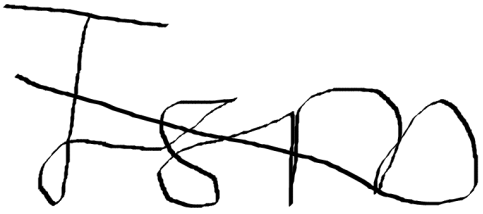
University of Exeter
College of Engineering Mathematics and Physical Sciences

Automated Quantification of Human Islet Structure and Geometry

Jordan Moore

Submitted by Jordan Moore, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Mathematics, November, 2020.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.



Signed: .....

Abstract

This thesis presents a novel toolkit for image analysis that seeks to improve our understanding of the structures of the islets of Langerhans within the human Pancreas. Our aim is to improve understanding of structural changes associated with age and diabetes.

The islets of Langerhans are structures within the pancreas which are the metabolic control centres for blood sugar (glycaemic) homeostasis, amongst other things. These islets contain several endocrine cell types including insulin producing beta cells and in diabetes this composition changes. One of the major problems with our understanding of what happens in human diabetes is the lack of available samples for study. In the University of Exeter medical school, we are able to access approximately 160 diabetic pancreases, making this the world's largest collection of such material.

Due to the progress of diseases such as diabetes being associated with changes in cellular composition of the islets, we examine the structure of islets both as a totality and at a cellular level. We have developed algorithms to automate the mapping of islets from microscopy images to spatially informative false colour map at a cellular resolution. We begin this analysis with a systematic analysis of the relevant cell numbers in the pancreas in *ex vivo* human subjects, with and without type 1 or type 2 diabetes.

We subsequently investigate this structure by constructing metrics that characterise the geometry of the islets (such as the isoparametric ratio) and how diabetes may change these. We analyse the topological properties of the islet structure using recently developed techniques in computational homology and in particular persistent homology upon individual islets. Specifically, this involves finding generating sets for the first and second Betti numbers for particular mapped microscopy images and investigating Betti number stability under perturbations of parameters used for the automated image analysis.

Finally, we explore cellular network structures within the islets, and how diabetes may change the pattern of intracellular connections within the islets. We do this via network metrics which characterise some aspects of the structure of the networks such as connectedness. We do this considering the graph as an overall islet network and the subnetworks between different cell types.

Acknowledgements

I wish to dedicate this thesis to both my late grand-fathers, Jack Freeman and Dudley Moore. Highly inquisitive, these men would have been easily capable of pursuing higher degrees, but alas came from a time when most were not encouraged to pursue academic betterment but instead pick up a trade. Indeed, one of Jack's final regrets in life was that he never went to university.

I would also like to thank my two main supervisors, Professor Peter Ashwin and Professor Noel Morgan - without their patience, support and input, completing a PhD would not have been possible. I'd like to also thank my parents, Michelle and Stephen Moore, without whose support, I would not have made it to university let alone undertake a PhD. Finally I'd like to thank my sisters, Dr. Rachel Morley and Megan Moore. In particular, Rachel's own PhD experience and help during my particularly difficult second year, helped me immensely.

Graviora Manent

Contents

List of Figures

1. Background and Motivation

1.1 Current Understanding of Human Islet Structure

1.1.1 Pancreas Anatomy

Nested in the duodenum, and anterior to the aorta, the macroanatomy of the pancreas is relatively well known [3]. Somewhat phallic looking, the pancreas can be divided into three main areas: The head- in which the majority of the mass lies, body and tail, which is illustrated in Figure 1.1. The pancreas is relatively unique because it consists of endocrine and exocrine tissues. The exocrine tissues come in the form of acini (singular acinus), which are berry like structures and makes up more that 82% of the pancreas by volume. The acinar cells secrete protein digestive enzymes into the duct system. Commencing from intralobular ducts, this secretion flows to the interlobular ducts, and finally into the duodenum, through the main pancreas duct (duct of Wirsung). The ducts are made of thick collagen which prevent the pancreas from auto-digesting. Damage to the ducts can cause a leak, which triggers pancreatitis [4].

The endocrine tissue is dispersed throughout the pancreas, in structures called islets of Langerhans. This dispersion is not uniform; there is increased density of islets in the pancreas tail [5]. There also exist a small number of beta cells outside of any islet structure (in groups of 1-3) cells. It is not definitively known why these cells exist; some authors have proposed that they might be islets in an early stage of formation.

In the pancreas, much like any other organ, the vasculature oxygenates the

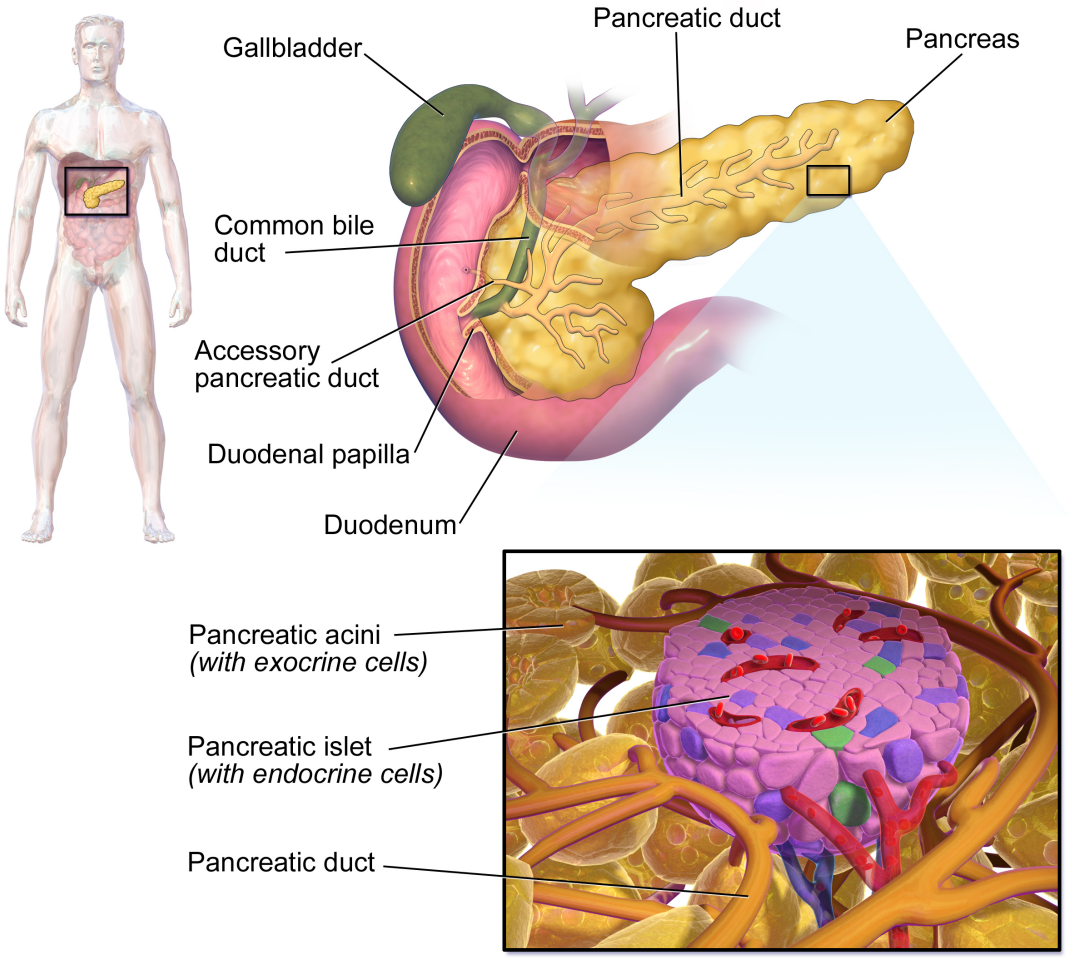


Figure 1.1: A diagram of the pancreas and a contained islet [1]. The blood vessels oxygenate the organ. Hormones, including insulin, glucagon and somatostatin are released into the blood by the islets, located throughout the pancreas which are represented by the dots. However, the islets are not distributed evenly; there is a similar density of islets in the pancreas head and body regions, but a greater than a two-fold increase in islet density in the tail.

tissues present, and supplies energy/nutrients. Blood vessels infuse islets, with a density five times that of the exocrine tissues [6]. The vasculature plays a key role in percolating hormones from islet cells, making them an important component in the islet anatomy.

The nerves present in the pancreas contain a mixture of sympathetic and parasympathetic types, including sensory nerves. Insulin secretion is stimulated by the parasympathetic system and inhibited by the sympathetic system. This enables the islet cells to respond to environmental stimuli such as those triggered by eating and digestion. The exact synaptic connections within the islet, however, are not

known to date [7].

1.1.2 Islet Anatomy

The islets are ovular structures, containing a mix of at least five different types of endocrine cells [2]. α cells secrete glucagon (GCG), which is a hormone that raises the blood glucose levels. α-cells comprise between 15% and 20% of the number of islet cells. β cells secrete insulin which lowers blood glucose level. Beta cells comprise between 50% and 80% of the islet mass. Beta cells also secrete islet amyloid polypeptide (IAPP), also known as amylin, which works to prevent spikes in glycemic levels by slowing gastric emptying. δ cells secrete a hormone called somatostatin (SST), which inhibits glucagon and insulin release. δ cells comprise between $5 - 10\%$ of the volume of islets. γ cells secrete pancreatic polypeptide (PP) which regulates pancreatic secretions (exocrine and endocrine). ε cells secrete a hormone called ghrelin and was the most recent cell type to be found (2002) [8]. Functionally, ghrelin regulates appetite [2].

The islets are surounded by a protective extracellular matrix (ECM) comprised of two components: The basement membrane (BM) and the interstitial membrane (IM) [9]. Primarily constituted of collagen and laminins, the BM limits the movement of cells and soluble molecules. The IM allows the tissues flexibility and elasticity, and underlies the BM. In T1D, the basement membrane has been shown to undergo severe degradation [10].

Whilst size and distribution of islets are highly variable, there are two major types of islets recognised- diffuse and compact islets [11]. Diffuse islets are located predominantly in the head of the pancreas, and contain significantly more γ cells. In human islets, endocrine cells are distributed throughout the islet, seemingly without structure.

1.1.3 Islet Cell Interactions

Within islets, different cell interactions influence the quantity of endocrine hormone secretion. Most interactions take place via chemical and electrochemical signalling.

These signals either travel via the islet space (interstitum) or via the local islet microcirculatory network. All signalling mechanisms that are mediated via the vasculature can only travel downstream. Here, we are merely outlining some of the better known mechanisms at a high level. This is a highly active area of research, and is fundamental to understanding and eventually quantifying how islets work - both healthy and diabetic ones [12]. The mechanisms are represented in Figure 1.2.

A key interaction within the context of diabetes research are beta-beta cell interactions. These take place via cadherins, gap junction proteins, eph receptors - membrane bound proteins, ephrin ligands and secreted insulin. N-, R- and E-cadherins and a neural cell adhesion molecule are produced in the islets. The N- and E- variants of cadherin (calcium-dependent adhesion molecules) localize to the cell-cell contacts, whilst R-cadherin is present in the cytoplasm. These molecules play a role in glucose stimulated insulin secretion, and help to synchronise the secretion from adjacent beta cells. Gap junction proteins connect beta cells, and allow ionic exchanges between neighbouring beta cells which allows for synchronisation of their ionic cytoplasmic concentration and electrical activity. This provides beta cells further mechanisms by which to synchronise their insulin secretions [13]. These mechanisms work over all timescales; Betram et al. provide a mathematical model of how both fast and slow scale electrical activities combine to create a wide range of timescales over which insulin secretion happens [14]. Ephs and ephrins work together to promote insulin secretion in response to low glycaemic levels. In this context blood glucose concentration is the primary metabolic signalling agent. Secreted insulin acts to inhibit further insulin secretion, via a negative feedback mechanism [15].

Beta-beta cell interaction are not the only source, or even the primary driver of insulin secretion. Insulin secretion happens constitutively, though it it is increased primarily as a result of high blood sugar levels. High blood sugar levels trigger the closure of potassium sensitive channels in the beta cells causing membrane depolarisation [16]. This membrane depolarisation, in turn stimulates electrical activity, which in turn triggers calcium ion influx and insulin release [**?**].

Alpha cells produce glucagon at an increased rate when blood glucose levels

are low. Alpha - alpha cell connections cooperate positively as glucagon promotes its own secretion via an unknown positive feedback mechanism. Beta cells secrete ionic zinc as part of the mechanism for producing insulin. This ionic zinc acts as an electrochemical inhibitor for glucagon secretion from alpha cells. Glucagon acts as a signalling agent to promote beta cells insulin secretion and delta cell somatostatin secretion via membrane receptors [17].

Delta cells interact with both beta and alpha cells to inhibit their secretions. The production of somatostatin in delta cells is promoted by the presence of glucagon via both membrane receptors and membrane electrochemical potassium ion signals.

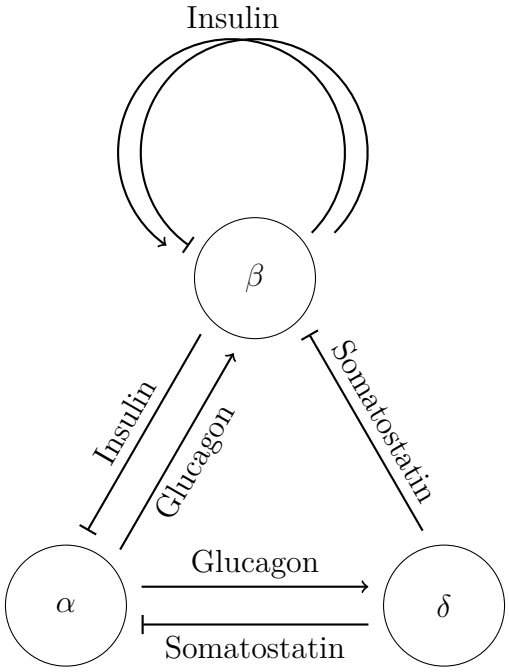


Figure 1.2: Network diagram displaying different kinds of cooperative or inhibitive relations between cells in the islets. Beta cells synchronise with each other. Insulin also inhibits the production of glucagon in the alpha cells. Alpha cells also produce glucagon which simulates both insulin and somatostatin production. Delta cells produce somatostatin which inhibits the other two hormone productions.

1.2 What is Diabetes Mellitus?

1.2.1 Medical Definition of Diabetes Mellitus

Diabetes Mellitus is characterised by poor control of blood glucose (glycaemic) levels. The World Health Organisation (WHO) defines diabetes as a blood glucose level of $> 11.1 mmol/l$ 2 hours post-prandially (following a meal), whereas, normal blood

sugar is $< 7.8 mmol/l$ [18]. Diabetes can be characterised as a disruption within the glucose homoeostasis system, illustrated in figure 1.3.

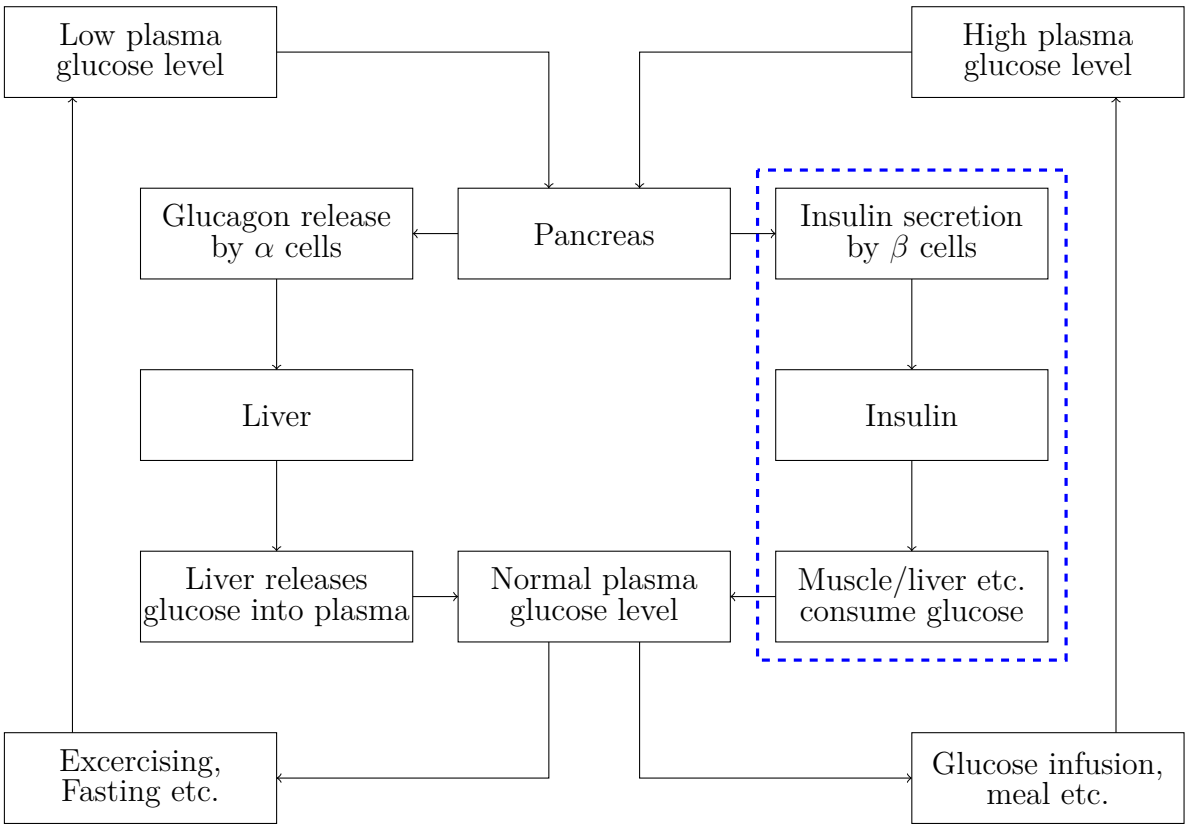


Figure 1.3: Simplified representation of glucose homeostasis system. Glucagon and insulin are the two primary hormones which modify glycaemic levels, thereby maintaining blood glucose homoeostasis in healthy individuals. Insulin works to lower blood sugar levels and glucagon raises blood sugar levels. Diabetes occurs when the subsystem encompassed within the blue square malfunctions.

Whilst many types of diabetes exist and will be explained in due course, there are two main types; type 1 diabetes (T1D) and type 2 diabetes (T2D). There is no concise definition for either, but typically, T1D is an autoimmune disease, occurs in young patients - the mean T1D onset age in US is 14 - and requires exogenous insulin to manage [19, 20]. T2D is typically a disease of older patients - the mean onset age in the US is 45. It is thought to be mediated by reduced insulin sensitivity in peripheral tissues and a decline in insulin secretion and can often be managed by lifestyle changes as well as orally activated agents and/or injected insulin where this proves to be insufficient [21, 22].

There is a third classification, though not universally accepted, amongst this spectrum, called type 1.5 diabetes or latent autoimmune diabetes in adults (LADA).

Patients typically present with characteristics associated with both T1D, such as autoimmune beta cell destruction, and T2D, such as older age of onset [23, 24].

Due to imprecision about the classification and heterogeneity of how diabetes presents, a recent study identified 5 different clusters of diabetes patients. This data driven study, using K-means classifications and hierarchical clustering identified the following, based on factors including genetic similarities, treatment and disease progression. Cluster 1, or severe autoimmune diabetes (SAID) was characterised by early-onset disease, relatively low BMI, poor metabolic control, insulin deficiency and the presence of glutamic acid decarboxylase antibodies (GADA). Antibodies to GADA are an autoimmune marker that indicates substantial autoimmune islet inflammation and is discussed in section 1.3. Cluster 2, or severe insulin-deficient diabetes (SIDD), was GADA negative but otherwise similar to cluster 1: low age at onset, relatively low BMI, low insulin secretion, and poor metabolic control. Cluster 3, or severe insulin-resistant diabetes (SIRD) was characterised by insulin resistance and high BMI. Cluster 4, or mild obesity-related diabetes (MOD) was also characterised by obesity but not by insulin resistance. Finally, cluster 5, or mild age-related diabetes (MARD) were older than patients in other clusters, but showed, similar to cluster 4, only modest metabolic derangements [25]. Though more work is required on this classification methodology, a similar endotype classification has been attempted with respect to type 1 diabetes [26]. As we increase our understanding of the disease and are able to classify it better, it is hoped that we may be able to treat and possibly even cure or prevent both main types diabetes in humans.

As a brief aside, it is worth mentioning that there are other types of diabetes, though these will fall outside the scope of this project [27, 28, 29].

1.2.2 A Brief History of Diabetes

The term diabetes mellitus comes from Greek and Latin. ‘Diabetes’ comes from the Greek work for syphon and ‘mellitus’ comes from the Latin for honey-sweet. Together, these describe one of the main symptoms of the disease, which is excessive urination of glucose containing urine.

Diabetes has been known to exist since ancient times [30]; a symptomatic description appeared on the Ebers papyrus in 1550 BC. Subsequently, a distinction between what is now called type 1 and type 2 diabetes was made in 5^{th} and 6^{th} century AD in India by physicians such as Sushrut. Whilst the sweetness of urine was rediscovered in Europe by the English court physician, Thomas Willis in the 17^{th} century, it took until the following century for another English physician, Matthew Dobson, to show that the diabetic sweetness of urine was a result of sugar [31].

Perhaps the most famous name associated with diabetes physiology is Paul Langerhans. In his doctoral thesis, 1869, he was the first person to describe clusters of cells, later known as islets of Langerhans. The next great leap in our understanding of diabetes came in 1899. Oskar Minkowski and Josef von Mering, removed the pancreas of a dog inducing diabetic symptoms, which for the first time showed that the pancreas was a vital component with respect to diabetes. Insulin, from the Latin for 'island' was discovered in 1921. Following testing on dogs, the first diabetic human was treated with injected insulin in 1922. [32].

Determining the amino acid structure of insulin and its three dimensional structure using X-ray crystallography was achieved during the 50's and 60's by two Nobel prize winners, Frederick Sanger and Dorothy Hodgkin, respectively. In the 80's, major advances occurred in the production of insulin. Using additives such as zinc, it has become possible for the absorption of insulin to be controlled from the sub-cutaneous insulin injection, mitigating some of the side effects caused by injected insulin [33].

More recently, advances in 'big data' and mathematical techniques, such as machine learning are allowing us to develop better treatments for diabetes, such as automated insulin pumps - using glycaemic sensors to calculate the quantity of insulin to be injected, and be able to better predict risk of developing diabetes [34].

1.2.3 Epidemiology of Diabetes

The risk factors for developing diabetes are multivariate, and include both environmental and genetic causes. In general the global prevalence of diabetes stands

at 9.3% as of 2019, which is expected to rise to 10.2% in 2030 [35]. There exists, however, significant regional variability within both types of diabetes, due to the compounding of different epidemiological geographical risk factors. For T1D, this ranges from incidences of 0.0001% in China and Venuzuala, to 0.0365% in Finland [36]; For T2D, this ranges from about 4% amongst the Mapunche Indians in Chile [37], to 38% among the Pima Indians [38].

Age is a significant epidemiological factor determining type of diabetes that one is likely to develop. Over 85% of diabetes patients under the age of 20 have T1D. Diagnosis of T1D peaks around 10-14 years of age [39]. On the other hand, T2D is more commonly a disease of elderly people, though there is a large increase in cases of young onset T2D, thought to be as a result of the increase in global incidence of childhood obesity [40, 41].

Obesity is considered the primary cause of T2D; about 80% of new cases of T2D present as obese [42]. The risk of developing T2D increases with respect to BMI; medically obese people are 6-10 times more likely to contract diabetes in their lifetime according to the NHNANES survey in the US, compared to the general population [43]. Obesity is considered to be a significant risk factor in T1D [44], with a particular study in Finland linking a 10% age relative weight increase to a 60% increase in the risk of developing T1D [45].

Independently of obesity, other lifestyle factors are considered a substantial risk factor for T2D. As a result, urban populations are more likely to develop T2D [46]. Exercise lowers the probability of developing type 2 diabetes. Patients who exercise the most have a $25-60\%$ lower risk of developing type 2 diabetes compared with the background population, notwithstanding other factors [47]. Furthermore diet appears to be another lifestyle factor; a diet, containing high levels of sugars and fats, combined with low fibre increased insulin resistance by 5.7%. In the case of T1D, there appears to be increased prevalence diabetes for those born in winter months [48]. There also appears to be some risk factors associated with diet, such as how long a child is breast fed for [49].

In terms of genetic risks, two specific genetic markers (haplotypes) have been

identified on the 6^{th} chromosome, specifically the human leukocyte antigen (HLA) complex. Whilst $90 - 95\%$ of young children who have T1D have either or both of these haplotypes, specifically HLA-risk genotypes DR3/4-DQ8 or DR4/DR4, about 5% of these develop T1D [50]. There are over 80 genetic loci strongly associated with type 2 diabetes, and whilst each loci appears to have only a minor impact on T2D heritability, this indicates that there is a strong genetic risk to the inheritance of T2D [51].

1.2.4 Symptoms of Diabetes

Diabetes causes many symptoms, as a result of the increased blood sugar levels (hyperglycaemia). Hyperglycaemia causes changes to capillary architecture and permeability, causing many complications, including organ failure , which if uncontrolled can be fatal. A key motivation for studying diabetes is to enable us to prevent these complications can be extremely life limiting and painful.

A particularly dangerous outcome of diabetes is a condition called diabetic ketoacidosis (DKA), which is caused by uncontrolled blood sugar levels. It is characterised by hyperglycaemia and hyperketonaemia (high levels of ketones in the blood). The high levels of ketones lowering the pH of blood, also known as metabolic acidosis. In turn, this can lead to include nocturia and polyuria, weight loss, weakness, blurred vision and respiratory acidosis; a medical emergency in which oxygen cannot be effectively transported by the haemoglobin. More symptoms include abdominal pain, leg cramps, nausea and vomiting, confusion and drowsiness and in 10% of cases, a coma. DKA is most associated with type 1 diabetes, though not exclusively. [52]

Diabetic eye disease, one form of which is diabetic retinopathy may be experienced by most people with diabetes over a 25 year period [53]. Hyperglycaemica triggers the production of more capillaries in the retina. However, these are weak and prone to bursting, causing a condition called diabetic macular edema which, if untreated, can cause total blindness.

Diabetic nephropathy is characterised by proteinuria (high concentrations

of protein in the urine), which is a sign of impaired renal (kidney) function [54]. Nephropathy includes a large number of disorders, with symptoms including muscular wasting and deformity.

Diabetic neuropathy is diabetic induced damage to the nervous system, especially in the limbs. It is caused by diabetic induced microvascular damage. Symptoms can include loss of senses and extreme pain [55].

Other diabetic complications include increased risk of macrovascular disease, such as strokes by 3-5 times [56]. The risk of diabetic patients having to have legs, toes and feet amputated is increased 10-30 times. Erectile dysfunction occurs 10-15 year earlier in men with diabetes as without [57]. Further, diabetes can also cause a wide spectrum of gastrointestinal problems. Diabetes has been linked to skin and connectivity disorders, as well as psychological and psychiatric disorders.

1.2.5 Cost of Diabetes

According to a 2019 Diabetes UK report, 4.7 million people in the UK had diabetes, with a new diagnosis every two minutes. Of these, 90% had T2D, 8% had T1D and the remainder have rarer types of diabetes [58]. The number of people in the UK, diagnosed with diabetes has doubled in the last 20 years and by 2030, it is estimated that more than 5.5 million people will have diabetes in the UK.

The NHS spends about £10 billion a year on diabetes, which accounts for 10% of its entire budget. Of this, £8 billion is spent on treating complications of diabetes, whilst an estimated 3 in 5 cases of T2D can be prevented [59]. People with diabetes are twice as likely to be admitted to hospital and one in six people in hospital have diabetes. Diabetes leads to more than 8500 leg, toe or foot amputations every year and 1700 people become visually impaired as result of diabetes. This leads to a loss of independence, which is not accounted for in the £10 billion. On top of this, fewer than 1 in 5 T1D patients are meeting their recommended treatment targets that will lower their risk of complications, whilst the corresponding figure for T2D is fewer than 2 in 5.

In the US, the indirect and direct cost of treating diabetes was estimated

to be $245 billion per annum [60]. Aggregated together, the global annual cost of diabetes is $825 billion and set to rise [61]. This economic burden, along with the suffering caused by diabetes as illustrated throughout this section, provides sufficient motivation to study and understand diabetes better, with the eventual aim to cure the disease.

1.3 Current Understanding of Diabetes Aetiology

Despite large amounts of research, the underlying causes of diabetes remain poorly understood. This is problematic as it is limiting our ability to treat the diseases, in both type 1 and type 2 cases. Much of our understanding comes from animal models, including the non-obese diabetic (NOD) mouse and the biobreeding (BB) rat model.

1.3.1 Current Understanding of Type 1 Aetiology

The main underlying cause of T1D is thought to be an autoimmune mediated destruction of beta cells, known as insulitis, within the islets of Langerhans. In younger onset cases, this destruction causes inflammation and lesions in the pancreas. T1D causes approx 90% of beta cell destruction in lab animals and a 50% to 80% reduction in plasma insulin concentrations. Creating a toolkit to quantify this destruction in humans as opposed to animals is a key aim of this project.

The clinical presentation is the final stage of the onset of T1D. In the months before this, it has been hypothesised that circulating immune antibodies build up to a threshold, indicating that white blood cells have been causing autoimmune destruction, at which point the disease develops. However, not all individuals who have these antibody markers progress to developing T1D. For a more complete timeline of autoimmune mediated destruction, we have to rely on animal models. For insulitis, it seems that an acute bacterial or viral infection of the pancreas triggers a strong autoimmune response, in which the leucocytes misidentify beta cells as non-self agents [62]. In non insulitis cases of T1D, chronic infections of the islets trigger a slower destruction of the beta cells, leading to diabetes [63].

Beta cells are assumed to secrete chemokines, attracting white blood cells, as

outlined by Wilcox *et al.* [64]. However, various authors have proposed different white blood cells taking an active role in T1D aetiology. Within the literature, there is a lack of consensus over these autoimmune mechanisms underlying T1D [65, 66, 67, 68].

Another key factor during the period T1D onset, of which we are uncertain about is the so-called 'honeymoon' period. This term refers to the time after the destruction of beta cells in which the diabetes goes into temporary remission. It is uncertain how this happens, though one hypothesis is that there is a positive feedback loop regarding the destruction of beta cells increasing the strain on the remaining ones.

1.3.2 Current Understanding of Type 2 Aetiology

Whilst there is no consensus on how T2D develops, the literature suggests there are two main components involved in type 2 diabetes pathophysiology - beta cell dysfunction and insulin resistance, though it is still debated as to which one is the primary driver of T2D [69].

1.4 Morphometry

Morphometry comes from Greek and roughly translates to measurement of shape/form. In this thesis, we use morphometry to better describe how the islets change as a result of diabetes. We are particularly interested in the islets as a structure as there is a large body of evidence to suggest that the environment of the beta cells affects function [70, 71].

1.4.1 Use of Morphometry in Mathematical Biology Research

To date the brain is the main theatre in which Morphometry has been used. There are three main types of morphometry used; Voxel-based, deformation-based and tensor-based morphometry. The data used in these morphometry methods are

magnetic resonance imagery (MRI) scans as MRI scans can quantify grey and white brain matter with a highly reliability. [72]

Whilst it is largely suitable to use 3D morphometry to investigate brain structures, we do not have access to 3D data in this project, and as such we will be using 2D morphometry. There is currently a growing body of work in which feature detection and image analysis are being used in biological research [73, 74].

Published in June 2020, Samber et al. used synchrotron X-ray fluorescence to conduct an analysis of an islet. In this study, they investigated the presence of K, Ca, Fe and Zn ions present in the islet, with a view to analyse the shape and distribution of cells within islet to a spatial resolution of $500nm$ [75]. Ca and K ions play a key role in insulin secretion moderation, Zn ions play a role in insulin formation and Fe ions are implicated in beta cell dysfunction and death. By understanding the prevalence and spatial distribution of these ions, both under low and high insulin conditions, Samber et al. hoped to add new insights into islet functions.

1.4.2 Gestalt Principles

A particular form of 2D morphometry that has been recently used in diabetes research are Gestalt principles. Gestalt is German for 'shape' and refers to how the mind groups similar images and fills in the gaps in order to perceive an amenable version of reality [76]. In this research, Dybala et al. use the idea of a stereotypic islet architecture to 'fill in the blanks' and complete the missing data from an islet microscopy scan to get a full picture of what a human and mouse islets looks like. This could prove useful as there is no stereotypic human islet, so maximising the use out of existing data is highly important [77].

1.4.3 Persistent Homology in Mathematical Biology Research

Persistent homology is the method to study structural variation in data over a spatial parameter. Whilst the study and measure of abstract structure has not been conducted yet in islet research, in the wider field of mathematical biology research,

it is a new but promising field. Thus far, it has been used in fields such as protein structure [78] and branching topologies of plants [79] to advance our understanding of these concepts. In this piece of work, we will be using it to further our understanding of islets structure and diabetes.

1.4.4 Networks in Diabetes research

Networks theory applied to cellular composition in islets is relatively new but it is promising because the function of the islets depends on a network of the component parts. Stožer et al. modelled beta cell islet subnetworks, primarily considering calcium ions in mouse islets [80]. They investigate how mouse T2D affects the subnetwork connectivity and provides evidence that T2D lowers the connectivity of the islets.

Published in 2020 by Xavier et al., beta cell subnetworks have been identified as drivers for insulin secretion, and 'hub' or highly connected beta cells coordinate the action of other beta cells in the network. [81]. The research has been done using calcium imaging of mouse islets, from which it was deduced that these 'hub' cells had increased levels of the enzyme, glucokinase present. Understanding how the beta cells are connected within the islets is the next step to increasing our knowledge of beta cell functionality.

1.5 Overview of Thesis

Having outlined the problems that we wish to explore, above, this body of work will seek to primarily create a toolkit to add to the current knowledge of human islets of Langerhans. Broadly speaking, we will be comparing islet samples of young onset T1D with their respective controls, older onset T1D with their respective controls, and finally T2D and their respective controls. We then seek to draw some initial conclusion about changes, if any, that diabetes has on the islets of Langerhans, based on the small number of samples that we have. Our aim is to seek to demonstrate the feasibility of this methodology, which would allow us to draw stronger conclusions, should we gain access to a larger quantity of human data.

In chapter 2, we assemble an algorithm to convert micrographs to useable spatial data. The chapter considers a simple aggregation of the cell numbers between diabetic groups and their controls, and statistically compares them. This algorithm is novel in that, to our knowledge, is the first uses of feature detection and cell segmentation in an automated manner for islets of Langerhans.

Subsequently, in chapter 3, we use the spatial segmentation generated in chapter 2 to construct and utilise various spatial data to investigate various changes in the islets as a result of diabetes. The novelty here is that this is the first time that aspects of the geometry of the islets of Langerhans have been automatically analysed and quantified in order to assess the impact of diabetes.

Geometry can be considered as a spatial manifestation of topology. In chapter 4, we utilise persistent homology to analyse the underlying structures of the islets, and how diabetes may change this. This is the first time that topological analysis has been used in the field of diabetes research, and should be considered more as a proof of concept.

By considering the geometry and topology, in chapter 3 and chapter 4, we are analysing the macrostructure of the islets. We are also interested in further understanding how the microstructure of the islets changes as a result of diabetes. In chapter 5, we use the processed data from chapter 2 to generate networks based on the Voronoi partition between the identified cells in the islets. We then analyse how diabetes affects this underlying structure.

Chapter 6 finishes with a critical discussion of the results in the thesis. In this chapter, we collate the results from the thesis and use this to pose future hypothesis that can be assessed with additional data. We then suggest some direction for future research question that result from the work in this thesis.

2. Extracting Islet Structures from Micrographs

The aim of this chapter is to enable us to develop a more accurate understanding of the islets of Langerhans' cellular composition, in different subjects before and after the onset of diabetes. We compare the test subjects' cell counts with the control groups and the results have been validated by comparing our results to those obtained manually by the University of Exeter Medical School (UoEMS). We will then use this processed data in subsequent chapters, to further investigate how the islet structure, from geometric, topological and network perspectives, is altered by diabetes.

The Chapter is structured as follows: in section 2, we outline the algorithm used to analyse the micrographs as obtained from the UoEMS, followed by statistical analysis of these results. In section 3, we outline the *in silico* model that we will use in conjunction with the work in section 2 to improve our understanding of the aetiology of T1D and in section 4 we discuss and summarise the results of this chapter.

In future work, we would aim to process the images of T1D controls to serve as an input for a computational model for insulitis (further details are provided in Appendix C). The model will then simulate beta cell destruction using islet geometry extracted from the control subjects' images and output an *in silico* post-insulitis image. This image will be statistically compared to the processed T1D patient images, in order to draw conclusions about the underlying mechanisms of T1D beta cell destruction as stated in the thesis aims.

Despite the prevalence and serious nature of diabetes, very little is known about

the aetiology of this disease, which hinders effective treatment. Current methods of treatments include insulin injections, mostly for T1D [82], whilst for T2D, lifestyle changes are the preferred initial treatment [83]. Treating diabetes patients with exogenous (artificial, injected) insulin has severe side effects, including hypoglycaemia and hypokalemia (low blood potassium) [84]. Therefore a lot of research effort is concentrated on finding alternative ways for treating T1D patients that would negate the need for exogenous insulin, including islet transplants [85] and autoimmune therapy [86]. However, these approaches have been unsuccessful, despite Serup *et al.* predicting in the BMJ in 2001, "islet cell transplantation is likely to become the treatment of choice for most insulin dependent diabetics within the next five to ten years" [87]. Whilst islet transplantation works well, the treatment does not due to the limited supply of islets to transplant. The lack of success in all of these experimental treatments outlined in this paragraph can be partially attributed to the poor understanding of the architecture of the human pancreas and the islets [88].

