71	27.52/ 11.2	32.61(4.39),28.87(5.89), 31.76(6.27),34.85(7.83)
qchild1A	193.33/ 189.45	191.87/ 186.54	0.548/ 0.531	0.287	0.424	0.231	(1.11/ 3.69)x10^{-9}	(8.01(0.925),7.79(0.761), 7.77(0.219),8.27(0.661))x10^{-4}
vchild1A	42.45/ 41.35	45.36/ 42.81	0.178/ 0.198	0.503	0.711	0.546	0.861/ 0.715	10.15(0.727),9.28(1.42), 9.86(1.33),10.69(1.41)
Rechild1A	100.11/ 98	103.02/ 99.45	0.274/ 0.301	<0.000	0.276	<0.000	<0.000/ 135.7	159.95(8.74),150.9(16.96), 154.73(11.65),166.57(6.52)
wsschild2A	83.64/ 80.92	86.56/ 82.37	0.351/ 0.389	<0.000	0.238	<0.000	<0.000/ 34.4	39.29(3.86),36.14(4.67), 35.61(7.75),42.29(6.37)
qchild2A	193.82/ 189.41	191.86/ 186.5	0.552/ 0.524	0.29	0.425	0.233	(1.25/ 3.7)x10^{-9}	(5.82(0.926),6.04(0.762), 6.13(0.219),5.56(0.665))x10^{-4}
vchild2A	41.07/ 37.56	43.98/ 39.01	0.478/ 0.516	0.277	0.442	0.235	0.2494/ 0.8105	10.33(0.48),9.89(0.41), 9.84(1.33),10.72(1.42)
Rechild2A	90.93/ 91.32	92.78/ 93.83	0.094/ 0.102	0.096	0.562	0.189	12.53/ 53.5	132.63(29.28),140.16(26.13), 141.53(29.89),147.41(26.5)

PinQoutA	90.66/ 90.62	93.57/ 92.01	0.113/ 0.131	0.607	0.674	0.794	73.39/ 35.35	16.56(9.20),24.05(6.43), 15.56(12.01),24.31(12.82)
wssparentV	86.08/ 81.71	88.99/ 83.16	0.653/ 0.678	0.749	0.777	0.7	54.93/ 23.51	17.97(5.74),19.86(5.3), 22.21(10.46),20.63(11.96)
pressureV	54.64/ 54.03	57.15/ 55.45	0.144/ 0.167	0.078	0.378	0.012	0.039/ 3.020	0.86(0.78),0.79(0.72), 0.71(0.49),1.12(0.98)
vparentaV	55.36/ 50.83	58.27/ 52.29	0.689/ 0.734	0.759	0.783	0.708	4.377/ 1.801	7.87(1.75),8.44(1.47), 9(2.8), 8.47(3.4)
ReparentV	112.33/ 107.7	115.2/ 109.15	0.712/ 0.765	0.764	0.784	0.711	509.5/ 206.9	175.22(20.37),181.84(15.61), 186.77(28.57),180.08(37.29)
wsschild1V	72.58/ 69.04	75.47/ 70.49	0.476/ 0.497	0.419	0.553	0.376	6.065/ 10.04	16.56(1.27),17.65(3.48), 19.97(3.56),17.39(6.15)
qchild1V	186.5/ 181.93	185.04/ 179.02	0.699/ 0.713	0.746	0.768	0.759	(2.89/ 0.98)x10^{-8}	(6.7(2.51),6.54(2.162), 6.3(2.71),6.4(2.75))x10^{-4}
vchild1V	44.71/ 40.34	47.62/ 41.7	0.653/ 0.687	0.789	0.812	0.746	2.092/ 0.709	6.023(0.498),6.324(1.516), 6.747(1.785),6.216(2.338)
Rechild1V	109.75/ 104.08	112.65/ 105.53	0.94/ 0.943	0.937	0.933	0.911	1227.7/ 118.7	112.08(26.06),114.24(32.53), 115.52(40.33),112(44.95)
wsschild2V	79.42/ 74.47	82.33/ 75.92	0.789/ 0.823	0.951	0.951	0.934	125.56/ 8.754	21.8(8.53),21.96(11.29), 20.10(11.74),21.51(14.09)
qchild2V	187.3/ 180.95	184.15/ 178.92	0.598/ 0.623	0.756	0.779	0.771	(3.89/ 1.18)x10^{-8}	(5.72(2.51),5.91(2.17), 6.16(2.7),6.06(2.74))x10^{-4}
vchild2V	43.48/ 38.35	46.39/ 39.81	0.834/ 0.856	0.852	0.854	0.806	2.463/ 0.591	6.61(1.06),6.64(1.57), 6.24(1.65),6.32(2.42)
Rechild2V	98.97. 93.61	101.88/ 95.07	0.887/ 0.928	0.681	0.687	0.593	111.36/ 76.32	106.52(19.97),107.95(13.01), 105.77(9.45),103.32(9.64)
PinQoutV	83.75/ 81.27	86.66/ 82.72	0.318/ 0.345	0.863	0.9	0.862	96.12/ 15.37	26.88(9.28),27.7(9.26), 26.26(10.51),31.21(12.77)

^{a}: Full versus Restricted (null) model, ^{b}: chi-squared p-value and p-value based on Satterthwaite's approximation, ^{c}: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, ^{d}: Variance of Subjects/Variance of Residuals

Table A.10: Analysis of the non-significant main features for category 4

Feature name	AIC^a	BIC^a	(p-(χ^2)/ p-(s))^b	ICC	Ω^2	F^c	(V_R/V_E)^d	Group Means(SD) (Y3, Y2, Y1, Y0)
AlphaA	163.4/ 160	147.8/ 133.44	0.326/ 0.325	0.109	0.152	0.112 nested:<0.000	0.0044/ 0.0349	0.767(0.181),0.763(0.206), 0.804(0.185),0.798(0.181)
BetaA	62.87/ 59.81	90.44/ 75.56	0.401/ 0.406	0.015	0.01	<0.000 nested:<0.000	<0.000/ 0.0673	1.285(0.202),1.345(0.214), 1.312(0.198),1.321(0.178)
LambdaA	556.65/ 553.78	541.65/ 526.4	0.377/ 0.378	0.11	0.175	0.111 nested:0.013	0.0015/ 0.0122	0.869(0.106),0.864(0.127), 0.878(0.109),0.875(0.101)
Lambda1A	828.76/ 827.5	813.24/ 799.87	0.23/ 0.234	0.015	0.069	0.014 nested:0.018	(9.09/ 623.6)x10^{-5}	0.896(0.086),0.866(0.074), 0.845(0.084),0.85(0.071)
Lambda2A	600.9/ 596.9	585.2/ 569.4	0.576/ 0.575	0.064	0.09	0.065 nested:<0.000	0.0007/ 0.0113	0.819(0.103),0.832(0.121), 0.826(0.109),0.827(0.104)
ThetaA	436.16/ 432.63	463.72/ 448.38	0.48/ 0.486	0.09	0.133	0.085 nested:0.01	0.016/ 0.169	142.8(40.1),140.26(38.57), 149.4(47.4),149.6(47.6)
Adj.GammaA	909.68/ 907.21	893.45/ 879.65	0.316/ 0.319	<0.000	0.01	<0.000 nested:<0.000	<0.000/ 0.0052	0.823(0.045),0.825(0.056), 0.833(0.052),0.838(0.061)
Width_parentA	765.28/ 763.34	793.88/ 778.83	0.373/ 0.376	0.17	0.25	0.16 nested:0.032	0.075/ 0.383	3.823(0.677),3.958(0.710), 4.004(0.717),3.96(0.623)
Width_child1A	735.51/ 735.27	762.83/ 751.25	0.101/ 0.104	0.173	0.322	0.161 nested:0.073	0.071/ 0.339	3.599(0.619),3.814(0.712), 3.766(0.668),3.762(0.659)
Width_child2A	550.21/ 546.42	573.98/ 565.75	0.121/ 0.122	0.151	0.233	0.145 nested:0.023	0.038/ 0.218	3.100(0.535),3.244(0.527), 3.297(0.463),3.285(0.496)
Width_allA	602.17/ 601.77	629.33/ 618.22	0.082/ 0.081	0.189	0.302	0.176 nested:0.05	0.055/ 0.242	3.508(0.56),3.672(0.578), 3.689(0.55),3.669(0.53)
Angle.BC_A	288.99/ 285.55	316.55/ 301.28	0.467/ 0.474	0.046	0.071	0.0055/ nested:0.118	0.0056/ 0.119	115.34(39.33),111.12(35.21), 119.23(44.53),117.87(40.04)
AlphaV	157.69/ 154.94	144.53/ 128.62	0.355/ 0.354	0.159	0.179	0.161 nested:0.001	0.0079/ 0.0415	0.647(0.222),0.668(0.223), 0.677(0.221),0.645(0.216)
BetaV	36.65/ 31.33	62.92/ 44.94	0.877/ 0.88	0.054	0.066	0.052 nested:0.012	0.0034/ 0.0593	1.141(0.145),1.159(0.122), 1.154(0.156),1.134(0.175)

LambdaV	706.57/ 703.82	693.4/ 677.5	0.354/ 0.355	0.165	0.185	0.167 nested:0.051	0.0033/ 0.0164	0.792(0.141),0.805(0.138), 0.811(0.14),0.791(0.139)
Lambda1V	1343/ 1337	1330/ 1311	0.927/ 0.927	0.045	0.062	0.045 nested:<0.000	0.0002/ 0.0059	0.836(0.058),0.836(0.059), 0.831(0.062),0.833(0.063)
Lambda2V	847.72/ 844.22	834.56/ 817.9	0.475/ 0.476	0.146	0.158	0.147 nested:<0.000	0.0022/ 0.0131	0.737(0.122),0.751(0.118), 0.751(0.122),0.734(0.127)
ThetaV	411.26/ 409.12	437.58/ 422.28	0.277/ 0.279	0.129	0.152	0.131 nested:<0.000	0.0163/ 0.1087	129.9(33.9),133.3(38.2), 134.9(35.1),127.7(32.6)
Adj.GammaV	1474/ 1469	1461/ 1442	0.875/ 0.879	0.04	0.05	0.036 nested:0.043	0.0002/ 0.0047	0.754(0.069),0.767(0.056), 0.774(0.034),0.773(0.063)
Width_parentV	1555/ 1551	1582/ 1565	0.515/ 0.519	0.197	0.201	0.194 nested:<0.000	0.177/ 0.735	4.456(0.97),4.436(0.875), 4.541(0.91),4.63(1.001)
Width_child1V	1465/ 1461	1491/ 1474	0.679/ 0.682	0.2	0.208	0.197 nested:0.084	0.1553/ 0.6311	4.156(0.901),4.141(0.809), 4.216(0.865),4.299(0.933)
Width_child2V	808.06/ 805.76	404.03/ 402.88	0.221/ 0.222	0.188	0.216	0.195 nested:0.078	0.051/ 0.208	3.204(0.523),3.262(0.517), 3.332(0.498),3.303(0.474)
Width_allV	1200/ 1197	1226/ 1210	0.349/ 0.351	0.208	0.215	0.207 nested:0.092	0.1053/ 0.4028	3.939(0.732),3.946(0.667), 4.031(0.678),4.077(0.731)
Angle.BC_V	286.56/ 281.81	312.88/ 294.97	0.741/ 0.744	0.093	0.114	0.093 nested:<0.000	0.0092/ 0.0891	115.93(36.15),118.12(36.92), 118.88(36.34),116.4(35.49)

^{a}: Full versus Restricted (null) model, ^{b}: chi-squared p-value and p-value based on Satterthwaite's approximation, ^{c}: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, ^{d}: Variance of Subjects/Variance of Residuals

Table A.11: Analysis of the non-significant main features for category 5

Feature name	AIC^a	BIC^a	(p-(χ^2)/ p-(s))^b	ICC	Ω^2	F^c	(V_R/V_E)^d	Group Means(SD) (Prog, Non-Prog, DR)
AlphaA	234.19/ 230.28	218.91/ 208.15	0.352/ 0.368	0.161	0.215	0.161 nested:<0.000	0.0068/ 0.036	0.75(0.212),0.785(0.203) 0.792(0.184)
BetaA	133.85/ 135.7	151.89/ 148.98	0.054/ 0.064	0.031	0.735	0.023 nested:0.463	0.0015/ 0.037	1.253(0.171),1.317(0.162) 1.27(0.176)
LambdaA	881.51/ 879.54	868.23/ 857.40	0.362/ 0.278	0.169	0.225	0.168 nested:<0.000	0.0025/ 0.0125	0.836(0.117),0.878(0.101) 0.872(0.109)
Lambda1A	1306/ 1304	1292/ 1282	0.372/ 0.391	0.072	0.109	0.073 nested:<0.000	(5.23/66.39) x10^{-4}	0.912(0.057),0.922(0.052) 0.927(0.044)
Lambda2A	950.77/ 950.37	480.39/ 478.18	0.111/ 0.113	0.112	0.154	0.104 nested:<0.000	0.0013/ 0.0115	0.803(0.116),0.832(0.111) 0.824(0.113)
ThetaA	806.5/ 804.62	826.75/ 819.78	0.053/ 0.054	0.108	0.151	0.099 nested:<0.000	0.022/ 0.198	147.61(50.47),143.73(43.48) 158.77(48.56)
Adj.GammaA	1429/ 1427	1414/ 1407	0.058/ 0.071	0.031	0.742	0.022 nested:<0.470	(1.29/56.15) x10^{-4}	0.782(0.068),0.800(0.063) 0.788(0.068)
Angle.BC_A	6329/ 6328	6350/ 6343	0.076/ 0.087	0.112	0.148	0.102 nested:<0.000	173/ 1508	122.17(49.54),114.42(39.72) 125.73(44.77)
AlphaV	49.06/ 47.76	78.98/ 76.54	0.16/ 0.167	0.576	0.703	0.081 nested:<0.000	0.0044/ 0.0502	0.672(0.231),0.635(0.233) 0.699(0.234)
BetaV	95.45/ 91.12	207.83/ 204.34	0.23/ 0.256	0.558	0.688	0.111 nested:<0.000	0.0077/ 0.0619	1.141(0.163),1.182(0.141) 1.175(0.178)
LambdaV	677/54/ 674.34	663.45/ 660.34	0.601/ 0.616	0.141	0.099	0.103 nested:<0.000	0.0023/ 0.0204	0.806(0.149),0.797(0.155) 0.822(0.148)
Lambda1V	1525/ 1522	1512/ 1499	0.867/ 0.869	0.048	0.077	0.054 nested:<0.000	(3.54/61.45) x10^{-4}	0.908(0.069),0.912(0.066) 0.912(0.074)
Lambda2V	853.44/ 850.97	839.82/ 828.28	0.465/ 0.481	0.123	0.165	0.126 nested:<0.000	0.0022/ 0.0156	0.724(0.123),0.717(0.109) 0.722(0.125)
ThetaV	577.06/ 576.57	599.26/ 590.68	0.105/ 0.108	0.098	0.132	0.089 nested:<0.000	0.0123/ 0.1257	135.91(37.66),125.55(32.26) 137.7(39.7)
Adj.GammaV	1661/ 1659	1647/ 1636	0.475/ 0.491	0.09	0.123	0.094 nested:<0.000	(5.12/49.34) x10^{-4}	0.753(0.064),0.753(0.055) 0.762(0.068)
Fractal	628.32/ 625.48	618.53/ 609.16	0.559/ 0.569	0.837	0.759	0.761	0.0029/ 0.0009	1.652(0.058),1.65(0.054), 1.637(0.064)
Lacunarity	543.34/ 539.6	533.56/ 523.29	0.878/ 0.882	0.806	0.717	0.722	0.0040/ 0.0015	0.167(0.071),0.171(0.063), 0.176(0.078)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.12: Analysis of the non-significant features of the different areas for category 5

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Prog, Non-Prog, DR)
AVmid_AVbeg	88.74/ 87.14	81.08/ 74.35	0.303/ 0.33	0.579	0.711	0.576	0.016/ 0.011	-0.024(0.156),-0.109(0.168), 0.005(0.179)
AVend_AVbeg	84.54/ 82.04	76.87/ 69.22	0.471/ 0.498	0.734	0.609	0.612	0.018/ 0.011	0.011(0.166),0.094(0.146), 0.006(0.193)
AVend_AVmid	67.71/ 63.88	60.04/ 51.11	0.912/ 0.918	0.502	0.64	0.524	0.018/ 0.016	-0.013(0.161),-0.014(0.166), 0.01(0.223)
ellipse1parV	317.29/ 311.13	323.9/ 314.95	0.925/ 0.931	0.666	0.525	0.544	0.9784/ 0.8169	5.804(1.169),5.802(1.51), 5.946(1.508)
ellipse1ch1V	309.07/ 305.19	321.84/ 312.85	0.941/ 0.944	0.405	0.254	0.274	0.408/ 1.078	4.24(1.256),4.283(0.974), 4.353(1.227)
ellipse1ch2V	307.52/ 304.12	320.29/ 311.79	0.736/ 0.751	0.361	0.517	0.377	0.579/ 0.954	4.652(1.142),4.382(1.334), 4.431(1.36)
ellipse1ThetaV	725.17/ 722.46	737.94/ 730.12	0.523/ 0.545	0.419	0.581	0.429	55.19/ 73.23	83.49(11.53),82.24(10.56), 86.56(10.88)
ellipse1parA	361.6/ 357.64	374.37/ 365.30	0.977/ 0.979	0.683	0.802	0.696	2.54/ 1.11	6.789(1.829),6.819(2.135), 6.684(1.872)
ellipse1ch1A	362.96/ 359.55	375.73/ 367.21	0.745/ 0.759	0.601	0.735	0.612	2.031/ 1.285	5.679(1.895),5.208(1.05), 5.809(1.874)
ellipse1ch2A	375.47/ 372.46	388.24/ 380.12	0.611/ 0.628	0.575	0.721	0.582	2.135/ 1.529	5.42(1.756),6.069(2.613), 5.259(1.755)
ellipse1ThetaA	754.66/ 750.82	767.43/ 758.48	0.925/ 0.931	0.284	0.443	0.301	49.15/ 114.46	86.38(13.24),87.63(13.92), 85.71(10.22)
ellipse2parV	363.37/ 359.63	376.14/ 367.31	0.786/ 0.794	0.667	0.786	0.677	5.904/ 2.808	8.199(3.021),9.001(2.104), 8.12(3.005)
ellipse2ch1V	331.3/ 329.32	341.07/ 336.98	0.265/ 0.279	0.275	0.429	0.279	0.404/ 1.031	5.878(1.244),5.868(0.907), 5.368(1.197)
ellipse2ch2V	352.02/ 348.39	364.79/ 356.06	0.831/ 0.841	0.285	0.449	0.289	0.414/ 1.041	6.646(3.234),7.308(2.339), 6.804(3.195)
ellipse2ThetaV	725.48/ 722.65	738.75/ 731.21	0.716/ 0.734	0.752	0.844	0.759	143.1/ 45.3	84.11(14.05),79.99(6.39), 85.05(14.54)
ellipse2parA	363.37/ 359.32	376.14/ 367.98	0.878/ 0.893	0.677	0.803	0.688	2.524/ 1.144	7.094(1.883),6.686(2.362), 6.941(1.694)
ellipse2ch1A	333.3/ 329.32	346.07/ 336.98	0.895/ 0.867	0.677	0.796	0.691	1.852/ 0.831	5.704(1.531),5.781(1.935), 5.707(1.618)
ellipse2ch2A	352.02/ 348.39	364.79/ 356.06	0.831/ 0.842	0.217	0.359	0.231	0.533/ 1.768	5.557(1.6),5.295(1.366), 5.378(1.324)
ellipse2ThetaA	722.48/ 719.44	735.25/ 727.11	0.616/ 0.634	0.525	0.664	0.538	72.98/ 62.56	85.21(10.81),84.11(10.38), 81.99(13.24)
ellipse3parV	484.11/ 480.27	496.88/ 487.93	0.928/ 0.929	0.409	0.568	0.429	4.357/ 5.794	14.99(3.39),15.43(1.15), 15.21(3.34)
ellipse3ch1V	398.25/ 394.56	410.79/ 401.91	0.894/ 0.901	0.818	0.891	0.825	5.756/ 1.216	11.59(2.63),12.14(2.09), 11.61(2.71)
ellipse3ch2V	399.46/ 395.75	412.23/ 403.43	0.862/ 0.871	0.835	0.891	0.842	6.26/ 1.174	10.95(2.38),10.61(1.93), 10.53(3.21)
ellipse3ThetaV	691.42/ 689.43	704.19/ 697.1	0.365/ 0.385	0.478	0.619	0.481	44.95/ 48.38	80.05(9.02),83.14(8.13), 83.69(11.21)
ellipse3parA	437.58/ 437.51	450.26/ 445.24	0.129/ 0.134	0.347	0.497	0.321	1.878/ 3.981	11.48(2.44),9.77(2.57), 11.52(2.18)
ellipse3ch1A	449.64/ 447.16	462.41/ 454.82	0.467/ 0.487	0.472	0.631	0.474	3.452/ 3.829	9.425(2.866),8.124(2.098), 9.211(2.497)
ellipse3ch2A	401.39/ 398.67	414.16/ 406.34	0.527/ 0.547	0.608	0.744	0.613	3.051/ 1.923	9.413(2.175),8.367(2.361), 9.115(2.188)
ellipse3ThetaA	692.33/ 688.45	705.1/ 696.11	0.941/ 0.944	0.467	0.632	0.484	45.63/ 48.72	86.82(9.52),87.54(12.01), 86.21(8.39)
AVR_KNUD	293.4/ 291.9	283.5/ 275.4	0.291/ 0.3	0.092	0.165	0.091	0.0012/ 0.0121	0.698(0.088),0.726(0.159), 0.723(0.093)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Ran-

dom effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, ^{d}: Variance of Subjects/Variance of Residuals

Table A.13: Analysis of the non-significant tortuosity features for category 5

Feature name	AIC^a	BIC^a	(p-(χ^2)/ p-(s))^b	ICC	Ω^2	F^c	(V_R/V_E)^d	Group Means(SD) (Prog, Non-Prog, DR)
mean_phi	925.1/ 924.6	915.1/ 907.9	0.197/ 0.199	0.434	0.532	0.419	(3.06/ 4.23)x10^{-4}	0.229(0.029),0.218(0.023), 0.227(0.027)
median_phi	799.7/ 797.5	789.1/ 780.9	0.406/ 0.411	0.487	0.584	0.482	(7.02/ 7.51)x10^{-4}	0.147(0.043),0.135(0.032), 0.143(0.037)
75th_phi	650.1/ 649.9	640.3/ 632.7	0.213/ 0.214	0.395	0.495	0.376	(1.02/ 1.69)x10^{-3}	0.361(0.056),0.338(0.046), 0.356(0.053)
mean_tau	1335/ 1336	1325/ 1319	0.107/ 0.111	0.382	0.503	0.367	(3.66/ 5.79)x10^{-5}	0.041(0.011),0.036(0.009), 0.037(0.009)
median_tau	1780/ 1778	1770/ 1762	0.284/ 0.294	0.405	0.512	0.397	(4.21/ 6.41)x10^{-6}	0.012(0.003),0.011(0.003), 0.012(0.004)
75th_tau	650.2/ 649.2	640.2/ 632.2	0.063/ 0.06	0.361	0.484	0.348	(2.24/ 4.19)x10^{-5}	0.037(0.008),0.033(0.007), 0.036(0.009)
sd_tau	916.8/ 915.1	906.8/ 898.5	0.313/ 0.317	0.305	0.415	0.294	(2.04/ 4.88)x10^{-4}	0.089(0.026),0.081(0.027), 0.083(0.024)
mean_psi	1364/ 1363	1354/ 1347	0.202/ 0.205	0.374	0.488	0.369	(2.96/ 5.05)x10^{-5}	0.047(0.010),0.044(0.008), 0.046(0.009)
median_psi	1586/ 1584	1576/ 1568	0.297/ 0.3	0.542	0.645	0.541	(1.74/ 1.47)x10^{-5}	0.018(0.007),0.016(0.005), 0.018(0.005)
75th_psi	1225/ 1222	1215/ 1205	0.669/ 0.675	0.453	0.549	0.446	(7.71/ 9.55)x10^{-5}	0.055(0.015),0.051(0.011), 0.053(0.013)
sd_psi	1173/ 1170	1163/ 1154	0.638/ 0.639	0.226	0.329	0.231	(4.39/ 14.55)x10^{-5}	0.075(0.015),0.072(0.012), 0.073(0.014)
Tort.blend	333.4/ 334.9	322.1/ 323.3	0.061/ 0.06	0.379	0.481	0.344	0.0043/ 0.0081	1.056(0.113),0.991(0.105), 1.031(0.115)
Tort_ellipse1	224.62/ 222.89	215.8/ 208.2	0.321/ 0.339	0.018	0.03	0.008	<0.000/ 0.011	0.051(0.085),0.081(0.114), 0.083(0.112)
Tort_ellipse2	111.22/ 109.08	102.4/ 94.38	0.395/ 0.408	0.189	0.312	0.191	0.0049/ 0.0213	0.057(0.152),0.099(0.184), 0.055(0.086)
Tort_ellipse3	336.6/ 336.4	327.76/ 326.89	0.148/ 0.153	0.248	0.401	0.245	0.013/ 0.004	0.067(0.096),0.054(0.069), 0.027(0.026)

^{a}: Full versus Restricted (null) model, ^{b}: chi-squared p-value and p-value based on Satterthwaite's approximation, ^{c}: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, ^{d}: Variance of Subjects/Variance of Residuals