A major reason for this lack of understanding is a shortage of human data as the pancreas rapidly self digests after death as a result of the enzymatic action of the organ's acinar tissues [3]. Furthermore, it is also difficult to access the pancreas due to its recessed location, anterior to the duodenum [3]. For this reason, much of the previous research has focused on non human islets as they are easier to obtain. Hence, our understanding of insulitis in animal models, such at the BioBreeding rat, is much better than in humans. [89]. Additionally, human islets behave differently to animal models [90], explaining why making further progress towards understanding human type 1 diabetes is challenging.

We are particularly interested in quantifying the number of beta cells in an islet in order to better understand the progression of diabetes. Brissova *et al.* [91] used murine islets to quantify alpha, delta and beta cells. Their method summed the intensity (on an 8-bit scale) after immunolabelling each of the hormones, per pixel, within the islet. From this, the percentage of an islet that each of these hormones occupied was recorded and compared to each other. However, the main drawback of this approach is that it does not identify individual cells, which limits the information

available, such as quantifying the number of cells in an islet. We are interested in this detail, particularly the number of beta cells, to precisely quantify the effects of diabetes and age on the islets via measuring the amount of hormones and pancreatic cell numbers.

An algorithm to analyse 3D mouse and monkey islet images has been developed in 2017 by Nhu *et al.* [92] . They apply this in a case study to analyse the structure and composition of 3D animal islet images. Their conclusions relate to the structure of animal islets, such as 3D clustering of alpha cells in murine islets. The specifics of their algorithm differs from ours, for example, they used a watershed algorithm [93] to identify the nuclei, which by their own acknowledgement, resulted in dense clustering and overlaps of nuclei in the islets, hence making cell counting problematic. By contrast, we used a local filter combined with an erosion algorithm (see image analysis section section 2.2) to help resolve this inaccuracy in nucleus identification.

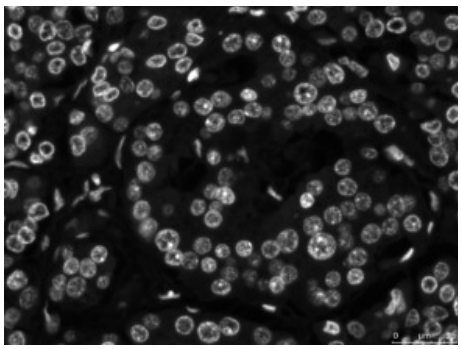
Despite the difficulty in acquiring *ex vivo* human islets and the subsequent sparse literature, it is known that human islet architecture changes as a result of diabetes [94]. A deeper knowledge of the morphology of the human islets of Langerhans is required to better understand the significance and role of insulitis with respect to diabetes. This includes a comprehensive knowledge of the ratios of different cell types throughout the islets, both before and after the onset of diabetes. Bosco *et al.* [95] have performed some analysis of human islet morphology. Their method involved comparing the ratios of the different hormones to each other, in the same manner as Brissova *et al.* They then further assessed the architecture by manually comparing the cells exclusively around its periphery and vasculature. By focusing on cells only in this region, this study was limited in the conclusions that could be drawn as these parts of the islet are not necessarily representative of the totality of the islets. Their work also demonstrated that human islets have a unique architecture - important for islet functionality, when compared to animals. However, to date, no studies have been undertaken to understand directly how diabetes affects human islet morphology, and existing animal models of T1D aetiology cannot be simply translated for humans [96].

Historically, image processing techniques on islets have often been performed by hand on individual images. This has included human islet images such as that done by Kilimnik *et al.* [97]. However, only limited conclusions could be drawn from islet architecture, due to the methodology limitations as identified above. By way of example, an important factor with respect to the onset of diabetes (especially T1D) is beta cell loss; this quantification has not been done in a systematic manner to date. Additionally, manual implementation of algorithms is subject to human biases. Therefore processes that are fully automated are desirable in order to extract reliable quantitative information from the available islet microscopy images.

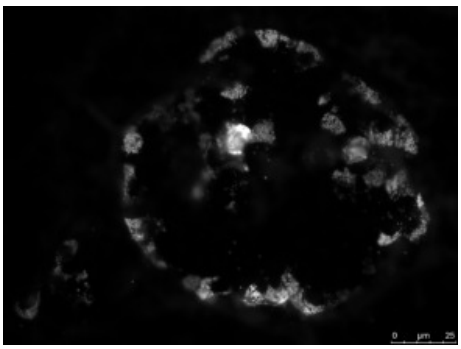
2.1 Description of Data Set

Recently the group of Professor Noel Morgan in the University of Exeter Medical School took microscopy images on a large number (250) of *ex vivo* human islet sets (example of image given in Figure 2.1). Each islet image set consists of 5 images: A DAPI stained image which we use to extract information about where the nuclei are in the islets. DAPI is a chemical which binds to DNA and fluoresces when exposed to certain wavelength of light; the three hormone stained images (insulin, glucagon and somatostatin) which we use to extract information about the boundaries of the islets and the types of cells present; and the final image is an overlay of the previous 4 images. This gives us a total of 1250 images for analysis. The image sets themselves are divided into three: Older onset T1D and their respective controls; young onset T1D and their respective controls; finally, T2D and their respective controls. These cases are age and sex matched with their respective control groups, based on the material available from the Exeter Archival Diabetes Biobank (EADB). Young onset usually means below 12 years of age and older onset is above this age.

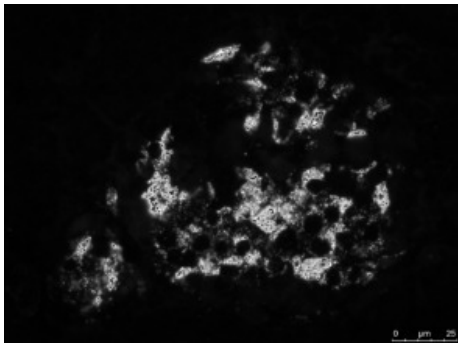
The microscopy images come in the form of tagged image file format (.TIF) and the image resolution is 1296 by 966 pixels. The images were taken at a 20$\times$ magnification on a LEICA DM4000 B-LED fluorescence microscope using LAS X image software, with a conversion scale between pixels and units of 25μm per 92 pixels. islets were selected at random, by searching the tissue under a particular



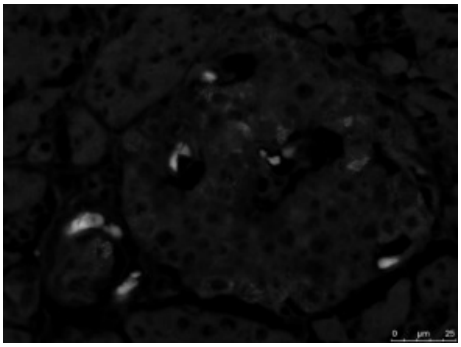
(a) DAPI stain of example islet image



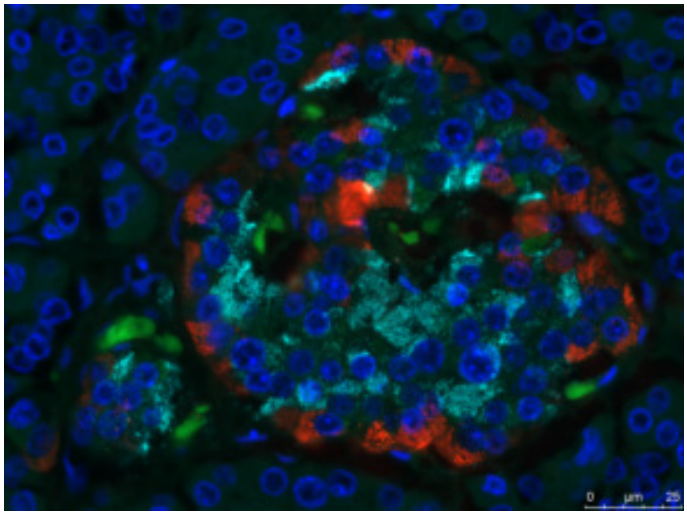
(b) Glucagon stain of example islet image



(c) Insulin stain of example islet image



(d) Somatostatin stain of example islet image



(e) Coloured overlay of all the above hormones

Figure 2.1: An islet immunofluorescence microscopy image set. Figure 2.1a represents DAPI (dark blue in Figure 2.1e), which fluoresces to indicate the presence of nuclei. Figure 2.1b shows glucagon (red in Figure 2.1e), which is a hormone that raises the blood glucose levels. Glucagon is secreted by alpha cells, which make up between 15% and 20% of the number of islet cells. Figure 2.1c is insulin (light blue in Figure 2.1e), which is a hormone that works to lower blood sugar levels. Insulin is secreted by beta cells, which comprises between 50% and 80% of the islet mass. Finally, Figure 2.1d is somatostatin (green in Figure 2.1e), which is a hormone that inhibits glucagon and insulin release. Somatostatin is secreted by delta cells which comprises between 5% and 10% of the islet mass. There are also two other cell types that the microscopy image does not display; gamma cells which secrete a hormone called pancreatic polypeptide (PP) that regulates pancreatic secretions and epsilon cells, which secrete a hormone called ghrelin that regulates appetite [2].

wavelength at which glucagon fluoresces (555 μm) to identify the islets. Numbers of patients and total islets per group are included in Appendix A.

2.2 Image Analysis of Islet Micrographs

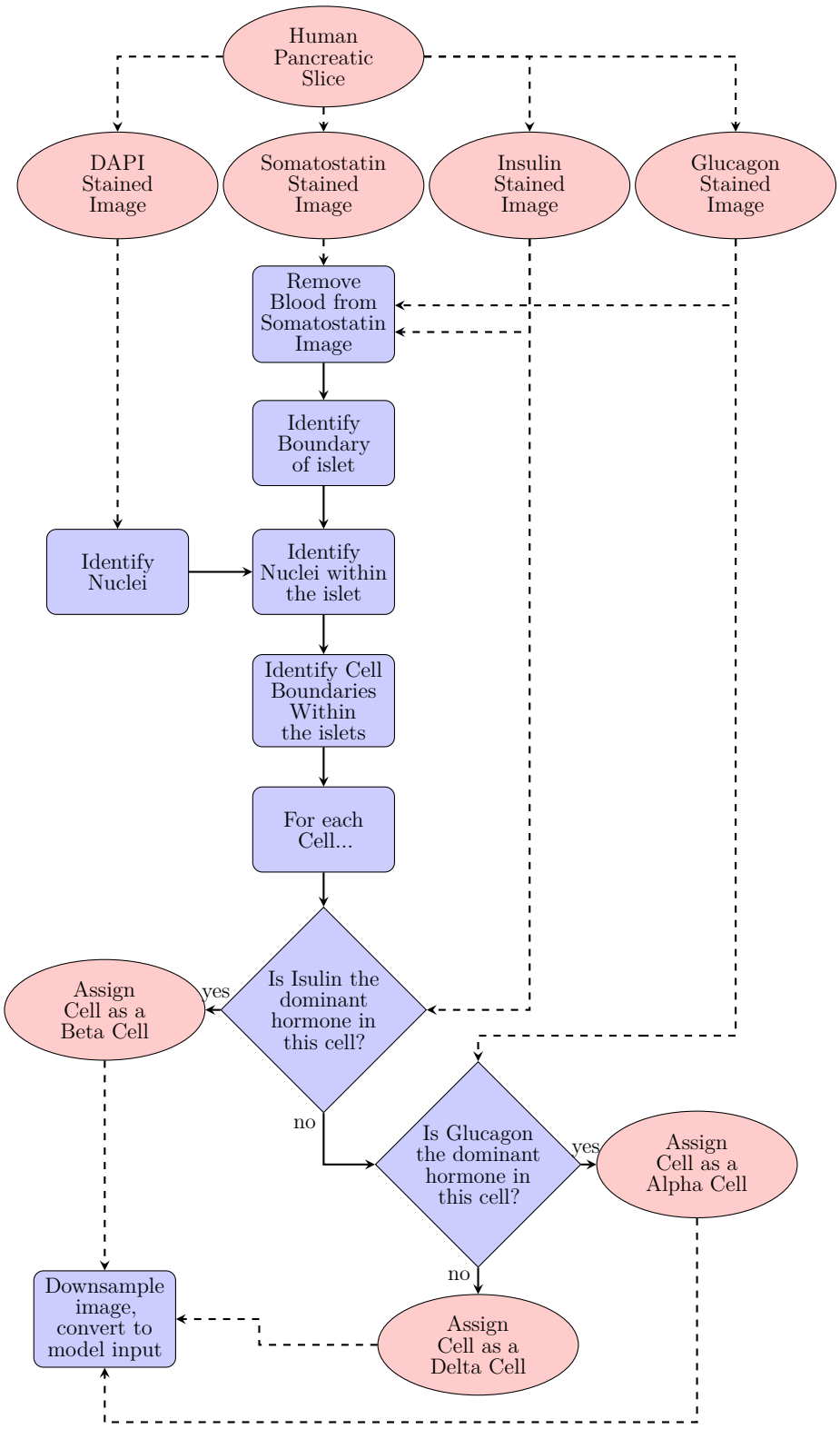


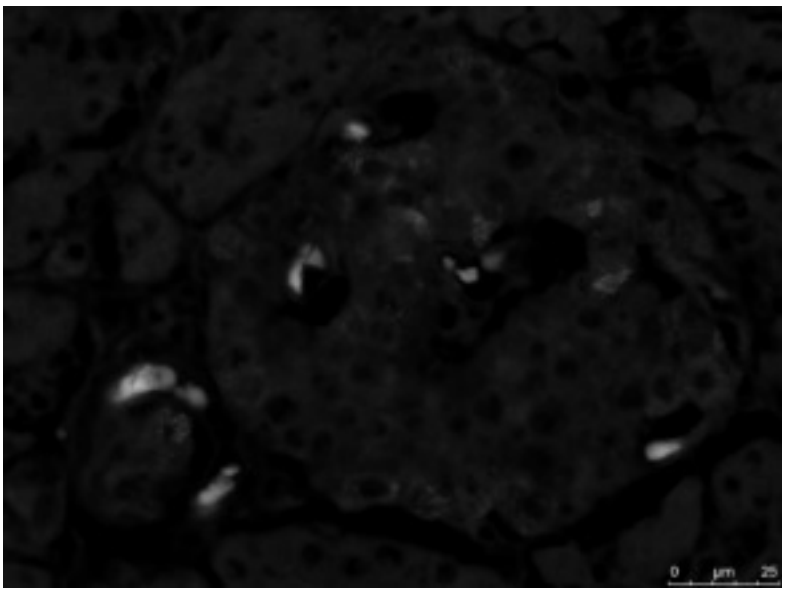
Figure 2.2: The flowchart outlines the image processing workflow. Briefly, the images are imported into an array and processed using primarily image filters. Subsequently, we use these processed images to decide what cells the image data corresponds to. The images are then further processed so that they can be used as an input for a cellular automata model of type 1 diabetes Appendix C.

The programming components of this work were conducted in Python3. Standard image processing techniques were employed using the scikit image library [98]. All of the linear algebra operations used the NumPy library [99]. Statistics were done using the scipy stats library [100]. Each of the images were read into a predimensioned array object to $1296 \times 966 \times 3$, corresponding to the image dimensions and the overlay or stack of red, green and blue images, respectively. The objective of the image processing was to convert the microscopy images to quantifiable data with respect to the types of cells present, the outline is displayed in Figure 2.2. We then downsample (reduce the resolution of) the images so that they may be used as an input for a *in silico* insulitis model. Outlined below is the methodology via which this is achieved. The scale on these images is $25\mu m$ per 92 pixels - which we denote as S_c. It is also useful to define here the processed islet outline black and white (binary) images B which is the member of the set of processed islet outline binary images $\mathcal{B}$.

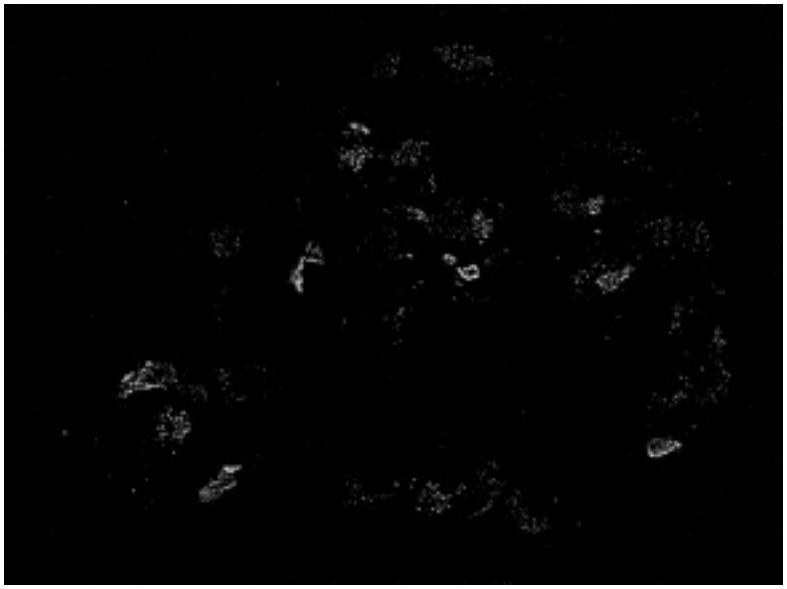
2.2.1 Removing Blood Artefacts From the SST Image

The first image processing step was to remove blood artefacts from the somatostatin (SST) images, which we denote with T. Under the wavelength that the SST was imaged at, blood fluoresces. Subject specialists can tell the difference between the blood and somatostatin as somatostatin appears as granular objects, whereas blood appears as a blurry circular shapes. To differentiate the two, we took the Laplacian of the somatostatin image. A Laplacian is an operator that measures the magnitude of the second order derivatives, blurry objects have a low magnitude under the Laplacian filter, whereas granular objects have a much higher magnitude. Acting on the SST image array, the operation is:

$$\Delta T = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) T, \tag{2.1}$$



(a) This SST image comes from a T2D control patient. There is a large quantity of dull background green within the islet, representing the blood and some small grain blobs in the foreground, representing the presence of SST cells.



(b) This is the mask as generated from Figure 2.3a using a Laplacian, Gaussian blur, and triangle filter as outlined above

Figure 2.3: Demonstration of the process for preprocessing the SST images.

where T is the gray-scale SST image array; This is approximated and implemented by scikit image using the five point finite difference method [101]:

$$\Delta T_{i,j} \approx \frac{1}{S_c^2} \left(T_{i+1,j} + T_{i-1,j} + T_{i,j+1} + T_{i,j-1} - 4 \cdot T_{i,j} \right), \tag{2.2}$$

where, i and j are indexes representing the horizontal or x and vertical or y pixel positions respectively. We then apply a Gaussian blur, using the scikit image library

[102], with a parameter of 1 pixel to smooth the data, followed by a triangle filter, from the scikit image library [103] to create a mask for the SST image. A mask is a binary array where values take either true or false. Both the triangle filter and the Gaussian blur are explained in more detail below. By applying the Hadamard product (also called entry-wise product) between the original image and the mask, the original image is allowed through where the mask is true (1) and where false (0), the value of the new image array is also zero. This is displayed below in Figure 2.3.

A Gaussian blur is used to smooth an image to account for noise, obtained during the microscopy process, although this reduces the detail of the image. The blur works by convolving each pixel upon its corresponding discrete Gaussian kernel. The Gaussian kernel in a two dimensional space, i, j, can be expressed as:

$$f(i,j) = \frac{1}{2\pi\sigma^2} \cdot \exp\left[-\frac{i^2+j^2}{2\sigma^2}\right], \tag{2.3}$$

where i and j are the pixel location in the 'x' and 'y' directions. σ is the standard deviation, which represents a blurring coefficient. For us, in the SST image, this coefficient was 1, which was arrived at by a parameter search combined with visually inspecting the image. > 1 resulted in outlines that lost too much detail regarding the areas of the mask. < 1 resulted in images that were disjointed and did not sufficiently capture the macroscopic details of the SST mask.

Subsequently, we needed to identify the islets from the background of the image. The triangle filter is an image processing method to separate a foreground from a background. The method is appropriate for images that have a histogram which is unimodal. The histograms of our islet images are highly unimodal histograms, as shown in Figure 2.4a. This algorithm has been used in bioimage processing since 1977 to split a foreground and a background [104].

Firstly, the SST, Insulin and GCG channels are combined and converted to grayscale. The intensity of each pixel is recorded and used to generate a histogram. The horizontal axis is categorised into bins, denoted by x_i, $i \in 0..256$, corresponding to the possible quanta of intensity from 8 bit images ($2^8 = 256$). Each x_i has corresponding frequency, H_i. Secondly, we find the both index, $i = M$, corresponding

to the max histogram frequency (2.4) and the index, $i = m$, corresponding to the max bin with non zero value (2.5).

$$M = \underset{i}{\text{argmax}}(H_i) \tag{2.4}$$

$$m = \max\{j \mid H_j \neq 0\} \tag{2.5}$$

The triangle filter [104] is applied to this image by drawing a straight line (L_1) between the Cartesian points (x_M, H_M) and (x_m, H_m). Following this, the distance, d_j is calculated perpendicularly from L_1 to H_j (2.6), $j \in M...m$, and the argument $j = T$ is that at which d_j is maximal (2.7).

$$d_j = d(H_j \perp L_1) \mid j \in (M...m) \tag{2.6}$$

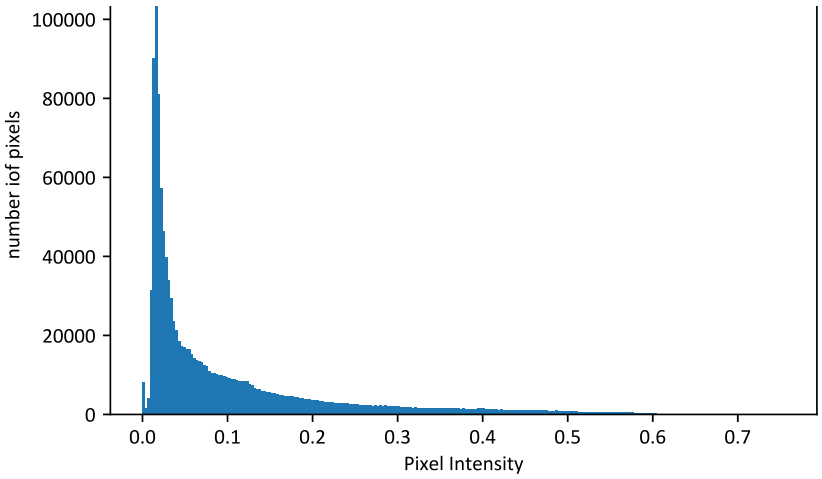
$$T = \underset{j}{\text{argmax}}(d_j) \tag{2.7}$$

x_T is determined to be the threshold value. This is illustrated below in Figure 2.4b.

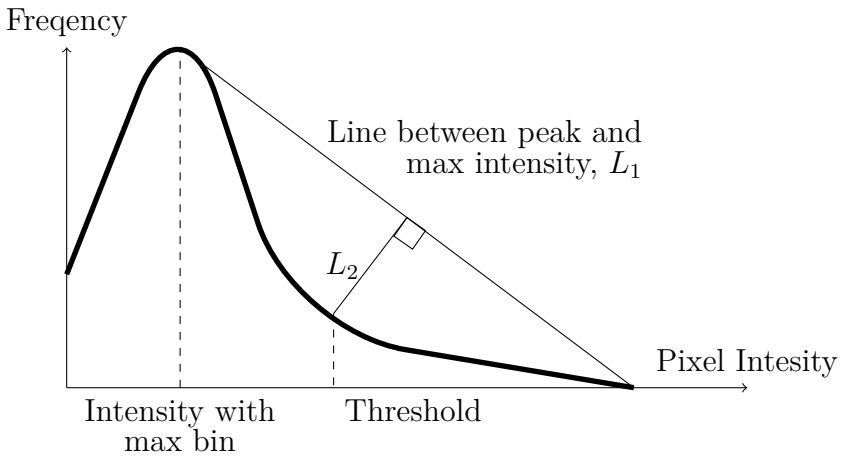
2.2.2 Finding Islet Boundary

To find the boundary of an islet, we aggregated the 3 hormone stains, insulin, glucagon and somatostatin to form a grayscale image, via addition. From that, to smooth the image, we took the Gaussian blur, with a smoothing parameter (σ value) of 10. This parameter value balanced the need on one hand to smooth the noisy data (the higher the parameter the better the smoothing) with the need to preserve important information about the shape of the islets (the lower the parameter the less data lost). We then applied a triangle filter on this Gaussian blurred image to create a binary image in which 1 represents a point in the islet and 0 corresponds to a point in the background tissue. However, despite smoothing the image, there were still small regions of pixels outside the islet and small holes in the islets (this can be seen in Figure 2.5d), which was resolved by removing small objects and holes with areas under 1000 pixels; a balance between removing visual artefacts and not removing "legitimate" features, arrived at by using a visual inspection and a parameter search. This was done by implementing an algorithm from scikit-image

Figure 2.4: (a) Typical image histogram; (b) A diagram presenting the idea of the triangle filter.



(a) A histogram of the example image in Figure 2.1. There are 256 bins, with each bin corresponding to a level of intensity in the image. Each Column is the number of pixels with that particular intensity. Note the highly uni-modal nature of the histogram, making it suitable for processing with a triangle filter.

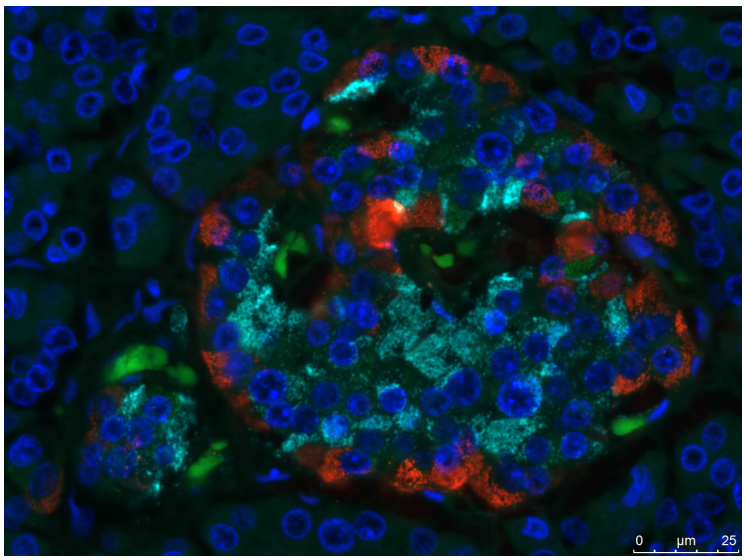


(b) A diagram illustrating how a triangle filter works on a histogram of an image to generate a threshold. Step 1 is to generate a 256 binned histogram of the image, the peaks being represented by the thick black line on the diagram. Step 2 is to locate the bin with the maximum intensity and draw a straight line from this peak to the highest pixel intensity present (L_1). Step 3 is to draw a perpendicular line from the previous line to the histogram peak line, with the maximum Euclidean distance (L_2). Finally, the point on the histogram line at which the perpendicular line gives the threshold between foreground and background.

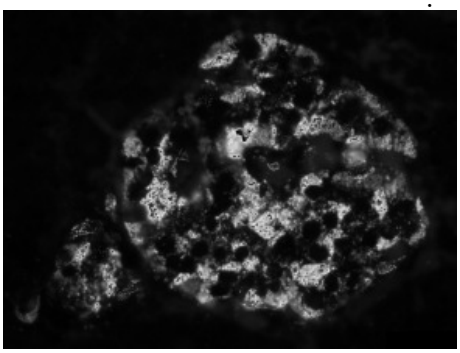
[105] and a comparison of an image and its processed boundary are provided below, in Figure 2.5. We denote binary image matrices of the islets as $\mathcal{B}$.

2.2.3 Finding the Nuclei of the Islet Cells

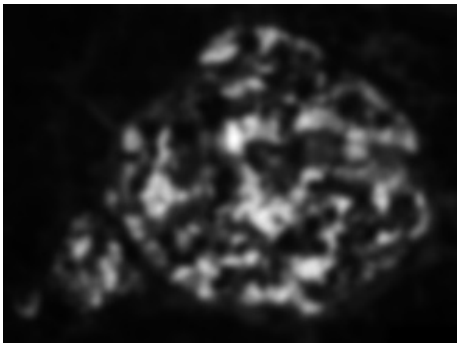
We wished to find points, unique to each cell, in order to segment the cells within the islets. Nuclei are unique to each cell, so by finding points near the centres of



(a) This is the overlay of the images that we are using in this example



(b) This is the result of aggregating the insulin stain, the glucagon stain, and the post processed somatostatin stain, as shown in Figure 2.5a. This a false colour grayscale image



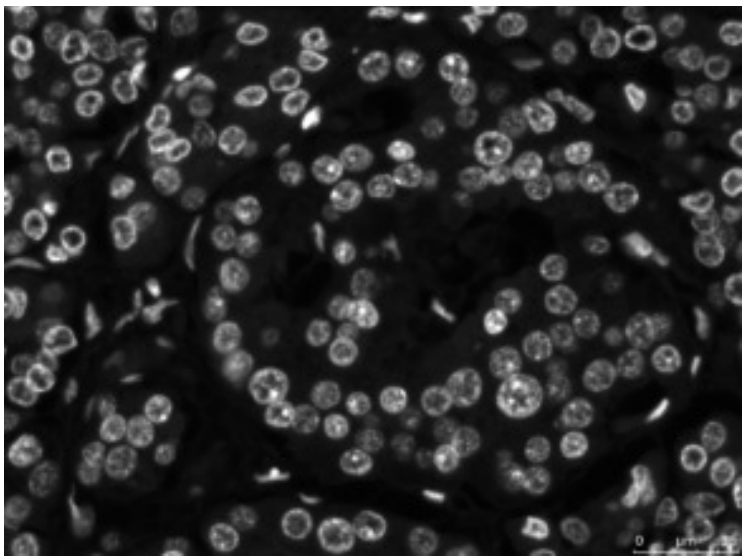
(c) This was the result of applying a Gaussian filter with a smoothing parameter of 10 pixels on Figure 2.5b

(d) This was the result of applying a triangle filter to separate the foreground from the background in Figure 2.5c

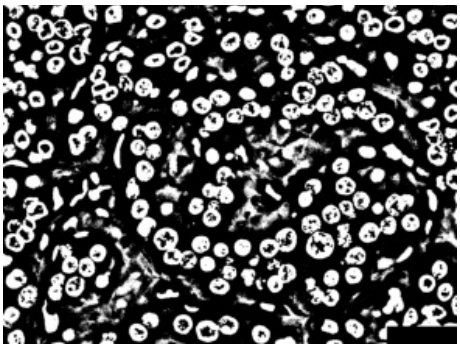
(e) This was the result of removing the small objects, which we deemed to be 'noise' from Figure 2.5d

Figure 2.5: The 4 image processing steps taken to deduce the islet boundary from the three endocrine hormone stained images.

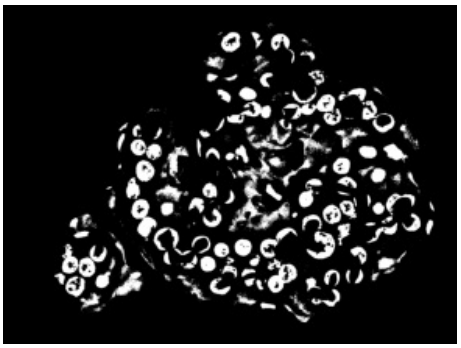
nuclei, we have points around which to segment cells. For this, we used the DAPI stains, as DAPI binds to DNA, which appear as luminous circular objects on the



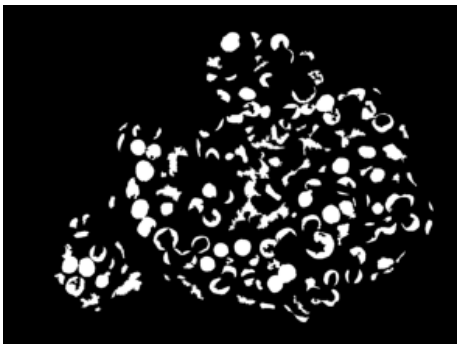
(a) Original DAPI stained image. Note the nuclei overlap in highly dense regions, causing a 'glowing' effect.



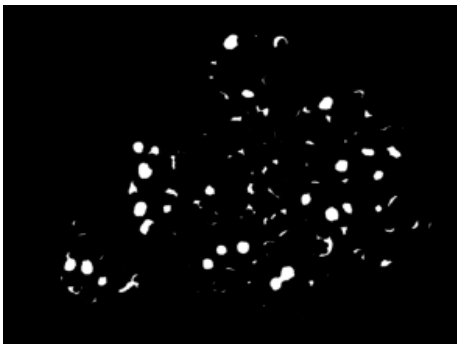
(b) DAPI processing step 1, apply a local filter to Figure 2.6a



(c) Apply the islet mask as generated in Figure 2.5 upon the image as generated in Figure 2.6b. This removes nuclei that are not in the islet.



(d) Removal of small objects and holes from Figure 2.6c



(e) Erosion algorithm on Figure 2.6d, to remove noise and disentangle nuclei. Displayed in the image are blob fragments of nuclei. We are only interested in where the centres of these blobs and by extension, nuclei are, with a modelling assumption that these are the centres of the cells.

Figure 2.6: The image processing steps outlined above are required to identify where the centres of the nuclei are.

images. However, these circular objects express a halo [106], which means in high nuclei density regions, such as the islets, the non nuclei parts of the islet can be just as bright as the nuclei in the background region. Consequently, using a filter with a uniform threshold such as a triangle filter would be inappropriate here. We therefore used an adaptive threshold filter, or local filter to achieve this. This works by determining the mean pixel intensity of a surrounding 101 by 101 square pixel object, and using the resulting value as the threshold value for that pixel (shown in Figure 2.6b). Objects with a smaller pixel area resulted in areas outside the nuclei being above the threshold, and a larger pixel area resulted in large swathes of the islets in their entirety being above the threshold value.

This process resulted in a noisy image, so to further rectify this, we removed small holes and objects under 100 pixels (shown in Figure 2.6d), making the assumption that objects smaller than 100 pixels represent image noise and anything bigger represents nuclei. Subsequently, we applied the islet shape as a mask to the image (shown in Figure 2.6c). However, the resultant image still had overlapping nuclei, and some artefacts from above and below the 2-dimensional microscopy image plane. To fix this, we applied an erosion algorithm, using a structuring element of a disk with radius of 7 pixels (shown in Figure 2.6e). Erosion algorithms work by tracing around the edge of an object with a structuring element so that overlapping points are removed. The point of contact between the image and the structuring element is the centre of the element. In our case, this is the geometric centre of the disk. From this, the x and y positions of the centres of the remaining areas were stored as the centres of the nuclei within the islets, and the points around which we dissect the region into cells. The erosion object was selected in order to provide a balance to be a big enough shape to disentangle the nuclei, whilst not removing them in their entirety, obtained through visual inspection and a parameter search.

2.2.4 Identifying the Cell Boundaries

The next step in our analysis was to identify the cell boundaries within the islets. We did this via an implementation of a Voronoi algorithm, shown in Figure 2.7. We

had to post process the output from the Voronoi algorithm from the SciPy spatial library [107]. The reason why it could not be used directly was that it was designed to be used with the matplotlib library, for graphing, whereas we wished to implement it with the Scikit Image library, for image processing and therefore we had to modify the code accordingly.

The Voronoi algorithm represents a decomposition of a metric space ($\mathbb{M}$), equipped with metric function $\mu : \mathbb{M} \times \mathbb{M} \rightarrow [0, \infty)$ around a finite set of points ($p_i \in \mathbb{M} : i \in 1...n$) [108].