Table A.14: Analysis of the non-significant haemodynamic features for category 5

Feature name	AIC^a	BIC^a	(p-(χ^2)/ p-(s))^b	ICC	Ω^2	F^c	(V_R/V_E)^d	Group Means(SD) (Prog, Non-Prog, DR)
wssparentA	937.4/ 934.5	950.4/ 942.7	0.574/ 0.591	0.823	0.902	0.826	862.1/ 180.5	54.08(34.50),57.83(26.48), 52.81(29.16)
pressureA	655.4/ 654.3	667.4/ 666.3	0.089/ 0.095	0.516	0.693	0.343	10.84/ 20.68	3.403(4.688),3.741(3.942), 3.368(3.471)
vparentaA	603.2/ 600.35	616.56/ 608.23	0.543/ 0.563	0.859	0.923	0.862	37.82/ 6.06	16.42(7.04),17.34(5.06), 16.33(6.12)
ReparentA	1018/ 1016	1032/ 1024	0.521/ 0.538	0.886	0.938	0.887	2540.6/ 321.4	261.96(56.59),271.65(48.34), 262.63(50.17)
wsschild1A	811.9/ 807.4	824.4/ 815.4	0.765/ 0.781	0.654	0.771	0.664	147.29/ 74.26	26.93(14.72),30.12(13.23), 28.75(16.14)
vchild1A	509.9/ 505.3	522.8/ 513.7	0.708/ 0.726	0.664	0.781	0.673	7.799/ 3.788	8.618(3.44),9.997(3.195), 9.026(3.571)
Rechild1A	988.2/ 985.3	1004/ 996.1	0.665/ 0.682	0.588	0.726	0.597	712.7/ 480.8	143.62(36.44),158.04(33.54), 147.97(33.81)
wsschild2A	799.4/ 796.9	814.5/ 805.8	0.775/ 0.789	0.711	0.818	0.721	154.9/ 60.17	32.98(15.03),34.45(16.56), 32.83(15)
vchild2A	448.7/ 445.6	461.4/ 445.3	0.668/ 0.684	0.741	0.841	0.748	5.389/ 1.821	8.991(2.775),9.03(2.74), 8.892(2.72)
Rechild2A	923.88/ 920.78	937.65/ 928.65	0.638/ 0.655	0.599	0.739	0.609	379.1/ 244.3	127.46(26.73),128.34(25.93), 125.32(23.61)

PinQoutA	734.44/ 731.75	747.88/ 739.31	0.71/ 0.724	0.683	0.691	0.808	74.92/ 33.51	24.38(10.85),22.42(9.85), 24.67(9.46)
wssparentV	672.05/ 668.42	684.55/ 675.92	0.832/ 0.842	0.889	0.924	0.894	205.79/ 24.39	24.01(12.8),23.34(12.12), 25.43(17.23)
pressureV	421.53/ 420.45	434.03/ 427.95	0.233/ 0.263	0.177	0.311	0.181	1.09/ 4.92	2.17(1.78),2.43(1.86), 2.21(1.87)
vparentaV	423.94/ 420.21	436.44/ 427.71	0.871/ 0.879	0.884	0.923	0.889	12.73/ 1.584	9.35(3.31),9.23(3.89), 9.59(4.19)
ReparentV	837.25/ 833.44	849.8/ 840.99	0.91/ 0.916	0.881	0.923	0.886	1236.3/ 158.4	188.91(33.54),184.98(34.53), 190.25(40.56)
wsschild1V	605.74/ 602.71	618.24/ 610.21	0.615/ 0.633	0.882	0.926	0.884	93.55/ 12.18	20.22(9.85),17.99(9.12), 22.69(10.82)
Rechild1V	756.28/ 752.49	768.78/ 759.99	0.902/ 0.908	0.862	0.926	0.868	458.25/ 69.48	110.55(21.36),113.46(31.32), 113.78(22.53)
wsschild2V	610.8/ 606.43	623.11/ 614.33	0.894/ 0.901	0.913	0.857	0.863	88.61/ 14.04	20.54(9.24),21.34(9.91), 22.25(10.48)
Rechild2V	749.88/ 746.49	762.38/ 753.87	0.739/ 0.753	0.928	0.957	0.931	627.26/ 46.78	112.14(25.94),112(20.13), 116.55(25.63)
PinQoutV	554.43/ 551.1	566.92/ 558.6	0.715/ 0.732	0.794	0.873	0.801	36.83/ 9.14	31.05(6.41),33.61(5.98), 30.51(7.06)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Table A.15: Analysis of the non-significant main features for category 6

Feature name	AIC[a]	BIC[a]	(p-(χ^2)/ p-(s))[b]	ICC	Ω^2	F[c]	(V_R/V_E)[d]	Group Means(SD) (Y3, Y2, Y1, DR)
AlphaA	115.86/ 113.24	102.97/ 90.66	0.336/ 0.344	<0.000	0.417	<0.000 nested:0.226	<0.000/ 0.0234	0.779(0.184),0.831(0.165), 0.791(0.145),0.813(0.156)
BetaA	125.04/ 120.06	148.76/ 133.62	0.795/ 0.808	0.489	0.804	0.494 nested:0.142	0.065/ 0.048	1.265(0.301),1.25(0.324), 1.33(0.256),1.351(0.231)
LambdaA	101.83/ 96.01	88.27/ 72.92	0.978/ 0.981	0.789	0.965	0.795 nested:0.101	0.0671/ 0.0087	0.807(0.103),0.818(0.106), 0.812(0.122),0.805(0.105)
Lambda1A	428.52/ 424.15	414.97/ 400.43	0.653/ 0.684	0.04	0.586	0.043 nested:0.316	<0.000/ 0.0052	0.902(0.067),0.902(0.061), 0.922(0.059),0.926(0.061)
Lambda2A	364.38/ 360.76	350.82/ 337.03	0.497/ 0.521	0.164	0.865	0.149 nested:0.497	0.0016/ 0.0039	0.815(0.106),0.851(0.095), 0.835(0.081),0.841(0.097)
ThetaA	301.68/ 297.93	325.41/ 311.49	0.523/ 0.543	0.316	0.431	0.321 nested:<0.000	0.081/ 0.171	1.236(0.351),1.247(0.371), 1.176(0.367),1.347(0.387)
Adj.GammaA	22.78/ 17.04	46.52/ 30.61	0.967/ 0.974	0.326	0.992	0.381 nested:0.546	0.029/ 0.0055	0.838(0.093),0.842(0.097), 0.839(0.097),0.829(0.095)
Width_parentA	587.68/ 581.75	611.4/ 595.31	0.995/ 0.995	0.752	0.835	0.76 nested:<0.000	1.295/ 0.409	3.408(1.346),3.255(1.55), 3.507(1.318),3.383(1.139)
Width_child1A	441.67/ 436.48	465.4/ 450.04	0.847/ 0.853	0.435	0.991	0.445 nested:0.486	0.244/ 0.037	3.506(0.764),3.613(0.724), 3.638(0.77),3.431(0.604)
Width_child2A	318.86/ 315.35	342.45/ 328.91	0.476/ 0.492	0.446	0.738	0.443 nested:0.129	0.133/ 0.128	3.072(0.516),3.3(0.609), 3.241(0.589),3.08(0.416)
Width_allA	414.74/ 409.17	438.46/ 422.72	0.934/ 0.937	0.602	0.902	0.607 nested:0.173	0.349/ 0.125	3.329(0.742),3.39(0.812), 3.462(0.792),3.297(0.608)
Angle.BC_A	2284/ 2281	2307/ 2294	0.432/ 0.455	0.044	0.116	0.052 nested:0.008	99.36/ 1787.47	128.64(54.54),129.45(50.65), 122.45(42.54),128.78(48.32)
AlphaV	88.68/ 88.56	74.56/ 64.93	0.111/ 0.109	0.014	0.659	0.011 nested:0.407	<0.000/ 0.021	0.732(0.212),0.713(0.197), 0.686(0.186),0.771(0.179)
BetaV	42.39/ 42.27	66.05/ 55.98	0.106/ 0.117	0.05	0.459	0.025 nested:0.235	0.0018/ 0.051	1.136(0.136),1.116(0.153), 1.081(0.182),1.166(0.174)
LambdaV	307.79/ 307.71	294.11/ 284	0.107/ 0.129	0.023	0.662	0.022 nested:0.401	0.0003/ 0.0081	0.846(0.125),0.835(0.123), 0.821(0.117),0.872(0.105)
Lambda1V	480.63/ 476.89	467.03/ 453.11	0.519/ 0.522	<0.000	0.479	<0.000 nested:0.279	<0.000/ 0.0047	0.912(0.061),0.91(0.063), 0.892(0.086),0.908(0.075)
Lambda2V	326.45/ 327.62	303.12/ 304.56	0.067/ 0.086	0.076	0.576	0.053 nested:0.297	0.0006/ 0.0084	0.794(0.111),0.783(0.117), 0.753(0.116),0.816(0.105)
ThetaV	219.88/ 220.56	234.45/ 235.65	0.056/ 0.067	0.045	0.218	0.026 nested:0.072	0.0039/ 0.1349	145.75(43.5),136.4(36.18), 133.68(40.64),125.77(33.95)

Adj.GammaV	507.39/ 507.3	493.34/ 492.65	0.115/ 0.154	0.051	0.428	0.028 nested:0.211	0.0001/ 0.0043	0.779(0.065),0.773(0.07), 0.753(0.077),0.783(0.068)
Width_parentV	595.66/ 593.48	619.45/ 607.07	0.281/ 0.317	0.307	0.345	0.282 nested:<0.000	0.269/ 0.684	3.873(0.906),4.155(1.091), 4.306(0.996),3.783(0.718)
Width_child1V	515.45/ 513.08	539.24/ 526.68	0.304/ 0.345	0.332	0.369	0.321 nested:<0.000	0.219/ 0.466	3.63(0.8),3.871(0.924), 3.931(0.831),3.516(0.534)
Width_child2V	339.97/ 334.89	363.76/ 348.48	0.82/ 0.864	0.293	0.322	0.304 nested:<0.000	0.093/ 0.212	3.025(0.644),3.157(0.544), 3.155(0.468),3.034(0.396)
Width_allV	461.69/ 458.37	485.47/ 471.96	0.443/ 0.491	0.362	0.385	0.348 nested:<0.000	0.192/ 0.361	3.509(0.74),3.727(0.81), 3.797(0.706),3.444(0.502)
Angle.BC_V	2190/ 2188	2212/ 2204	0.056/ 0.064	0.053	0.430	0.032 nested:0.205	36.42/ 860.29	119.33(37.67),116.57(38.96), 122.01(45.67),108.87(35.23)
Fractal	103.45/ 99.19	98.51/ 89.06	0.656/ 0.619	0.837	0.891	0.668	0.0027/ 0.0013	1.676(0.085),1.669(0.031), 1.644(0.077),1.642(0.053)
Lacunarity	110.34/ 109.43	85.12/ 83.43	0.345/ 0.355	0.745	0.456	0.653	0.0034/ 0.0018	0.163(0.131),0.156(0.102), 0.177(0.098),0.164(0.065)

[a]: Full versus Restricted (null) model, [b]: chi-squared p-value and p-value based on Satterthwaite's approximation, [c]: Random effect variance (for subjects and, when applicable, the nested factor), divided by the total variance under the full model, [d]: Variance of Subjects/Variance of Residuals

Bibliography

Adamis, A. (2002). Is diabetic retinopathy an inflammatory disease? *British Journal of Ophthalmology,* 86(4)*,* 363–365.

Aiello, L. M. (2003). Perspectives on diabetic retinopathy. *American journal of ophthalmology,* 136(1)*,* 122–135.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer.

Al-Diri, B. (2008). *Analysis of Blood Vessels for Characterization of Diabetic Retinopathy*. PhD thesis*,* University of Lincoln.

Al-Diri, B.*,*Hunter, A.*,* & Steel, D. (2009). An active contour model for segmenting and measuring retinal vessels. *IEEE Transactions on Medical imaging,* 28(9)*,* 1488–1497.

Al-Diri, B.*,*Hunter, A.*,*Steel, D.*,* & Habib, M. (2010). Manual measurement of retinal bifurcation features. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology* (pp. 4760–4764).: IEEE.

Alamouti, B. & Funk, J. (2003). Retinal thickness decreases with age: an oct study. *British journal of ophthalmology,* 87(7)*,* 899–901.

Alberti, K. G. M. M. & Zimmet, P. f. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. part 1: diagnosis and classification of diabetes mellitus. provisional report of a who consultation. *Diabetic medicine,* 15(7)*,* 539–553.

Aletti, M.*,*Gerbeau, J.-F.*,* & Lombardi, D. (2015). Modeling autoregulation in three-dimensional simulations of retinal hemodynamics. *Journal for Modeling in Ophthalmology,* 1.

AmDiabAss *et al.* (2010). Standards of medical care in diabetes—2010. *Diabetes care,* 33(Supplement 1), S11–S61.

Anderson, D. R.,Burnham, K. P., & Thompson, W. L. (2000). Null hypothesis testing: problems, prevalence, and an alternative. *The journal of wildlife management,* (pp. 912–923).

Antcliff, R. & Marshall, J. (1999). The pathogenesis of edema in diabetic maculopathy. In *Seminars in ophthalmology,* volume 14 (pp. 223–232).: Taylor & Francis.

Antonetti, D. A.,Barber, A. J.,Bronson, S. K.,Freeman, W. M.,Gardner, T. W.,Jefferson, L. S.,Kester, M.,Kimball, S. R.,Krady, J. K.,LaNoue, K. F., *et al.* (2006). Diabetic retinopathy seeing beyond glucose-induced microvascular disease. *Diabetes,* 55(9), 2401–2411.

Arend, O.,Remky, A.,Plange, N.,Martin, B., & Harris, A. (2002). Capillary density and retinal diameter measurements and their impact on altered retinal circulation in glaucoma: a digital fluorescein angiographic study. *British journal of ophthalmology,* 86(4), 429–433.

Ashton, N. (1949). Vascular changes in diabetes with particular reference to the retinal vessels: Preliminary report. *The British journal of ophthalmology,* 33(7), 407.

Baayen, R. H.,Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language,* 59(4), 390–412.

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of mathematical psychology,* 12(4), 387–415.

Bank, A. J.,Wilson, R. F.,Kubo, S. H.,Holte, J. E.,Dresing, T. J., & Wang, H. (1995). Direct effects of smooth muscle relaxation and contraction on in vivo human brachial artery elastic properties. *Circulation research,* 77(5), 1008–1016.

Barber, A. J.,Gardner, T. W., & Abcouwer, S. F. (2011). The significance of vascular and neural apoptosis to the pathology of diabetic retinopathy. *Investigative ophthalmology & visual science,* 52(2), 1156–1163.

Bates, D. (2007). Linear mixed model implementation in lme4. *Manuscript, University of Wisconsin,* 15.

Bates, D. (2014). Computational methods for mixed models. *LME4: Mixed-Effects Modeling with R,* (pp. 99–118).

Bates, D.,Mächler, M.,Bolker, B., & Walker, S. (2014a). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bates, D.,Maechler, M.,Bolker, B.,Walker, S., *et al.* (2014b). lme4: Linear mixed-effects models using eigen and s4. *R package version,* 1(7).

Bates, D.,Maechler, M.,Bolker, B.,Walker, S.,Christensen, R. H. B.,Singmann, H.,Dai, B.,Grothendieck, G.,Green, P., & Bolker, M. B. (2016). Package 'lme4'.

Baynes, J. W. (1991). Role of oxidative stress in development of complications in diabetes. *Diabetes,* 40(4), 405–412.

Beckman, J. A.,Creager, M. A., & Libby, P. (2002). Diabetes and atherosclerosis: epidemiology, pathophysiology, and management. *Jama,* 287(19), 2570–2581.

Beckman, J. A.,Goldfine, A. B.,Gordon, M. B., & Creager, M. A. (2001). Ascorbate restores endothelium-dependent vasodilation impaired by acute hyperglycemia in humans. *Circulation,* 103(12), 1618–1623.

Bek, T. (1999). Diabetic maculopathy caused by disturbances in retinal vasomotion. a new hypothesis. *Acta Ophthalmologica Scandinavica,* 77(4), 376–380.

Bek, T. & Helgesen, A. (2001). The regional distribution of diabetic retinopathy lesions may reflect risk factors for progression of the disease. *Acta Ophthalmologica Scandinavica,* 79(5), 501–505.

Ben-David, A. (2008). Comparison of classification accuracy using cohen's weighted kappa. *Expert Systems with Applications,* 34(2), 825–832.

Bertelsen, G.,Erke, M. G.,von Hanno, T.,Mathiesen, E. B.,Peto, T.,Sjølie, A. K., & Njølstad, I. (2013). The tromsø eye study: study design, methodology and results on visual acuity and refractive errors. *Acta ophthalmologica,* 91(7), 635–642.

Besharse, J. C. & Bok, D. (2011). *The retina and its disorders*. Academic Press.

Blackwelder, W. C. (1982). "proving the null hypothesis" in clinical trials. *Controlled clinical trials,* 3(4), 345–353.

Bland, J. M. & Altman, D. G. (1995). Multiple significance tests: the bonferroni method. *Bmj,* 310(6973), 170.

Bourne, R. R.,Stevens, G. A.,White, R. A.,Smith, J. L.,Flaxman, S. R.,Price, H.,Jonas, J. B.,Keeffe, J.,Leasher, J.,Naidoo, K., *et al.* (2013). Causes of vision loss worldwide, 1990–2010: a systematic analysis. *The lancet global health,* 1(6), e339–e349.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern recognition,* 30(7), 1145–1159.

Breiman, L. (2001). Random forests. *Machine learning,* 45(1), 5–32.

Burgansky-Eliash, Z. (2012). *The Effect of Diabetes Mellitus on Retinal Function*. INTECH Open Access Publisher.

Burgansky-Eliash, Z.,Barak, A.,Barash, H.,Nelson, D. A.,Pupko, O.,Lowenstein, A.,Grinvald, A., & Rubinstein, A. (2012). Increased retinal blood flow velocity in patients with early diabetes mellitus. *Retina,* 32(1), 112–119.

Burgansky-Eliash, Z.,Nelson, D. A.,Bar-Tal, O. P.,Lowenstein, A.,Grinvald, A., & Barak, A. (2010). Reduced retinal blood flow velocity in diabetic retinopathy. *Retina,* 30(5), 765–773.

Bursell, S.-E.,Clermont, A. C.,Kinsley, B. T.,Simonson, D. C.,Aiello, L. M., & Wolpert, H. A. (1996). Retinal blood flow changes in patients with insulin-dependent diabetes mellitus and no diabetic retinopathy. *Investigative ophthalmology & visual science,* 37(5), 886–897.

Bursell, S.-E.,Takagi, C.,Clermont, A. C.,Takagi, H.,Mori, F.,Ishii, H., & King, G. L. (1997). Specific retinal diacylglycerol and protein kinase c beta isoform modulation mimics abnormal retinal hemodynamics in diabetic rats. *Investigative ophthalmology & visual science,* 38(13), 2711–2720.

Cai, J. & Boulton, M. (2002). The pathogenesis of diabetic retinopathy: old concepts and new questions. *Eye,* 16(3), 242–260.

Calivá, F.,Aletti, M.,Al-Diri, B., & Hunter, A. (2015). A new tool to connect blood vessels in fundus retinal images. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4343–4346).: IEEE.

Calles-Escandon, J. & Cipolla, M. (2001). Diabetes and endothelial dysfunction: a clinical perspective. *Endocrine reviews,* 22(1), 36–52.

Callewaert, B.,Willaert, A.,Kerstjens-Frederikse, W.,De Backer, J.,Devriendt, K.,Albrecht, B.,Ramos-Arroyo, M.,Doco-Fenzy, M.,Hennekam, R.,Pyeritz, R., *et al.* (2008). Arterial tortuosity syndrome: clinical and molecular findings in 12 newly identified families. *Human mutation,* 29(1), 150–158.

Causin, P.,Guidoboni, G.,Malgaroli, F.,Sacco, R., & Harris, A. (2016). Blood flow mechanics and oxygen transport and delivery in the retinal microcirculation: multiscale mathematical modeling and numerical simulation. *Biomechanics and modeling in mechanobiology,* 15(3), 525–542.

Chandrinos, K.,Pilu, M.,Fisher, R., & Trahanias, P. (1998). *Image processing techniques for the quantification of atherosclerotic changes*. Department of Artificial Intelligence, University of Edinburgh.

Chapman, N.,Dell'Omo, G.,Sartini, M.,Witt, N.,Hughes, A.,Thom, S., & Pedrinelli, R. (2002). Peripheral vascular disease is associated with abnormal arteriolar diameter relationships at bifurcations in the human retina. *Clinical Science,* 103(2), 111–116.

Cheung, C. Y.-l.,Lamoureux, E.,Ikram, M. K.,Sasongko, M. B.,Ding, J.,Zheng, Y.,Mitchell, P.,Wang, J. J., & Wong, T. Y. (2012). Retinal vascular geometry in asian persons with diabetes and retinopathy. *Journal of diabetes science and technology,* 6(3), 595–605.

Cheung, M. W.-L. (2013). Implementing restricted maximum likelihood estimation in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal,* 20(1), 157–167.

Chew, S. K.,Colville, D.,Canty, P.,Hutchinson, A.,Wong, A.,Luong, V.,Wong, T. Y.,McDonald, C., & Savige, J. (2016). Hypertensive/microvascular disease and copd: a case control study. *Kidney and Blood Pressure Research,* 41(1), 29–39.

Christodoulidis, A.,Hurtut, T.,Tahar, H. B., & Cheriet, F. (2016). A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images. *Computerized Medical Imaging and Graphics*.

Ciulla, T. A.,Amador, A. G., & Zinman, B. (2003). Diabetic retinopathy and diabetic macular edema pathophysiology, screening, and novel therapies. *Diabetes care,* 26(9), 2653–2664.

Clermont, A. C. & Bursell, S.-E. (2007). Retinal blood flow in diabetes. *Microcirculation,* 14(1), 49–61.

Cogan, D. G.,Toussaint, D., & Kuwabara, T. (1961). Retinal vascular patterns: Iv. diabetic retinopathy. *Archives of Ophthalmology,* 66(3), 366–378.

Corbeil, R. R. & Searle, S. R. (1976). Restricted maximum likelihood (reml) estimation of variance components in the mixed model. *Technometrics,* 18(1), 31–38.

Creager, M. A.,Lüscher, T. F.,Cosentino, F.,Beckman, J. A., *et al.* (2003). Diabetes and vascular disease pathophysiology, clinical consequences, and medical therapy: part i. *Circulation,* 108(12), 1527–1532.

Cui, Y.,Kim, D.-S., & Park, K.-C. (2005). Antioxidant effect of inonotus obliquus. *Journal of Ethnopharmacology,* 96(1), 79–85.

Cunha-Vaz, J. G.,Fonseca, J. R.,de Abreu, J. R., & Lima, J. J. (1978). Studies on retinal blood flow: Ii. diabetic retinopathy. *Archives of ophthalmology,* 96(5), 809–811.

Cutler, D. R.,Edwards, T. C.,Beard, K. H.,Cutler, A.,Hess, K. T.,Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology,* 88(11), 2783–2792.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap methods and their application,* volume 1. Cambridge university press.

Daxer, A. (1993). The fractal geometry of proliferative diabetic retinopathy: implications for the diagnosis and the process of retinal vasculogenesis. *Current eye research,* 12(12), 1103–1109.

De La Rubia, G.,Oliver, F. J.,Inoguchi, T., & King, G. L. (1992). Induction of resistance to endothelin-1's biochemical actions by elevated glucose levels in retinal pericytes. *Diabetes,* 41(12), 1533–1539.

Del Corso, L.,Moruzzo, D.,Conte, B.,Agelli, M.,Romanelli, A. M.,Pastine, F.,Protti, M.,Pentimone, F., & Baggiani, G. (1998). Tortuosity, kinking, and coiling of the carotid artery: expression of atherosclerosis or aging? *Angiology,* 49(5), 361–371.

Deng, H. & Runger, G. (2012). Feature selection via regularized trees. In *The 2012 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–8).: IEEE.

Deng, H. & Runger, G. (2013). Gene selection with guided regularized random forest. *Pattern Recognition,* 46(12), 3483–3489.

Devijver, P. A. & Kittler, J. (1982). *Pattern recognition: A statistical approach*. Prentice hall.

DiabConCompTr (1993). The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl j Med,* 1993(329), 977–986.

DiCiccio, T. J. & Efron, B. (1996). Bootstrap confidence intervals. *Statistical science,* (pp. 189–212).

Do, M. T. H. & Yau, K.-W. (2010). Intrinsically photosensitive retinal ganglion cells. *Physiological reviews,* 90(4), 1547–1581.

Dobree, J. (1970). Simple diabetic retinopathy. evolution of the lesions and therapeutic considerations. *The British journal of ophthalmology,* 54(1), 1.

Donnelly, R. & Horton, E. (2008). *Vascular complications of diabetes: current issues in pathogenesis and treatment*. John Wiley & Sons.

Dowling, J. E. (1987). *The retina: an approachable part of the brain*. Harvard University Press.

Dubrofsky, E. (2009). *Homography estimation*. PhD thesis, UNIVERSITY OF BRITISH COLUMBIA (Vancouver.

Efron, B. & Tibshirani, R. (1997). Improvements on cross-validation: the 632+ bootstrap method. *Journal of the American Statistical Association,* 92(438), 548–560.