Let $x \in \mathbb{M}$ be a point in the metric space, then let the bisection between p_i and p_j be a subset of points in $\mathbb{M}$ denoted by:

$$L_{i,j} = \{x \mid d(p_i, x) = d(p_j, x)\} \tag{2.8}$$

This bisection divides the metric space, and denote the subspace containing p_i as:

$$D_{i,j} = \{x \mid d(p_i, x) < d(p_j, x)\} \tag{2.9}$$

Each Voronoi region VR_i can be considered as an intersection of all the $D_{i,j}$ subspaces generated over the index j:

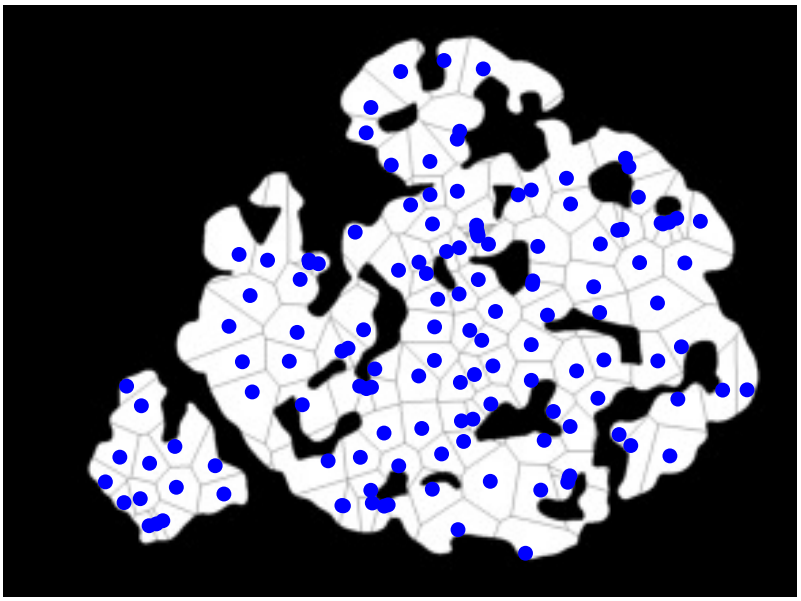
$$VR_i(\mathbb{M}) = \bigcap_{j \in 1...n, j \neq i} D_{i,j} \tag{2.10}$$

The union of the intersection of all the Voronoi region gives the voronoi diagram upon $\mathbb{M}$:

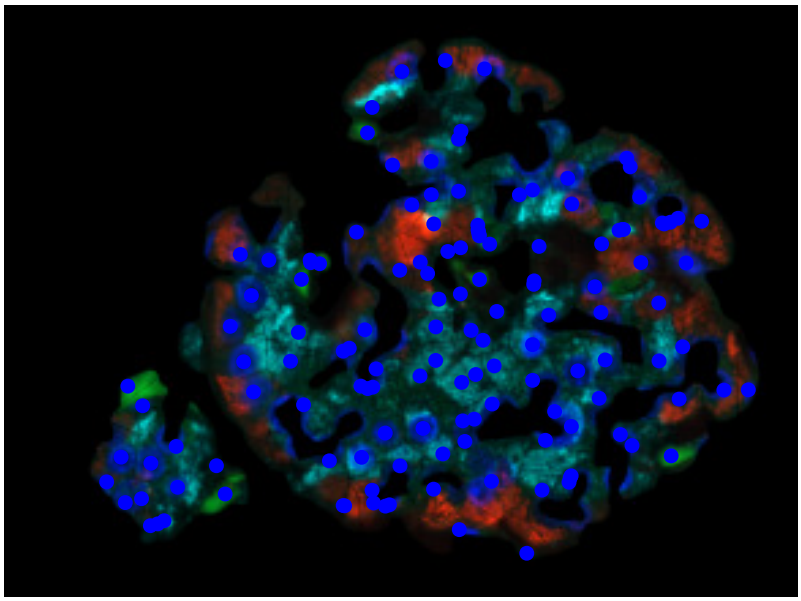
$$V(\mathbb{M}) = \bigcup_{i,j \in 1...n, i \neq j} VR_i \cap VR_j \tag{2.11}$$

In two dimensional Euclidean space, the most efficient way of computing this is to find the dual graph of the Delaunay triangulation of a set of point in the plane ($a_1...a_n$). A Delaunay triangulation is a way of triangulating $a_1...a_n$, such that no point, $q \in a_1...a_n$ lies within the circumcircle of any triangle. A circumcircle of a triangle is the circle defined by the three vertexes of the triangle. The dual graph upon a graph maps planes to vertexes; mapped vertexes are connected by edges if

the corresponding faces are adjacent. The vertexes in this dual graph are calculated to be at the circumcentres of the Delaunay triangles.



(a) Image displaying the cell boundaries, calculated from a Voronoi algorithm. The nuclei centres, given by the blue spots are based on Figure 2.6. The white region is identical to the one displayed in Figure 2.5e



(b) Here the white in Figure 2.7a is overlayed with the hormones from the original islet image in Figure 2.1.

Figure 2.7: Presented here are firstly the Voronoi split cells, followed by an overlay with the hormones from the original islet.

2.2.5 Identifying the Type of Cell

In each cell, defined by the Voronoi algorithm, as shown in Figure 2.7a, we needed to assign it to a type (alpha, beta or delta). This was done by summing the levels of

each hormone (glucagon, somatostatin and insulin) respectively within each identified cell (shown in Figure 2.7b) and recording this. The cell type was then determined from the highest level of hormone within each cell. Due to the effect of the Laplacian filter on the SST preprocessing, as outlined in Figure 2.3, some of the hormone staining was removed, so a corrective factor was required. By manually analysing the ratio of SST mask area to the shape overlayed by the SST hormone mask, in imageJ, combined with a visual inspection and a small parameter search, this ratio was calculated to be approximately 1.4. The result is shown in Figure 2.8.

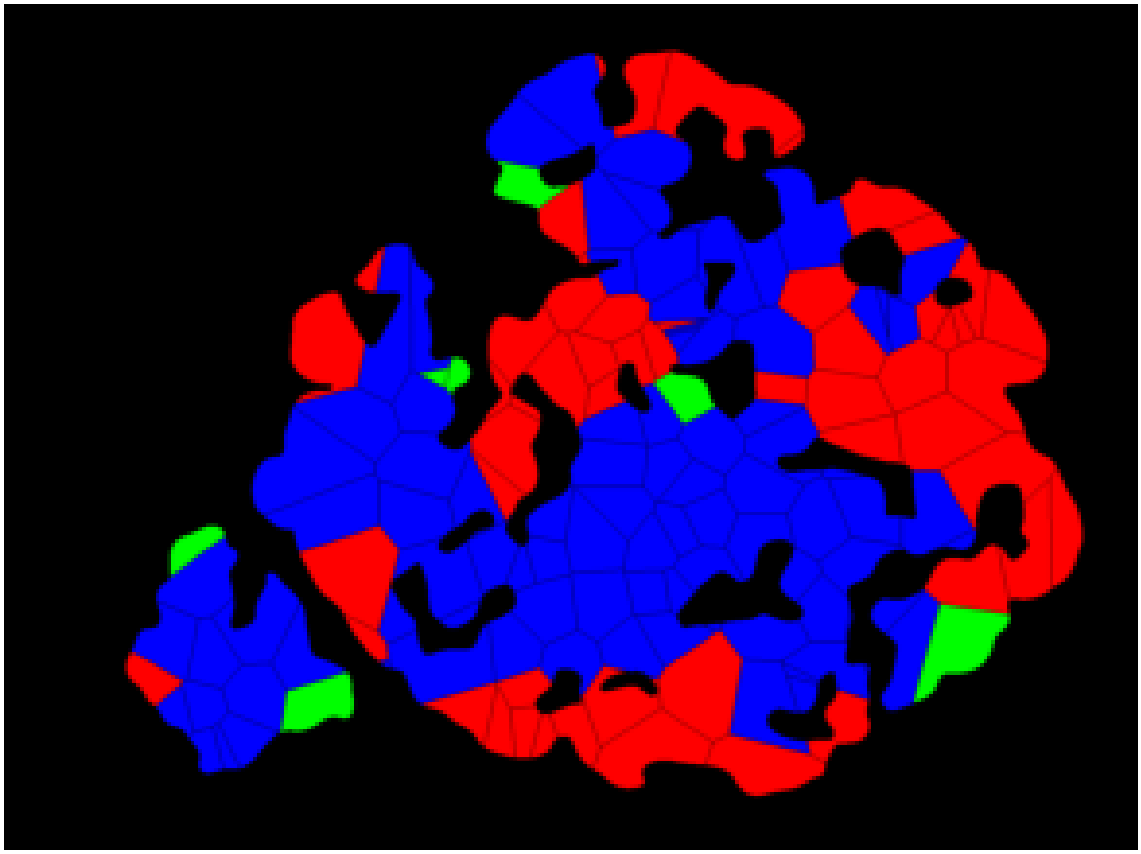


Figure 2.8: Cells were assigned to either alpha, beta or delta, based on their relative amounts of hormones that they contained. Due to the preprocessing of the SST stains, a corrective weighting had to be used when assessing these levels.

2.3 Method Validation

We have to test that our results from the image processing matches up to a small, random selection of cell counts, done by hand, at the UoE Medical School

2.3.1 Validation of Results

In order to validate the image processing methods, we compared our results with expert manual analysis of 5 islets from both the T1D and T2D cases and took the absolute percent difference as the comparison metric. These islets were randomly selected from both these groups. We then compared our work against results obtained through the University of Exeter Medical School (UoEMS). Their algorithm was to firstly record percentage of each islet occupied by the 3 endocrine hormones (SST, GCG and Insulin). Secondly, they recorded the total cell count. Both these processes were conducted manually in ImageJ. The results of this algorithm were recorded in recorded and displayed in the tables below.

After being given the patient identifiers and image numbers, we compiled our own table from our own results, via our own algorithm. We then compared the two by calculating the difference in percent composition between their measurements and ours, and taking the mean. The maximum discrepancy was in the alpha cell counts, with an absolute percent difference of 12.93%. Given that cell identification within an islet image is subjective, in general, the fact that our results correspond reasonably well to those obtained from the UoEMS, provides evidence that the algorithm that we conducted is accurate.

Presented here are 5 tables. The first is are the cell counts as done by an expert in UoEMS and the second are the cell counts as done using our algorithm Figure 2.9. On the second page, these are displayed as percentages of the islet, and the final table finds the absolute difference in percent composition between the two Figure 2.10.

2.3.2 Parameter Sensitivity Analysis

Parameter sensitivity analysis is the measure by which the change in parameter in an algorithm input affects the outputs from the system. In other words it can tell us how much the uncertainty of the input can affect the reliability of the output.

We have done a parameter sensitivity analysis firstly from the individual sample image, used throughout this piece of work from Figure 2.1, followed by a parameter

Result from UoEMS						
Patient type	Patient identifier	Image group	Total cell count	Alpha cell count	Beta cell count	Delta cell count
T1D	nPOD 6038-01PB	1	96	34	57	5
T1D	nPOD 6038-01PB	5	39	16	0	15
T1D	SC116	7	89	42	0	12
T1D	SC116	2	33	17	0	7
T1D	nPOD 6086-02PB	9	29	16	0	8
T2D	PM138 66 4	2	71	19	38	3
T2D	PM138 66 4	3	46	12	27	5
T2D	PM184 66 2	10	62	26	29	9
T2D	PM192 69	2	42	10	23	6
T2D	PM192 69	3	106	29	26	14

Our Results						
Patient type	Patient identifier	Image group	Total cell count	Alpha cell count	Beta cell count	Delta cell count
T1D	nPOD 6038-01PB	1	115	32	80	3
T1D	nPOD 6038-01PB	5	47	26	3	18
T1D	SC116	7	68	51	1	16
T1D	SC116	2	28	21	0	7
T1D	nPOD 6086-02PB	9	32	23	2	7
T2D	PM138 66 4	2	71	22	45	4
T2D	PM138 66 4	3	46	13	28	5
T2D	PM184 66 2	10	71	17	48	6
T2D	PM192 69	2	38	10	19	9
T2D	PM192 69	3	121	52	58	11

Figure 2.9: In these two tables, we present the firstly the medical school results followed by our for analysing the cell counts in a random selection of islets from our sample sets. In the medical school results, the total cells and the breakdown don't always add up. This is due to either the presence of other cell types or uncertainty of cell type.

sensitivity analysis on the population of islet images. We have done analyses to investigate the stability of the algorithm under perturbations of:

1. The blood Gaussian blur parameter
2. The islet Gaussian blur parameter
3. The local filter window size
4. And finally, the radius of the erosion parameter for the DAPI stains

Subsequently, we plot in Figure 2.11 how varying the parameter changes the cell count in the individual analysis and the mean number of cells per islet image for the population analysis.

In Figure 2.11a, we have very little change in cell count with respect to parameter variation throughout the graph. We would expect the cell count to be

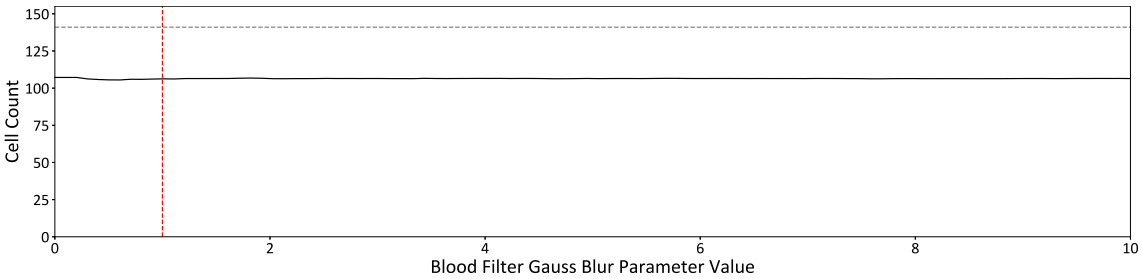
Result from UoEMS						
Patient type	Patient identifier	Image group	Total cell count	Alpha cell count	Beta cell count	Delta cell count
T1D	nPOD 6038-01PB	1	96	35.42%	59.38%	5.21%
T1D	nPOD 6038-01PB	5	39	51.61%	0.00%	48.39%
T1D	SC116	7	89	77.78%	0.00%	22.22%
T1D	SC116	2	33	70.83%	0.00%	29.17%
T1D	nPOD 6086-02PB	9	29	66.67%	0.00%	33.33%
T2D	PM138 66 4	2	71	31.67%	63.33%	5.00%
T2D	PM138 66 4	3	46	27.27%	61.36%	11.36%
T2D	PM184 66 2	10	62	40.63%	45.31%	14.06%
T2D	PM192 69	2	42	25.64%	58.97%	15.38%
T2D	PM192 69	3	106	42.03%	37.68%	20.29%

Our Results						
Patient type	Patient identifier	Image group	Total cell count	Alpha cell count	Beta cell count	Delta cell count
T1D	nPOD 6038-01PB	1	115	27.83%	69.57%	2.61%
T1D	nPOD 6038-01PB	5	47	55.32%	6.38%	38.30%
T1D	SC116	7	68	75.00%	1.47%	23.53%
T1D	SC116	2	28	75.00%	0.00%	25.00%
T1D	nPOD 6086-02PB	9	32	71.88%	6.25%	21.88%
T2D	PM138 66 4	2	71	31.67%	63.33%	5.00%
T2D	PM138 66 4	3	46	27.27%	61.36%	11.36%
T2D	PM184 66 2	10	71	23.94%	67.61%	8.45%
T2D	PM192 69	2	38	26.32%	50.00%	23.68%
T2D	PM192 69	3	121	42.98%	47.93%	9.09%

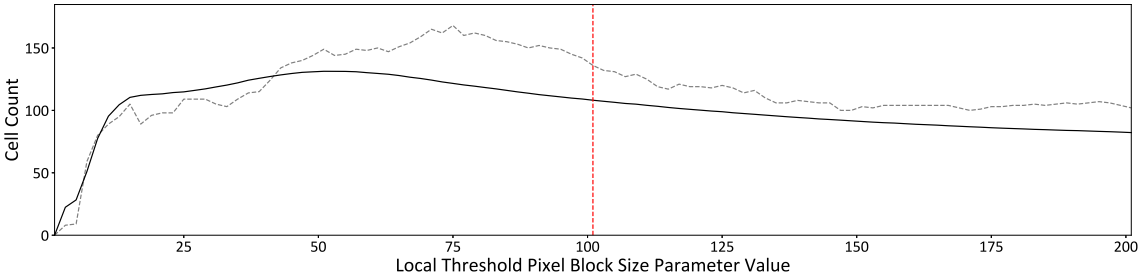
Composition Percent Difference Table						
Patient type	Patient identifier	Image group	Total cell count	Alpha cell count	Beta cell count	Delta cell count
T1D	nPOD 6038-01PB	1	19.79%	7.59%	10.19%	2.60%
T1D	nPOD 6038-01PB	5	20.51%	3.71%	6.38%	10.09%
T1D	SC116	7	23.60%	2.78%	1.47%	1.31%
T1D	SC116	2	15.15%	4.17%	0.00%	4.17%
T1D	nPOD 6086-02PB	9	10.34%	5.21%	6.25%	11.46%
T2D	PM138 66 4	2	0.00%	0.00%	0.00%	0.00%
T2D	PM138 66 4	3	0.00%	0.00%	0.00%	0.00%
T2D	PM184 66 2	10	14.52%	16.68%	22.29%	5.61%
T2D	PM192 69	2	9.52%	0.67%	8.97%	8.30%
T2D	PM192 69	3	14.15%	0.95%	10.25%	11.20%
		Average	12.76%	4.18%	6.58%	5.47%

Figure 2.10: In these three tables we present firstly the two tables from Figure 2.9 but as a percent composition of the islets so that they can easily be compared as a percent difference in table three.

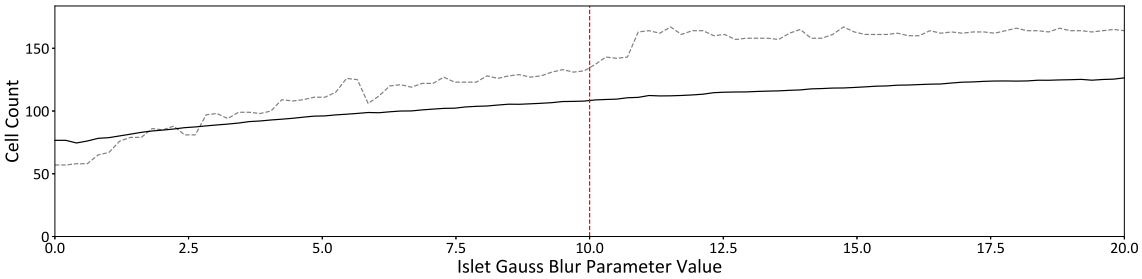
relatively unaffected by this parameter as it primarily affects the designation of cells. In Figure 2.11b , we vary the pixel edge length of square for the local filter parameter as outlined in subsection 2.2.3. Between 1 and 13, we see a rapid increase in the cell count, followed by a plateau and then a slow decline in the mean cell count



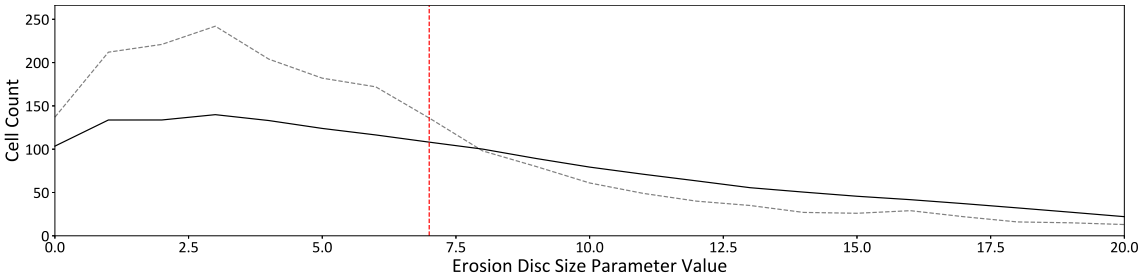
(a) Here, we have on the 'x' axis, the Gaussian blur parameter for the SST stained image to remove the blood. The red dashed line is the default value which is 1. The parameter was tested from 0 to 10 with a stepsize of 0.01.



(b) Here, we have on the 'x' axis, the size of the pixel edge length for the square that forms local filter parameter. The red dashed line is the default value which is 101. This parameter can only be an odd number. We vary this from 1 to 201 with a step-size of 2.



(c) Here, along the 'x' axis, we have the value for the Gauss blur parameter for the overall islets. The red dashed line is the default parameter which is 10. We vary it from 0 to 20 with a step size of 0.2.



(d) Here, the 'x' axis is the parameter of the erosion element disk radius. The red dashed line is the default parameter which has a value of 7, and must be a discrete value. We vary it from 0 to 20 with a step size of 1.

Figure 2.11: Here are the parameter stability analysis for the key parameters for assessing the number of cells per image. Along the 'y' axis, we have the mean cell count per image. The black solid line represents the mean across all islet images and the dashed line represents the parameter sensitivity analysis performed on the image in Figure 2.1

within each image. From a parameter sensitivity perspective, our parameter choice is sensible as there is a low parameter variability in the region of a length of 101 pixels square edge length, which is to say that in the region of 101 pixels the graph

remains reasonably constant. In Figure 2.11c, we observe a small increase in the mean number of cells in an image as we vary the Gauss blur parameter for the islet image. The graph is reasonably constant around our choice of parameter, so this is a sensible choice from the perspective of sensitivity analysis. Finally, in Figure 2.11d, we see an initial plateau of values, followed by a moderate decline of cell count with respect to variation in the parameter of the erosion element. There is a negative gradient in the graph around our choice of parameter value; however, this is expected, as when we perform an erosion algorithm on a number of small objects, we would expect that the larger the erosion element size, the more objects will disappear.

The results from the population average from our parameter search in Figure 2.11 seem to agree almost exactly with our individual sample analysis. The main difference is that in the population analysis, the uncertainties are lesser. That is to say, in all the parameter values in our model, the graphs are approximately locally constant in the region of our parameter choice lie which is denoted by the dashed red line in Figure 2.11.

2.4 Statistical Analysis

The primary objective of the statistical analysis is to assess what changes occur in the islet cellular composition between subject groups and controls as a result of diabetes. The islets' cell counts are assumed to depend on 3 significant factors: the subject group that they come from, variability of cell counts between different patients within their subject group, along with individual islet variability within each patient. Subject group refers to whether the patient or islet image belongs to either the old onset T1D, the young onset T1D, the T2D group or these groups' respective controls. Each subject group contains several patients (between 3 and 7, full details are in Appendix A), and each of these patients have several islet images associated with them (from 6-12).

Firstly, We need to test the variability of the islet images' cell counts between the patients within each subject group to assess whether an islet image has a cell count dependence on the patient to which it belonged. If this variability is statistically

significant, then this will need to be accounted for in subsequent tests, by adjusting the degrees of freedom. By way of example, a patients' mean beta cell counts within each islet could be dependent on the type of T1D that they have (recent authors have suggested that there are several phenotypes of T1D [25]). If this were the case, the number of independent samples would be the number of T1D patients, as opposed to the number of T1D islets. Secondly, we then need to test which cell counts are different between subjects' islets and their respective controls. Thirdly, we wish to ascertain the magnitude of the cell changes between the subject islet groups and their controls. This can be achieved by finding the effect size for these significant changes.

2.4.1 Testing Islet Variability Between Patients Within Subject Groups

In order to test the variability of an islet's cell count between patients, we need to establish what type of statistical test to use. If the data was normally distributed, this would be a simple procedure of applying Student's t-test. However, via application of the D'Agostino's K-squared test [109], it was found that the data was not normally distributed. The D'Agostino K-squared test works by taking the square sum of the kuritosis and skewness of the data, to calculate a K statistic. If normally distributed, this K statistic will result in a Chi squared distribution with 2 degrees of freedom. A Chi squared test then returns a P-value, which can be used to determine if the data is normally distributed. The mean probability of the groups of data being normally distributed is 9.27%, so we can reject the hypothesis that the data is normally distributed at a 95% threshold. We therefore chose to use a non-parametric method to test islet variability between patients within subject groups.

Given that we wish to test if there is a correlation between islet cell count and the patient (within subject groups) to whom it belongs, we used a non-parametric analysis of variance (ANOVA) test, or a Kruskal Wallis (KW) H test. This test works by ranking the data, taking a square sum penalty for each subjects' deviation from the average rank, with a corrective factor for sample size. The mean P value for this

test was calculated to 0.431% over all the groups, so we reject the null hypothesis that the cell counts exhibit a dependency on the patient within the subject group. This suggests that we do not need to correct for clustered data as islets are determined to have an insignificant association between patients within subject groups. The KW test is implemented by [110] and the results are recorded in section A.1.

2.4.2 Testing for Cell Type Change Between Subject Groups

Next, we determine whether there is a change in the cell counts between types of patients and their respective controls. As the data was non-normally distributed, it was appropriate to use a non-parametric test to compare the samples. The Mann Whitney U two-tailed test allows for testing if a randomly selected item from one list would be equal to a randomly selected item from a second list, via ranking the data and measuring the rank deviation from the mean rank. This was implemented using a SciPy stats function [111], and the results were recorded in section A.2. Below, the overall cell counts are shown in Figure 2.12 and cell counts, broken down into the three endocrine cell types are depicted in Figure 2.13.

The results from the statistical analysis could be summarised as follows:

- In the older onset T1D group, we find significant increase in the delta cell count between the patients and control, with a decrease in beta cell count. Total cell count is also decreased between patients and control.

- In the young onset T1D group, we find a significant delta cell count decrease between the patients and control, with a decrease in beta cell count. Total cell count is also decreased between patients and control.

- In the T2D we find a significant delta cell count increase between the patients and control, however total cell count did not show a significant change.

2.4.3 Measuring the Effect Size

We wished to compute the magnitude of change between subjects and their controls. We did this via the effect size for the Mann Whitney test (called theta statistic [112])

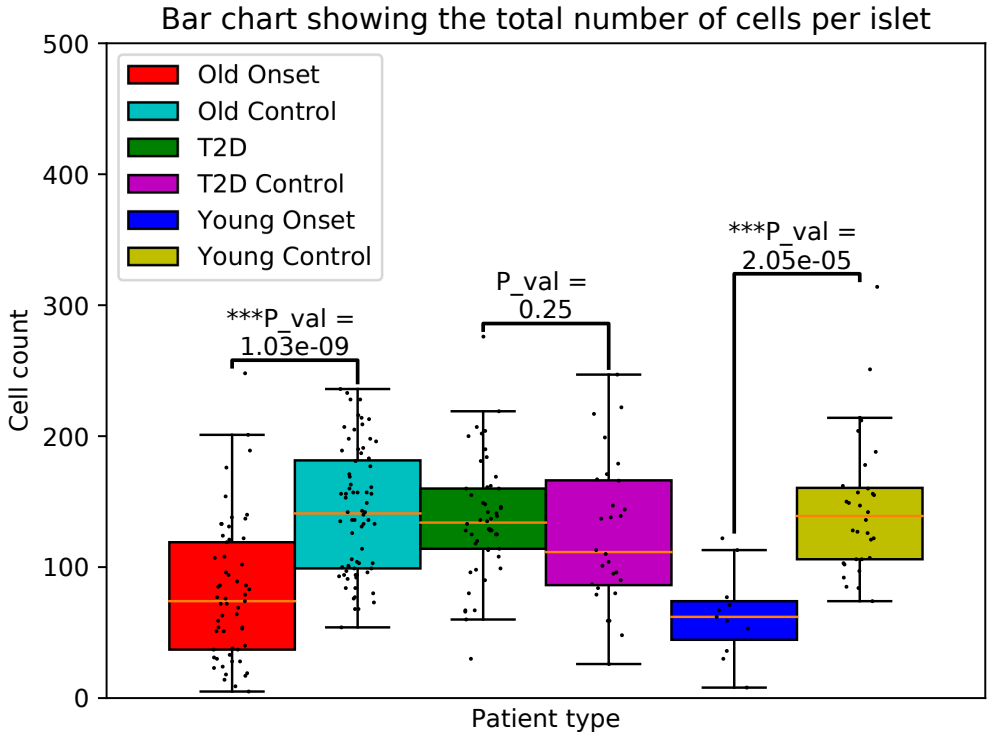


Figure 2.12: Bar chart illustrating cell count changes between patient type and their controls.

on the samples. Specifically, the test assesses the probability, P_{ab} of one random item from list 'a' is greater than one from list 'b', which is given by:

$$P_{ab} = \frac{U}{n_a n_b}. \tag{2.12}$$

Here, U is the Mann-Whitney U statistic between sample 'a' and 'b' and n_a and n_b are the size of the samples of 'a' and 'b' respectively. This was calculated and recorded section A.2. If the two groups were identical, we would expect an effect size of 0.5. We see substantial effect size deviations from this in the beta cell counts between the young onset and older onset T1D cases, corresponding to autoimmune mediated destruction of beta cells, as expected. We also see substantial effect sizes for delta cell counts between the groups, suggesting that there is an increase in delta cells as a result of diabetes. These effect sizes are illustrated graphically above the bars in Figure 2.13.

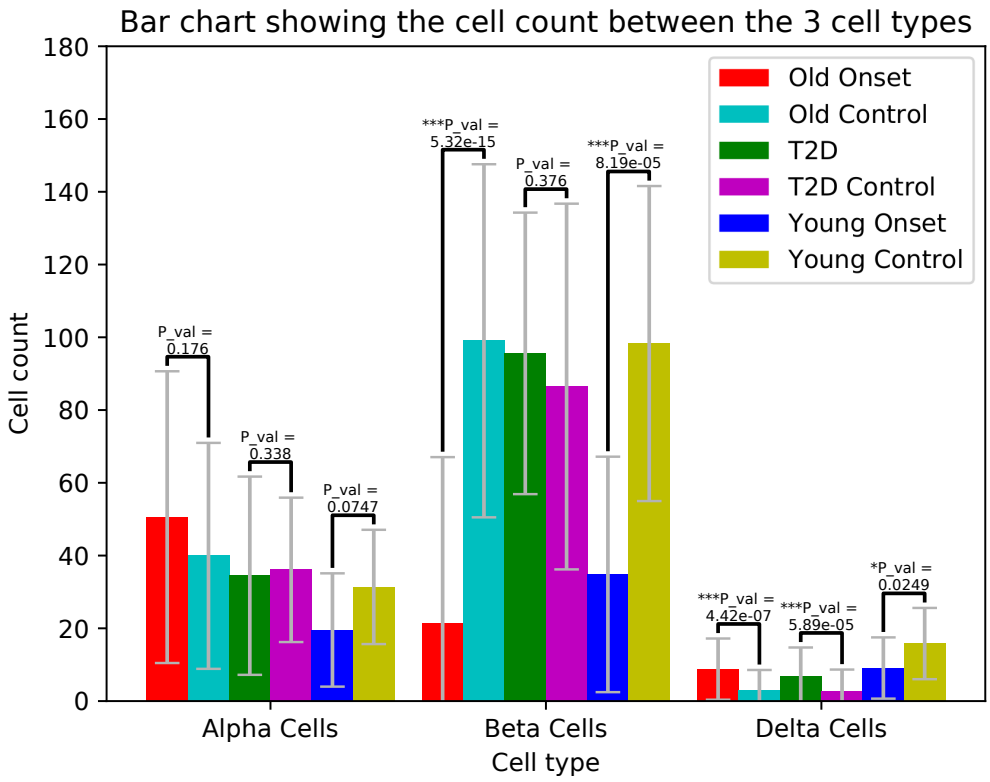


Figure 2.13: Bar chart depicting cell count changes between patients and their respective controls, split by cell type.

2.5 Summary & Discussion

In summary, we have analysed a set of islet images and compared the different number of cells between subject and control data; this was done for young onset T1D, old onset T1D and T2D. Subsequent image processing allows T1D control images to serve as an input to our model for T1D. Moreover, this additional image processing allows the T1D subject images to be used for validation of the proposed mechanisms within the model.

Delta cell type counts appear to have increased as a result of older onset T1D and T2D, along with the expected reduction in beta cell counts between the two T1D subject groups and their controls. With respect to the change in the number of delta cells, we know that there is minimal human post-natal endocrine cell proliferation [113], so this would imply one of two possibilities: either our previous understanding of human endocrine cell proliferation is incomplete or we could suggest that the beta cells might trans-differentiate into other cell types. In fact, the latter has been

previously suggested by authors such as Weir *et al.* [114].

In clinical T1D, there is often a delay between diagnosis and need for supplementary insulin injections. This has be referred to in the literature as a 'honeymoon phase' [115]. Cellular trans-differentiation would help to explain this as delta cells secrete SST, which works to inhibit beta cell function [116]. A loss of beta cells, reducing insulin secretion capacity, combined with an increase in SST, through an increase in delta cells could result in a positive feedback loop. This would result in the sharp decline in glycaemic control, which is frequently observed with T1D onset.

The above results might be useful in providing evidence towards verifying the trans-differentiation hypothesis; research that has not been previously undertaken through quantifying and comparing cell numbers within human islets.

2.6 Critical Evaluation

Given that we are working with a small sample size, this could induce relatively high margins of error. For a given sample probability, p, and a sample size, n, standard error, S_e, is typically given by:

$$S_e \approx \sqrt{\frac{p\,(1-p)}{n}}. \tag{2.13}$$

For example, if we have 50 islet samples of a particular subject group and let one third of the cells be alpha cells. This would yield a margin of error of 6.67%. This is quite high, and it limits firm conclusions that can be drawn from this work. For example in Figure 2.13, observationally, it looks as though there are changes in the alpha cell counts in both the younger and older T1D groups. However, more data would be required in order to be able to draw a definitive conclusion from this.

There was some subjectivity in which filters were chosen and the parameters used. For example, when considering the DAPI stain, one could have considered an ellipse detection algorithm, such as those proposed by Nelson et al. [74]. However, due to the density of the nuclei and as discussed in the image analysis section, and that the nuclei overlapped due to the high luminescence of DAPI, this would have

made machine learning ellipse challenging and time consuming to implement. In order to account for the high density and overlapping nuclei, we used an erosion algorithm, and the structuring element was a disc of pixel radius seven as it yielded results that best accounted for this excess overlap.

There are a variety of algorithms to separate a series of 2D points, corresponding to centres of nuclei, into cells. This included for example, Sharif *et al.*, who used a watershed algorithm to segment red blood cells [117]. The review by Irshad *et al.* into methods for nuclei detection and cell segmentation provides further reference into many of these methods [118]. Ultimately, as the images are quadchromatic and the cells are assumed to be tightly packed, a Voronoi algorithm was used to segment the cells. Voronoi algorithms have precedence in bio-imaging, for example, Demir *et al.* use one for identifying tumours [119].

The constructed false colour maps of the cells (chloropleth) suffered from random spatially dependent statistical errors. A chloropleth is a map that is shaded or coloured according to a test statistic and spatial boundaries. In our system, the test statistic was the level of hormone in a cell and the spatial boundaries were the cell boundaries as constructed by the Voronoi algorithm. This is clearly an approximation of the cell boundaries which could cause issues such as cells getting incorrectly assigned. This kind of statistical error is called the modifiable areal unit problem (MAUP) [120]. There is no evidence that islet endocrine cells have predispositions to be particular shapes within the islet, so there is no reason to believe that any systematic errors have occurred due to MUAP. A worked example illustrating a MAUP is given below, in Figure B.1.

When compared to the cell count results, obtained by a collaborator in the UoE Medical School, the maximum discrepancy was in the alpha cell counts, with a absolute percent difference of 12.93%. As neither cell count necessarily represents the absolute truth (identifying cells within a human islets is subjective), this is considered an acceptable error; though our method does eliminate most human biases.

3. Quantifying the Geometry of Islets

3.1 Measuring an Islet

In order to make serious progress in understanding any changes as a result of diabetes, it is important to understand how the overall structure, size and shape of the islet changes as a result of diabetes. A particular study, by Serion *et al.*, compares the area of islets before and after diabetes [121], offering contradictory conclusions to this thesis, in which they say that T1D does not radically change an islets' geometry: Only using six T1D islets for comparison, and identifying the islet area by hand as opposed to an automated algorithm. Despite knowing that the islet secretions are dependent on the shape and structure of the islet [122], to date, no systematic study has been performed on islets to assess their structure, either from a qualitative or quantitative perspective - both in healthy and diabetic patients.

It is therefore important to assess the effect that diabetes has on the shape of the islets and the effect that shape has on diabetes- to better understand how the islet secretions change. Furthermore, any future modelling work that would seek to model proposed spatio temporal mechanisms for diabetes must match the spatial metric results, outlined in this chapter.

In the previous chapter, we generated binary islet images, B from raw micrograph images (Figure 2.5d). In this chapter we examine the geometry and macrostructures of islets, using the binary image obtained via the steps taken in chapter 2. We do not, however, consider the cellular location and decomposition of islets in this section. This will be further examined in chapter 5.

For these images, we quantify the islets using the following metrics: area, perimeter, diameter, pseudo-incircle, pseudo-circumcircle, average radius, isoparametric radius and perimeter to radius ratio. We formally define these measures below, and the result are presented in section 3.3.

3.2 Calculating the Different Geometric Metrics of Islets

From processing the images in the section above, we have a set of two dimensional images embedded within $\mathbb{R}^2$. The following algorithms were written and implemented in Python3. Where appropriate, morphology definitions follow [123]. We now introduced some notation which will be useful for formally defining the metrics used to measure our islets.

Firstly, we recall that the set of binary islet images were defined in the previous chapter as $\mathcal{B}$, and an individual binary islet image is B. We can index this image array with $p = (p_1, p_2)$. We define the coordinate set I as:

$$I = \{1...m\} \times \{1...n\}, \tag{3.1}$$

where m, n is the size of the image in the 'x' and 'y' direction. In other words B is the set of pixel points within the image,

$$B = \{B_p\}_{p \in I} \in \mathcal{B} \tag{3.2}$$

We now define a subset of indexes, $I_B \subset I$, which corresponds to the set of indexes over which the binary image has a value of 1:

$$I_B = \{p \in I \mid B_p = 1\} \subset I. \tag{3.3}$$

3.2.1 Filled Image, Erosion and Halo

Given an image $B \in \mathcal{B}$ we may wish to consider only the 'external' boundary of the shape represented by the filled image. To this end we consider the union $S(B)$ of connected components of the complement of B that do not intersect the edge of the image. The filled image is then defined as:

$$F(B)_p = \begin{cases} 1 & \text{if } B_q = 1 \text{ or } S(B)_q = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

for all $p \in I$. An example of $F(B)$ is given by either of the two disconnected components in binary image Figure 2.5e.

The erosion of B by structuring element k is denoted by $\varepsilon_k(B) : \mathcal{B} \to \mathcal{B}$. Essentially, the erosion algorithm tracks around the edge of an image with a structuring element and removes the overlap between structuring element and image. When k is a disk of radius 1 pixel, we can express the 1-Erosion, $\varepsilon_1 : \mathcal{B} \to \mathcal{B}$ of $B \in \mathcal{B}$ as:

$$\varepsilon_1(B)_p = \begin{cases} 1 & \text{if } B_q = 1 \text{ for all } q \text{ s.t. } d(p,q) \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

for all $p \in I$.