El-Remessy, A. B.,Bartoli, M.,Platt, D. H.,Fulton, D., & Caldwell, R. B. (2005). Oxidative stress inactivates vegf survival signaling in retinal endothelial cells via pi 3-kinase tyrosine nitration. *Journal of cell science,* 118(1), 243–252.

Engerman, R. L. & Kern, T. S. (1984). Experimental galactosemia produces diabetic-like retinopathy. *Diabetes,* 33(1), 97–100.

Engerman, R. L. & Kern, T. S. (1993). Aldose reductase inhibition fails to prevent retinopathy in diabetic and galactosemic dogs. *Diabetes,* 42(6), 820–825.

Eva, N. H. A. W. Z. & Mansour, F. A. (1998). Prevention of pericyte loss by trolox in diabetic rat retina. *Journal of Toxicology and Environmental Health Part A,* 54(6), 467–475.

Family, F. & Vicsek, T. (1991). *Dynamics of fractal surfaces*. World Scientific.

Feke, G. T.,Tagawa, H.,Deupree, D. M.,Goger, D. G.,Sebag, J., & Weiter, J. (1989). Blood flow in the normal human retina. *Investigative ophthalmology & visual science,* 30(1), 58–65.

Fischler, M. A. & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM,* 24(6), 381–395.

Fisher, R. A. (1925). *Statistical methods for research workers*. Genesis Publishing Pvt Ltd.

Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics,* 7(2), 179–188.

Foroutan-Pour, K.,Dutilleul, P., & Smith, D. (1999). Advances in the implementation of the box-counting method of fractal dimension estimation. *Applied mathematics and computation,* 105(2), 195–210.

Fox, J. (2002). *Linear Mixed Models: Appendix to An R and S-Plus companion to applied regression*. Sage.

Frank, R. N. (1995). Diabetic retinopathy. *Progress in Retinal and Eye Research,* 14, 361–392.

Friedman, J.,Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software,* 33(1), 1.

Fuchsjäger-Mayrl, G.,Polak, K.,Luksch, A.,Polska, E.,Dorner, G. T.,Rainer, G.,Eichler, H.-G., & Schmetterer, L. (2001). Retinal blood flow and systemic blood pressure in healthy young subjects. *Graefe's archive for clinical and experimental ophthalmology,* 239(9), 673–677.

Funk, R. (1997). Blood supply of the retina. *Ophthalmic research,* 29(5), 320–325.

Garcia Jr, J. P.,Garcia, P. T., & Rosen, R. B. (2002). Retinal blood flow in the normal human eye using the canon laser blood flowmeter. *Ophthalmic research,* 34(5), 295–299.

Gardner, T. W.,Antonetti, D. A.,Barber, A. J.,LaNoue, K. F.,Levison, S. W.,Group, P. S. R. R., *et al.* (2002). Diabetic retinopathy: more than meets the eye. *Survey of ophthalmology,* 47, S253–S262.

Garhofer, G.,Zawinka, C.,Resch, H.,Huemer, K.,Schmetterer, L., & Dorner, G. (2004a). Response of retinal vessel diameters to flicker stimulation in patients with early open angle glaucoma. *Journal of glaucoma,* 13(4), 340–344.

Garhofer, G.,Zawinka, C.,Resch, H.,Kothy, P.,Schmetterer, L., & Dorner, G. (2004b). Reduced response of retinal vessel diameters to flicker stimulation in patients with diabetes. *British Journal of Ophthalmology,* 88(7), 887–891.

Gariano, R.,Iruela-Arispe, M.,Sage, E., & Hendrickson, A. (1996). Immunohistochemical characterization of developing and mature primate retinal blood vessels. *Investigative ophthalmology & visual science,* 37(1), 93–103.

Geisser, S. (1993). *Predictive inference,* volume 55. CRC press.

Girden, E. R. (1992). *ANOVA: Repeated measures*. Number 84. Sage.

Gobel, W. & Lieb, W. (1995). Changes in orbital hemodynamics caused by nitroglycerin and nifedipine. a study using color duplex ultrasound. *Ophthalmologe,* 92, 206–211.

Golpe, R.,Mateos-Colino, A.,Testa-Fernández, A.,Pena-Seijo, M.,Rodríguez-Enríquez, M.,González-Juanatey, C.,Martín-Vázquez, F. J.,Pose-Reino, A.,Domínguez-Pin, N.,Garnacho-Gayarre, N., *et al.* (2016). Blood pressure profile and hypertensive organ damage in copd patients and matched controls. the retapoc study. *PloS one,* 11(6), e0157932.

Grisan, E.,Foracchia, M., & Ruggeri, A. (2008). A novel method for the automatic grading of retinal vessel tortuosity. *IEEE transactions on medical imaging,* 27(3), 310–319.

Gross, E. R.,LaDisa, J. F.,Weihrauch, D.,Olson, L. E.,Kress, T. T.,Hettrick, D. A.,Pagel, P. S.,Warltier, D. C., & Kersten, J. R. (2003). Reactive oxygen species modulate coronary wall shear stress and endothelial function during hyperglycemia. *American Journal of Physiology-Heart and Circulatory Physiology,* 284(5), H1552–H1559.

Group, E. D. P. R. *et al.* (2004). The prevalence of diabetic retinopathy among adults in the united states. *Archives of ophthalmology,* 122(4), 552.

Group, E. T. D. R. S. R. *et al.* (1987). Treatment techniques and clinical guidelines for photocoagulation of diabetic macular edema: Early treatment diabetic retinopathy study report number 2. *Ophthalmology,* 94(7), 761–774.

Group, E. T. D. R. S. R. *et al.* (1991). Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified airlie house classification: Etdrs report number 10. *Ophthalmology,* 98(5), 786–806.

Grunwald, J. E.,Maguire, A. M., & Dupont, J. (1996). Retinal hemodynamics in retinitis pigmentosa. *American journal of ophthalmology,* 122(4), 502–508.

Grunwald, J. E.,Riva, C. E.,Baine, J., & Brucker, A. J. (1992). Total retinal volumetric blood flow rate in diabetic patients with poor glycemic control. *Investigative ophthalmology & visual science,* 33(2), 356–363.

Guo, Y.,Logan, H. L.,Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC medical research methodology,* 13(1), 100.

Guran, T.,Zeimer, R. C.,Shahidi, M., & Mori, M. T. (1990). Quantitative analysis of retinal hemodynamics using targeted dye delivery. *Investigative ophthalmology & visual science,* 31(11), 2300–2306.

Habib, M. S.,Al-Diri, B.,Hunter, A., & Steel, D. H. (2014). The association between retinal vascular geometry changes and diabetic retinopathy and their role in prediction of progression–an exploratory study. *BMC ophthalmology,* 14(1), 1.

Hammer, M.,Vilser, W.,Riemer, T.,Mandecka, A.,Schweitzer, D.,Kühn, U.,Dawczynski, J.,Liemt, F., & Strobel, J. (2009). Diabetic patients with retinopathy show increased retinal venous oxygen saturation. *Graefe's Archive for Clinical and Experimental Ophthalmology,* 247(8), 1025–1030.

Hammes, H.-P.,Lin, J.,Renner, O.,Shani, M.,Lundqvist, A.,Betsholtz, C.,Brownlee, M., & Deutsch, U. (2002). Pericytes and the pathogenesis of diabetic retinopathy. *Diabetes,* 51(10), 3107–3112.

Han, H.-C. (2012). Twisted blood vessels: symptoms, etiology and biomechanical mechanisms. *Journal of vascular research,* 49(3), 185–197.

Hanley, J. A. & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology,* 143(1), 29–36.

Hardarson, S. H. & Stefánsson, E. (2012). Retinal oxygen saturation is altered in diabetic retinopathy. *British journal of ophthalmology,* 96(4), 560–563.

Harhaj, N. S. & Antonetti, D. A. (2004). Regulation of tight junctions and loss of barrier function in pathophysiology. *The international journal of biochemistry & cell biology,* 36(7), 1206–1237.

Hart, W. E.,Goldbaum, M.,Côté, B.,Kube, P., & Nelson, M. R. (1999). Measurement and classification of retinal vascular tortuosity. *International journal of medical informatics,* 53(2), 239–252.

Hartnett, M. E.,Stratton, R. D.,Browne, R. W.,Rosner, B. A.,Lanham, R. J., & Armstrong, D. (2000). Serum markers of oxidative stress and severity of diabetic retinopathy. *Diabetes care,* 23(2), 234–240.

Hayes, T. C. & Horowitz, P. (1989). *Student manual for the art of electronics*. Cambridge university press Cambridge.

Hayreh, S. S. (1969). Blood supply of the optic nerve head and its role in optic atrophy, glaucoma, and oedema of the optic disc. *The British journal of ophthalmology,* 53(11), 721.

Hayreh, S. S.,Zimmerman, B.,McCarthy, M. J., & Podhajsky, P. (2001). Systemic diseases associated with various types of retinal vein occlusion. *American journal of ophthalmology,* 131(1), 61–77.

Heitmar, R.,Blann, A. D.,Cubbidge, R. P.,Lip, G. Y., & Gherghel, D. (2010). Continuous retinal vessel diameter measurements: the future in retinal vessel assessment? *Investigative ophthalmology & visual science,* 51(11), 5833–5839.

Heneghan, C.,Flynn, J.,O'Keefe, M., & Cahill, M. (2002). Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis. *Medical image analysis,* 6(4), 407–429.

Hildebrand, G. D. & Fielder, A. R. (2011). Anatomy and physiology of the retina. In *Pediatric retina* (pp. 39–65). Springer.

Hill, M. A.,Meininger, G. A.,Davis, M. J., & Laher, I. (2009). Therapeutic potential of pharmacologically targeting arteriolar myogenic tone. *Trends in pharmacological sciences,* 30(7), 363–374.

Hiroki, M.,Miyashita, K., & Oda, M. (2002). Tortuosity of the white matter medullary arterioles is related to the severity of hypertension. *Cerebrovascular Diseases,* 13(4), 242–250.

Hove, M. N.,Kristensen, J. K.,Lauritzen, T., & Bek, T. (2004). Quantitative analysis of retinopathy in type 2 diabetes: identification of prognostic parameters for developing visual loss secondary to diabetic maculopathy. *Acta Ophthalmologica Scandinavica,* 82(6), 679–685.

Hove, M. N.,Kristensen, J. K.,Lauritzen, T., & Bek, T. (2006). The relationships between risk factors and the distribution of retinopathy lesions in type 2 diabetes. *Acta Ophthalmologica Scandinavica,* 84(5), 619–623.

Howell, D. C. (2006). Multiple comparisons with repeated measures. *Retrieved Jan,* 8, 2007.

Hubbard, L. D.,Brothers, R. J.,King, W. N.,Clegg, L. X.,Klein, R.,Cooper, L. S.,Sharrett, A. R.,Davis, M. D.,Cai, J.,in Communities Study Group, A. R., *et al.* (1999). Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study. *Ophthalmology,* 106(12), 2269–2280.

Hubbert, M. K. *et al.* (1956). *Darcy's law and the field equations of the flow of underground fluids*. Shell Development Company, Exploration and Production Research Division.

Hunter, A.,Lowell, J.,Ryder, R.,Basu, A.,Steel, D., *et al.* (2005). Tram-line filtering for retinal vessel segmentation.

Ikram, M. K.,Cheung, C. Y.,Lorenzi, M.,Klein, R.,Jones, T. L.,Wong, T. Y., *et al.* (2013). Retinal vascular caliber as a biomarker for diabetes microvascular complications. *Diabetes Care,* 36(3), 750–759.

Islam, F.,Nguyen, T.,Wang, J.,Tai, E.,Shankar, A.,Saw, S.,Aung, T.,Lim, S.,Mitchell, P., & Wong, T. (2009). Quantitative retinal vascular calibre changes in diabetes and retinopathy: the singapore malay eye study. *Eye,* 23(8), 1719–1724.

Jaccard, J.,Becker, M. A., & Wood, G. (1984). Pairwise multiple comparison procedures: a review. *Psychological Bulletin,* 96(3), 589.

Jorgensen, C. M.,Hardarson, S. H., & Bek, T. (2014). The oxygen saturation in retinal vessels from diabetic patients depends on the severity and type of vision-threatening retinopathy. *Acta ophthalmologica,* 92(1), 34–39.

Joussen, A. M.,Poulaki, V.,Le, M. L.,Koizumi, K.,Esser, C.,Janicki, H.,Schraermeyer, U.,Kociok, N.,Fauser, S.,Kirchhof, B., *et al.* (2004). A central role for inflammation in the pathogenesis of diabetic retinopathy. *The FASEB journal,* 18(12), 1450–1452.

Joussen, A. M.,Poulaki, V.,Mitsiades, N.,Kirchhof, B.,Koizumi, K.,Döhmen, S., & Adamis, A. P. (2002). Nonsteroidal anti-inflammatory drugs prevent early diabetic retinopathy via tnf-α suppression. *The FASEB journal,* 16(3), 438–440.