The halo is a subset of the filled binary image. It is defined to be the edge pixels of the filled binary image. This is expressed as $H : \mathcal{B} \to \mathcal{B}$ of the image $B \in \mathcal{B}$ to be

$$H(B) = F(B) - \epsilon(F(B)). \tag{3.6}$$

An example of a halo is given in pink in Figure 3.4. We define a subset of indexes $I_H \subset I$, which corresponds to the location of the pixels in the halo image which equal 1.

$$I_H = \{p \in I \mid H(B)_p = 1\} \subset I. \tag{3.7}$$

3.2.2 Area and Perimeter

A measurement of the area, $a : B \in \mathcal{B} \rightarrow [0, \infty)$, can be found simply by taking the count of pixels in the image and dividing by the square of the micrograph scaling constant:

$$a = \frac{1}{S_c^2}|F(B)| \tag{3.8}$$

There is no such thing as the exact perimeter of a shape in an image. When calculating from the edge pixels of a shape, one can imagine all of the following as reasonable interpretations of perimeter distance. Whilst subjective, the approach

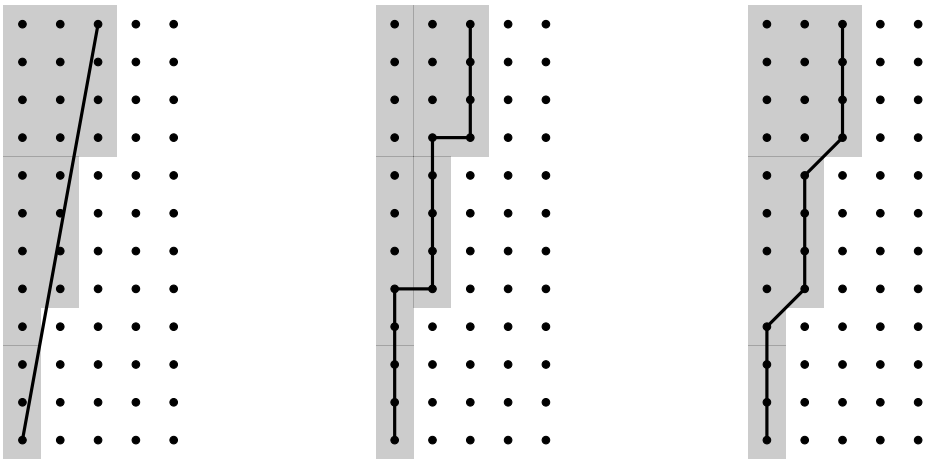


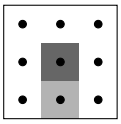
Figure 3.1: The above represents the edge of a shape, and three different ways of measuring the perimeter contribution. On the left, is a simple diagonal line between upper right and bottom left points on the edge of the image, giving a perimeter contribution of $3\sqrt{17}$. In the middle, we use an algorithm whereby we can only travel around the perimeter in the cardinal directions, giving a perimeter contribution of 13. Finally, on the right, is the approach that the algorithm that we are using takes in which you can travel around the perimeter in the cardinal and inter-cardinal directions. this gives a perimeter contribution of $9 + 2\sqrt{2}$.

that we took was to use the right hand option within Figure 3.1, as this provided a balance between following the edge of the image tightly, whilst also being easy to compute. Once we have a perimeter in pixels lengths, we convert to micrometers-dividing by the micrograph scaling constant, S_c [124]. We approximate the perimeter, $p : B \in \mathcal{B} \rightarrow [0, \infty)$ as:

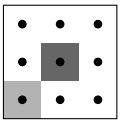
$$p = \frac{1}{S_c}\left(|p_1| + |p_3| + |p_2|\sqrt{2} + |p_5|\sqrt{2} + |p_4|\frac{1+\sqrt{2}}{2}\right) \tag{3.9}$$

$$I_H = p_1 \cap p_2 \cap p_3 \cap p_4 \cap p_5 \tag{3.10}$$

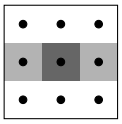
Where $p1 \subset I_H$ is the set of halo pixels with one cardinally adjacent pixel, as shown in Figure 3.2a; $p_2 \subset I_H$ is the set of halo pixels with one inter-cardinally adjacent pixel, as shown in Figure 3.2b; $p3 \subset I_H$ is the set of halo pixels with two cardinally adjacent pixels, as shown in Figure 3.2c and Figure 3.2d; $p_4 \subset I_H$ is the set of pixels with one cardinally adjacent and one inter-cardinally adjacent pixel as shown in Figure 3.2e and Figure 3.2f. $p_5 \subset I_H$ is the set of pixels with two inter-cardinally adjacent pixels, as shown in Figure 3.2g and Figure 3.2h.



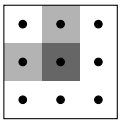
(a) An example pixel from the set p_1. p_1 is the set of halo pixels with one cardinally adjacent perimeter pixel.



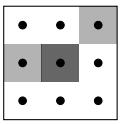
(b) An example pixel from the set p_2. p_2 is the set of halo pixels with one inter-cardinally adjacent perimeter pixel.



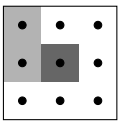
(c) An example pixel from the set p_3. p_3 is the set of halo pixels with two cardinally adjacent pixel.



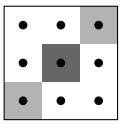
(d) A second example pixel from the set p_3.



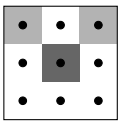
(e) An example pixel from the set p_4. p_4 is the set of perimeter pixels with once cardinally adjacent pixel and one inter-cardinally adjacent point.



(f) A second example pixel from the set p_4



(g) An example pixel from the set p_5. p_5 is the set of halo pixels with two inter-cardinally adjacent pixels



(h) A second example pixel from the set p_5

Figure 3.2: Presented here are eight examples of the possible configurations of Perimeter pixels with surrounding perimeter pixels. To generate all 60 permutations, one simply rotates the above generator images by $\{\pi/4, \pi/2, 3\pi/4\}$

3.2.3 Diameter

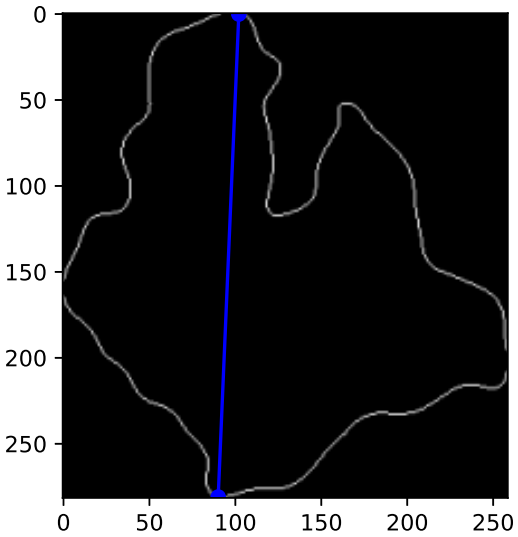
Another important measure of shape that we can make is to calculate the diameter of each islet image, $D : \mathcal{B} \to [0, \infty)$. We define this as the maximal distance between any two pixels in the image.

$$D(B) = \frac{1}{S_c} \cdot \max_{(p,q) \in {I_B}^2} \{d(p, q)\} \tag{3.11}$$

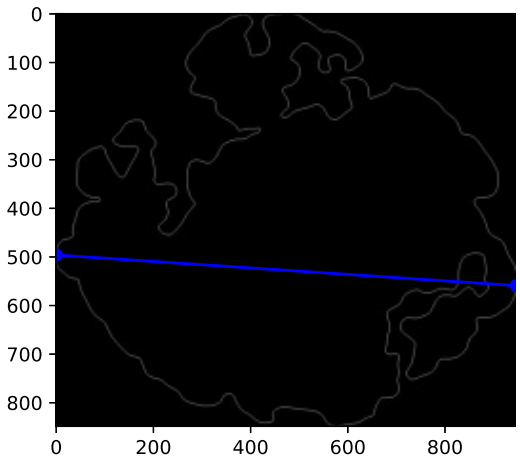
Where d is the Euclidean distance function. There is however, no algorithm to automatically do this, but we can construct our own to do this. However, the naive algorithm would be dependent on the square of the number of pixels in the islet, for calculating the Euclidean distance between every single pixel in the islet. The number of pixels in the islet is dependent on the size of the image, so we have a quartic order of operation, $O(m^2n^2)$. We observe that an equivalent definition for the diameter would be the maximal distance between any two points on the halo.

$$D(B) = D(H(B)), \tag{3.12}$$

for all $B \in \mathcal{B}$. The algorithm for the above would be dependent on the number of the square of the number pixels in the perimeter - so we have an quadratic order of operations given by, $O(mn)$. The halo and diameter are shown for our example islet in Figure 3.3.



(a) Halo of first islet identified in the sample image from Figure 2.1 with the diameter overlayed in blue



(b) Halo of second islet identified in the sample image from Figure 2.1 with the diameter overlayed in blue

Figure 3.3: Displayed here are the halo of both islets identified in our sample image. From this, we calculated a diameter, which has been displayed and overlayed in blue.

3.2.4 Minor and Major Axis

One particularly useful primary measure that can be applied to an image are the moments of the image. For a binary islet image, we can express the raw moments as:

$$M_{\{o,p\}} = \sum_i \sum_j i^o j^p B_{i,j} \tag{3.13}$$

Where $i \in \{0...m\}$ and $j \in \{0...n\}$. From this, M_{00} is the area of the islet, and the centroid, can be given as:

$$F(B) = \frac{1}{M_{\{0,0\}}} \{M_{\{1,0\}}, M_{\{0,1\}}\} \tag{3.14}$$

An finally, we can give the central moments as:

$$\bar{M}_{\{o,p\}} = \sum_i \sum_j (i - F_0)^o (j - F_1)^p B_{i,j} \tag{3.15}$$

The second order moments ($\bar{M}_{\{2,0\}}, \bar{M}_{\{1,1\}}, \bar{M}_{\{0,2\}}$) gives us the the non-standardized covariance of the image, the third order moments ($\bar{M}_{\{3,0\}}$... etc.) give the coskewness and the fourth order gives the cokurtosis. An image, much like a probability distribution, can be fully defined through its moments. The major and minor axis lengths, m_a, m_i, respectively are defined to be the major and minor axis lengths of the elipse with the same second order moments as our images. We can subsequently approximate a radius, R_a by taking the mean of the major and minor axis length, and dividing through by the micrograph scaling constant, S_c.

$$R_a = \frac{m_a + m_i}{2S_c} \tag{3.16}$$

3.2.5 Pseudo-Incircle and Pseudo-Circumcircle

The incentre of a shape is the point that is equidistant from each side [125]. It is well defined for a triangle, regular polygons and some irregular polygons, such as rhombuses. The shortest line connecting the incentre to each edge is a radius of the incircle. The incircle is defined as the circle which tangentially touches every side of

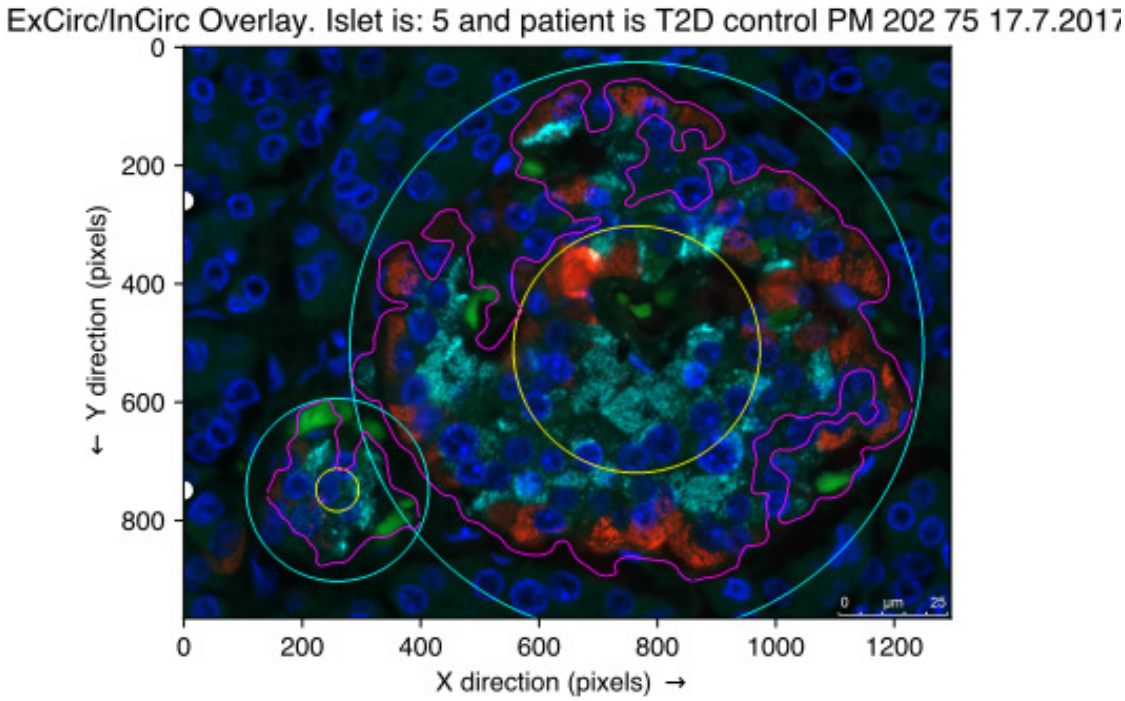


Figure 3.4: Displayed here are the pseudo-incircles, displayed as yellow circles and the pseudo-circumcircles, displayed as a cyan circle for the two islets identified within this image - the islet halo is in pink.

a shape. A property of the incircle is that it is the largest circle that can be fully encompassed within a triangle or regular polygon.

Similarly, the circumcentre of a shape is a point that is equidistant from each vertex. The line between the vertex and circumcentre is a radius of the circumcircle. The circumcentre and incentre are in the same position for a regular polygon. In addition to being well defined for triangles and regular polygons, the circumcentre can be well defined for certain irregular shapes, such as an isosceles trapezium. The circumcircles is defined as the circle which passes every vertex of a shape. A property of a circumcircle is that it is the smallest circle to fully encompass a shape. A polygon that possesses both an incircle and circumcircle are bicentric.

Concave polygons can never have well defined circumcircle or incircle. In this piece of work, we are inducing shapes on digital images, and 2D islet shapes which happen to have a concave perimeter- due to the folds in the membranes. We define our pseudo-incircle as the largest disk, centred upon the centroid ($F(B)$), of $B \in \mathcal{B}$, that can be inscribed in the halo. The radius of this circle is the pseudo-inradius, r_{in}. Equivalently, r_{in} can be defined as the minimum distance between the centroid and

the islet halo. Similarly, the pseudo-circumcircle is the smallest circle, centred at the centroid of the islet, which fully encompasses the islet. The radius of this circle is the pseudo-circumradius, r_{out}, which could also be defined as the maximum distance between the centroid and the islet halo.

We can express this by first supposing that $c \in I$ and $r > 0$. We define the disk of radius r centred at c, $\tilde{C}(r, c) \in \mathcal{B}$, by:

$$\tilde{C}(r,c)_p = \begin{cases} 1 & \text{if } d(p,c) \leq r \\ 0 & \text{otherwise.} \end{cases} \tag{3.17}$$

and define the circle of radius r and centre c to be

$$C(r, c) = H(\tilde{C}(r, c)). \tag{3.18}$$

Suppose $B \in \mathcal{B}$ and $c \in I_B$. We define the pseudo-inradius r_{in} of B centred at $F(B)$ to be

$$r_{in}(B) = \min\{r \mid C(r, F(B)) \odot H(B) \neq 0\}. \tag{3.19}$$

Where $\odot$ is the Hadamard or entry-wise product, and the corresponding

$$C(r_{in}(B), F(B)) \tag{3.20}$$

is the pseudo-incircle centred at $F(B)$. Correspondingly, we define the pseudo-circumradius r_{out} of B centred at $F(B)$ to be:

$$r_{out}(B) = \max\{r \mid C(r, F(B)) \odot H(B) \neq 0\}. \tag{3.21}$$

and the corresponding

$$C(r_{out}(B), F(B)) \tag{3.22}$$

is the pseudo-circumcircle centred at $F(B)$.

3.2.6 Dimensionless Measures of Shape

From the basic primary measures above, we can construct a further series of metrics to analyse the islets. A simple one is the ratio, Q_r between the incircle and circumcircle radii, as this is one potential way of measuring the "roundness" of an islet.

$$Q_r = \frac{r_{in}}{r_{out}} \tag{3.23}$$

From the radius and the perimeter, we can get a measure of the perimeter to radius ratio:

$$P_r = \frac{P}{2\pi R_a} \tag{3.24}$$

We can further take the isoparametric ratio, which is a dimensionless area to perimeter constant:

$$I_{rat} = \frac{4\pi A}{P^2} \tag{3.25}$$

3.3 Results

Graphical results from the metric calculations are displayed here, as a set of 6 histograms. These graphs have 10 equal bins, so the shape and profile of the data can be easily compared. The graphs on the left are those of the patients and the mirroring graph on the right are those of their respective controls. If an islet present touched the edge of the image, it was discounted as it would be a partial islet image and therefore give an inaccurate representation of the islet. This left us with 294 islets for comparison, of which 16 belong to the young onset group, 41 belong to the young onset control group, 89 belong to the T1D group, 47 belong to the T1D control group, and finally 61 belong to the T2D group and 40 belong to the T2D control group. These results are analysed in section 3.5

In both Figure 3.5 and Figure 3.6, displaying the histogram of islet areas and perimeters, respectively, we observe a decrease in the mean of the area in both the case types of T1D, whilst we don't see a change in T2D. On the right hand side as a subject ages.

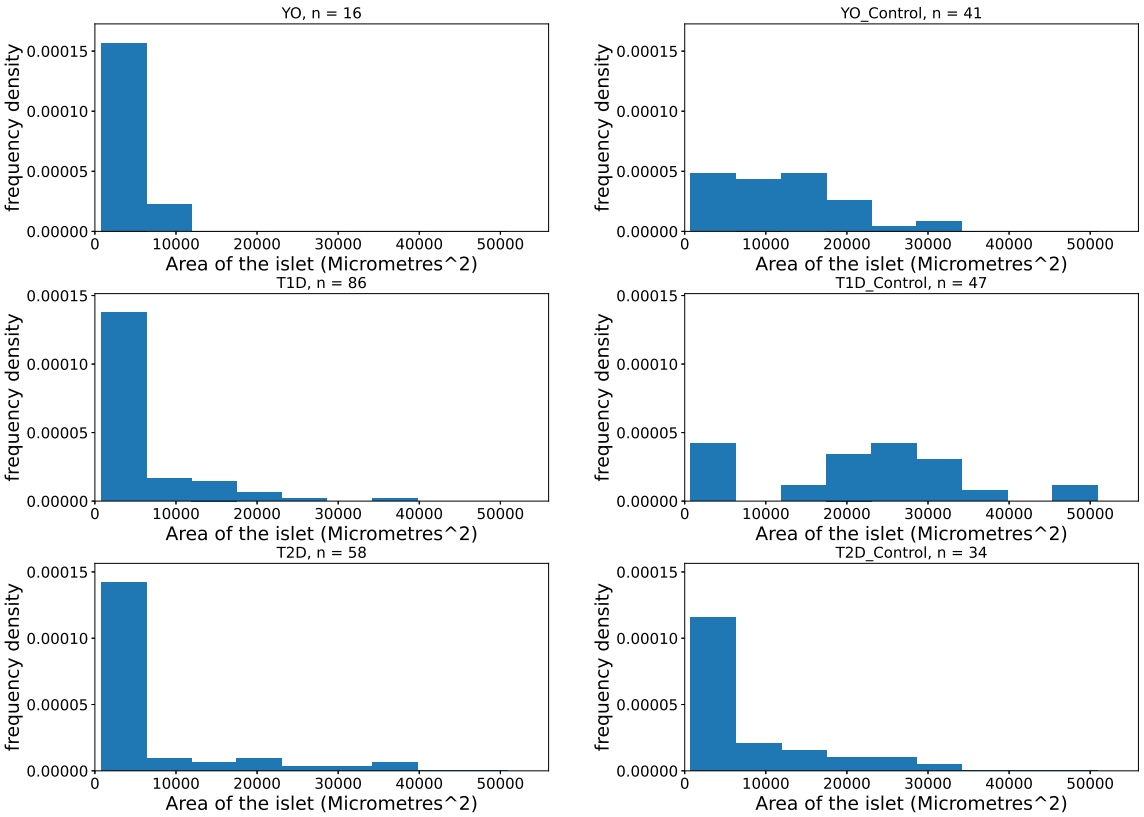


Figure 3.5: Presented here are the histograms of the area of the islets, as calculated in subsection 3.2.2. The horizontal axis represents the area of the islets, given in metres squared, and the vertical is is the frequency density, normalised so the total area under the graph is 1.

In Figure 3.7, displaying the histograms of the islet diameter, we see similar results to Figure 3.5, but the change in shapes look different - we appear to see some changes in the skewness.

In Figure 3.8, we see a reduction in the mean of the ratio for the pseudo-incircle and pseudo-circumcircle between the young onset T1D and their respective controls. We also see this change in the older onset, but there also seems to be a change in second and third order moment in addition to the mean. We don't appear to see a significant change with the T2D and respective controls. We seem to see a reduction in this ratio with respect to age.

In Figure 3.9 we see a reduction in the mean of the isoparametric radius between young onset T1D controls and their patients. It's not clear if this is the case with the older onset cases or the T2D cases. By comparing the young and old onset control cases, we can suggest that age seems to lower the mean of the isoparametric radius

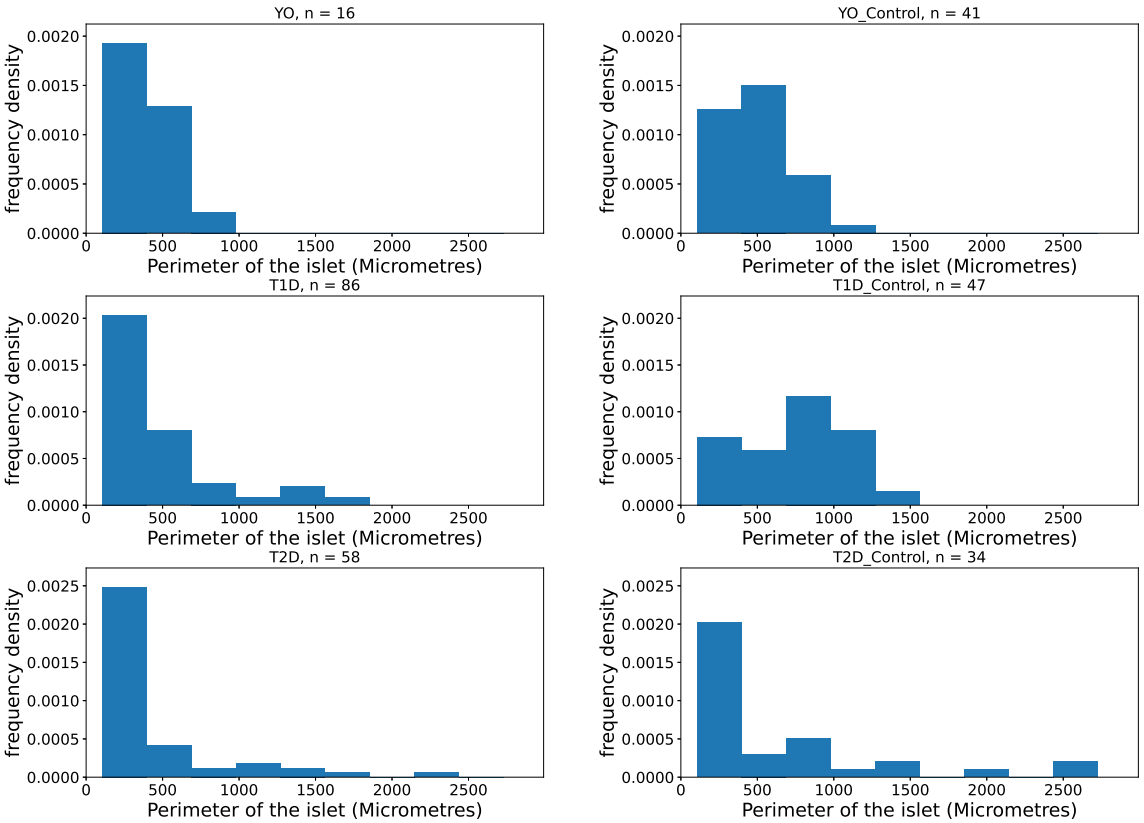


Figure 3.6: Presented here are the histograms of the perimeter of the islets, as calculated in subsection 3.2.2. The horizontal axis represents the perimeter of the islets, given in metres, and the vertical is is the frequency density, normalised so the total area under the graph is 1.

In Figure 3.10, we see a reduction in mean of the perimeter to radius ration in the young onset case, but this is not clear in the older onset T1D and T2D cases and their controls. Age appears to lower the mean of this ratio.

3.4 Summary and Discussion

In this chapter we have used the segmentation of the images in Chapter 2 to quantify the geometry of islets, and applied it to the samples of human islets that we have available. This is with a view to apply this to a large number of human islets which have yet to be imaged at the UoE medical school. Using this data, a full case study could be made to rigorously analyse how diabetes affects the islets, and also how ageing affects the islets too - to date, this has not been widely analysed.

Little research has been conducted into the geometry of the islets and how diabetes affects this. The only studies to date look at the macro-structure of the

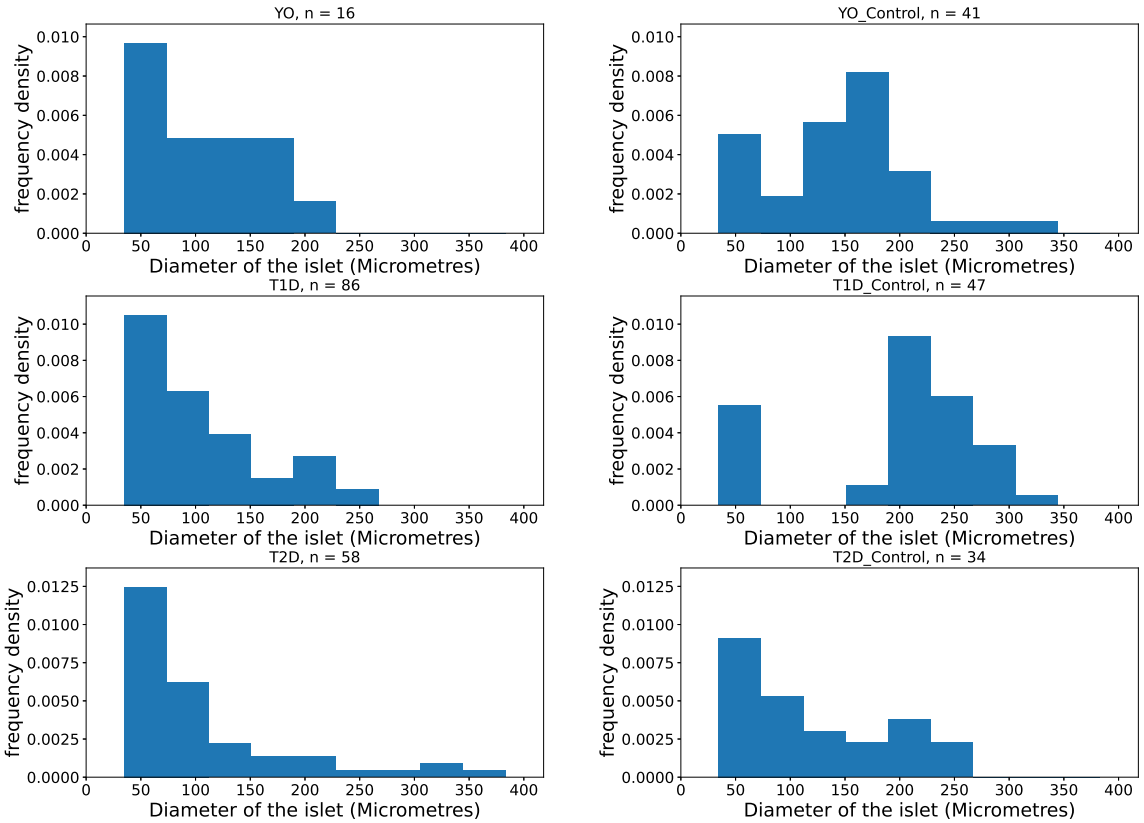


Figure 3.7: Presented here are the histograms of the diagram of the islets, as calculated in subsection 3.2.3. The horizontal axis represents the diameter of the islets, given in metres, and the vertical is is the frequency density, normalised so the total area under the graph is 1.

islets have focussed on the area, and stereo-graphically observing how this changes as one traverses from the head to the tail of the pancreas.

Despite this, there is significant evidence that the geometry affects the functioning of the islets. A greater surface area (represented by perimeter here), provides greater opportunity for Leucocyte infiltration in the context of the interstital infiltration theory of diabetes aetiology [126]. A greater volume (represented by area here) would indicate a greater number of cell and/or more sparse cells [127]. More circular islets would indicate the cells are close together, and better connected, as opposed to less circular islets [128].

3.5 Critical Analysis of Results

A change of area within the islet between control and patient groups would indicate a diabetes associated volume change. A proportional change in area of k between

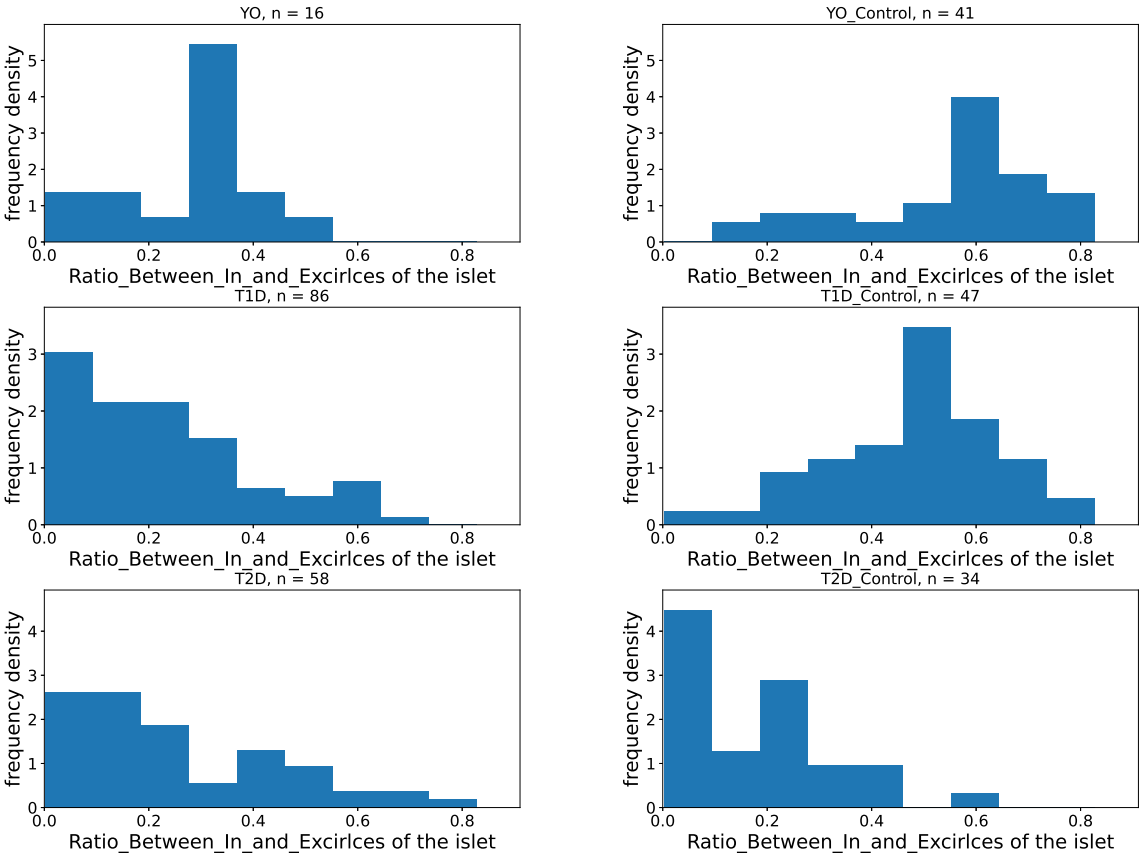


Figure 3.8: Presented here are the histograms of the ratio of the radius of the pseudo-incircle to the pseudo-circumcircle of the diagram of the islets, as calculated in subsection 3.2.5. The horizontal axis represents the ratio on the islets, given in metres, and the vertical is the frequency density, normalised so the total area under the graph is 1.

the area of these two groups would suggest an approximate $k^{3/2}$ change in volume. The perimeter is also an important measure of an islet to calculate. A change of perimeter betwixt control and patient groups would indicate that the islets basement membranes are changing as a consequence of diabetes.

It would be an interesting study to be able to track the individual cells throughout both the diabetes process and the ageing process. Perhaps, both their locations an the impact on the islet geometry. However, with current technology, using *ex vivo* samples, this would not be possible to directly assess this. That being said, with future technology, this might become possible.

From our small case study, there appears to be a significant shrinking of the islets in terms of area, perimeter and diameter as a result of young onset type 1 diabetes. This is expected as beta cells are removed from the system, the cross sectional area and perimeter would shrink. However, it is not necessarily expected that the isoparametric ratio, the perimeter to radius ratio and the ratio between the

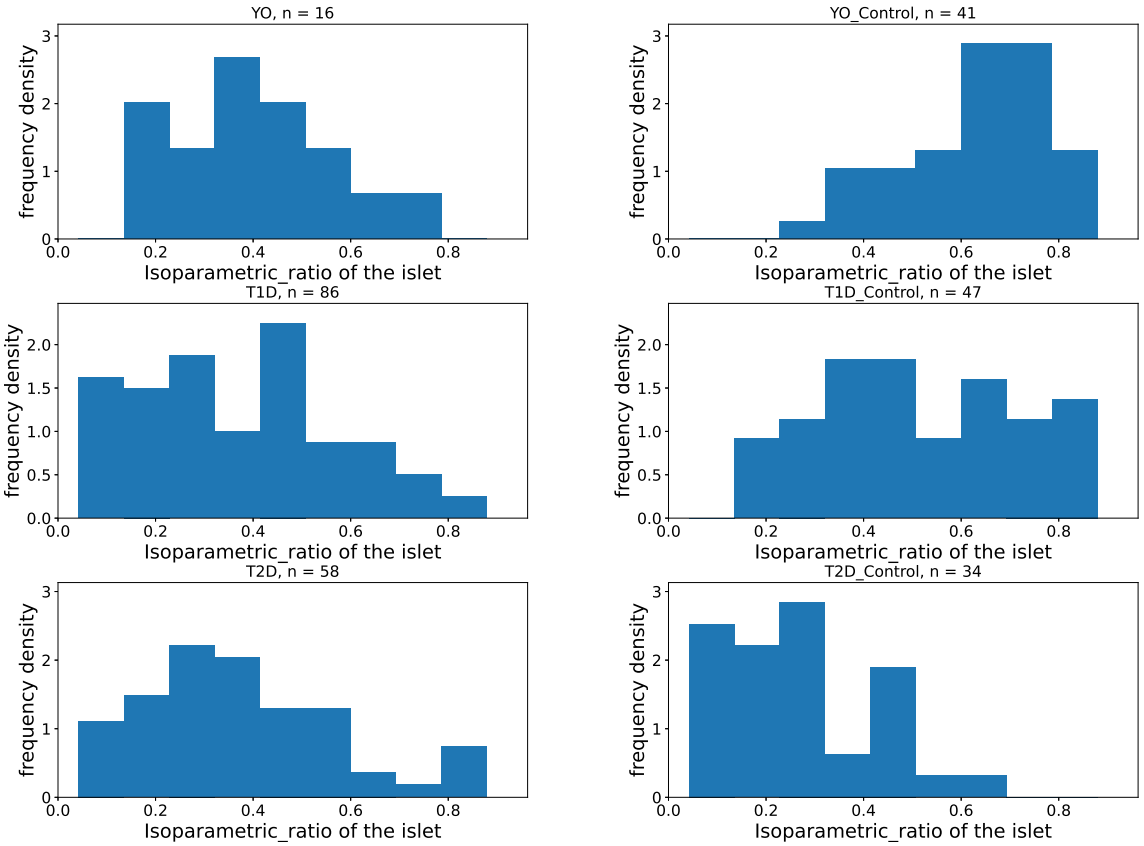


Figure 3.9: Presented here are the histograms of the isoparametric radius of the islets as , as calculated in subsection 3.2.6. The horizontal axis represents the isoparametic ratio of the islets, which is unitless, and the vertical is is the frequency density, normalised so the total area under the graph is 1.

incircle and the circumcircle would shrink. Whilst much more data would be needed to confirm these results, if it were true, it would imply that the islets become more folded as a result of diabetes, or that diabetes is causing the islets to fold.

We appear to see similar, if less pronounced results, between older onset T1D and their respective controls. However, in order to quantify the size of this with significant degree of meaning, we would like to have an order of magnitude more data. Such data will likely be available in the next year or so, and will be analysed in due course. This would indicate either a different mechanism or that the mechanism for older onset T1D is less potent than the younger onset patients.

However, in type 2 diabetes, there does not appear to be significant changes in terms of the geometry of the islets in type 2 diabetes apart from the ratio between the incircle and circumcircle. This could imply something about the islet folds becoming less pronounced in type 2 diabetes, but more research would need to be conducted on this.