Kador, P. F.,Akagi, Y.,Takahashi, Y.,Ikebe, H.,Wyman, M., & Kinoshita, J. H. (1990). Prevention of retinal vessel changes associated with diabetic retinopathy in galactose-fed dogs by aldose reductase inhibitors. *Archives of ophthalmology,* 108(9), 1301–1309.

Kaiser, H. J.,Schoetzau, A., & Flammer, J. (1997). Blood flow velocity in the extraocular vessels in chronic smokers. *British journal of ophthalmology,* 81(2), 133–135.

Kalitzeos, A. A.,Lip, G. Y., & Heitmar, R. (2013). Retinal vessel tortuosity measures and their applications. *Experimental eye research,* 106, 40–46.

Kamiya, A. & Takahashi, T. (2007). Quantitative assessments of morphological and functional properties of biological trees based on their fractal nature. *Journal of Applied Physiology,* 102(6), 2315–2323.

Kamiya, A. & Togawa, T. (1972). Optimal branching structure of the vascular tree. *The Bulletin of mathematical biophysics,* 34(4), 431–438.

Karch, R.,Neumann, F.,Neumann, M., & Schreiner, W. (2000). Staged growth of optimized arterial model trees. *Annals of biomedical engineering,* 28(5), 495–511.

Kassab, G. S. & Fung, Y.-C. B. (1995). The pattern of coronary arteriolar bifurcations and the uniform shear hypothesis. *Annals of biomedical engineering,* 23(1), 13–20.

Kawagishi, T.,Nishizawa, Y.,Emoto, M.,Konishi, T.,Maekawa, K.,Hagiwara, S.,Okuno, Y.,Inada, H.,Isshiki, G., & Morii, H. (1995). Impaired retinal artery blood

flow in iddm patients before clinical manifestations of diabetic retinopathy. *Diabetes Care,* 18(12), 1544–1549.

Kenward, M. G. & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics,* (pp. 983–997).

Kern, T. S. & Engerman, R. L. (1996). Capillary lesions develop in retina rather than cerebral cortex in diabetes and experimental galactosemia. *Archives of ophthalmology,* 114(3), 306–310.

Keyser, B. J.,Flaharty, P. M.,Sergott, R. C.,Brown, G. C.,Lieb, W. E., & Annesley, W. H. (1994). Color doppler imaging of arterial blood flow in central retinal vein occlusion. *Ophthalmology,* 101(8), 1357–1361.

Khoobehi, B.,Firn, K.,Thompson, H.,Reinoso, M., & Beach, J. (2013). Retinal arterial and venous oxygen saturation is altered in diabetic patientsretinal oxygen saturation trend in diabetes. *Investigative ophthalmology & visual science,* 54(10), 7103–7106.

Kifley, A.,Wang, J. J.,Cugati, S.,Wong, T. Y., & Mitchell, P. (2007). Retinal vascular caliber, diabetes, and retinopathy. *American journal of ophthalmology,* 143(6), 1024–1026.

Kim, D. Y.,Fingler, J.,Werner, J. S.,Schwartz, D. M.,Fraser, S. E., & Zawadzki, R. J. (2011). In vivo volumetric imaging of human retinal circulation with phase-variance optical coherence tomography. *Biomedical optics express,* 2(6), 1504–1513.

Kim, K. E.,Kim, D. M.,Flammer, J., & Kim, K. N. (2015). Central retinal venous pressure in eyes of normal-tension glaucoma patients with optic disc hemorrhage. *PloS one,* 10(5), e0127920.

King, G. L.,Shiba, T.,Oliver, J.,Inoguchi, T., & Bursell, S.-E. (1994). Cellular and molecular abnormalities in the vascular endothelium of diabetes mellitus. *Annual review of medicine,* 45(1), 179–188.

Kirby, B. J. (2010). *Micro-and nanoscale fluid mechanics: transport in microfluidic devices*. Cambridge University Press.

Kitabchi, A. E.,Umpierrez, G. E.,Miles, J. M., & Fisher, J. N. (2009). Hyperglycemic crises in adult patients with diabetes. *Diabetes care,* 32(7), 1335–1343.

Klein, R.,Klein, B. E., & Moss, S. E. (1984a). Visual impairment in diabetes. *Ophthalmology,* 91(1), 1–9.

Klein, R.,Klein, B. E.,Moss, S. E., & Cruickshanks, K. J. (1995). The wisconsin epidemiologic study of diabetic retinopathy xv: the long-term incidence of macular edema. *Ophthalmology,* 102(1), 7–16.

Klein, R.,Klein, B. E.,Moss, S. E., & Cruickshanks, K. J. (1998). The wisconsin epidemiologic study of diabetic retinopathy: Xvii: The 14-year incidence and progression of diabetic retinopathy and associated risk factors in type 1 diabetes11proprietary interest: none. *Ophthalmology,* 105(10), 1801–1815.

Klein, R.,Klein, B. E.,Moss, S. E.,Davis, M. D., & DeMets, D. L. (1984b). The wisconsin epidemiologic study of diabetic retinopathy: Ii. prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years. *Archives of ophthalmology,* 102(4), 520–526.

Klein, R.,Klein, B. E.,Moss, S. E.,Davis, M. D., & DeMets, D. L. (1984c). The wisconsin epidemiologic study of diabetic retinopathy: Iii. prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years. *Archives of ophthalmology,* 102(4), 527–532.

Klein, R.,Klein, B. E.,Moss, S. E.,Davis, M. D., & DeMets, D. L. (1984d). The wisconsin epidemiologic study of diabetic retinopathy: Iv. diabetic macular edema. *Ophthalmology,* 91(12), 1464–1474.

Klein, R.,Klein, B. E.,Moss, S. E.,Wong, T. Y.,Hubbard, L.,Cruickshanks, K. J., & Palta, M. (2004). The relation of retinal vessel caliber to the incidence and progressionof diabetic retinopathy: Xix: The wisconsin epidemiologic study of diabetic retinopathy. *Archives of ophthalmology,* 122(1), 76–83.

Klein, R.,Knudtson, M. D.,Lee, K. E.,Gangnon, R., & Klein, B. E. (2008). The wisconsin epidemiologic study of diabetic retinopathy xxii: the twenty-five-year progression of retinopathy in persons with type 1 diabetes. *Ophthalmology,* 115(11), 1859–1868.

Klein, R.,Moss, S. E.,Klein, B. E.,Dams, M. D., & DeMets, D. L. (1989). The wisconsin epidemiologic study of diabetic retinopathy: Xi. the incidence of macular edema. *Ophthalmology,* 96(10), 1501–1510.

Klein, R.,Myers, C. E.,Lee, K. E.,Gangnon, R., & Klein, B. E. (2012). Changes in retinal vessel diameter and incidence and progression of diabetic retinopathy. *Archives of Ophthalmology,* 130(6), 749–755.

Klein, R.,Sharrett, A. R.,Klein, B. E.,Chambless, L. E.,Cooper, L. S.,Hubbard, L. D., & Evans, G. (2000). Are retinal arteriolar abnormalities related to atherosclerosis? the atherosclerosis risk in communities study. *Arteriosclerosis, Thrombosis, and Vascular Biology,* 20(6), 1644–1650.

Knowler, W. C.,Bennett, P. H., & Ballintine, E. J. (1980). Increased incidence of retinopathy in diabetics with elevated blood pressure: a six-year follow-up study in pima indians. *New England Journal of Medicine,* 302(12), 645–650.

Knudtson, M. D.,Lee, K. E.,Hubbard, L. D.,Wong, T. Y.,Klein, R., & Klein, B. E. (2003). Revised formulas for summarizing retinal vessel diameters. *Current eye research,* 27(3), 143–149.

Kohavi, R. *et al.* (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai,* volume 14 (pp. 1137–1145).

Kohner, E. M.,Patel, V., & Rassam, S. M. (1995). Role of blood flow and impaired autoregulation in the pathogenesis of diabetic retinopathy. *Diabetes,* 44(6), 603–607.

Konno, S.,Feke, G. T.,Yoshida, A.,Fujio, N.,Goger, D. G., & Buzney, S. M. (1996). Retinal blood flow changes in type i diabetes. a long-term follow-up study. *Investigative ophthalmology & visual science,* 37(6), 1140–1148.

Kowluru, R. A. & Chan, P.-S. (2007). Oxidative stress and diabetic retinopathy. *Journal of Diabetes Research,* 2007.

Kriegman, D. (2007). Homography estimation. *UCSD Computer Science and Engineering*.

Krueger, C. & Tian, L. (2004). A comparison of the general linear mixed model and repeated measures anova using a dataset with multiple missing data points. *Biological research for nursing,* 6(2), 151–157.

Kuhn, M. (2016). *caret: Classification and Regression Training*. R package version 6.0-68.

Kunicki, A.,Oliveira, A.,Mendonça, M.,Barbosa, C., & Nogueira, R. (2009). Can the fractal dimension be applied for the early diagnosis of non-proliferative diabetic retinopathy? *Brazilian Journal of Medical and Biological Research,* 42(10), 930–934.

Kursa, M. B. & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software,* 36(11), 1–13.

Kursa, M. B.,Rudnicki, W. R., *et al.* (2010). Feature selection with the boruta package.

Kuznetsova, A.,Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmertest'. *R package version,* (pp. 2–0).

Laird, N. M. & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics,* (pp. 963–974).

Lee, C. S.,Lee, A. Y.,Sim, D. A.,Keane, P. A.,Mehta, H.,Zarranz-Ventura, J.,Fruttiger, M.,Egan, C. A., & Tufail, A. (2015a). Reevaluating the definition of intraretinal microvascular abnormalities and neovascularization elsewhere in diabetic retinopathy using optical coherence tomography and fluorescein angiography. *American journal of ophthalmology,* 159(1), 101–110.

Lee, R.,Wong, T. Y., & Sabanayagam, C. (2015b). Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss. *Eye and Vision,* 2(1), 1.

Leishman, R. (1957). The eye in general vascular disease hypertension and arteriosclerosis. *British Journal of Ophthalmology,* 41(11), 641–701.

Leontidis, G.,Al-Diri, B., & Hunter, A. (2014). Diabetic retinopathy: current and future methods for early screening from a retinal hemodynamic and geometric approach. *Expert Review of Ophthalmology,* 9(5), 431–442.

Leontidis, G.,Al-Diri, B., & Hunter, A. (2016a). Exploiting the retinal vascular geometry in identifying the progression to diabetic retinopathy using penalized logistic regression and random forests. In *Emerging Trends and Advanced Technologies for Computational Intelligence* (pp. 381–400). Springer.

Leontidis, G.,Al-Diri, B., & Hunter, A. (2016b). Summarising the retinal vascular calibres in healthy, diabetic and diabetic retinopathy eyes. *Computers in biology and medicine,* 72, 65–74.

Leontidis, G.,Al-Diri, B.,Wigdahl, J., & Hunter, A. (2015a). Evaluation of geometric features as biomarkers of diabetic retinopathy for characterizing the retinal vascular changes during the progression of diabetes. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (pp. 5255–5259).

Leontidis, G.,Wigdahl, J.,Al-Diri, B.,Ruggeri, A., & Hunter, A. (2015b). Evaluating tortuosity in retinal fundus images of diabetic patients who progressed to diabetic retinopathy. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*.

Li, J.,Du, Q., & Sun, C. (2009). An improved box-counting method for image fractal dimension estimation. *Pattern Recognition,* 42(11), 2460–2469.

Li, Q.,Feng, B.,Xie, L.,Liang, P.,Zhang, H., & Wang, T. (2016). A cross-modality learning approach for vessel segmentation in retinal images. *IEEE transactions on medical imaging,* 35(1), 109–118.

Libby P, Ridker PM, M. A. (2002). Inflammation and atherosclerosis. *Circulation,* 105(9), 1135.

Little, R. J. & Rubin, D. B. (2014). *Statistical analysis with missing data*. John Wiley & Sons.

Locker, L.,Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods,* 39(4), 723–730.

Lorenzi, M.,Feke, G. T.,Pitler, L.,Berisha, F.,Kolodjaschna, J., & McMeel, J. W. (2010). Defective myogenic response to posture change in retinal vessels of well-controlled type 1 diabetic patients with no retinopathy. *Investigative ophthalmology & visual science,* 51(12), 6770–6775.

Lotmar, W.,Freiburghaus, A., & Bracher, D. (1979). Measurement of vessel tortuosity on fundus photographs. *Albrecht von Graefes Archiv für klinische und experimentelle Ophthalmologie,* 211(1), 49–57.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on,* volume 2 (pp. 1150–1157).: Ieee.

Lowell, J.,Hunter, A.,Steel, D.,Basu, A.,Ryder, R.,Fletcher, E., & Kennedy, L. (2004a). Optic nerve head segmentation. *IEEE Transactions on medical Imaging,* 23(2), 256–264.

Lowell, J.,Hunter, A.,Steel, D.,Basu, A.,Ryder, R., & Kennedy, R. L. (2004b). Measurement of retinal vessel widths from fundus images based on 2-d modeling. *IEEE transactions on medical imaging,* 23(10), 1196–1204.

Lu, X. & Kassab, G. (2004). Nitric oxide is significantly reduced in ex vivo porcine arteries during reverse flow because of increased superoxide production. *The Journal of physiology,* 561(2), 575–582.

Lupascu, C. A. (2010). Human visual perception and retinal diseases. In *proc. of CREATE 2010 Conference, Gjøvik, Norway* (pp. 102–106).

Mainster, M. A. (1990). The fractal properties of retinal vessels: embryological and clinical implications. *Eye,* 4(1), 235–241.

Mandecka, A.,Dawczynski, J.,Blum, M.,Müller, N.,Kloos, C.,Wolf, G.,Vilser, W.,Hoyer, H., & Müller, U. A. (2007). Influence of flickering light on the retinal vessels in diabetic patients. *Diabetes care,* 30(12), 3048–3052.

Mandelbrot, B. B. (1977). *Fractals*. Wiley Online Library.

Mandelbrot, B. B. (1983). *The fractal geometry of nature,* volume 173. Macmillan.

Matlab (2014). *version 8.4.0 (R2014b)*. Natick, Massachusetts: The MathWorks Inc.

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics,* 11(2), 204–209.

McCulloch, C. E. & Neuhaus, J. M. (2001). *Generalized linear mixed models*. Wiley Online Library.

McLean, R. A.,Sanders, W. L., & Stroup, W. W. (1991). A unified approach to mixed linear models. *The American Statistician,* 45(1), 54–64.

Miyamoto, K. & Ogura, Y. (1999). Pathogenetic potential of leukocytes in diabetic retinopathy. In *Seminars in ophthalmology,* volume 14 (pp. 233–239).: Taylor & Francis.

Murray, C. D. (1926a). The physiological principle of minimum work applied to the angle of branching of arteries. *The Journal of general physiology,* 9(6), 835–841.

Murray, C. D. (1926b). The physiological principle of minimum work: I. the vascular system and the cost of blood volume. *Proceedings of the National Academy of Sciences of the United States of America,* 12(3), 207.

Navarro, R. (2009). The optical design of the human eye: a critical review. *Journal of optometry,* 2(1), 3–18.

Nguyen, T. T.,Wang, J. J.,Islam, F. A.,Mitchell, P.,Tapp, R. J.,Zimmet, P. Z.,Simpson, R.,Shaw, J., & Wong, T. Y. (2008a). Retinal arteriolar narrowing predicts incidence of diabetes the australian diabetes, obesity and lifestyle (ausdiab) study. *Diabetes,* 57(3), 536–539.

Nguyen, T. T.,Wang, J. J.,Sharrett, A. R.,Islam, F. A.,Klein, R.,Klein, B. E.,Cotch, M. F., & Wong, T. Y. (2008b). Relationship of retinal vascular caliber with diabetes and retinopathy the multi-ethnic study of atherosclerosis (mesa). *Diabetes care,* 31(3), 544–549.

Nguyen, T. T.,Wang, J. J., & Wong, T. Y. (2007). Retinal vascular changes in pre-diabetes and prehypertension new findings and their research and clinical implications. *Diabetes care,* 30(10), 2708–2715.

Nilsson, G. E.,Tenland, T., & Oberg, P. A. (1980). Evaluation of a laser doppler flowmeter for measurement of tissue blood flow. *IEEE Transactions on Biomedical Engineering,* (10), 597–604.

Nilsson, R.,Peña, J. M.,Björkegren, J., & Tegnér, J. (2007). Consistent feature selection for pattern recognition in polynomial time. *Journal of Machine Learning Research,* 8(Mar), 589–612.

Nolan, S. A. & Heinzen, T. (2010). *Essentials of statistics for the behavioral sciences*. Macmillan.

O'Driscoll, G.,Green, D.,Rankin, J.,Stanton, K., & Taylor, R. (1997). Improvement in endothelial function by angiotensin converting enzyme inhibition in insulin-dependent diabetes mellitus. *Journal of Clinical Investigation,* 100(3), 678.

O'Grady, K. E. (1982). Measures of explained variance: Cautions and limitations. *Psychological Bulletin,* 92(3), 766.

Ohkubo, Y.,Kishikawa, H.,Araki, E.,Miyata, T.,Isami, S.,Motoyoshi, S.,Kojima, Y.,Furuyoshi, N., & Shichiri, M. (1995). Intensive insulin therapy prevents the progression of diabetic microvascular complications in japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. *Diabetes research and clinical practice,* 28(2), 103–117.

Oliveira, W. S.,Teixeira, J. V.,Ren, T. I.,Cavalcanti, G. D., & Sijbers, J. (2016). Unsupervised retinal vessel segmentation using combined filters. *PloS one,* 11(2), e0149943.

Owen, C. G.,Newsom, R. S.,Rudnicka, A. R.,Barman, S. A.,Woodward, E. G., & Ellis, T. J. (2008). Diabetes and the tortuosity of vessels of the bulbar conjunctiva. *Ophthalmology,* 115(6), e27–e32.

Oyster, C. W. (1999). *The human eye: structure and function*. Sinauer Associates.

Pancera, P.,Ribul, M.,Presciuttini, B., & Lechi, A. (2000). Prevalence of carotid artery kinking in 590 consecutive subjects evaluated by echocolordoppler. is there a correlation with arterial hypertension? *Journal of internal medicine,* 248(1), 7–12.

Papaioannou, T. G. & Stefanadis, C. (2005). Vascular wall shear stress: basic principles and methods. *Hellenic J Cardiol,* 46(1), 9–15.

Parr, J. & Spears, G. (1974a). General caliber of the retinal arteries expressed as the equivalent width of the central retinal artery. *American journal of ophthalmology,* 77(4), 472–477.

Parr, J. & Spears, G. (1974b). Mathematic relationships between the width of a retinal artery and the widths of its branches. *American journal of ophthalmology,* 77(4), 478–483.

Patton, N.,Aslam, T.,MacGillivray, T.,Dhillon, B., & Constable, I. (2006a). Asymmetry of retinal arteriolar branch widths at junctions affects ability of formulae to predict trunk arteriolar widths. *Investigative ophthalmology & visual science,* 47(4), 1329–1333.

Patton, N.,Aslam, T. M.,MacGillivray, T.,Deary, I. J.,Dhillon, B.,Eikelboom, R. H.,Yogesan, K., & Constable, I. J. (2006b). Retinal image analysis: concepts, applications and potential. *Progress in retinal and eye research,* 25(1), 99–127.

Pemp, B.,Cherecheanu, A.-P.,Garhofer, G., & Schmetterer, L. (2013). Calculation of central retinal artery diameters from non-invasive ocular haemodynamic measurements in type 1 diabetes patients. *Acta ophthalmologica,* 91(5), e348–e352.

Perez-Rovira, A.,MacGillivray, T.,Trucco, E.,Chin, K.,Zutis, K.,Lupascu, C.,Tegolo, D.,Giachetti, A.,Wilson, P.,Doney, A., *et al.* (2011). Vampire: vessel assessment and measurement platform for images of the retina. In *2011 Annual International*

Conference of the IEEE Engineering in Medicine and Biology Society (pp. 3391–3394).: IEEE.

Polak, K.,Dorner, G.,Kiss, B.,Polska, E.,Findl, O.,Rainer, G.,Eichler, H.-G., & Schmetterer, L. (2000). Evaluation of the zeiss retinal vessel analyser. *British Journal of Ophthalmology,* 84(11), 1285–1290.

Poletti, E.,Grisan, E., & Ruggeri, A. (2011). Image-level tortuosity estimation in wide-field retinal images from infants with retinopathy of prematurity. In *Conference proceedings:... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference,* volume 2012 (pp. 4958–4961).

Posada, D. & Buckley, T. R. (2004). Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic biology,* 53(5), 793–808.

Pournaras, C. J.,Rungger-Brändle, E.,Riva, C. E.,Hardarson, S. H., & Stefansson, E. (2008). Regulation of retinal blood flow in health and disease. *Progress in retinal and eye research,* 27(3), 284–330.

Pries, A.,Ley, K.,Claassen, M., & Gaehtgens, P. (1989). Red cell distribution at microvascular bifurcations. *Microvascular research,* 38(1), 81–101.

Pries, A.,Neuhaus, D., & Gaehtgens, P. (1992). Blood viscosity in tube flow: dependence on diameter and hematocrit. *American Journal of Physiology-Heart and Circulatory Physiology,* 263(6), H1770–H1778.

Provis, J. M.,Penfold, P. L.,Cornish, E. E.,Sandercoe, T. M., & Madigan, M. C. (2005). Anatomy and development of the macula: specialisation and the vulnerability to macular degeneration. *Clinical and Experimental Optometry,* 88(5), 269–281.

Provost, F. J.,Fawcett, T., & Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. In *ICML,* volume 98 (pp. 445–453).

Qaum, T.,Xu, Q.,Joussen, A. M.,Clemens, M. W.,Qin, W.,Miyamoto, K.,Hassessian, H.,Wiegand, S. J.,Rudge, J.,Yancopoulos, G. D., *et al.* (2001). Vegf-initiated blood–retinal barrier breakdown in early diabetes. *Investigative ophthalmology & visual science,* 42(10), 2408–2413.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramsay, R. C.,Knobloch, W. H., & Cantrill, H. L. (1986). Timing of vitrectomy for active proliferative diabetic retinopathy. *Ophthalmology,* 93(3), 283–289.

Rashid, M. H. (2010). *Power electronics handbook: devices, circuits and applications*. Academic press.

Rassam, S.,Patel, V., & Kohner, E. (1995). The effect of experimental hypertension on retinal vascular autoregulation in humans: a mechanism for the progression of diabetic retinopathy. *Experimental physiology,* 80(1), 53–68.

RCOphth (2012). Diabetic retinopathy guidelines.

Reynolds, O. (1883). An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. *Proceedings of the royal society of London,* 35(224-226), 84–99.

Ribeiro, M. L.,Nunes, S. G., & Cunha-Vaz, J. G. (2013). Microaneurysm turnover at the macula predicts risk of development of clinically significant macular edema in persons with mild nonproliferative diabetic retinopathy. *Diabetes Care,* 36(5), 1254–1259.

Rice, M. J.,Sweat Jr, R. H.,Rioux, J. M.,Williams, W. T., & Routt, W. (2002). Non-invasive measurement of blood components using retinal imaging. US Patent 6,477,394.

Rifkin, R. & Klautau, A. (2004). In defense of one-vs-all classification. *Journal of machine learning research,* 5(Jan), 101–141.

Rimmer, T.,Fallon, T., & Kohner, E. (1989). Long-term follow-up of retinal blood flow in diabetes using the blue light entoptic phenomenon. *British journal of ophthalmology,* 73(1), 1–5.

Riordan-Eva, P. & Whitcher, J. (2008). *Vaughan & Asbury's general ophthalmology*. Wiley Online Library.

Riva, C. & Petrig, B. (1980). Blue field entoptic phenomenon and blood velocity in the retinal capillaries. *JOSA,* 70(10), 1234–1238.

Riva, C. E.,Grunwald, J. E.,Sinclair, S. H., & Petrig, B. (1985). Blood velocity and volumetric flow rate in human retinal vessels. *Investigative ophthalmology & visual science,* 26(8), 1124–1132.

Riva, C. E.,Sinclair, S. H., & Grunwald, J. E. (1981). Autoregulation of retinal circulation in response to decrease of perfusion pressure. *Investigative ophthalmology & visual science,* 21(1), 34–38.

Rodieck, R. W. (1973). *The vertebrate retina: principles of structure and function.* WH Freeman.

RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Ruta, L.,Magliano, D.,LeMesurier, R.,Taylor, H.,Zimmet, P., & Shaw, J. (2013). Prevalence of diabetic retinopathy in type 2 diabetes in developing and developed countries. *Diabetic medicine,* 30(4), 387–398.

Sasongko, M.,Wong, T.,Nguyen, T.,Cheung, C.,Shaw, J., & Wang, J. (2011). Retinal vascular tortuosity in persons with diabetes and diabetic retinopathy. *Diabetologia,* 54(9), 2409–2416.

Sasongko, M. B.,Wang, J. J.,Donaghue, K. C.,Cheung, N.,Benitez-Aguirre, P.,Jenkins, A.,Hsu, W.,Lee, M.-L., & Wong, T. Y. (2010). Alterations in retinal microvascular geometry in young type 1 diabetes. *Diabetes Care,* 33(6), 1331–1336.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics bulletin,* 2(6), 110–114.

Scheie, H. G. (1953). Evaluation of ophthalmoscopic changes of hypertension and arteriolar sclerosis. *AMA archives of ophthalmology,* 49(2), 117–138.