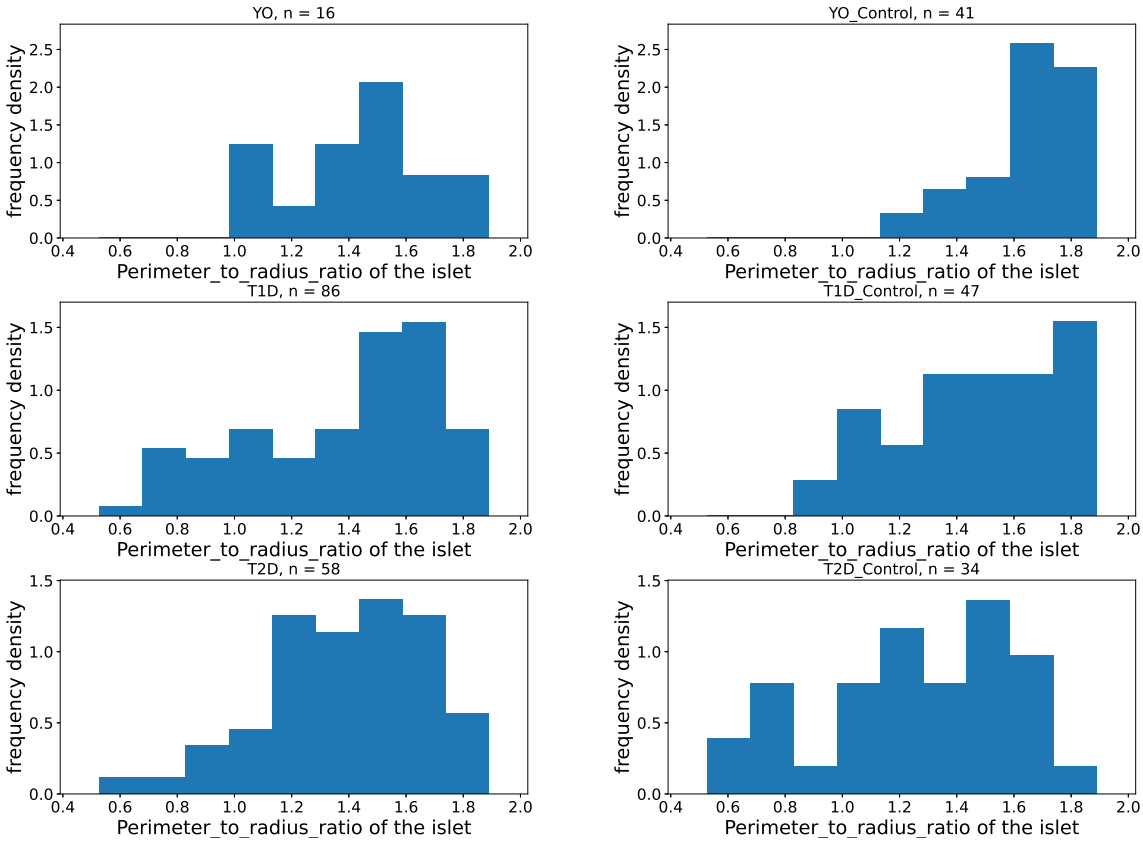


Figure 3.10: Presented here are the histograms of the perimeter to radius ratio of the islets as , as calculated in subsection 3.2.6. The horizontal axis represents the perimeter to radius ratio of the islets, which is unitless, and the vertical is the frequency density, normalised so the total area under the graph is 1.

In terms of as the islet ages (Tracking down the control groups from young onset T1D to older onset T1D and finally T2D) it does not appear that there is a significant change in the areas, perimeter or diameter, but potentially the three ratio metrics appears to change, which could tell us about how the islets change shape as one ages.

4. Persistent Homology Analysis of islets

In the previous chapter, we were considering the association between the geometry of the islets and diabetes and age. In this chapter, we are looking at a more fundamental classification of an object - its underlying structure, which is independent of the geometry. For this we will be using a relatively recent branch of mathematics, to quantify the underlying structure of the islets, called persistent homology.

Persistent homology is the study of the underlying structure of a set of data and is formally defined in section 4.1. In this body of work, we have generated from the data given a set of sets of 2D data points, namely the centres of all the nuclei within the sets of islet images. However, we know that this is a subset of a 3D data set, namely a 3D islet. If we can gain some insight into what this underlying structure might look like, we might be able to draw some association between the islet structures and the impact of both age and diabetes.

Use of topological persistence in biosciences is a relatively new idea. In 2015, Topaz *et al.* proposed using barcodes - graphical representations of the homology - to analyse mathematical models representing biological aggregations such as one for bird flocks, fish schools and insect swarms [129]. The Vicsek model is a discrete time, spatially stochastic model which predicts position of a set of particles with respect to time [130]. Along with a noise term, there is an attraction/repulsion term to simulate the clustering. By performing persistent homology analysis, it allowed for the analysis of the connectedness and clustering of biological aggregations, for example periodicity within murmurings of starlings.

In a similar manner to our work, in 2019 Atienza *et al.* used a Voronoi

algorithm, followed by persistent topology analysis to analyse the difference between cell clustering in the the epithelium of the wings of different stages of development of Drosophila [131]. This work potentially has applications in wider usage in cell tissue organisation, as the measurements taken are invariant to noise and scale invariance-whilst measuring the global properties. Whilst the use of topology in biomedical sciences is relatively recent, it may have important applications in studying the underlying properties of tissues, and hence we are investigating it in this chapter.

Within diabetes research Garside *et al.* have used topological data from retina scans to investigate and quantify diabetic retinopathy [132]. In this study, they have created a toolkit for the automatics diagnosis of eye disease based on the topological changes in a retina scan.

There are two main reasons for using homology with respect to analysing islet images. Firstly, the number of islets per image (0^{th} Betti number) indicates the packing density of the islets, and any changes, will give us some idea of how this might change as a result of diabetes. Secondly, The number of 'holes' in each islet (1^{st} Betti number), could represent either folds in the 3D structure (especially if close to the edge) or could represent vascularisation. A change would indicate important changes in vascularisation and folds as a result of diabetes, which would be functionally significant. This is all the information that we can obtain using homology on 2D images.

4.1 Definition of Homology

This section largely follows chapter 2 from [133].

Homology is a component of the field of topology. *Topology* is the study of structure, two objects are considered topologically equivalent or homeomorphic if there exists a one to one continuous map between them. Homology is a method of studying topological equivalence by associating structures, such as simplicial complexes to Abelian groups.

A simplicial complex is a collection of simplicies. One definition (in the last paragraph on page 102 of [133]) of a n-simplex, Δ^n is a set of $n+1$ points $[v_0...v_n]$,

embedded in $\mathbb{R}^m$, such that the points don't lie in a hyperplane of dimension of less than n (which requires $m > n$). Informally, an *n-simplex* is an '$n+1$' dimensional generalization of triangle. The ordering of a simplex's vertices is important and an 'n-simplex' denotes an n-simplex with an ordering of its vertices. As a result of an ordering of of the n-simplex, $[v_0...v_n]$, we have a natural ordering of the edges $[v_i, v_j]$, with respect to increasing subscript index. [133]. Illustrated below, in Figure 4.1, is an example of a n-simplex:

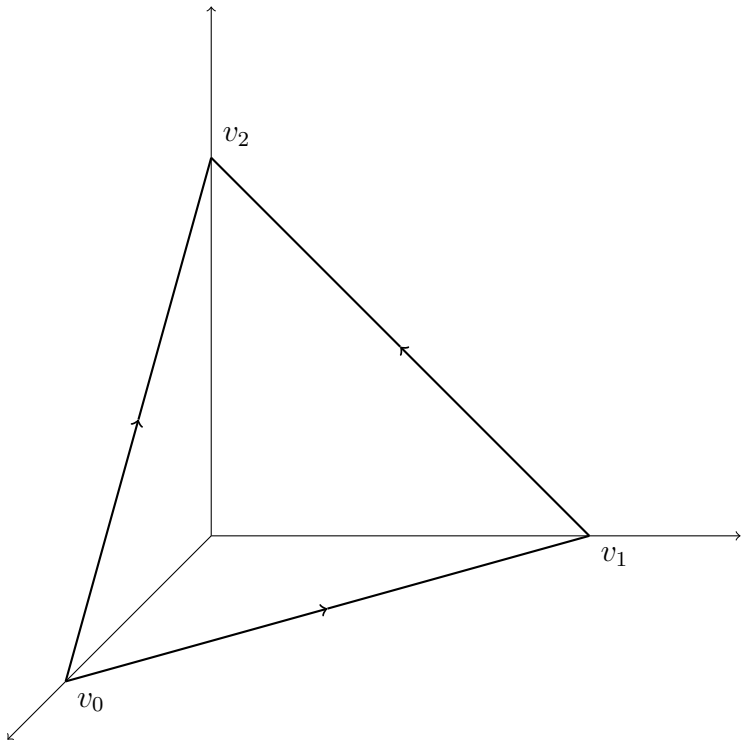


Figure 4.1: Here is an example of the standard n-simplex in $\mathbb{R}^{n+1}$ for $n = 2$ - a triangle; we illustrate the standard 2-simplex $[v_0, v_1, v_2]$ where v_0, v_1, v_2 are the unit vectors as shown here. This corresponds to the space: $\Delta^n = \{(t_0, ..., t_n) \in \mathbb{R}^{n+1} | \sum_i t_i = 1,\ t_i \geq 0,\ \forall i\}$, with the orientation as shown.

The i^{th} face of a n simplex is obtained by removing the i^{th} vertex, denoted by $[v_0, ..., \hat{v}_i, ..., v_n]$; the remaining n vertices span an $n-1$ simplex. Thus, the vertices of a face, or of any subsimplex spanned by a subset of the vertices, will always be ordered according to their order in the larger simplex. The *boundary* of the n-simplex, $\partial\Delta^n$, is obtained by taking the union of the faces of the n-simplex, and the *open simplex* $\overset{\circ}{\Delta}{}^n$ is defined to be $\Delta^n - \partial\Delta^n$, or the interior of the n-simplex.

A *simplicial complex*, $X \in \mathbb{R}^n$, is a disjoint collection of simplices and can be defined by the following two properties:

1. Every face of a simplex of X is in X.

2. The intersection of any two simplices of X is a face of each of them.

Homology associates an Abelian group structure with each simplicial complex X and each dimension n. Let $\Delta_n(X)$ be the free Abelian group with basis the open n-simplices, e^n_α of X. Elements of Δ_n are called *n-chains*, can be written as a sum $\sum_\alpha n_\alpha e^n_\alpha$, where each $n_\alpha \in \mathbb{Z}$. Alternatively, this n chain could be written as $\sum_\alpha n_\alpha \sigma_\alpha$, where σ_α is the map from the various n-simplicies to X. Informally, this *n-chain* can be considered as a weighted sum over the simplicial components of a simplicial complex. The *boundary homomorphism* can be defined as a map $\delta_n : \Delta_n(X) \to \Delta_{n-1}(X)$ by assigning values to the basis elements:

$$\delta_n(\sigma_\alpha) = \sum_i (-1)^i \sigma_\alpha[v_0, ..., \hat{v}_i, ...v_n] \tag{4.1}$$

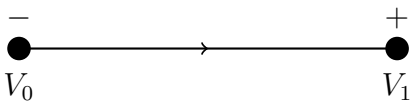
An example of this operator is illustrated on the first three simplices in Figure 4.2. As a consequence of Equation (4.1), the composite map of the boundary of the boundary is equal to zero.

$$\begin{aligned}\delta_{n-1}\delta_n(\sigma) = &\sum_{j<i} (-1)^i (-1)^j \sigma[v_0, ..., \hat{v}_j, ..., \hat{v}_i, ..., v_n] \\ &+ \sum_{j>i} (-1)^i (-1)^{j-1} \sigma[v_0, ..., \hat{v}_i, ..., \hat{v}_j, ..., v_n]\end{aligned} \tag{4.2}$$

which equals zero by arbitrarily relabelling the indexes. *A chain complex* is the sequence of homomorphisms of abelian groups between the n-chains of a simplicial complex:

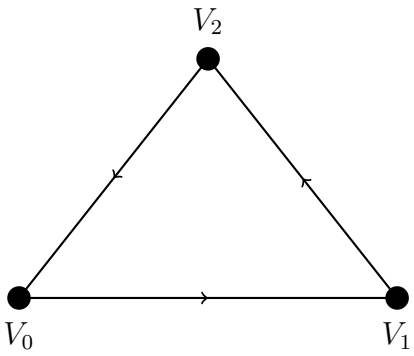
$$... \to \Delta_{n+1} \overset{\delta_{n+1}}{\to} \Delta_n \to ... \to \Delta_1 \overset{\delta_1}{\to} \Delta_0 \overset{\delta_0}{\to} 0 \tag{4.3}$$

Equation (4.2) implies that Im $\delta_{n+1} \subset$ Ker δ_n. Thence, we define the n^{th} *homology group* as the quotient group $H_n =$ Ker δ_n/Im δ_{n+1}. The number of independent generators for the n^{th} homology group corresponds the n^{th} *Betti number*. For two dimensional objects, embedded in $\mathbb{R}^2$, this is sufficient to determine the structure of the object. In the rest of this chapter, an element in a set of independent generators of the 0^{th} homology group and the 1^{st} homology group are referred to as B_0 and B_1



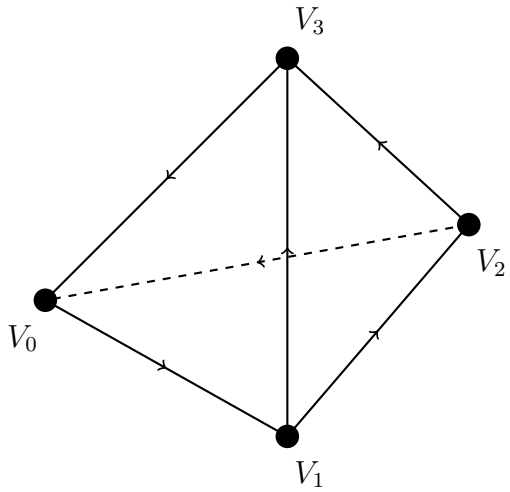
Boundary of 1 -simplex:

$$\delta_1[v_0, v_1] = [v_1] - [v_0]$$



Boundary of 2 -simplex:

$$\delta_2[v_0, v_1, v_2] = [v_1, v_2] - [v_0, v_2] + [v_0, v_1]$$



Boundary of 3 -simplex:

$$\begin{aligned}\delta_3[v_0, v_1, v_2, v_3] =& [v_1, v_2, v_3] - [v_0, v_2, v_3] \\ &+ [v_0, v_1, v_3] - [v_0, v_1, v_2]\end{aligned}$$

Figure 4.2: Illustration of the definition of the boundary operator, applied on the first 3 simplices

components respectively.

4.2 Point Clouds and Complexes

We have discussed previously about how we analyse a simplicial complex using algebra to classify it. We will now discuss how to generate a simplicial complex from a point cloud - such as we have for the position of the nuclei in the islet images. These include, but are not limited to, the Rips complex and the Čech complex [134]. These complexes can only be formed on spaces equipped with a metric, as they require a proximity parameter, denoted with ε. In the Rips complex, the 0-complexes are generated on the points themselves. The 1-simplexes are generated between

any 2 points which lie within ε of each other; the 2-complexes and 3-complexes are generated whenever 3 and 4 points are pairwise within ε of each other respectively. In the Čech complex k-simplices are constructed where the $\varepsilon/2$ balls around $k+1$ simplexes have a non empty intersection. We could use other complexes, such as the witness complex, for example, but due to its utility and ease of computation, we will be using the Rips complex. We will now illustrate a Rips complex and a barcode via an example in Figure 4.3.

The complexes, are largely dependent on the value of ε. Small ε values result in a cluster of 0-complexes, whilst large ε values result in a monolithic structure. Both these extremes would tell us nothing about the underlying topological structure of the image, which leads onto the question: Is there an optimal value for ε, and if so what is it?

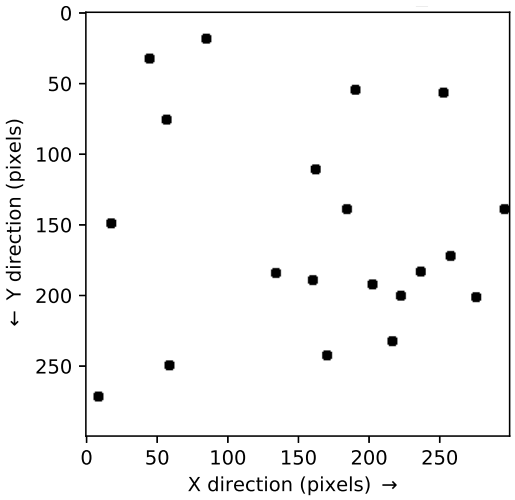
4.3 Persistence and Barcodes

This brings us to persistence. We do not know a priori what the value of epsilon should be, but as we increase the value of ε, the generated Rips complex grows in size. That is to say, as $\varepsilon_1 \leq \varepsilon_2 \leq \cdots \leq \varepsilon_M$, so we have the following inclusion relationship for simplicial complexes,

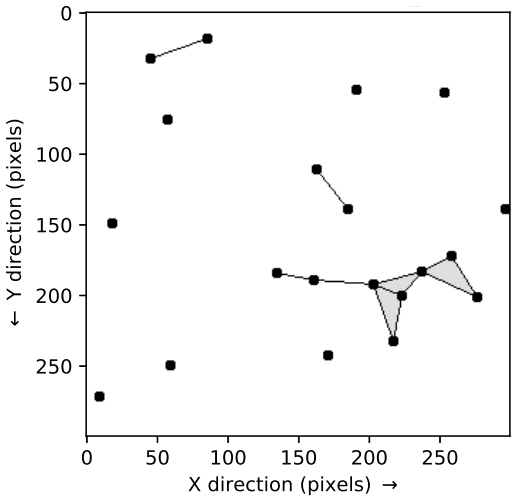
$$S_{\epsilon_1} \subseteq S_{\epsilon_2} \subseteq \cdots \subseteq S_{\epsilon_M} \tag{4.4}$$

This sequence is called a filtration. Persistent homology tracks features that persist along a wide range of values of ε, whilst transient features are considered noise. Barcodes are a graphical way of displaying the persistent topological features present throughout a point cloud. In order to do this, we need to track the B_0 and B_1 components, also called the connected components and 1 dimensional holes throughout the parameter search. This tracking is illustrated in Figure 4.4.

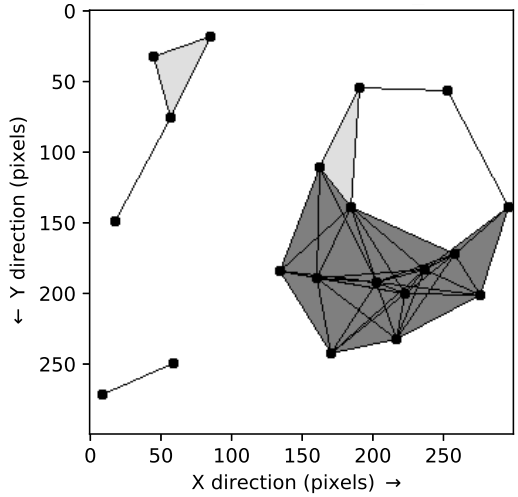
The zeroth Betti number corresponds to the connected components, and the first Betti number reduces to the number of independent cycles within the network [135]. Computationally tracking the B_0 components is relatively trivial. We can assign indexes to the nuclei from 0 to the total cell number, which is consistent



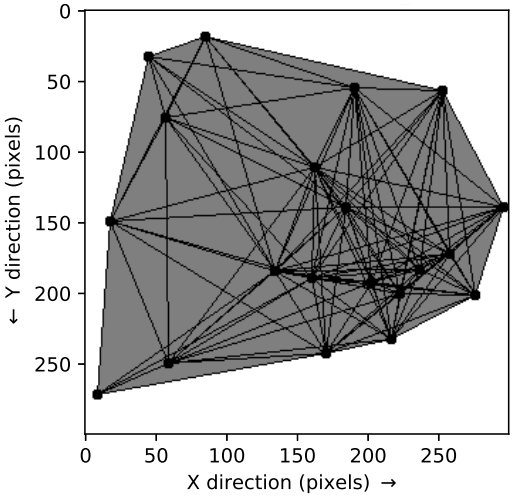
(a) Presented here is an example of a point cloud of 15 points, pseudo-randomly located - using a seeded Mersenne twister prime algorithm. From this, we generate a Rips complex with various distance parameters, ε, to investigate the underlying structure.



(b) This is the Rips complex for an epsilon value of 43 pixels. Note there are now some linked B_0 components. This is the complex for an epsilon value of 0 pixels



(c) for certain values of epsilon there is a B_1 component. Illustrated here is the Rips complex for an epsilon value of 93 pixels

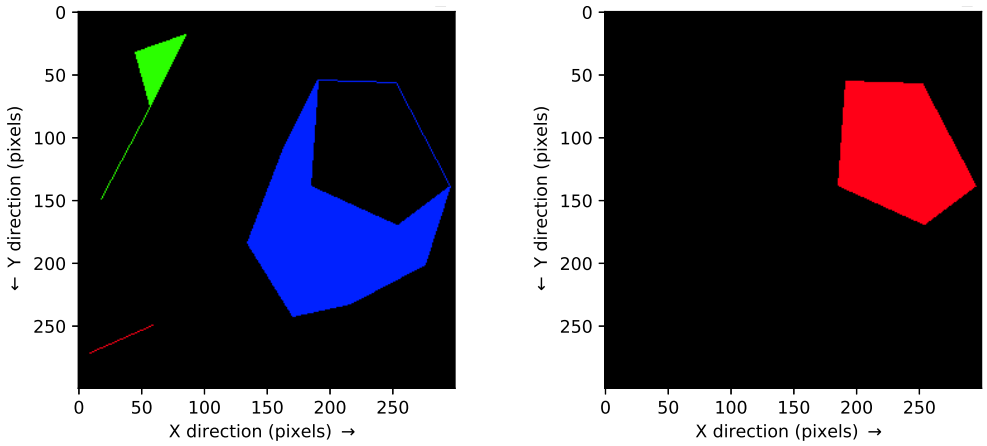


(d) As epsilon becomes large enough the point cloud coalesces into one B_0 component. This is the complex for and epsilon value of 181 pixels.

Figure 4.3: Presented here are a series of Rips complex with increasing values of ϵ We go from a series of point, which tells us nothing of the underlying structure at one extreme, to a blob, which similarly tells us nothing of the structure at the other extreme. This leads to the obvious question of what value should ϵ be?

across the parameter space. We then take the index with the minimum number within each B_0 component and aggregate it to that indexes bar within the barchart in Figure 4.5a. We then sort the bars by length order in order to get a sense of persistence. Presented in Figure 4.5a

For the B_1 Barcodes, this is a little more complicated. Each independent 'hole' corresponds to an independent cycle. We could, similarly to the B_0 components, track

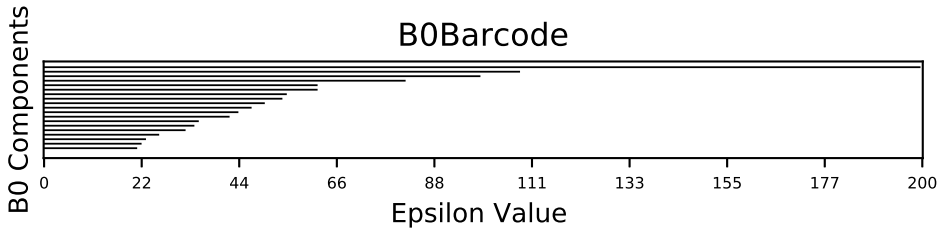


(a) Here we have an illustration of the B_0 components, for the complex in Figure 4.3c

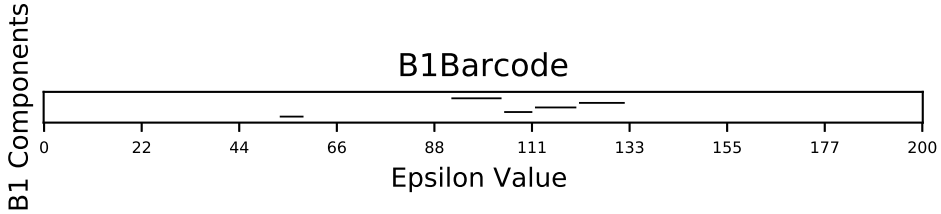
(b) Here we have an illustration of the B_1 components, for the complex in Figure 4.3c

Figure 4.4: An illustration of how we are tracking the 0^{th} and 1^{st} Betti numbers throughout the complex.

the vertex with the minimum value, but there could be more than one component with the same minimum index. Therefore, at each epsilon we value track the constituent vertexes in each B_1 component. If these change, we assign these to a new B_1 component. and record these onto the barchart as in Figure 4.5b.



(a) Here we have the barcode of the B_0 components for Figure 4.3. We track the components by the lowest numbered vertex in the B_0 object. As we increase the epsilon values, we notice a decrease in the number of bars present as we increase the distance parameter. The B_0 barcodes are sorted by length order. The higher up, the less transient the tracked component is.



(b) Here we have the barcode for the B_1 complex for Figure 4.3. In this particular example, the B_1 components appear to be relatively transient. We track the the B_1 components by the lowest unique vertex, and if the vertexes change, we label it as a new component

Figure 4.5: These are the B_0 and B_1 barcodes for the point cloud as presented in Figure 4.3a.

4.4 Barcode for Sample Islet

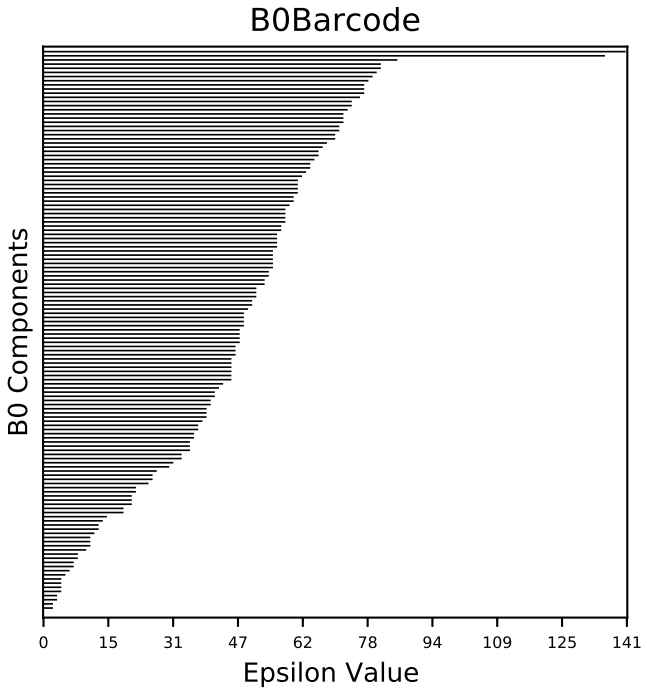


Figure 4.6: Here we have the barcode of the B_0 components. We track the components by the lowest numbered vertex in the B_0 object. As we increase the epsilon values, we notice a decrease in the number of bars present as we increase the distance parameter. The B_0 barcodes are sorted by length order. The higher up, the less transient the tracked component is.

Interpreting barcodes is certainly not an exact science, though, in our case, we believe that there are two islet components to the islet microscopy slide in Figure 2.1. However, one could reasonably make an argument that there exist only one islet sample in this image. Therefore for this particular islet image, a reasonable epsilon value must be above 85 for this condition to hold. Below, in Figure 4.7, we see the number of B_0 groups, plotted against the epsilon value.

To illustrate this inexactitude within the range of epsilon further, we present the connected components for a range of 4 epsilon values in Figure 4.8.

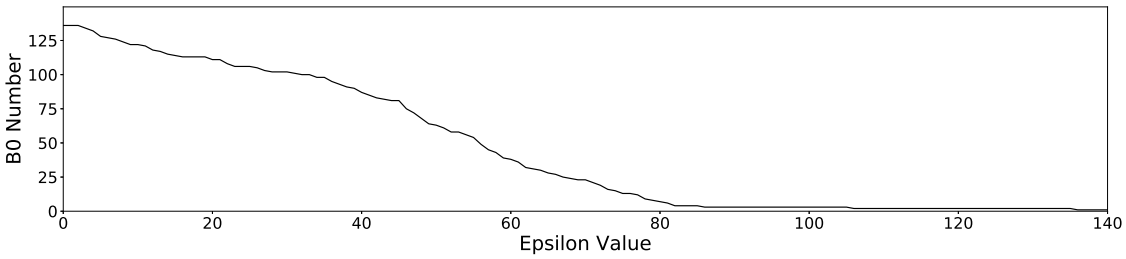
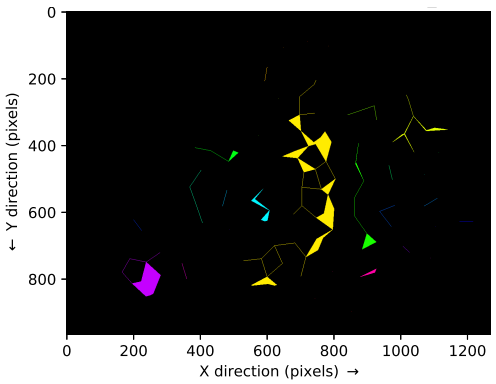
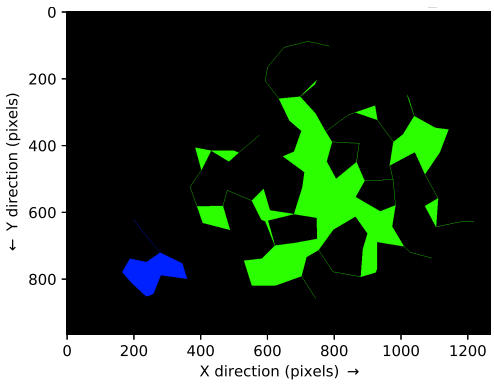


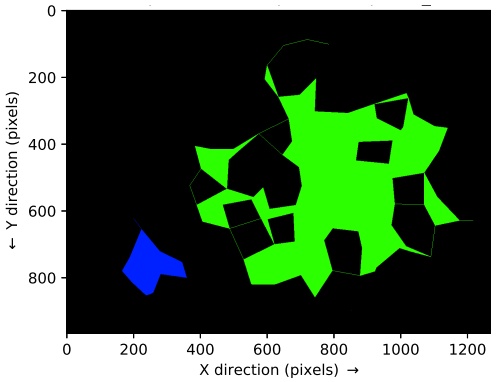
Figure 4.7: Displayed here is the zeroth Betti number, or disjoint components in the islet image from Figure 2.1 as a function of the distance parameter value. Note the relative stability between values of 86 and 136. This gives us a good indication of where the true epsilon parameter should be.



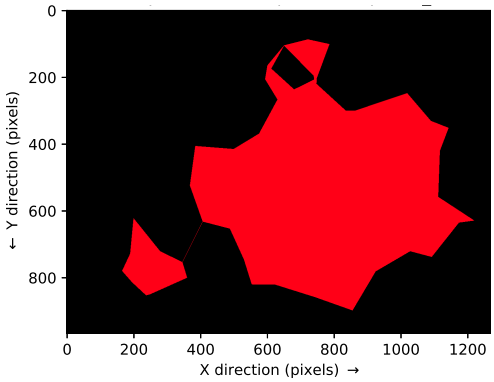
(a) These are the connected components for an epsilon value of 63. Here there are many B_0 components - this clearly does not correspond to the sample microscopy image. so the epsilon value is too low here.



(b) Here, the epsilon value is 88. Here, the point-cloud has amalgamated into just 2 disjoint components



(c) Here, the epsilon value is 105. We still have 2 disjoint components, illustrating the stability of this region.



(d) At an epsilon value of 105, the 2 disjoint components stop being so.

Figure 4.8: Presented here are the connected components, coloured, in order to illustrate the effects of varying the epsilon parameter value has on the B_0 number.

We see on the chart in Figure 4.9 that we have several persistent components, as well as transient components. To get a better sense of where these lie, we look at the graph of the count of B_1 components against the value of the epsilon parameter, presented in Figure 4.10.

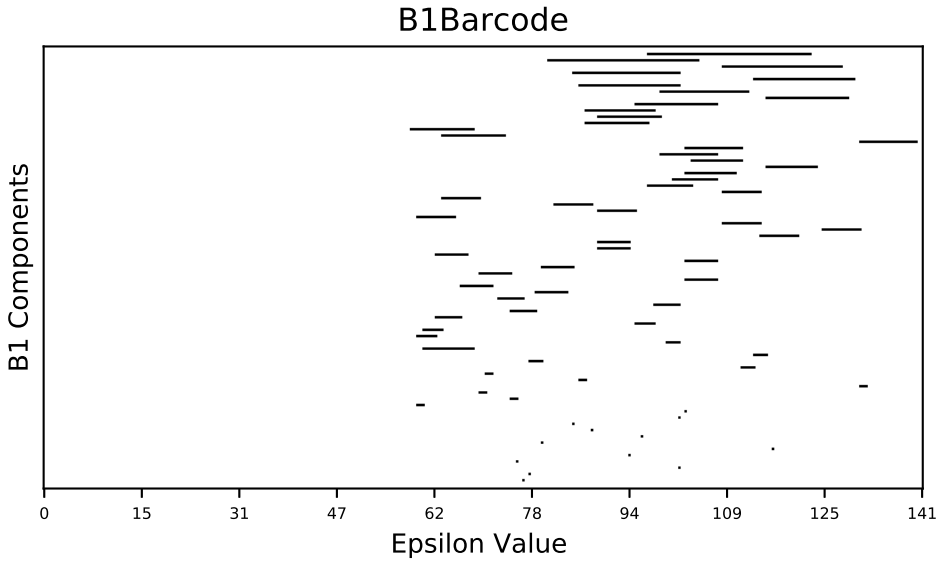


Figure 4.9: Here we have the barcode of the B1 components in our sample islet. We track the components by their vertex identifiers. Going up on the 'y' axis, we order barcode components by length, and along the 'x' axis, we have the epsilon value.

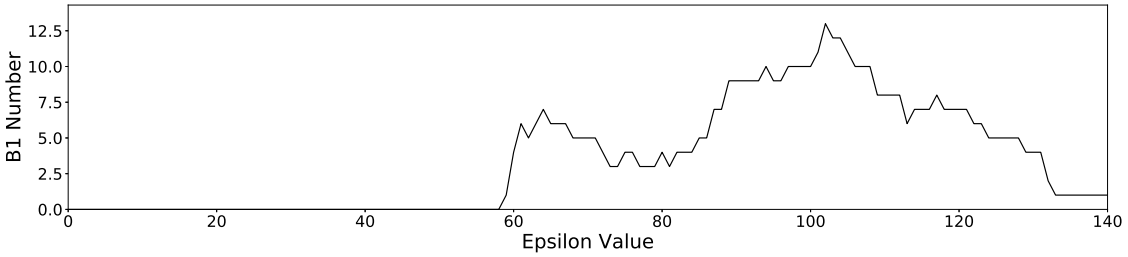
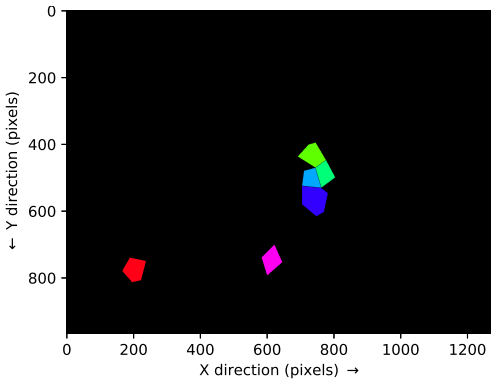


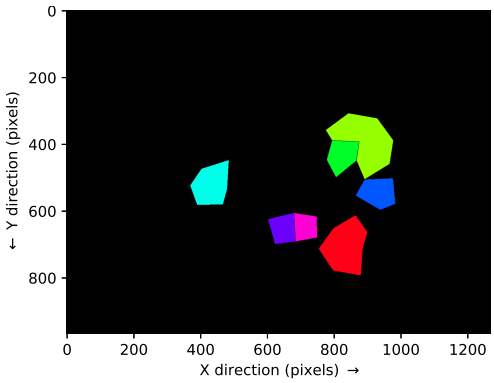
Figure 4.10: Here we have a line graph, reducing the barcode to the number of B_1 components against the epsilon value

We see clearly in Figure 4.10 that there is a high degree of variability with respect to the epsilon values on the graph. We will illustrated them for 4 particular value of epsilon

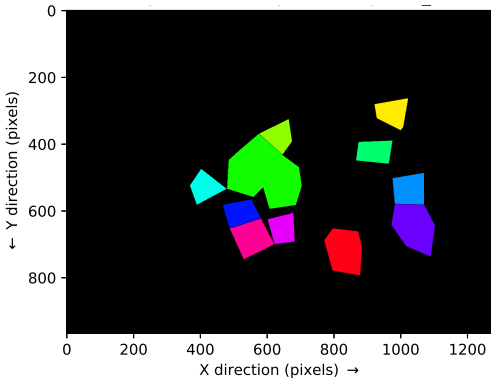
We see that whilst the number of B_0 components remains relatively stable between 85 and 135, we also see that in this region, from Figure 4.10 that the number of B_1 components does not. In its current form therefore, persistent homology is not a suitable tool for further understanding the structure and nature of human islets of Langerhans and more work on it would be required to get additional insight.



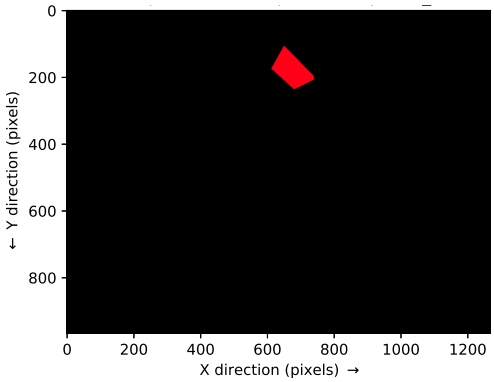
(a) Here, the epsilon value is 66. The B_1 components here appear to be somewhat clustered.



(b) Here with the epsilon value at 88, we see an increase in the number of B_1 components compared to at 66



(c) With a epsilon value of 105, we have a stagnation of the number of B_1 components compared to the epsilon value of 88



(d) With an epsilon value of 140, there is just one B_1 component present.

Figure 4.11: Graphical illustration of the B_1 components for a selection of epsilon components. Epsilon values are the same as in Figure 4.8

5. Quantifying the Islet Cell Network Structure

In the previous two chapters, we have been investigating the macrostructure of the islets, and how diabetes changes this. In this chapter, we are investigating the microstructure changes. We are doing this by inducing a network structure onto the images, as generated in section 2.2. From here on in, the nuclei centres of cells will be called nodes, and connections between cells will be called edges. For implementation, we are using a Python based network library, called NetworkX [136]. By using an induced network, we hope to test whether there is an association between the underlying connectedness within the islets and diabetes/age.

5.1 Inducing a Network from Processed Islet Images

Beginning from the Voronoi tessellation from Figure 2.5e, we take the dual graph of this. As outlined in, section 2.2 the geometric dual graph gives us the Delaunay triangulation. The vertices of this triangulation give us a set of nodes, $\mathcal{N} = \{p_1...p_n\} \in \mathbb{R}^2$. The set of lines connecting the Delaunay nodes are called edges are denoted with $\mathcal{E}_D = \{l_{p_i,p_j}\}$, where $\{l_{p_i,p_j}\}$ is the edge connecting the i^{th} and j^{th} node. However, this triangulation in a finite, bounded space has two main limitations- that need to be mitigated to enhance the accuracy.

The first issue is the edges between nodes straying outside the islet domain. This corresponds to cells being directly connected through the interstitial space, which would be a modelling error. It was not a satisfactory solution to simply

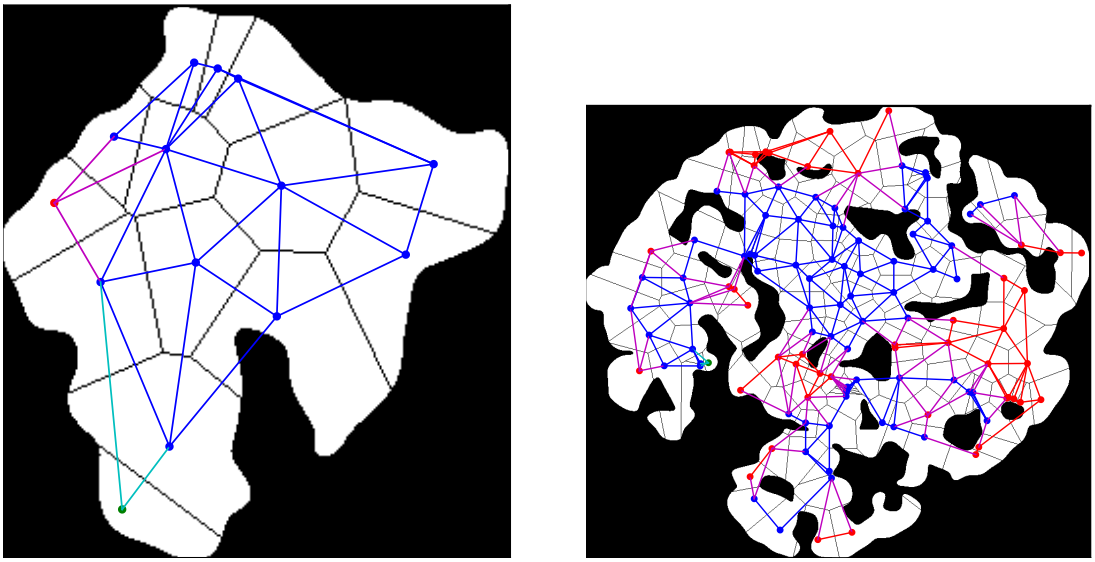
remove all the edges that overlapped into the empty space, as it removed edges that minimally strayed outside, so it did not take into account the concavity of the islet boundaries. It was neither satisfactory to measure the length of an edge outside the islet domain, as the folds in 2D are long and snaking, so this retained edges that should not have been there. A solution was to assess the area enclosed between the edge and islet. The second issue was particularly long connections at the edge of the domain as a naive Voronoi tesselation will stretch to infinity. This would result in cells being modelled as being directly connected over implausibly large distances. This can be mitigated by defining a new set of edges $\mathcal{E} \subseteq \mathcal{E}_D$, such that:

$$l_{p_i,p_j} \in \mathcal{E} \iff d(p_i, p_j) \leq d_{thresh} \tag{5.1}$$

$$l_{p_i,p_j} \in \mathcal{E} \iff Area(B) \leq A_{thresh} \tag{5.2}$$

$$B = \text{Bounded Components} \left\{ \mathcal{B} \cup l_{p_i,p_j} \right\}^c \tag{5.3}$$

Here, $d(,)$ is the Euclidean distance function; d_{thresh} and A_{thresh} are constants to be determined, and $\mathcal{B}$ is the binary image of the islet as defined in Chapter 3.



(a) This is the smaller of the two islet components within Figure 2.1

(b) This is the larger of the two components within Figure 2.1

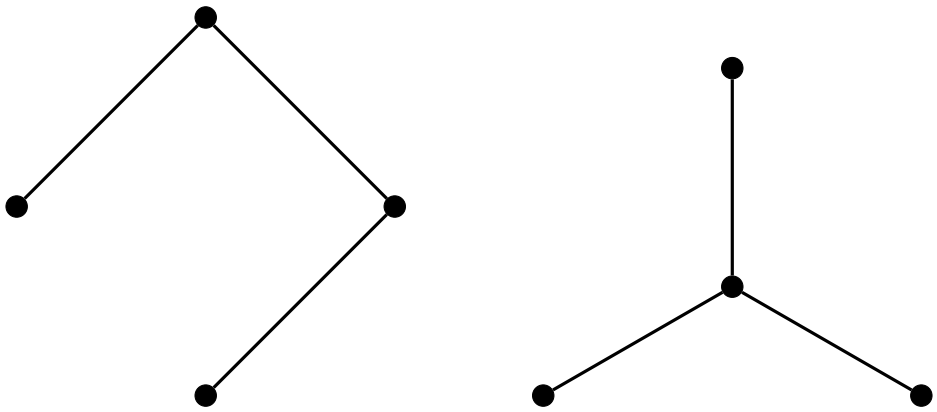
Figure 5.1: Induced network on sample microscopy image as from Figure 2.1. As in that image, red dots represent α cells, blue dots β cells, and finally green dots are the δ cells. Red, blue and green lines represent α-α, β-β and δ-δ connections respectively. The composites magenta, cyan, and yellow lines represent α-β, β-δ and δ-α connections respectively.

5.1.1 Discussion on Accuracy and Validity of Network Model of the Islets

The network is a priori, a mathematical construct, and an abstract representation of how the cells communicate and interact with each other. We make the assumption that cells communicate with their immediate neighbours. With gap junctions in between beta cells, this is definitely a valid assumption. Recent evidence has suggested that gap junctions might drive cell communications between other islet cell types [137]. Other types of communication, such as interstitial calcium and cytokines, are commonly modelled and assumed to follow inverse square law - this would make modelling via nearest neighbours accurate at a first order approximation [138]. We are not, however, going to be accurately modelling, for example, the capillary mode of communication between cells - a different type of model would be needed for this. Further, our assumption of a maximum distance for communication between cells was pragmatic, to make sure that cells that are clearly not adjacent are not connected to each other, but it almost certainly results in some level of inaccuracy. In previous chapters, we can test the parameters by doing a parameter search and comparing it to an medical expert's assessment of where the cells lie. We cannot do this here, as there is currently no way of telling from the scans by hand exactly which cells are connected to each other, to independently verify our algorithm.

5.2 Calculating Network Metrics

The motivation behind attempting to capture measurements of the network underpinning islets is to better understand certain aspects of the islets. Examining robustness, for example, can aid our understanding of autoimmune mediated destruction. The four metrics that we are proposing capture a wide range of characteristics of the islets so that we gather insight into how age and diabetes might affect the underlying structure.



(a) In this basic example, we have a mean betweeness centrality of this graph of 1 and a maximum centrality of 2

(b) In this basic example we have a mean betweeness centrality of 3/4 and a maximum betweeness centrality of 3

Figure 5.2: Both these networks have 4 nodes in 2 different configurations, which both have different configuration. We can see that in (a) we have 2 nodes that represent the key centre of the network, whereas in (b) this is 1 node, hence the different betweenness centrality value.

5.2.1 Betweeness Centrality

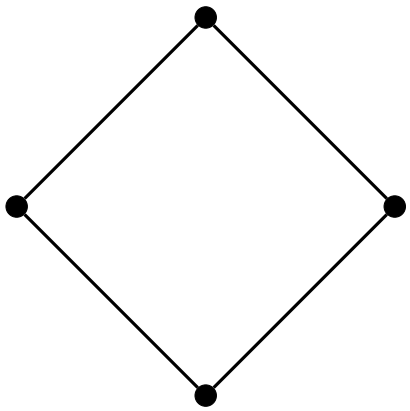
Betweenness centrality can be thought of as a measurement of the average 'importance' of each node within a network [139]. The betweenness centrality of each node in the network, $b_c(p_k)$, can be calculated as:

$$b_c(p_k) = \sum_{p_i, p_j \neq p_k \in \mathcal{N}} \frac{\sigma(l_{p_i,p_j}|p_k)}{\sigma(l_{p_i,p_j})} \tag{5.4}$$

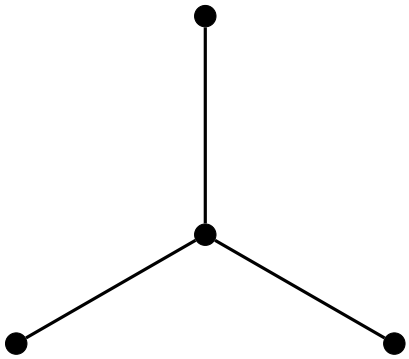
Here, $\sigma(l_{p_i,p_j})$ represents the number of shortest paths (each edge has a distance one) between p_i and p_j and $\sigma(l_{p_i,p_j}|p_k)$ are the number of shortest paths that pass through p_k. In order to get the mean and maximum betweenness centrality for the network, we take the mean and maximum values over all nodes. To understand this meaning further, let us examine two toy examples.

5.2.2 Degree Centrality

The mean degree centrality (m_d) is a measure of how connected the nodes are [140]. To calculate this, we take the sum of twice the total number of edges in the network, n_e, and divide it by the total number of nodes in the network, n_p. Max degree



(a) In this basic example, we have a mean degree centrality of this graph of 2 and a max degree centrality of 2



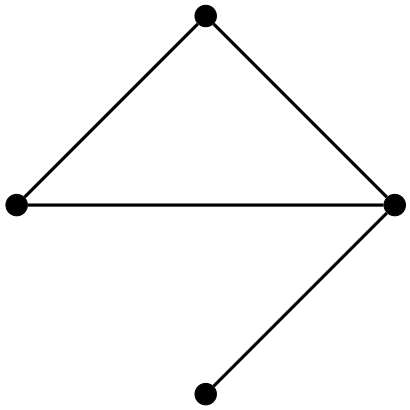
(b) In this basic example we have a mean degree centrality of 3/2 and a max degree centrality of 3

Figure 5.3: Both these networks have 4 nodes in 2 different configurations, which both have different configuration. We can see that in (a) All nodes are equal and have equal degree centrality (b) One node has a high degree centrality.

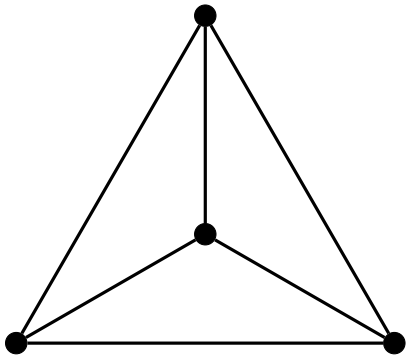
centrality is the maximum number of edges any particular node is connected to

$$m_d = 2\frac{n_e}{n_p} \tag{5.5}$$

5.2.3 Graph Transitivity



(a) In this basic example, we have a graph transitivity connectivity of this graph of 3/5



(b) In this basic example we have a graph transitivity of 1

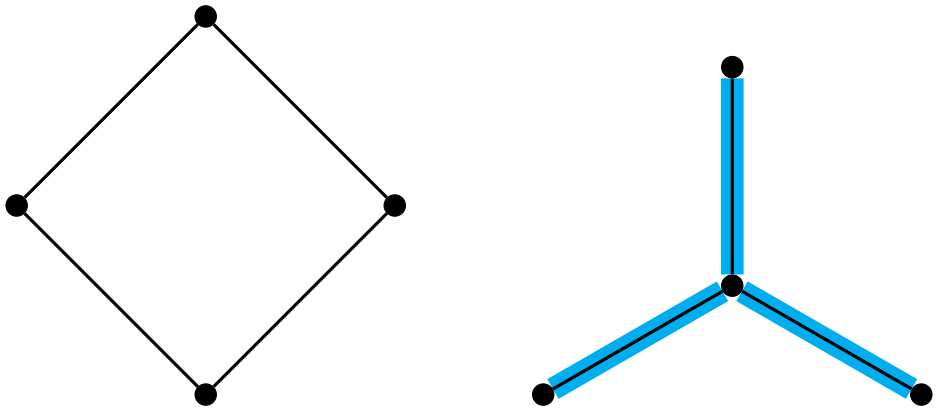
Figure 5.4: Both these networks have 4 nodes in 2 different configurations, which both have different configuration. We can see that in (a) Each node has less edges connecting it than in (b)

The graph transitivity, also called global clustering G_t is the measure of how triangulated a network is. Conceptually, this is a measure of a density of a networks connections [141]. It compares the number of possible triangles present in the network,

compared to the number of connected triples, or triads. Here, a triad are two edges with a shared node. This is expressed as:

$$G_t = 3\frac{\#\text{triangles}}{\#\text{triads}} \tag{5.6}$$

5.2.4 Bridges



(a) In this basic example, we have no bridges

(b) In this basic example we have 3 bridges

Figure 5.5: Both these networks have 4 nodes in 2 different configurations, which both have different configuration. We can see that in (a) all nodes are connected to 2 other nodes whereas in (b) 3 of the nodes have just one nearest neighbor

An edge of a graph is a bridge if and only if its deletion increases the number of connected components within a graph [142]. In the context of this research, it can be thought of as a measure of 'fragility' of the network.

5.3 Results

We first present the results for the overall networks, followed by the subnetworks of the beta cell contacts, followed by a brief mention of other potential subnetworks. Graphical results from the metric calculations are displayed here, as a set of 6 histograms with equal bins, so the shape and profile of the data can be easily compared. The graphs on the left are those of the patients and the right are those of their respective controls. n above each of these graphs refers to the number of islets with non-empty networks for comparison.

There are 294 islet images for comparison, of which 16 belong to the young

onset group, 41 belong to the young onset control group; 89 belong to the T1D group, 47 belong to the T1D control group; and finally 61 belong to the T2D group and 40 belong to the T2D control group. On the horizontal axis is the metric which is unit-less.

5.3.1 Overall Networks

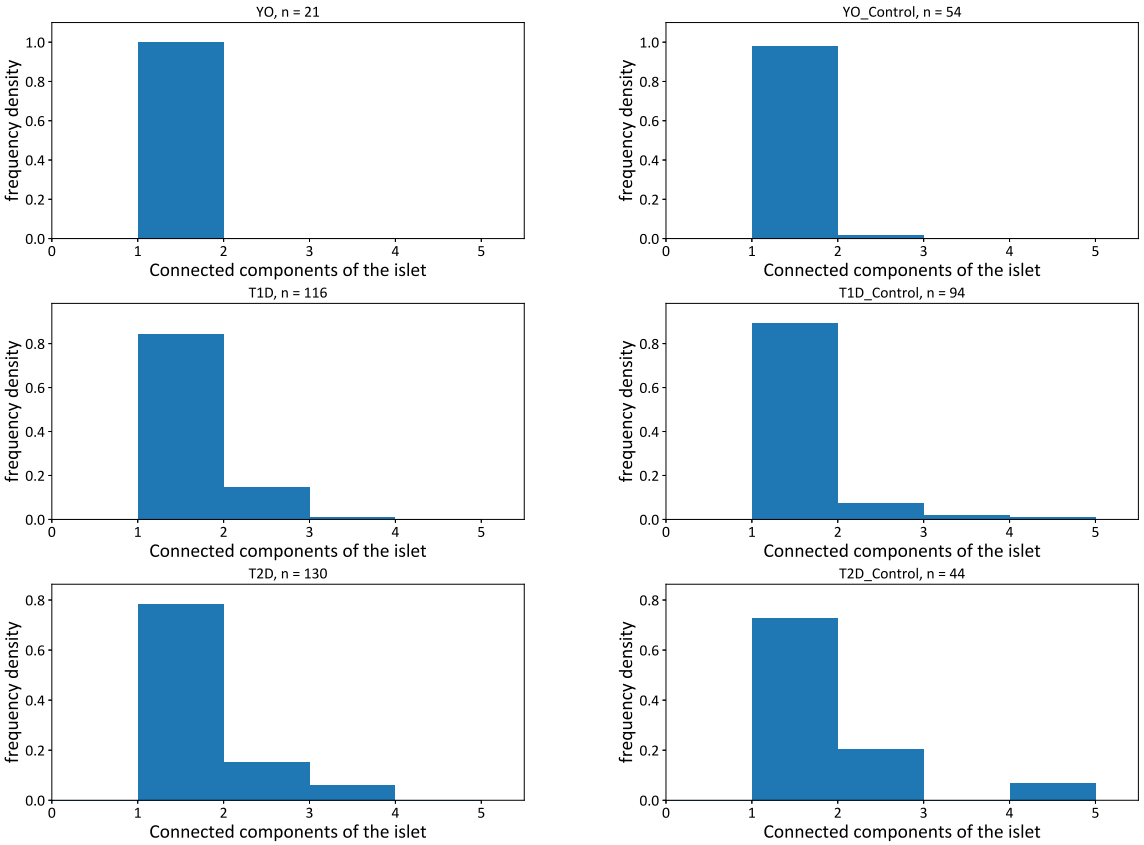


Figure 5.6: Presented here are the set of histograms for the number of connected components per islet. We clearly see an increase in the mean as we travel down the the right hand side of the histograms, but we don't necessarily see a change in the shape of the distributions when we compare diabetes cases to their controls.

In terms of the overall networks, between young onset type 1 and controls, the mean between centrality (Figure 5.7) does not change but the mean degree centrality does appear to decrease (Figure 5.8). Whilst the average importance of each node to keep the graph connected via minimum paths does not change, the average connectivity of the networks appear to decrease (Figure 5.6). The former point is supported by the fact that the number of bridges appears not to change (Figure 5.10). In addition to supporting that the number of minimum paths is not changing, it would appear that the graph fragility does not change. Somewhat

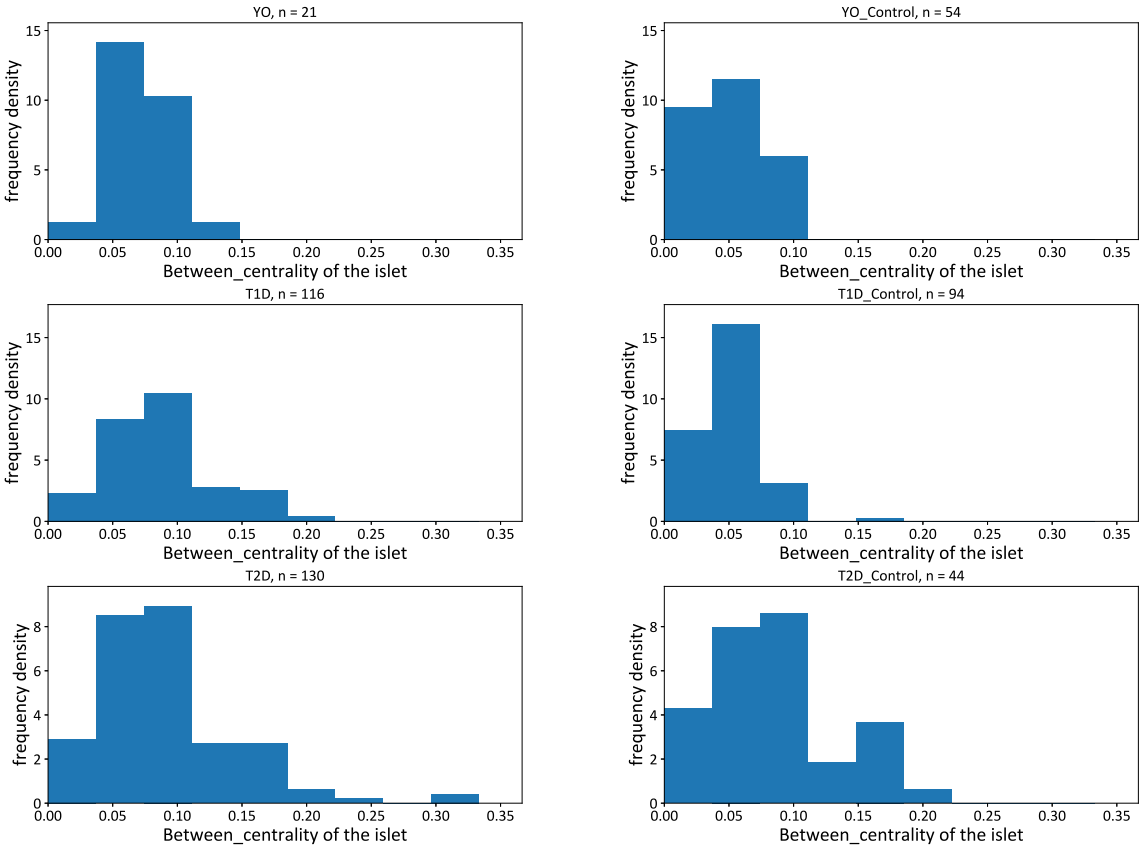


Figure 5.7: Presented here are the set of histograms for the mean between centrality metric for the overall islet networks as calculated in subsection 5.2.1. We appear to see an increase in skewness with age, and with both types of T1D, but it is unclear if the mean changes.

surprisingly, the global clustering metric does not change under diabetic conditions, the overall density of the connectivity of the network does not change (Figure 5.9). Between older onset type 1 cases and controls, we have remarkably similar results to the younger onset results.

Between type 2 and controls, the mean and max of both the between centrality and the degree centrality does not appear to change significantly. Neither does the global clustering metric or the number of bridges. Based on the measures that we have used, we don't see significant changes in the underlying networks in type 2 diabetes. Interestingly we also don't see a change in the number of cells, suggesting that in T2D, the underlying islet structure is robust.

As someone ages, the number of connected components per islet appears to increase. In addition, the kurtosis and mean in the mean and max between centrality distribution appears to increase, whilst the mean and maximum degree centrality appear to decrease. Suggesting that the importance of each cell to the networks

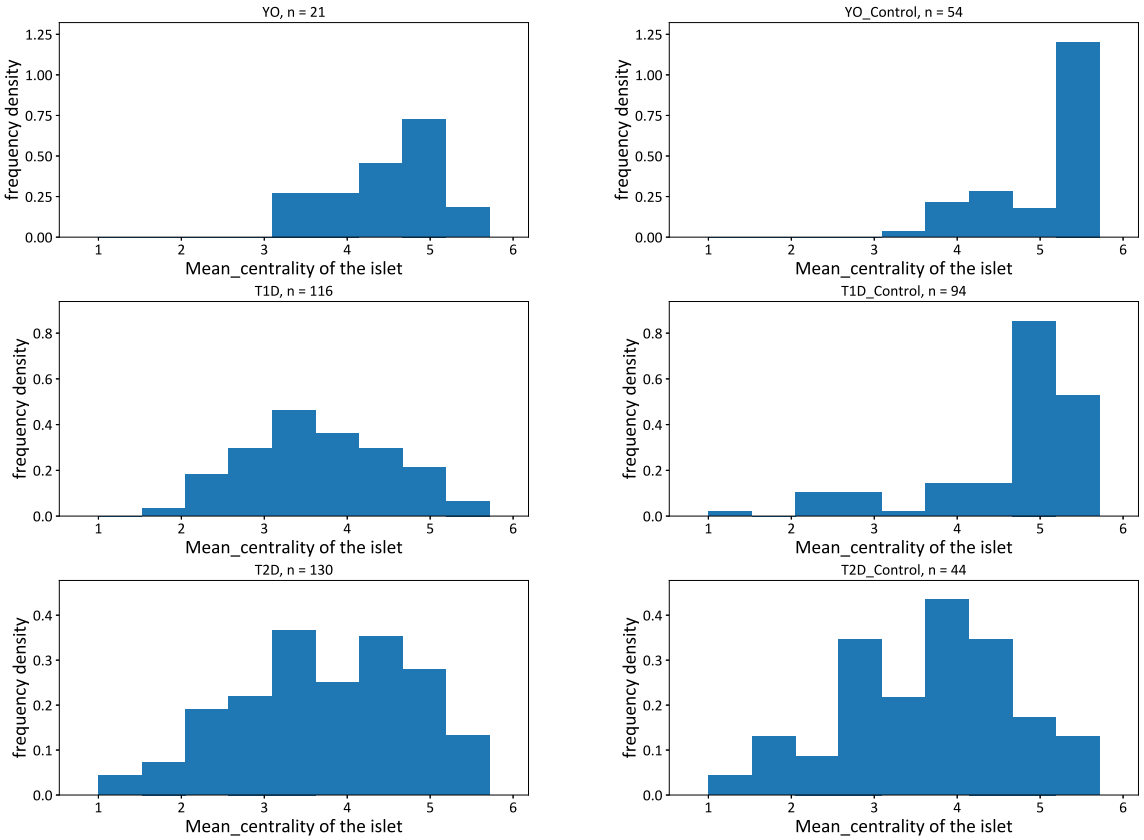


Figure 5.8: This is the set of histograms for the mean degree centrality metric for the overall islet networks. We see the mean appear to decrease as result of age and both forms of T1D.

appears to increase as one ages, and the number of cells each cell is connected to decreases. This is consistent with hypotheses about islet cells having very long half lives, but also not being regenerated in humans.

5.3.2 Beta Cell Subnetworks

In terms of the beta cell networks we see a lot of isolated cells and isolated triples, hence a number of zeroes and a smaller number of ones recorded in the global clustering metric histogram (Figure 5.14). These isolated components and somewhat random nature of the histograms for the T1D cases, is supported by the fact that many/most beta cells are killed with T1D. Most of the networks left are relatively isolated. Unless we have orders of magnitude more data, this is all we can conclude so far about how T1D changes the islets.

Between type 2 and controls, the mean of both the between centrality Figure 5.12 and the degree centrality (Figure 5.13) does not appear to change significantly.

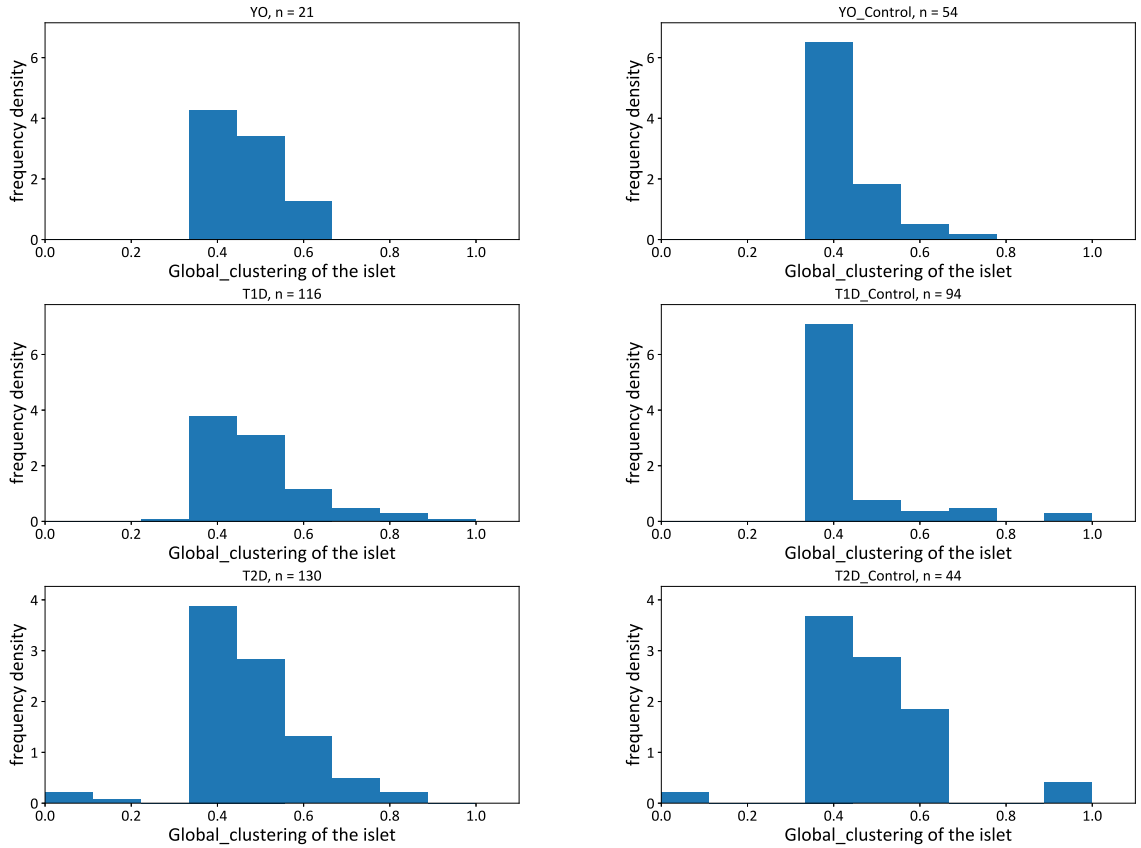


Figure 5.9: This is the set of histograms for the global clustering metric for the overall islet networks. It seems unclear what affect, if at all there is based on age and diabetes.

Neither does the global clustering metric or the number of bridges (Figure 5.15). This is much the same as the results that we see for the overall network. This thesis would hypothesise that this is for the same reasons, and overall in T2D, the underlying beta cell subnetworks are robust.

As someone ages, it appears that the number of connected components per islet increases (Figure 5.11). In addition, the mean and max of both between and degree centrality decreases, suggesting that the beta cells become less connected with each other, in addition to each beta cell in the subnetwork becoming less important. The global clustering and number of bridges does not appear to change, indicating that the density and fragility of the beta cell subnetwork does not change.

5.3.3 Other Cell Subnetworks

Here, in Figure 5.16 we can see that for the alpha alpha cell subnetwork, for the data that we have available, it does not appear that diabetes or ageing significantly impacts the underlying alpha cell subnetworks. That being said, as beta cells comprise the

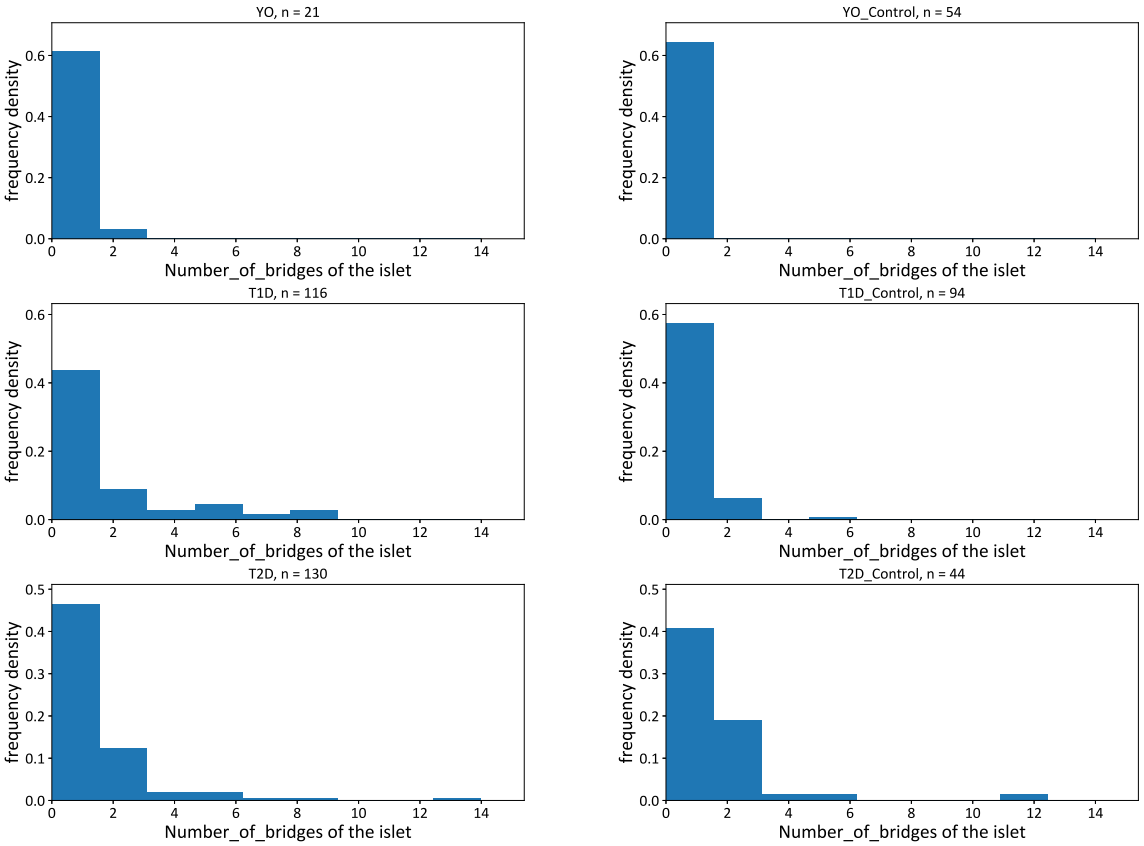


Figure 5.10: This is the set of histograms for the number of bridges metric for the overall islet networks. Age seems to increase the mean of the number of bridges, as does older onset T1D. It seems unclear in the other cases if this is the case too.

vast majority of cells in an islet, a significant amount of data is needed to make any conclusive remarks on this.

5.4 Discussion and Summary

In this chapter, we have created a set of algorithms to induce networks upon islets from images in chapter 2 and utilised metrics to quantify the resulting networks. This is with a view to analyse islets and how they are connected, and how this is associated with diabetes. In addition, network structures capture the underlying topological features of a structure that the graph approximates [73].

Previously there has been a large body of work, concerning how different cell types in the islets interact with each other. Beta-beta cells co-operate electrophysically to simulate islet secretion. Alpha and delta cells are also equipped with the same types of potassium, sodium and calcium channels as the beta cells, though the secretion dependencies on these are less clear [143].

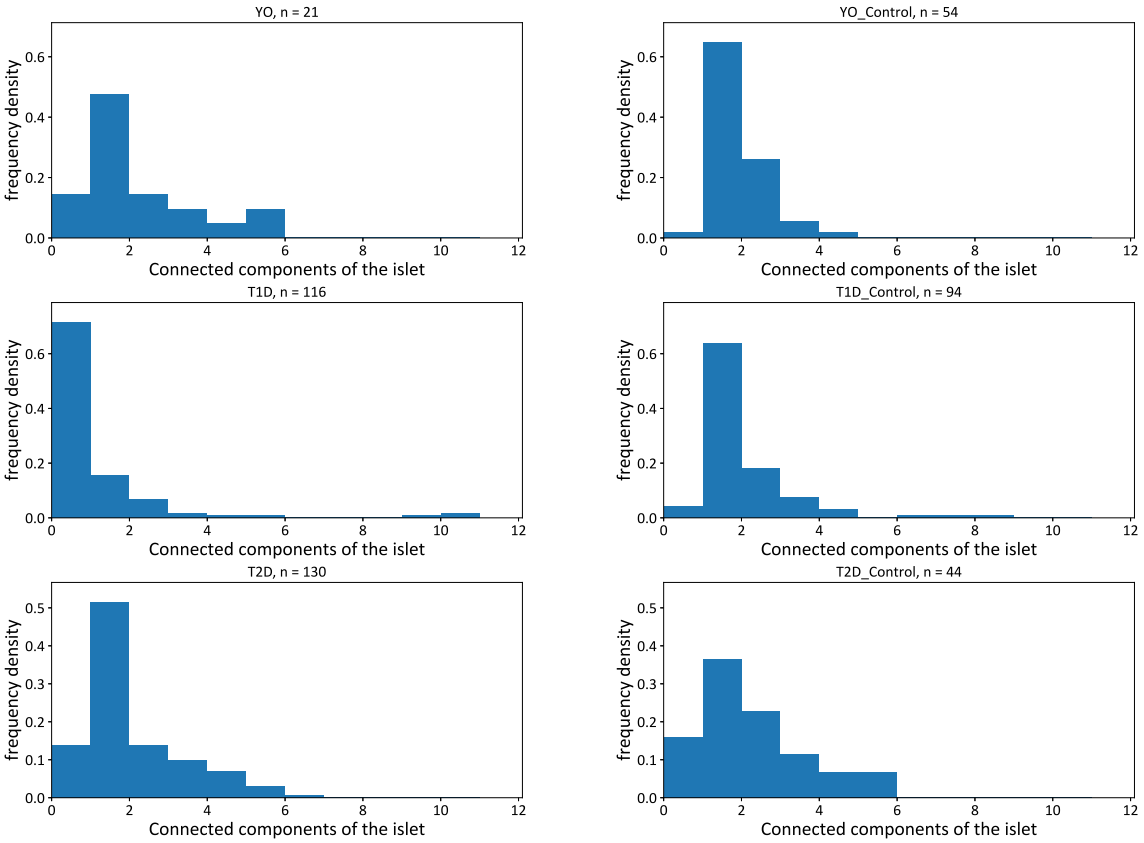


Figure 5.11: Presented here are the set of histograms for the number of connected components within the beta cell network per islet. There seems to be a decrease in mean and skewness in the older onset T1D and T2D cases compared to their controls but an increase in mean and skewness in the younger onset cases.