Schluchter, M. D. & Elashoff, J. T. (1990). Small-sample adjustments to tests with unbalanced repeated measures assuming several covariance structures. *Journal of Statistical Computation and Simulation,* 37(1-2), 69–87.

Schmidt, M. I.,Duncan, B. B.,Sharrett, A. R.,Lindberg, G.,Savage, P. J.,Offenbacher, S.,Azambuja, M. I.,Tracy, R. P.,Heiss, G.,investigators, A., *et al.* (1999). Markers of inflammation and prediction of diabetes mellitus in adults (atherosclerosis risk in communities study): a cohort study. *The Lancet,* 353(9165), 1649–1652.

Schreiner, W. & Buxbaum, P. F. (1993). Computer-optimization of vascular trees. *IEEE Transactions on Biomedical Engineering,* 40(5), 482–491.

Shapiro, S. S. & Francia, R. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association,* 67(337), 215–216.

Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika,* 52(3/4), 591–611.

Sherman, T. F. (1981). On connecting large vessels to small. the meaning of murray's law. *The Journal of general physiology,* 78(4), 431–453.

Simo, R. & Hernandez, C. (2014). Neurodegeneration in the diabetic eye: new insights and therapeutic perspectives. *Trends in Endocrinology & Metabolism,* 25(1), 23–33.

Sims, D. E. (1986). The pericyte—a review. *Tissue and Cell,* 18(2), 153–174.

Sinclair, S. H.,Grunwald, J. E.,Riva, C. E.,Braunstein, S. N.,Nichols, C. W., & Schwartz, S. S. (1982). Retinal vascular autoregulation in diabetes mellitus. *Ophthalmology,* 89(7), 748–750.

Sing, T.,Sander, O.,Beerenwinkel, N., & Lengauer, T. (2005). Rocr: visualizing classifier performance in r. *Bioinformatics,* 21(20), 7881.

Singh, V. P.,Bali, A.,Singh, N., & Jaggi, A. S. (2014). Advanced glycation end products and diabetic complications. *The Korean Journal of Physiology & Pharmacology,* 18(1), 1–14.

Snijders, T. A. (2011). *Multilevel analysis*. Springer.

Soares, J. V.,Leandro, J. J.,Cesar, R. M.,Jelinek, H. F., & Cree, M. J. (2006). Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on medical Imaging,* 25(9), 1214–1222.

Staal, J.,Abramoff, M. D.,Niemeijer, M.,Viergever, M. A., & van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE transactions on medical imaging,* 23(4), 501–509.

Starkweather, J. (2010). Linear mixed effects modelling using r. *Unpublished Manuscript*.

Steyerberg, E. W.,Pencina, M. J.,Lingsma, H. F.,Kattan, M. W.,Vickers, A. J., & Van Calster, B. (2012). Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. *European journal of clinical investigation,* 42(2), 216–228.

Stitt, A. W. & Curtis, T. M. (2005). Advanced glycation and retinal pathology during diabetes. *Pharmacological Reports,* 57, 156.

Stitt, A. W.,Lois, N.,Medina, R. J.,Adamson, P., & Curtis, T. M. (2013). Advances in our understanding of diabetic retinopathy. *Clinical science,* 125(1), 1–17.

Stitt, A. W.,McGoldrick, C.,Rice-McCaldin, A.,McCance, D. R.,Glenn, J. V.,Hsu, D. K.,Liu, F.-T.,Thorpe, S. R., & Gardiner, T. A. (2005). Impaired retinal angiogenesis in diabetes role of advanced glycation end products and galectin-3. *Diabetes,* 54(3), 785–794.

Stokoe, N. & Turner, R. (1966). Normal retinal vascular pattern. arteriovenous ratio as a measure of arterial calibre. *The British journal of ophthalmology,* 50(1), 21.

Sullivan, P. M.,Davies, G.,Caldwell, G.,Morris, A. C., & Kohner, E. M. (1990). Retinal blood flow during hyperglycemia. a laser doppler velocimetry study. *Investigative ophthalmology & visual science,* 31(10), 2041–2045.

Sutera, S. P. & Skalak, R. (1993). The history of poiseuille's law. *Annual Review of Fluid Mechanics,* 25(1), 1–20.

Suwa, N. & Takahashi, T. (1971). *Morphological and morphometrical analysis of circulation in hypertension and ischemic kidney*. Urban & Schwarzenberg.

Swaminathan, H. & Rogers, H. J. (2008). Estimation procedures for hierarchical linear models. *Multilevel modeling of educational data,* (pp. 469–519).

Takahashi, T.,Nagaoka, T.,Yanagida, H.,Saitoh, T.,Kamiya, A.,Hein, T.,Kuo, L., & Yoshida, A. (2009). A mathematical model for the distribution of hemodynamic parameters in the human retinal microvascular network. *Journal of biorheology,* 23(2), 77–86.

Tapp, R. J.,Shaw, J. E.,Harper, C. A.,De Courten, M. P.,Balkau, B.,McCarty, D. J.,Taylor, H. R.,Welborn, T. A., & Zimmet, P. Z. (2003). The prevalence of and factors associated with diabetic retinopathy in the australian population. *Diabetes care,* 26(6), 1731–1737.

Taylor, E. & Dobree, J. (1970). Proliferative diabetic retinopathy. site and size of initial lesions. *The British journal of ophthalmology,* 54(1), 11.

Terai, N.,Haustein, M.,Siegel, A.,Stodtmeister, R.,Pillunat, L. E., & Sandner, D. (2014). Diameter of retinal vessels in patients with diabetic macular edema is not altered by intravitreal ranibizumab (lucentis). *Retina,* 34(7), 1466–1472.

Thompson, D. W. (1917). On growth and form cambridge.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological),* (pp. 267–288).

Toda, N.,Imamura, T., & Okamura, T. (2010). Alteration of nitric oxide-mediated blood flow regulation in diabetes mellitus. *Pharmacology & therapeutics,* 127(3), 189–209.

Tolle, C. R.,McJunkin, T. R., & Gorsich, D. J. (2008). An efficient implementation of the gliding box lacunarity algorithm. *Physica D: Nonlinear Phenomena,* 237(3), 306–315.

Triantafyllou, A.,Anyfanti, P.,Gavriilaki, E.,Zabulis, X.,Gkaliagkousi, E.,Petidis, K.,Triantafyllou, G.,Gkolias, V.,Pyrpasopoulou, A., & Douma, S. (2014). Association between retinal vessel caliber and arterial stiffness in a population comprised of normotensive to early-stage hypertensive individuals. *American journal of hypertension,* 27(12), 1472–1478.

Tsai, A. S.,Wong, T. Y.,Lavanya, R.,Zhang, R.,Hamzah, H.,Tai, E. S., & Cheung, C. Y. (2011). Differential association of retinal arteriolar and venular caliber with diabetes and retinopathy. *Diabetes research and clinical practice,* 94(2), 291–298.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics,* (pp. 99–114).

UKPDS (1998a). Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (ukpds 33). *The Lancet,* 352(9131), 837–853.

UKPDS (1998b). Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: Ukpds 38. *BMJ: British Medical Journal,* (pp. 703–713).

Van Hecke, M.,Dekker, J.,Nijpels, G.,Moll, A.,Heine, R.,Bouter, L.,Polak, B., & Stehouwer, C. (2005). Inflammation and endothelial dysfunction are associated with retinopathy: the hoorn study. *Diabetologia,* 48(7), 1300–1306.

Vedaldi, A. & Fulkerson, B. (2008). VLFeat: An open and portable library of computer vision algorithms. `http://www.vlfeat.org/`.

Villegas, G. M. (1960). Electron microscopic study of the vertebrate retina. *The Journal of general physiology,* 43(6), 15–43.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society,* (pp. 307–333).

Wagener, H. & Wilder, R. (1921). The retinitis of diabetes mellitus. *Journal of the American Medical Association,* 76(8), 515–517.

Wagener, H. P. & Keith, N. M. (1939). Diffuse arteriolar disease with hypertension and the associated retinal lesions. *Medicine,* 18(3), 317–430.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic bulletin & review,* 14(5), 779–804.

Watkins, P. J. (2003). Abc of diabetes: The diabetic foot. *BMJ: British Medical Journal,* 326(7396), 977.

Weibel, J. & Fields, W. S. (1965). Tortuosity, coiling, and kinking of the internal carotid artery i. etiology and radiographic anatomy. *Neurology,* 15(1), 7–7.

Weiler, D. L.,Engelke, C. B.,Moore, A. L., & Harrison, W. W. (2015). Arteriole tortuosity associated with diabetic retinopathy and cholesterol. *Optometry & Vision Science,* 92(3), 384–391.

Weisstein, E. W. (2004). Bonferroni correction.

West, B. T.,Welch, K. B., & Galecki, A. T. (2014). *Linear mixed models: a practical guide using statistical software*. CRC Press.

WHO (2016). *Global report on diabetes*.

Wigdahl, J.,Guimarães, P.,Leontidis, G.,Triantafyllou, A., & Ruggeri, A. (2015). Automatic gunn and salus sign quantification in retinal images. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (pp. 5251–5254).: IEEE.

Wilkinson, C.,Ferris, F. L.,Klein, R. E.,Lee, P. P.,Agardh, C. D.,Davis, M.,Dills, D.,Kampik, A.,Pararajasegaram, R.,Verdaguer, J. T., *et al.* (2003). Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. *Ophthalmology,* 110(9), 1677–1682.

Williams, R.,Airey, M.,Baxter, H.,Forrester, J.,Kennedy-Martin, T., & Girach, A. (2004). Epidemiology of diabetic retinopathy and macular oedema: a systematic review. *Eye,* 18(10), 963–983.

Williamson, T. H. & Baxter, G. M. (1994). Central retinal vein occlusion, an investigation by color doppler imaging: blood velocity characteristics and prediction of iris neovascularization. *Ophthalmology,* 101(8), 1362–1372.

Winter, B. (2013). A very basic tutorial for performing linear mixed effects analyses. *arXiv preprint arXiv:1308.5499*.

Witt, N. W.,Chapman, N.,Thom, S. A. M.,Stanton, A. V.,Parker, K. H., & Hughes, A. D. (2010). A novel measure to characterise optimality of diameter relationships at retinal vascular bifurcations. *Artery research,* 4(3), 75–80.

Wolf, S.,Jung, F.,Kiesewetter, H.,Körber, N., & Reim, M. (1989). Video fluorescein angiography: method and clinical application. *Graefe's archive for clinical and experimental ophthalmology,* 227(2), 145–151.

Wong, T. Y.,Islam, F. A.,Klein, R.,Klein, B. E.,Cotch, M. F.,Castro, C.,Sharrett, A. R., & Shahar, E. (2006). Retinal vascular caliber, cardiovascular risk factors, and inflammation: the multi-ethnic study of atherosclerosis (mesa). *Investigative ophthalmology & visual science,* 47(6), 2341–2350.

Wong, T. Y.,Klein, R.,Klein, B. E.,Meuer, S. M., & Hubbard, L. D. (2003). Retinal vessel diameters and their associations with age and blood pressure. *Investigative ophthalmology & visual science,* 44(11), 4644–4650.

Wong, T. Y.,Klein, R.,Klein, B. E.,Tielsch, J. M.,Hubbard, L., & Nieto, F. J. (2001). Retinal microvascular abnormalities and their relationship with hypertension, cardiovascular disease, and mortality. *Survey of ophthalmology,* 46(1), 59–80.

Xu, R. (2003). Measuring explained variation in linear mixed effects models. *Statistics in medicine,* 22(22), 3527–3541.

Yang, X.,Deng, Y.,Gu, H.,Ren, X.,Lim, A.,Snellingen, T.,Liu, X.,Wang, N.,Pak, J. W.,Liu, N., *et al.* (2016). Relationship of retinal vascular calibre and diabetic retinopathy in chinese patients with type 2 diabetes mellitus: the desheng diabetic eye study. *British Journal of Ophthalmology,* (pp. bjophthalmol–2014).

Yau, J.,Kawasaki, R.,Islam, F.,Shaw, J.,Zimmet, P.,Wang, J., & Wong, T. (2010). Retinal fractal dimension is increased in persons with diabetes but not impaired glucose metabolism: the australian diabetes, obesity and lifestyle (ausdiab) study. *Diabetologia,* 53(9), 2042–2045.

Zamir, M. (1976). Optimality principles in arterial branching. *Journal of Theoretical Biology,* 62(1), 227–251.

Zamir, M. (1978). Nonsymmetrical bifurcations in arterial branching. *The Journal of general physiology,* 72(6), 837–845.

Zamir, M. (1999). On fractal properties of arterial trees. *Journal of theoretical biology,* 197(4), 517–526.

Zamir, M. (2016). *Flow in Branching Tubes*. Springer International Publishing.

Zou, H. & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* 67(2), 301–320.