Beta beta cell connections have been widely studied and there are many different types of models to assess and predict their behaviour base on known mechanisms, such as the calcium model from korošak *et al.* have created a 'spin' model - borrowed from statistical physics to try to describe this behaviour [144]. Further investigations have been conducted other islet cell types: Brereton *et al.* looked into how alpha., delta and PP cells might interact with each other via calcium electrochemical channels [145].

The literature demonstrates that islet cell functionality is linked to the cells that they are connected to, and whilst there are qualitative descriptions of human islet architecture such as that by Kim *et al.* [146], there are no studies which allow us to assess and quantify these cell connections and how diabetes and age might affect these. Hence, this chapter has provided a proof of concept toolkit to assess these connection, which should become more useful as the functionality of the cell connections becomes clearer.

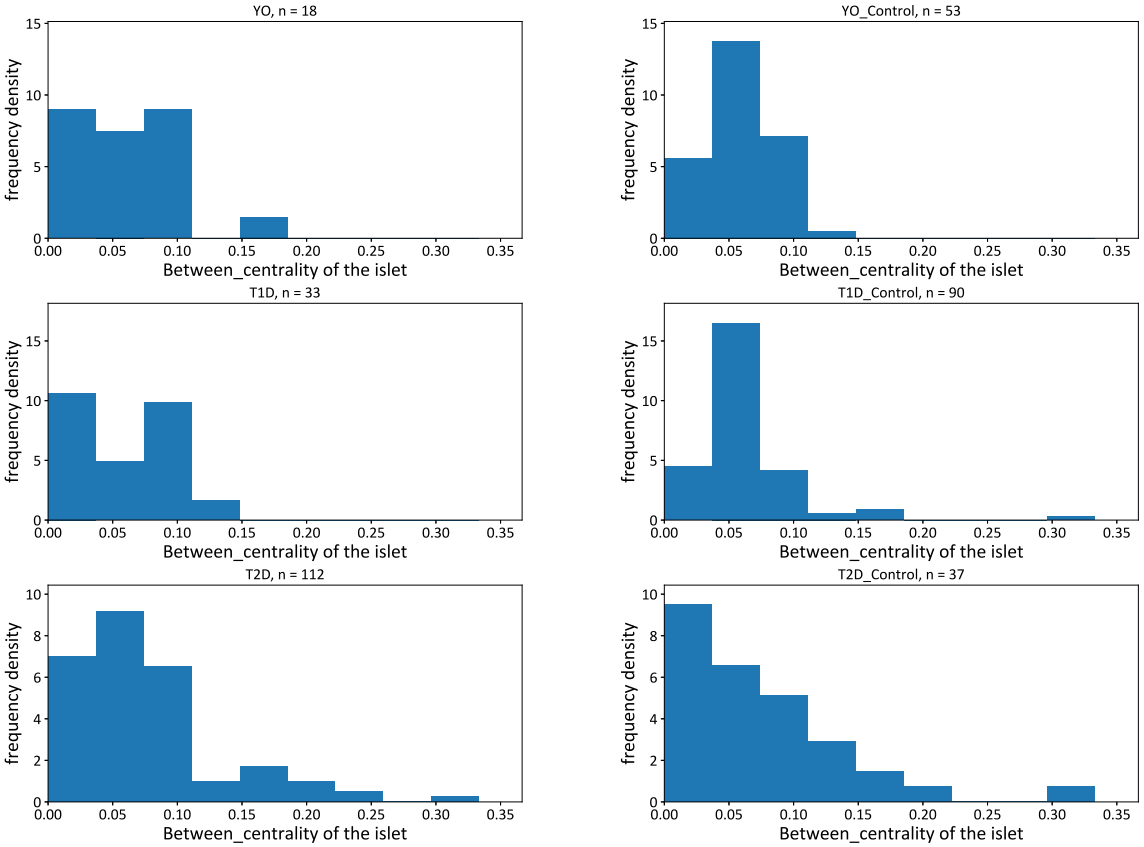


Figure 5.12: This is the set of histograms for the mean between centrality metric for the islet beta cell subnetworks as calculated in subsection 5.2.1. We appear to also see an increase in skewness with age, and with both types of T1D, but it is unclear if the mean changes.

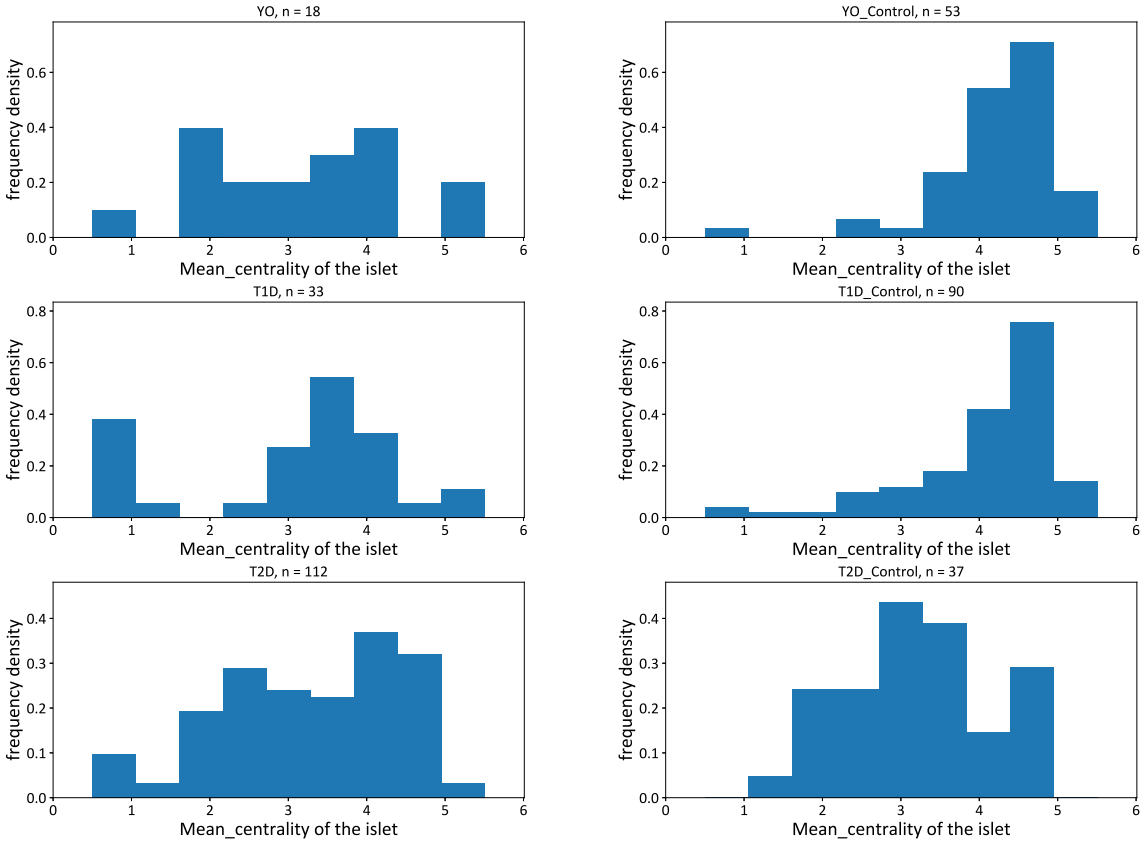


Figure 5.13: This is the set of histograms for the mean degree centrality metric for the islet beta cell subnetworks. It appears that mean degree centrality decreases in both younger and older onset T1D diabetes and age.

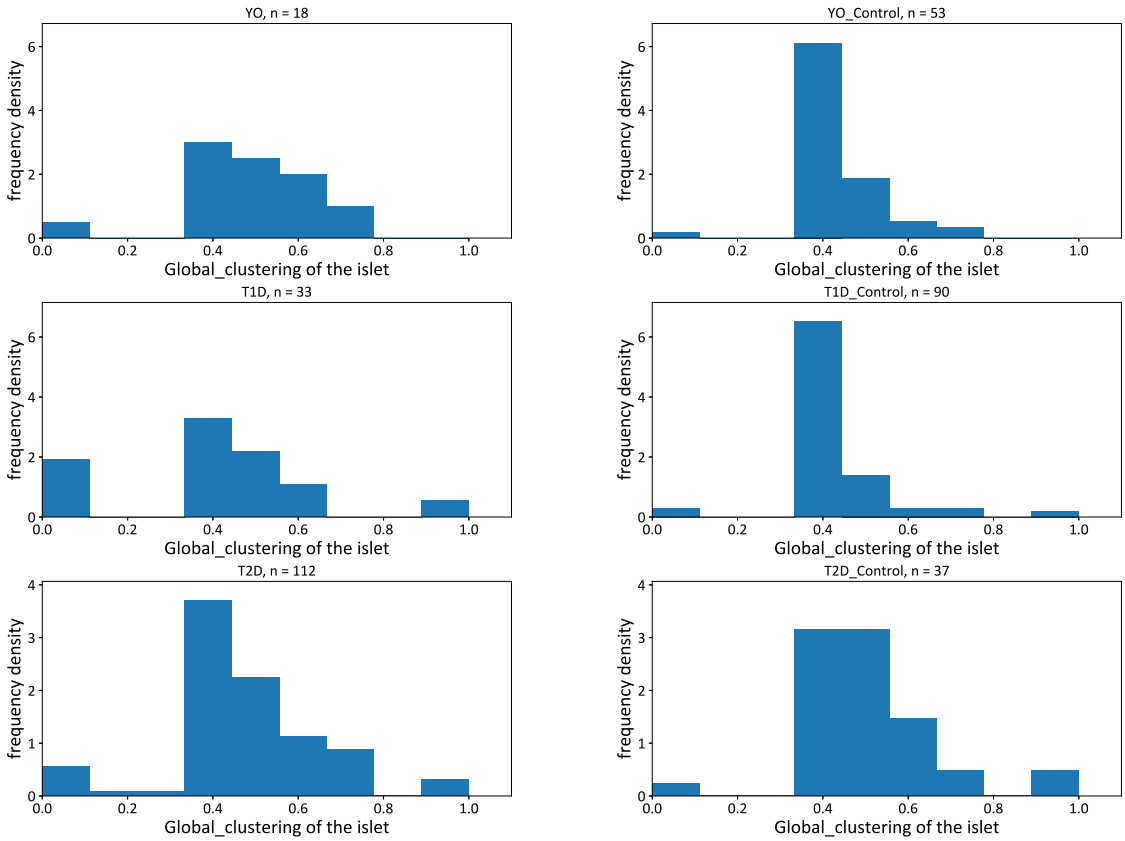


Figure 5.14: This is the set of histograms for the frequency density of the global clustering metric for the islet beta cell subnetworks. It seems unclear as to what is happening within the global clustering metric due to the number of isolated beta cells as highlighted by the individual cell present by the spikes on the left hand side of the graph. This leaves very few results from which to make any inference from.

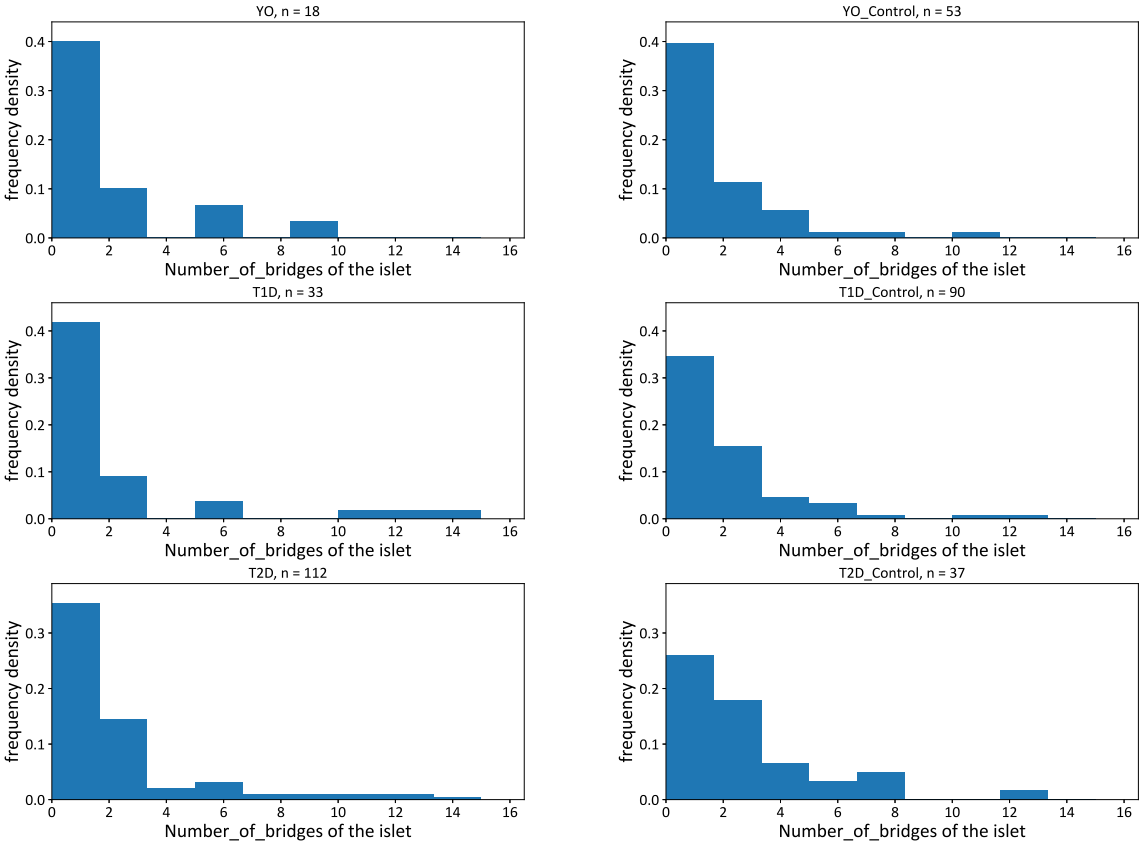


Figure 5.15: This is the set of histograms for the frequency density of the number of bridges metric for the islet beta cell subnetworks. It seems unclear whether there is a change in the number of bridges here in all cases.

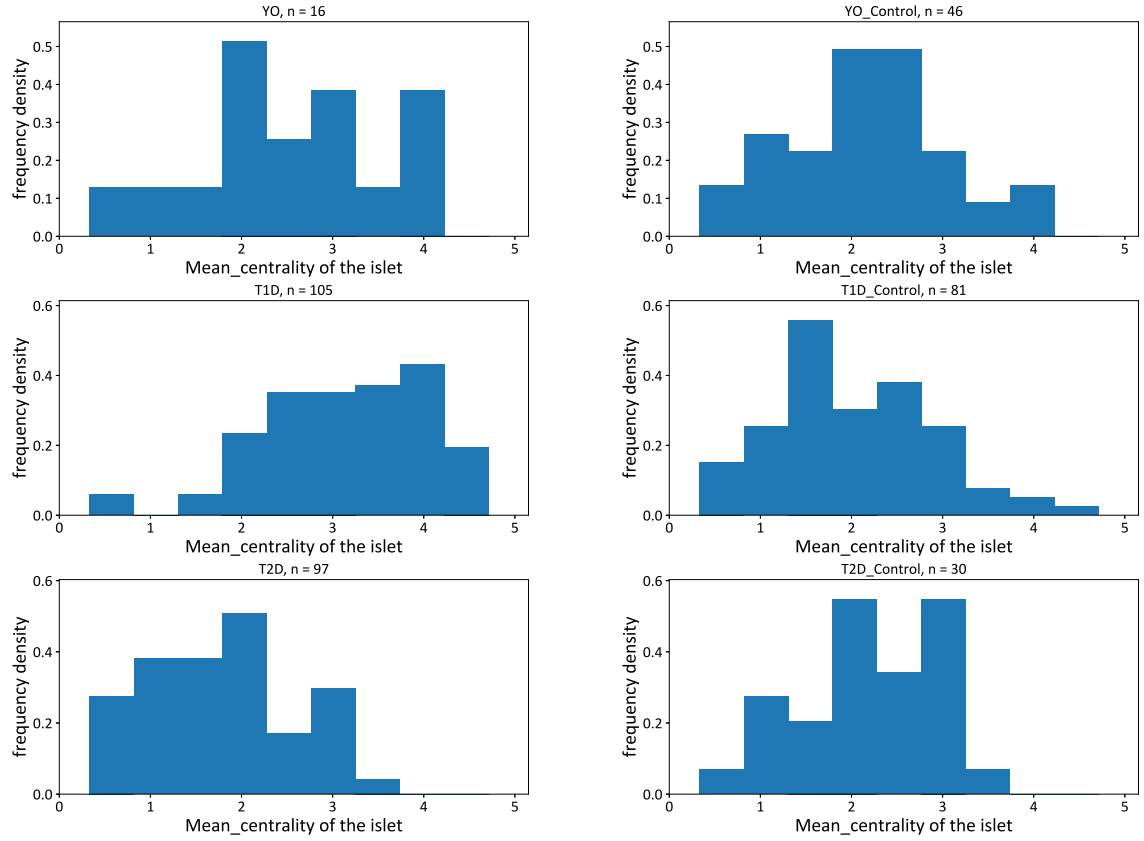


Figure 5.16: This is the set of histograms for the frequency density of the mean centrality metric for the islet alpha cell subnetworks as calculated in subsection 5.2.1.

6. Discussion and Conclusion

6.1 Results Summary

Here, we briefly highlight some of the main results from chapters 2-5 in Figure 6.1. We compare these results with the literature, in particular [147] and table 1 from [148].

6.2 Potential Hypotheses

A summary review table from 2015 with sources of how diabetes changes the cell counts, vasculature and structure within the islets can be found in table 1 of "New insights into the architecture of the islet of Langerhans: a focused cross-species assessment" [148].

6.2.1 Islet Plasticity

In Figure 2.12 we can see that our results seem to suggest that there is a significant reduction of cell counts as a result of diabetes in both the young onset T1D and the older onset T1D. However, our results seem to imply that whilst beta cells are killed some cells seem to become other cell types, called trans-differentiation and de-differentiation in the literature, as shown in Figure 2.13.

De-differentiation is the idea that a terminally differentiated cell can return to a more primitive state and trans-differentiation is the idea that a terminally differentiated cell type can switch lineages [149].

Islet characteristic	Effects of increasing age	Effects of young onset T1D	Effects of older onset T1D	Effects of T2D
Cell counts (from Figure 2.12)	Slight reduction in mean of cell numbers	Substantial reduction in cell count	no change in cell count	Substantial reduction of cell count
α, β, δ cell counts (from Figure 2.13)	Significant decrease in the number of δ cells	Decrease in numbers all cell types	Decrease in β cell counts, but an increase in α and δ cell counts	Slight increase in the number of δ cells
Islet shape (from section 3.3)	Islet size does not change - unclear whether the islets become more folded or not (Figure 3.10)	Islets shrink (Figure 3.5, Figure 3.6) and become more folded (Figure 3.9 and Figure 3.4)	Same as for young onset but less pronounced	No substantial changes to islet shape in T2D
Network connectivity (from subsection 5.3.1)	The kurtosis and mean of the between centrality (Figure 5.7) measures appears to increase whilst the degree centrality (Figure 5.8) appears to decrease whilst connected components per islet increases (Figure 5.6)	Decrease in degree centrality (Figure 5.8)	Same as for the younger onset cases	No significant changes to the network structure in T2D
β cell network connectivity (from subsection 5.3.2)	Between and degree centrality (Figure 5.12 and Figure 5.13) both decrease and connected components per islet increases (Figure 5.11)	Appearance of many isolated components (Figure 5.11)	Same as for younger onset cases	No significant changes to the β cell network structure in T2D

Figure 6.1: Table highlighting the key results from the thesis

For T1D, in Figure 2.13, our results may even suggest that for younger onset type 1 cases, there is little trans-differentiation happening and the insulitis is killing all cell types. This is supported in the literature as it has been widely reported that younger onset T1D tends to be more aggressive [150, 151, 152]. Therefore, it would be interesting to further investigate whether both alpha and delta cells are killed in addition to beta cells in younger onset T1D cases.

In older onset T1D, our results would seem to imply that less overall cells are killed than the younger onset, though potentially more beta cells are killed. The reason for this difference is appears to be that there are more alpha and delta cells. We hypothesise that this is due to cellular trans-differentiation, which is further supported in the literature [153, 154, 155].

In cases of T2D, we don't see a significant change in the total cell count, alpha cells and beta cells, though we see an increase in delta cells. It has been hypothesised by for example Wang *et al.* and Jain *et al.* that changes in beta cell numbers can vary highly between islets - location within the pancreas appears to be a factor [5], as does age of onset [156]. However the literature suggests that cellular dedifferentiation does occur in T2D. Once we have spatially and age aggregated image data, we could use this toolkit to test these hypotheses.

6.2.2 Islet Shape Changes

Within the literature, it appears that two major factors that influence the shape of islets is age and diabetes, though it appears that most research to date make weak implications with regards to shape change and islets. For older people their islets are less dense and have less beta cell mass [157]. One would therefore expect as a non-rigid body that the islets would become smaller as a result. To quote Matsuda, "Age related pathological changes in the pancreas have been unclear because they are often minor and non-specific. However, recent studies have shown that they are closely related to various pathological conditions such as pancreatic cancer and diabetes mellitus." [158]

From our results, as one ages, it does not necessarily appear that there is a

change in mean in the diameter or cross sectional areas of the islets (Figure 3.5) - though there could be some changes within the moments of the data which be interesting to analyse with more islet samples. However, we do see a reduction in the mean of the isoparametric ratio (Figure 3.9), the perimeter to radius ratio (Figure 3.10) and the incircle to circumcircle ratio (Figure 3.8), which means that we hypothesise as one ages, the islets are becoming more folded.

With regards to diabetes related shape changes, the literature regarding shape changes is similarly sparse - especially for T1D, but has also been implied as being important with regards to islet function . Within type 1 diabetes, among caprine models for example, Hani *et al.* suggest that type 1 diabetes can cause changes to the surrounding islet membrane, triggering islet deformities and affecting function [159]. Due to a reduction in beta cell mass, one would naturally expect a reduction in islet size but to date no quantitative study has been done on this.

As a result of T1D, we hypothesise that we see a significant reduction in the mean islet cross sectional areas (Figure 3.5), diameter (Figure 3.7) along with the isoparametric ratio (Figure 3.9), the ratio between the incircle and circumcircle (Figure 3.8) and the perimeter to radius ratio (Figure 3.10). In other words, we hypothesize that not only do the islets shrink as a result of T1D, but the islets also become more folded. We also hypothesise that the results are more pronounced and extreme in the younger onset patients than the older onset cases

With type 2 diabetes, it would appear that there is an increase in starfish shaped islets in rodent models [160]. There is mixed evidence of islet hypertrophying and atrophying in different circumstances, potentially dependent on the region in the pancreas. However this induced size difference likely has an affect on insulin secretion, so quantifying this is important to understand this better.

As a result of T2D, we don't see significant changes in between diabetic and their controls in terms of the geometry, however more data is required, especially at the extremes to assess the impact of atrophying and hypertrophying islets.

6.2.3 Vasculature and Islet Destruction

The literature is quite clear that vessel density is increased in diabetic patients as opposed to non-diabetic patients, though the vessels themselves are smaller in T1D cases [161, 162, 11, 163]

From our methodology, if it were to be made more robust, we would speculate that we would see an increase in the Betti numbers from the images as we moved from healthy patients to diabetic patients.

6.2.4 Islet Connectivity Changes

The cells within a pancreas interact with each other, electrochemically - a review by Koh *et al.* provides a succinct overview of what is known about the interaction between cell types in the pancreas [164].

Just as cells interact with each other ionically, it has become vogue in recent years to model cell clusters within islets as networks. For example, Johnston *et al.* considers beta cell hubs, which act as 'pacemakers' within beta cell networks [165]. Further there have been many ODE system models which describe the interaction between paracrine cell types. Among the most comprehensive to date is one that models the interactions of alpha, beta delta cells [166].

As a result of T1D, our results suggest the islets become less connected, with many cells being killed. With T2D, we don't find any significant changes to the underlying network for diabetes. On the other hand, as one ages, whilst we don't see high levels of cell loss, we do see the islets become less well connected - reflected especially in the increasing number of connected Figure 5.6 and Figure 5.11 components per islet.

6.3 Result and Impact Discussion and Summary

Chapter 1, is a literature review: Firstly, we outline a definition of diabetes, followed by an outline of the strucuture of the pancreas, the islets and cells contained within. Secondly, it contains a review of the current knowledge of the epidemiology and

aetiology of diabetes, both type 1 and type 2. Finally, there is a brief review of current literature of morphometry and the study of image analysis with application to biosciences.

In chapter 2, we firstly describe the data that we receive from the medical school, followed by a description of how the data was processed and analysed into spatial data. The algorithm developed to analyse the islets is a novel contribution to the field. Building upon this, we analysed the output of the algorithm to assess the effect of diabetes upon the number of insulin producing beta cells, along with the alpha and delta cells. This original analysis has yielded some interesting results - beta cell numbers are reduced in both classes of T1D, but stay the same in T2D, as has been previously proposed. Interestingly, it appears that the number of delta cells present increases in T2D and the older onset T1D, where as there is a decrease in T2D. Ghrelin producing delta cells regulate appetite as well as insulin release, so whilst we are not sure of the mechanism, this could have a significant impact on treatments.

In chapter 3, the spatial data generated as a result of chapter 2 to analyse how diabetes changes the shape of the islets. The form of an islets is functionally relevant, for example, folds in islets. By systematically analysing these features, we have assessed how the shape of the islets change associated with diabetes. our preliminary results would seem to indicate the islets shrink as a result of type 1 diabetes, especially the younger onset patients. This is corroborated in the literature and is called pancreatic atrophy [167, 168, 169]. On the other hand, in type 2 diabetes, we do not see similar results with regards to how the shape changes as a result of diabetes. This does not correspond to the literature - though the literature does focus on rat models [170, 171]. However, this could as a result of aging as opposed to the type 2 diabetes. We seem to see a growth in islets between the young onset type 1 controls to the older onset type 1 controls, followed by a shrinking of the islets to type 2 controls, with an increase in the variance throughout. This has not been previously studied in detail before, and whilst we would need more data to verify this, might provide some interesting insight into islet behaviour as a healthy

person ages.

In chapter 4 we build upon the use of the spatial data to analyse how the structure of the islets are affected by diabetes. However these structural results did not appear to be stable over a range of parameters in 2D. Analysing structure is an exciting new idea, which might be more viable in the 3D structure that islets actually inhabit.

In chapter 5 we firstly induced networks on islet images - processed in chapter 2. We then utilise several metrics from the NetworkX package to analyse these graphs with and have generated a toolkit to analyse the cell networks in islets, both as a totality and as a function of the different cell types. In general as one ages, our results would suggest that the networks become more disjointed, and whilst type 1 diabetes causes there to be less connections overall, type 2 does not. This is consistent with the literature, suggesting that cells are killed in type 1 diabetes, but there is a lack of evidence for this being the case in type 2 diabetes

A serious limitation of the work is that the data we are using in both the image analysis and model is a 2D projection of 3D data. Three dimensional data would allow additional degrees of freedom. Since we do not have access to 3D human islet images, this is beyond the scope of our investigations. When and if we are able to do this, 3D islet images could be used in an extension to this work.

6.4 Future Plans

6.4.1 Network Dynamic Model for Islet Secretion

By arbitrarily labelling the nodes representing cell centres within the islet, it would be possible to create a Markov chain islet secretion simulation model. The major diagonal values in the Markov chain would be the self-simulation which would allow us to understand spatial temporal behaviour, both under healthy and diabetic conditions. The stochastic matrix would be sparsely populated, allowing for easy identification of the egienvalues/vectors, so one could analyse the periodicity for short, medium and long term cycles as has been first identified and mathematically

analysed by Bertram and Sherman in 2000 [14].

Markov chain systems have been used in biosciences, specifically in neurosciences so model things such as ion-channel release modelling [172]. By taking the limit of time period between state changes in a Markov chain down to zero, we can create an ODE system, constrained to the region of the islet network connections. Describing the secretion state of the cells within the islet as a function of time.

These models would only include direct cell-cell contacts. We know that there are other modes of intracellular communication within the islets, such as via the vasculature [173]. We are, however, interested in how base behaviour of islets changes as a result of diabetes, but also how islets change as one ages - something that has not been widely researched in humans. Such models would be an interesting place to start.

6.4.2 Cellular Automata Model for Insulitis

We could construct a cellular automata to simulate the autoimmune destruction of beta as present within T1D and hypothesised to be present in T2D. There are several advantages and disadvantages to various model types - the feasibility and a basic outline for this potential project are given in Appendix C.

6.4.3 3D Islets

One of the biggest weaknesses of this research is that we do not look investigate 3 dimensional islets. Real 3D data at the resolution explored in this thesis does not exist for islets for humans. Whilst 3D scans have been taken of animal islets [174] it has been previously established that animal islets just are not similar enough to human islets to extrapolate results from them. If and when 3D human islet scans become available in significant enough quantities, they should be analysed in a similar fashion to what we have done here. As 3D islets would potentially have a lot more geometric and topological complexity than 2D islets, we would expect to be able to capture much more subtle aspects of the vasculature within the islets.

6.4.4 Constrained Voronoi

In our analysis, we have used a Voronoi algorithm in $\mathbb{R}^2$ and masked it using the cell outline. Research has been done by [175] in which more general Voronoi algorithms are developed and applied on restricted domains. In such cases a partition is found that considers only paths within the restricted domain. This would be difficult to implement, but directly applicable to our problem, and increase the accuracy of the cell designation and hence the overall results.

A. Results Tables

A.1 Intra Patient Type Table

Dataset	Kruskal-Wallis P-Value	Pt Count	Sample Size
Old onset #alpha cells per islet	2.05×10^{-2}	6	57
Old control #alpha cells per islet	1.12×10^{-3}	7	72
Old onset #beta cells per islet	2.87×10^{-3}	6	57
Old control #beta cells per islet	1.11×10^{-23}	7	72
Old onset #delta cells per islet	4.50×10^{-1}	6	57
Old control #delta cells per islet	1.40×10^{-17}	7	72
Young onset #alpha cells per islet	7.10×10^{-5}	5	11
Young control #alpha cells per islet	6.61×10^{-3}	4	32
Young onset #beta cells per islet	1.11×10^{-4}	5	11
Young control #beta cells per islet	1.44×10^{-3}	4	32
Young onset #delta cells per islet	3.27×10^{-3}	5	11
Young control #delta cells per islet	1.66×10^{-1}	4	32
T2D #alpha cells per islet	5.44×10^{-1}	5	50
T2D control #alpha cells per islet	5.48×10^{-1}	3	28
T2D #beta cells per islet	1.01×10^{-2}	5	50
T2D control #beta cells per islet	5.21×10^{-3}	3	28
T2D #delta cells per islet	5.27×10^{-1}	5	50
T2D control #delta cells per islet	7.64×10^{-6}	3	28

Presented here is a table comparing patients within each subject group, using the Kruskal Wallis H test. We use this to assess as to whether we need to correct for a dependency of the islet data on the individual patients.

A.2 Inter-Species Table

Dataset 1	Dataset 2	Mann Whitney U test P value	Effect size P value, P_{ab}
Old onset #alpha cells per islet	Old control #alpha cells	1.76×10^{-1}	5.70×10^{-1}
Old onset #beta cells per islet	Old control #beta cells	5.32×10^{-15}	9.92×10^{-2}
Old onset #delta cells per islet	Old control #delta cells	4.42×10^{-7}	7.58×10^{-1}
Old onset total cells per islet	Old control total cells	1.30×10^{-9}	1.86×10^{-1}
T2D #alpha cells per islet	T2D control #alpha cells	3.37×10^{-1}	4.34×10^{-1}
T2D #beta cells per islet	T2D control #beta cells	3.76×10^{-1}	5.61×10^{-1}
T2D #delta cells per islet	T2D control #delta cells	5.89×10^{-5}	7.72×10^{-1}
T2D total cells per islet	T2D control total cells	2.50×10^{-1}	4.34×10^{-1}
Young onset #alpha cells per islet	Young control #alpha cells	2.07×10^{-1}	5.83×10^{-1}
Young onset #beta cells per islet	Young control #beta cells	7.77×10^{-13}	2.91×10^{-2}
Young onset #delta cells per islet	Young control #delta cells	5.65×10^{-3}	6.82×10^{-1}
Young onset total cells per islet	Young control total cells	5.86×10^{-7}	1.71×10^{-1}

Presented here is a table comparing between patients groups and their respective control groups, using the Mann Whitney U test. We use this to assess if there has been a significant change in various cell counts and their respective controls.

B. Modifiable Unit Areal Problem (MUAP)

The modifiable unit areal problem is a problem involving the introduction of bias relating to spatial statistics in image analysis. As a spatial problem, it is perhaps best explained pictorially in Figure B.1 and is largely self explanatory on viewing the image. The most recognisable application in popular culture is 'gerrymandering'. Common in first past the post electoral type systems, gerrymandering is the practice of establishing district boundaries in order to gain more representation than support would typically allow. A typical example of this is in the US in which a candidate may lose the 'popular' vote but still win an election.

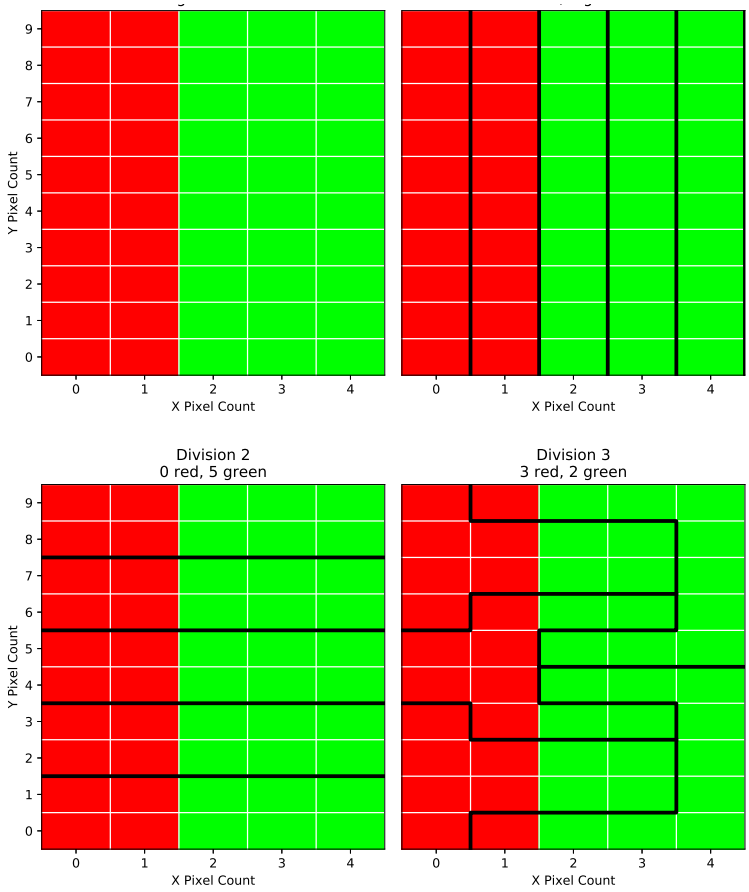


Figure B.1: An image plot demonstrating the modifiable areal unit problem. **In the top left**, we see an 'image' with 50 pixels. 20 are red hormone, whilst 30 are green hormone. From that, if we were to divide this in 5 cells, we would expect there to be 2 red cells and 3 green cells when assigning based on the hormone presence within the cell. **In the top right**, we see 2 red cells and 3 green cells, as expected. **In the bottom left**, the cell division yields no red cells and 5 green cells. However, **in the bottom right** there are 3 red cells and 2 green cells.

C. Cellular Potts Model for Insulitis

C.1 Converting to Image Data for Modelling

In order to make the islet images images suitable as an input for the computational model of beta cell destruction, it was necessary to reduce the resolution. This is in order to run simulations of the model in a reasonable amount of time for a large number of samples, especially during the development phase. Initially, the centroid of all the cells was rescaled from 1296 by 966 to 100 by 100 pixels. Subsequently, every islet pixel in the rescaled image was assigned to its nearest relevant centroid. The resulting image is shown in Figure C.1. For example, blue pixels looked for the nearest SST cell centre, via looping through all the centres and calculating the Euclidian distances. After assigning each islet pixel to a cell, each cell was recorded with its type and all its pixels as 1 dimension index array and saved to a .txt file; the 1 dimension index was recorded as row major index which is to say:

$$Index(i, j) = i + j \times i_{max}, \tag{C.1}$$

where i and j are index positions; i_{max} is the horizontal pixel size of the image.

C.2 Types of Cellular Automata Models

Our goal is to identify and test various mechanisms for beta cell destruction in the pancreas. Using the image processing techniques outlined above in section 2.2, we can obtain data regarding cell positioning within the islets for both the age

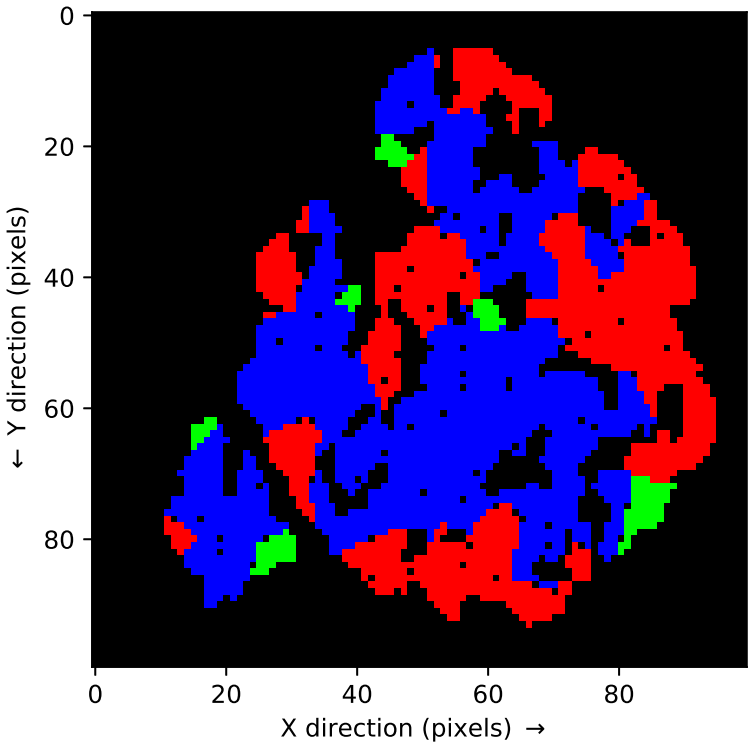


Figure C.1: Image of example islet with resolution reduced to 100 by 100 pixels.

matched controls and T1D subjects. From a modelling perspective, this is assumed to be before (control data) and after T1D (subject data), respectively. Using spatio-temporal models, as outlined in the thesis aims, will help us understand the underlying immune mediated T1D mechanisms better. This cannot be done via more conventional methods, such as dissections or animal models, due to the location and enzymatic nature of the pancreas, as discussed in the introduction.

According to a recent review by Osborne *et al.* [176] there are 5 major ways in which the motility, genesis and apoptosis of cells within tissues are currently computationally modelled. Two similar models in modern cell research are the Voronoi tessellation model and the vertex model. In a Voronoi tessellation model, the cells are assumed to be polygons, and a set of points corresponding to a bijective map with a unique point inside each cell. The positions of these points, also known as seeds, vary as a function of time with according to Newtonian mechanics [177]. A Voronoi tessellation is constructed about these points in order to represent cells. Cells are similarly assumed to be polygons in the vertex model. However, it is the vertices of cells that change position with respect to time, with each vertex representing a common point between 3 or more cells [178]. However, both of these

models assume a dense cell tissue structure, which is not representative of human islets within the pancreas. A further model type that we considered as a framework for solving our research questions is of a cellular automata type. This is an on-lattice model, which means that space and time are discretised [179]. In this model, each cell is represented by a pixel and these evolve with respect to time, explained further in Figure C.2. This model type was used by Ozturk *et al.* to construct a model for T1D insulitis in mouse islets [180]. However, given that we wish to investigate how geometry affects beta cell destruction, one uniform pixel per cell would not be sufficiently representative of the morphology of the cells and islets.

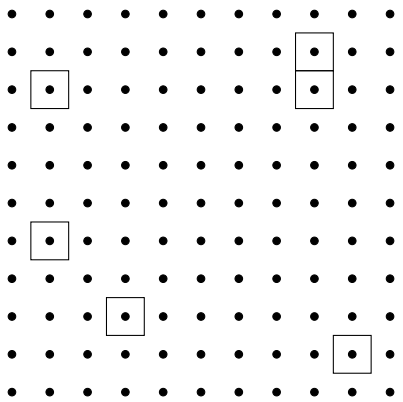


Figure C.2: An illustration of how cells would look in a cellular automaton model. Each cell (closed square) is anchored to a lattice site (black dot). A valid move on a time update can either be to stay put, or move to an adjacent lattice site, if that site is not already occupied by an existing cell.

Another modelling framework, the overlapping spheres model has been used by Wedgwood *et. al* [181] to develop a spatio-temporal model for insulitis. Overlapping spheres models are circular (in 2D) and interact with each other when the centres of the cell are within a certain range; this can be thought of as the cell radii. In the Wedgwood model, the white blood cells were assumed to be attracted to the beta cells, within the islets, via a freely diffusing chemokine gradient. This was modelled using a reaction-diffusion equation. Solving this equation analytically relied on the assumption that the islets were circular. From the data above, however, it is clear that the islets are not circular. In our model, we aim to use a realistic islet geometry that would require solving the reaction-diffusion equation numerically.

C.3 Cellular Potts Model

The framework that we would like to use to build our model is a cellular Potts framework. A cellular Potts model (CPM) is an on-lattice model where all cells consist of a number of pixels [182], illustrated in Figure C.3. Allowing multiple pixels per cell within a CPM, as opposed to cellular automaton models with a single pixel per cell, allows for more realistic cell and tissue geometry to be realised, which is a clear advantage.

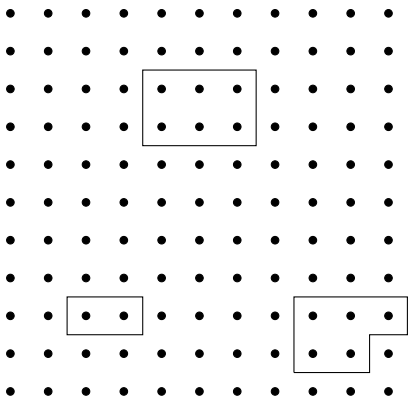


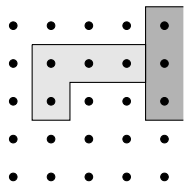
Figure C.3: An illustration of how cells (illustrated by the closed shapes) are modelled in a CPM. A cell is a set of connected lattice points. How they move on each time update, and what connected means in this context is outlined below.

A CPM works as follows: firstly, each cell has a set of valid moves that it can take which are calculated and recorded. A valid move (Figure C.4) for a cell is to extend into adjacent empty pixels, or contract, such that the cell is not split into two.

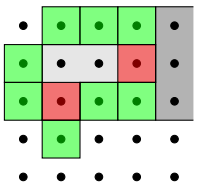
Secondly, a move from this set is chosen randomly. Thirdly, a Hamiltonian of the current lattice and the lattice after the proposed move is computed. A Hamiltonian in this context refers to the energy present due to the lattice configuration [**?**]. In 2D, this can be decomposed into three components:

$$H = H_{area} + H_{adhesion} + H_{chemical} \tag{C.2}$$

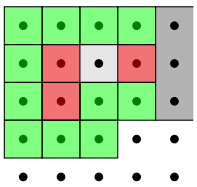
This overall Hamilton can be split into the Hamiltonian due to the area (H_{area}), a Hamiltonian due to a cell-cell contact ($H_{adhesion}$), and finally a Hamiltonian due to the presence of chemokines ($H_{chemical}$). The Hamiltonian contribution due to area



(a) Presented here is a 4 pixel, 'L' shaped cell, adjacent to another cell (comprising of 3 pixels).



(b) Illustration of the Neumann neighbourhood move set upon Figure C.4a



(c) Illustration of the Moore neighbourhood move set upon Figure C.4a

Figure C.4: An example of different sets of moves that a cell can make, depending on how a neighbourhood is defined. In green are cell expansion moves and in red are cell contraction moves. A cell cannot expand into another cell and whilst contracting, it cannot split to make two cells. In a CPM, the pixels constituting each cell must be either 4-connected (Neumann neighbourhood), which means that each pixel in a cell must have an adjacent cell pixel in one of the 4 cardinal directions, as illustrated in Figure C.4b. In the Moore neighbourhood, the 8 adjacent pixels in cardinal and inter-cardinal directions are said to be 8-connected, illustrated in Figure C.4c.

constraints can be written as the following:

$$H_{area} = \sum_{\sigma}^{\forall\, cells} \lambda_A \left(A\left(\sigma,\tau\right) - A_{target}\left(\sigma,\tau\right)\right)^2 + \sum_{\sigma}^{\forall\, cells} \lambda_P \left(P\left(\sigma,\tau\right) - P_{target}\left(\sigma,\tau\right)\right)^2, \tag{C.3}$$

where, σ denotes an individual cell, and τ is a type of cell. $A\left(\sigma,\tau\right)$ and $P\left(\sigma,\tau\right)$ are the cells current area and perimeter, respectively, and $A_{target}\left(\sigma,\tau\right)$ and $P_{target}\left(\sigma,\tau\right)$ are the cells target area and perimeter. λ_A and λ_P are area and perimeter elasticity constants, respectively. The area Hamiltonian is the square sum penalty for the deviation from each cells target area and perimeter. The adhesion contact Hamiltonian can be written as:

$$H_{adhesion} = \sum_{(i,j)(i',j')}^{\forall\, neighbours\, pixels} J_{\tau(\sigma)\tau'(\sigma')}\left(1 - \delta\left(\sigma(i,j), \sigma'(i',j')\right)\right), \tag{C.4}$$

where, as before σ denotes an individual cell, and τ is a type of cell. i, j denote cellular lattice positions around the edge of a cell; Chen *et al.* [183] suggest using a depth of 4 pixels from the edge of a cell. The Kronecker delta, $\delta(\sigma,\sigma')$ ensures that only surface sites between different cells are considered. $J_{\tau,\tau'}$ is the binding energy per unit perimeter. This binding energy constant can be positive or negative, depending on the type of cell-cell interactions. The chemical potential Hamiltonian

can be written as:

$$H_{chemical} = \sum_{\sigma}^{\forall\, cells} \mu(\tau)C(\sigma), \tag{C.5}$$

where, $C(\sigma)$ is the chemotactic gradient summed over the lattice positions of each cell. $\mu(\tau)$ is the chemotactic attraction constant over different cell types. The chemotactic gradient can be modelled via a reaction diffusion equation, which can be implemented discretely upon the lattice. The chemical potential Hamiltonian sums over the cell lattice positions and takes the product of the chemical potential at that position and the attraction constant of the cell. Finally, the energy difference, ∇H is calculated between the cells both before and after a proposed move is taken. If the proposed move has a negative energy change or exothermic change associated with it, then it is accepted unconditionally. If the proposed move has a positive energy change, then it is endothermic and is accepted with a probability of $\exp(-\nabla H/T)$, where T is the 'temperature' of the system; a number which when zero will allow a cell to only move in exothermic steps, but when infinite will allow the cell to move randomly, and not be bound by energy considerations.

The model will include beta cells which produce a chemotactic gradient that will recruit white blood cells to destroy the beta cells. This gradient will be based upon research by Christen *et al.* [184], and chemokines will be assumed to diffuse freely. There will be inert non-beta cells, representing the other cell types present in the pancreas. The location of the islet cells will be obtained from the processed microscopy images. There will be two types of white blood cells, CD20+ B cells and CD8+ T cells. The CD20+ and CD8+ will cooperate when close enough to activate the CD8+ T cells to secrete a β cell killing cytokine. This will be implemented by applying work by authors such as Scianna *et al.* [185] regarding implementation of cell motility in CPM.

We plan to build on this model further by introducing a basement membrane. This is a protein encapsulation of an islet, and has been implicated by various authors such as Kragl *et al.* [186] and Otokoski *et al.* [187] in protecting the islets from leukocyte infiltration (as discussed in the introduction). We would also like to add in CD48+ macrophages. Maree *et al.* have constructed a temporal model for this

specific infiltration [67], which we will use to inform modelling these leukocytes.

We are implementing this model in an library called CHASTE (standing for Cancer, Heart And Soft Tissue Environment; see [188]). Developed by programmers, mathematicians and biologists at Oxford university, this library, written in C++, is used in biomedical research to allow simulations of tissue behaviour *in silico*.

The first step to construct this model is to input realistic islets. For this, we plan to insert alpha, beta and delta cells into the model for the start of the simulation, the positions of which we obtained from the post processed images, as outlined in section C.1.

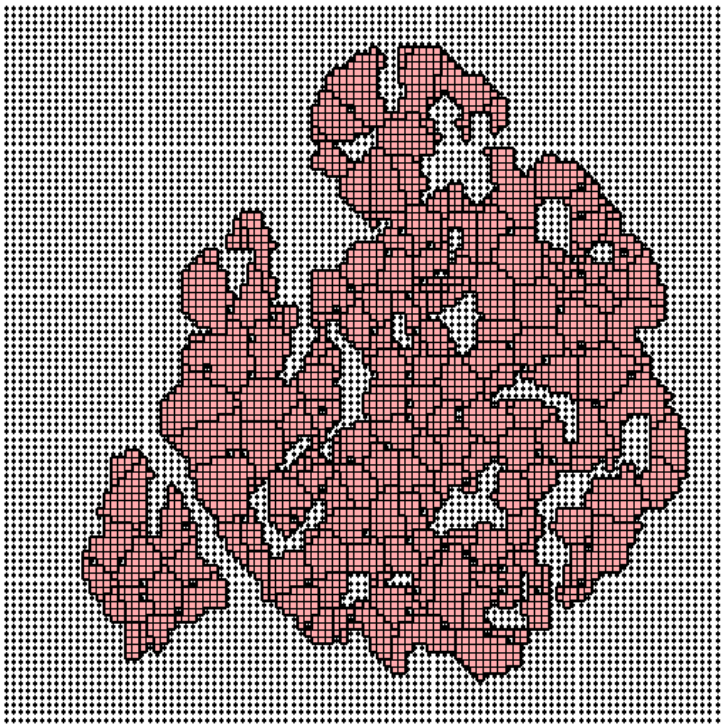


Figure C.5: Visual output from CHASTE, demonstrating the cells inserted in the model, processed from Figure 2.1 .

D. List of Abbreviations

The following table describes the significance of various abbreviations and acronyms used throughout the thesis. The page on which each one is defined or first used is also given.

Abbreviation	**Meaning**	**Page**
B	Processed, binary image of an islet	23
$\mathcal{B}$	The set of processed, binary images of the islets	23
GCG	Glucagon	3
SST	Somatostatin	3
T1D	Type 1 Diabetes	6
T2D	Type 2 Diabetes	6
UoEMS	University of Exeter Medical School	17

Bibliography

[1] B. Blaus, "Pancreatic tissue," *WikiJournal of Medicine*.

[2] I. M. Peter In't Veld, Silke Smeets, *Islets of Langerhans*, ch. 2. Islets of Langerhans, Springer Netherlands, 2014.

[3] D. E. Bockman, *The Pancreas: An Integrated Textbook of Basic Science, Medicine, and Surgery, Second Edition*, ch. 4. Wiley-Blackwell, 2008.

[4] D. S. Longnecker, "Anatomy and histology of the pancreas," *Pancreapedia: The Exocrine Pancreas Knowledge Base*, 2014.

[5] X. Wang, R. Misawa, M. C. Zielinski, P. Cowen, J. Jo, V. Periwal, C. Ricordi, A. Khan, J. Szust, J. Shen, *et al.*, "Regional differences in islet distribution in the human pancreas-preferential beta-cell loss in the head region in patients with type 2 diabetes," *PloS one*, vol. 8, no. 6, p. e67454, 2013.

[6] G. Camussi, M. M. Zanone, and E. Favaro, "From endothelial to β cells: insights into pancreatic islet microendothelium," *Current diabetes reviews*, vol. 4, no. 1, pp. 1–9, 2008.

[7] R. M. Buijs, S. J. Chun, A. Niijima, H. J. Romijn, and K. Nagai, "Parasympathetic and sympathetic control of the pancreas: a role for the suprachiasmatic nucleus and other hypothalamic centers that are involved in the regulation of food intake," *Journal of Comparative Neurology*, vol. 431, no. 4, pp. 405–423, 2001.

[8] C. L. Prado, A. E. Pugh-Bernard, L. Elghazi, B. Sosa-Pineda, and L. Sussel, "Ghrelin cells replace insulin-producing β cells in two mouse models of pancreas

development," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 9, pp. 2924–2929, 2004.

[9] A. Pugliese, M. Yang, I. Kusmarteva, T. Heiple, F. Vendrame, C. Wasserfall, P. Rowe, J. M. Moraski, S. Ball, L. Jebson, *et al.*, "The juvenile diabetes research foundation network for pancreatic organ donors with diabetes (npod) program: goals, operational model and emerging findings," *Pediatric diabetes*, vol. 15, no. 1, pp. 1–9, 2014.

[10] É. Korpos, N. Kadri, R. Kappelhoff, J. Wegner, C. M. Overall, E. Weber, D. Holmberg, S. Cardell, and L. Sorokin, "The peri-islet basement membrane, a barrier to infiltrating leukocytes in type 1 diabetes in mouse and human," *Diabetes*, vol. 62, no. 2, pp. 531–542, 2013.

[11] M. Marichal *et al.*, "Microscopic anatomy of the human islet of langerhans," in *The islets of Langerhans*, pp. 1–19, Springer, 2010.

[12] M. A. Kane, "Analysis, occurrence, and function of 9-cis-retinoic acid," *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, vol. 1821, no. 1, pp. 10–20, 2012.

[13] A. Wojtusciszyn, M. Armanet, P. Morel, T. Berney, and D. Bosco, "Insulin secretion from human beta cells is heterogeneous and dependent on cell-to-cell contacts," *Diabetologia*, vol. 51, no. 10, p. 1843, 2008.

[14] R. Bertram, J. Previte, A. Sherman, T. A. Kinard, and L. S. Satin, "The phantom burster model for pancreatic β-cells," *Biophysical journal*, vol. 79, no. 6, pp. 2880–2892, 2000.

[15] U. Dahl, A. Sjodin, and H. Semb, "Cadherins regulate aggregation of pancreatic beta-cells in vivo," *Development*, vol. 122, no. 9, pp. 2895–2902, 1996.

[16] Z. Fu, E. R Gilbert, and D. Liu, "Regulation of insulin synthesis and secretion and pancreatic beta-cell dysfunction in diabetes," *Current diabetes reviews*, vol. 9, no. 1, pp. 25–53, 2013.

[17] C. J. Barker, I. B. Leibiger, and P.-O. Berggren, "The pancreatic islet as a signaling hub," *Advances in biological regulation*, vol. 53, no. 1, pp. 156–163, 2013.

[18] D. Matthews, J. Hosker, A. Rudenski, B. Naylor, D. Treacher, and R. Turner, "Homeostasis model assessment: insulin resistance and β-cell function from fasting plasma glucose and insulin concentrations in man," *Diabetologia*, vol. 28, no. 7, pp. 412–419, 1985.

[19] K. G. M. M. Alberti and P. f. Zimmet, "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*, vol. 15, no. 7, pp. 539–553, 1998.

[20] K. L. Graham, R. M. Sutherland, S. I. Mannering, Y. Zhao, J. Chee, B. Krishnamurthy, H. E. Thomas, A. M. Lew, and T. W. Kay, "Pathogenic mechanisms in type 1 diabetes: the islet is both target and driver of disease," *The review of diabetic studies: RDS*, vol. 9, no. 4, p. 148, 2012.

[21] J. Tuomilehto, J. Lindström, J. G. Eriksson, T. T. Valle, H. Hämäläinen, P. Ilanne-Parikka, S. Keinänen-Kiukaanniemi, M. Laakso, A. Louheranta, M. Rastas, *et al.*, "Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance," *New England Journal of Medicine*, vol. 344, no. 18, pp. 1343–1350, 2001.

[22] S. E. Kahn, R. L. Hull, and K. M. Utzschneider, "Mechanisms linking obesity to insulin resistance and type 2 diabetes," *Nature*, vol. 444, no. 7121, pp. 840–846, 2006.

[23] P. P. Brahmkshatriya, A. A. Mehta, B. D. Saboo, and R. K. Goyal, "Characteristics and prevalence of latent autoimmune diabetes in adults (lada)," *ISRN pharmacology*, vol. 2012, 2012.

[24] K. S. O'Neal, J. L. Johnson, and R. L. Panak, "Recognizing and appropriately treating latent autoimmune diabetes in adults," *Diabetes Spectrum*, vol. 29, no. 4, pp. 249–252, 2016.

[25] E. Ahlqvist, P. Storm, A. Käräjämäki, M. Martinell, M. Dorkhan, A. Carlsson, P. Vikman, R. B. Prasad, D. M. Aly, P. Almgren, *et al.*, "Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables," *The Lancet Diabetes & Endocrinology*, vol. 6, no. 5, pp. 361–369, 2018.

[26] M. Battaglia, S. Ahmed, M. S. Anderson, M. A. Atkinson, D. Becker, P. J. Bingley, E. Bosi, T. M. Brusko, L. A. DiMeglio, C. Evans-Molina, *et al.*, "Introducing the endotype concept to address the challenge of disease heterogeneity in type 1 diabetes," *Diabetes care*, vol. 43, no. 1, pp. 5–12, 2020.

[27] A. D. Association *et al.*, "Gestational diabetes mellitus," *Diabetes care*, vol. 26, no. suppl 1, pp. s103–s105, 2003.

[28] P. A. Hart and S. T. Chari, "Diabetes mellitus and pancreatic cancer: why the association matters?," *Pancreas*, vol. 42, no. 8, 2013.

[29] H. Inoue, Y. Tanizawa, J. Wasson, P. Behn, K. Kalidas, E. Bernal-Mizrachi, M. Mueckler, H. Marshall, H. Donis-Keller, P. Crock, *et al.*, "A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (wolfram syndrome)," *Nature genetics*, vol. 20, no. 2, pp. 143–148, 1998.

[30] N. Nagaraju and K. Rao, "Folk–medicine for diabetes from rayalaseema of andhra pradesh," *Ancient science of life*, vol. 9, no. 1, p. 31, 1989.

[31] A. Das and S. Shah, "History of diabetes: from ants to analogs.," *The Journal of the Association of Physicians of India*, vol. 59, pp. 6–7, 2011.

[32] R. Luft, "Oskar minkowski: discovery of the pancreatic origin of diabetes, 1889," *Diabetologia*, vol. 32, no. 7, pp. 399–401, 1989.

[33] M. J. Adams, T. L. Blundell, E. J. Dodson, G. G. Dodson, M. Vijayan, E. N. Baker, M. M. Harding, D. Hodgkin, B. Rimmer, and S. Sheat, "Structure of rhombohedral 2 zinc insulin crystals," *Nature*, vol. 224, no. 5218, p. 491, 1969.

[34] Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, "Predicting diabetes mellitus with machine learning techniques," *Frontiers in genetics*, vol. 9, 2018.

[35] P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, S. Colagiuri, L. Guariguata, A. A. Motala, K. Ogurtsova, *et al.*, "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the international diabetes federation diabetes atlas," *Diabetes research and clinical practice*, vol. 157, p. 107843, 2019.

[36] D. P. Group, "Incidence and trends of childhood type 1 diabetes worldwide 1990–1999," *Diabetic Medicine*, vol. 23, no. 8, pp. 857–866, 2006.

[37] F. Pérez-Bravo, E. Carrasco, J. L. Santos, M. Calvillán, G. Larenas, and C. Albala, "Prevalence of type 2 diabetes and obesity in rural mapuche population from chile," *Nutrition*, vol. 17, no. 3, pp. 236–238, 2001.

[38] L. O. Schulz, P. H. Bennett, E. Ravussin, J. R. Kidd, K. K. Kidd, J. Esparza, and M. E. Valencia, "Effects of traditional and western environments on prevalence of type 2 diabetes in pima indians in mexico and the us," *Diabetes Care*, vol. 29, no. 8, pp. 1866–1871, 2006.

[39] N. G. Forouhi and N. J. Wareham, "Epidemiology of diabetes," *Medicine*, vol. 38, no. 11, pp. 602–606, 2010.

[40] M. M. Kelsey, A. Zaepfel, P. Bjornstad, and K. J. Nadeau, "Age-related consequences of childhood obesity," *Gerontology*, vol. 60, no. 3, pp. 222–228, 2014.

[41] T. Reinehr, "Type 2 diabetes mellitus in children and adolescents," *World journal of diabetes*, vol. 4, no. 6, p. 270, 2013.

[42] J. Mann, M. Lean, M. Toeller, G. Slama, M. Uusitupa, B. Vessby, *et al.*, "Recommendations for the nutritional management of patients with diabetes mellitus.," *European Journal of Clinical Nutrition*, vol. 54, no. 4, pp. 353–355, 2000.

[43] K. M. Nelson, G. Reiber, and E. J. Boyko, "Diet and exercise among adults with type 2 diabetes findings from the third national health and nutrition examination survey (nhanes iii)," *Diabetes care*, vol. 25, no. 10, pp. 1722–1728, 2002.

[44] S. Polsky and S. L. Ellis, "Obesity, insulin resistance, and type 1 diabetes mellitus," *Current Opinion in Endocrinology, Diabetes and Obesity*, vol. 22, no. 4, pp. 277–282, 2015.

[45] E. Hyppönen, S. M. Virtanen, M. G. Kenward, M. Knip, H. K. Akerblom, C. D. in Finland Study Group, *et al.*, "Obesity, increased linear growth, and risk of type 1 diabetes in children.," *Diabetes care*, vol. 23, no. 12, pp. 1755–1760, 2000.

[46] A. Sekikawa, H. Eguchi, M. Tominaga, K. Igarashi, T. Abe, H. Manaka, H. Sasaki, H. Fukuyama, T. Kato, Y. Kiyohara, *et al.*, "Prevalence of type 2 diabetes mellitus and impaired glucose tolerance in a rural area of japan: the funagata diabetes study," *Journal of Diabetes and its Complications*, vol. 14, no. 2, pp. 78–83, 2000.

[47] M. T. Hamilton, D. G. Hamilton, and T. W. Zderic, "Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease," *Diabetes*, vol. 56, no. 11, pp. 2655–2667, 2007.

[48] D. M. Maahs, N. A. West, J. M. Lawrence, and E. J. Mayer-Davis, "Epidemiology of type 1 diabetes," *Endocrinology and metabolism clinics of North America*, vol. 39, no. 3, pp. 481–497, 2010.

[49] E. Patelarou, C. Girvalaki, H. Brokalaki, A. Patelarou, Z. Androulaki, and C. Vardavas, "Current evidence on the associations of breastfeeding, infant

formula, and cow's milk introduction with type 1 diabetes mellitus: a systematic review," *Nutrition reviews*, vol. 70, no. 9, pp. 509–519, 2012.

[50] F. Pociot and Å. Lernmark, "Genetic risk factors for type 1 diabetes," *The Lancet*, vol. 387, no. 10035, pp. 2331–2339, 2016.

[51] C. Fuchsberger, J. Flannick, T. M. Teslovich, A. Mahajan, V. Agarwala, K. J. Gaulton, C. Ma, P. Fontanillas, L. Moutsianas, D. J. McCarthy, *et al.*, "The genetic architecture of type 2 diabetes," *Nature*, vol. 536, no. 7614, p. 41, 2016.

[52] D. W. Foster and J. D. McGarry, "The metabolic derangements and treatment of diabetic ketoacidosis," *New England Journal of Medicine*, vol. 309, no. 3, pp. 159–169, 1983.

[53] H. M. Leibowitz, D. Krueger, L. Maunder, R. Milton, M. Kini, H. Kahn, R. Nickerson, J. Pool, T. Colton, J. Ganley, *et al.*, "The framingham eye study monograph: an ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy, macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975.," *Survey of ophthalmology*, vol. 24, no. Suppl, pp. 335–610, 1979.

[54] E. J. Lewis, L. G. Hunsicker, R. P. Bain, and R. D. Rohde, "The effect of angiotensin-converting-enzyme inhibition on diabetic nephropathy," *New England Journal of Medicine*, vol. 329, no. 20, pp. 1456–1462, 1993.

[55] P. J. Dyck, P. Thomas, A. Asbury, A. Winegrad, and D. Porte, "Diabetic neuropathy," *Diabetic neuropathy*, pp. 255–295, 1999.

[56] P. Libby and J. Plutzky, "Diabetic macrovascular disease," 2002.

[57] B. Musicki and A. Burnett, "Endothelial dysfunction in diabetic erectile dysfunction," *International journal of impotence research*, vol. 19, no. 2, pp. 129–138, 2007.

[58] "Technical document for the diabetes prevalence model for england 2016," tech. rep., Wellington House, London, Jan. 2016.

[59] N. Hex, C. Bartlett, D. Wright, M. Taylor, and D. Varley, "Estimating the current and future costs of type 1 and type 2 diabetes in the uk, including direct health costs and indirect societal and productivity costs," *Diabetic Medicine*, vol. 29, no. 7, pp. 855–862, 2012.

[60] A. D. Association *et al.*, "Economic costs of diabetes in the us in 2017," *Diabetes care*, vol. 41, no. 5, pp. 917–928, 2018.

[61] N. R. F. Collaboration *et al.*, "Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with $4 \cdot 4$ million participants," *The Lancet*, vol. 387, no. 10027, pp. 1513–1530, 2016.

[62] A. O. De Beeck and D. L. Eizirik, "Viral infections in type 1 diabetes mellitus—why the β cells?," *Nature Reviews Endocrinology*, vol. 12, no. 5, pp. 263–273, 2016.

[63] M. Clark, C. J. Kroger, and R. M. Tisch, "Type 1 diabetes: a chronic anti-self-inflammatory response," *Frontiers in immunology*, vol. 8, p. 1898, 2017.

[64] A. Willcox, S. J. Richardson, A. J. Bone, A. K. Foulis, and N. G. Morgan, "Analysis of islet inflammation in human type 1 diabetes," *Clinical & Experimental Immunology*, vol. 155, no. 2, pp. 173–181, 2009.

[65] K. T. Coppieters, F. Dotta, N. Amirian, P. D. Campbell, T. W. Kay, M. A. Atkinson, B. O. Roep, and M. G. von Herrath, "Demonstration of islet-autoreactive cd8 t cells in insulitic lesions from recent onset and long-term type 1 diabetes patients," *Journal of Experimental Medicine*, pp. jem–20111187, 2012.

[66] S. Arif, P. Leete, V. Nguyen, K. Marks, N. M. Nor, M. Estorninho, D. Kronenberg-Versteeg, P. J. Bingley, J. A. Todd, C. Guy, *et al.*, "Blood and islet phenotypes indicate immunological heterogeneity in type 1 diabetes," *Diabetes*, p. DB_140365, 2014.

[67] A. F. Marée, R. Kublik, D. T. Finegood, and L. Edelstein-Keshet, "Modelling the onset of type 1 diabetes: can impaired macrophage phagocytosis make

the difference between health and disease?," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 364, no. 1842, pp. 1267–1282, 2006.

[68] A. Han, E. W. Newell, J. Glanville, N. Fernandez-Becker, C. Khosla, Y.-h. Chien, and M. M. Davis, "Dietary gluten triggers concomitant activation of cd4+ and cd8+ $\alpha\beta$ t cells and $\gamma\delta$ t cells in celiac disease," *Proceedings of the National Academy of Sciences*, vol. 110, no. 32, pp. 13073–13078, 2013.

[69] S. Kahn, "The relative contributions of insulin resistance and beta-cell dysfunction to the pathophysiology of type 2 diabetes," *Diabetologia*, vol. 46, no. 1, pp. 3–19, 2003.

[70] D. Hopcroft, D. Mason, and R. Scott, "Insulin secretion from perifused rat pancreatic pseudoislets," *In vitro cellular & developmental biology*, vol. 21, no. 8, pp. 421–427, 1985.

[71] P. A. Halban, C. B. Wollheim, B. Blondel, P. Meda, E. N. Niesor, and D. H. Mintz, "The possible importance of contact between pancreatic islet cells for the control of insulin release," *Endocrinology*, vol. 111, no. 1, pp. 86–94, 1982.

[72] J. Ashburner and K. J. Friston, "Voxel-based morphometry—the methods," *Neuroimage*, vol. 11, no. 6, pp. 805–821, 2000.

[73] S. Bonner, I. Kureshi, J. Brennan, G. Theodoropoulos, A. S. McGough, and B. Obara, "Exploring the semantic content of unsupervised graph embeddings: An empirical study," *Data Science and Engineering*, vol. 4, no. 3, pp. 269–289, 2019.

[74] C. J. Nelson, P. T. Jackson, and B. Obara, "Combining mathematical morphology and the hilbert transform for fully automatic nuclei detection in fluorescence microscopy," in *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*, pp. 532–543, Springer, 2019.

[75] B. De Samber, M. Bensellam, S. J. Van Malderen, F. Seiboth, D. Brückner, J. Garrevoet, G. Falkenberg, J.-C. Jonas, and L. Vincze, "Proof-of-concept

for 2d/ct element analysis of entire cryofrozen islets of langerhans using a cryoloop synchrotron x-ray fluorescence setup," *Journal of Analytical Atomic Spectrometry*, 2020.

[76] M. P. Dybala, J. K. Butterfield, B. K. Hendren-Santiago, and M. Hara, "Pancreatic islets and gestalt principles," *Diabetes*, vol. 69, no. 9, pp. 1864–1874, 2020.

[77] M. P. Dybala and M. Hara, "Heterogeneity of the human pancreatic islet," *Diabetes*, vol. 68, no. 6, pp. 1230–1239, 2019.

[78] K. Xia and G.-W. Wei, "Persistent homology analysis of protein structure, flexibility, and folding," *International journal for numerical methods in biomedical engineering*, vol. 30, no. 8, pp. 814–844, 2014.

[79] M. Li, K. Duncan, C. N. Topp, and D. H. Chitwood, "Persistent homology and the branching topologies of plants," *American journal of botany*, vol. 104, no. 3, 2017.

[80] A. Stožestovzer2013functional, M. Gosak, J. Dolenšek, M. Perc, M. Marhl, M. S. Rupnik, and D. Korošak, "Functional connectivity in islets of langerhans from mouse pancreas tissue slices," *PLoS Comput Biol*, vol. 9, no. 2, p. e1002923, 2013.

[81] G. D. S. Xavier and G. A. Rutter, "Metabolic and functional heterogeneity in pancreatic β cells," *Journal of molecular biology*, vol. 432, no. 5, pp. 1395–1406, 2020.

[82] M. A. Atkinson, G. S. Eisenbarth, and A. W. Michels, "Type 1 diabetes," *The Lancet*, vol. 383, no. 9911, pp. 69–82, 2014.

[83] W. C. Knowler, E. Barrett-Connor, S. E. Fowler, R. F. Hamman, J. M. Lachin, E. A. Walker, and D. M. Nathan, "Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin.," *The New England journal of medicine*, vol. 346, no. 6, pp. 393–403, 2002.

[84] A. D. Sanlioglu, H. A. Altunbas, M. K. Balci, T. S. Griffith, and S. Sanlioglu, "Clinical utility of insulin and insulin analogs," *Islets*, vol. 5, no. 2, pp. 67–78, 2013.

[85] H. de Kort, E. J. de Koning, T. J. Rabelink, J. A. Bruijn, and I. M. Bajema, "Islet transplantation in type 1 diabetes," *BMJ*, vol. 342, p. d217, 2011.

[86] N. A. Sherry, J. A. Kushner, M. Glandt, T. Kitamura, A.-M. B. Brillantes, and K. C. Herold, "Effects of autoimmunity and immune therapy on β-cell turnover in type 1 diabetes," *Diabetes*, vol. 55, no. 12, pp. 3238–3245, 2006.

[87] P. Serup, O. D. Madsen, and T. Mandrup-Poulsen, "Islet and stem cell transplantation for treating diabetes," *Bmj*, vol. 322, no. 7277, pp. 29–32, 2001.

[88] C. M. Cohrs, C. Chen, S. R. Jahn, J. Stertmann, H. Chmelova, J. Weitz, A. Bähr, N. Klymiuk, A. Steffen, B. Ludwig, *et al.*, "Vessel network architecture of adult human islets promotes distinct cell-cell interactions in situ and is altered after transplantation," *Endocrinology*, vol. 158, no. 5, pp. 1373–1385, 2017.

[89] A. J. King, "The use of animal models in diabetes research," *British journal of pharmacology*, vol. 166, no. 3, pp. 877–894, 2012.

[90] J. Dolenšek, M. S. Rupnik, and A. Stožer, "Structural similarities and differences between the human and the mouse pancreas," *Islets*, vol. 7, no. 1, p. e1024405, 2015.

[91] M. Brissova, M. J. Fowler, W. E. Nicholson, A. Chu, B. Hirshberg, D. M. Harlan, and A. C. Powers, "Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy," *Journal of Histochemistry & Cytochemistry*, vol. 53, no. 9, pp. 1087–1097, 2005.

[92] H. T. T. Nhu, R. A. E. Drigo, P.-O. Berggren, and T. Boudier, "A novel toolbox to investigate tissue spatial organization applied to the study of the islets of langerhans," *Scientific reports*, vol. 7, p. 44261, 2017.

[93] K. Parvati, P. Rao, and M. Mariya Das, "Image segmentation using gray-scale morphology and marker-controlled watershed transformation," *Discrete Dynamics in Nature and Society*, vol. 2008, 2008.

[94] S. S. Roscioni, A. Migliorini, M. Gegg, and H. Lickert, "Impact of islet architecture on β-cell heterogeneity, plasticity and function," *Nature Reviews Endocrinology*, vol. 12, no. 12, p. 695, 2016.

[95] D. Bosco, M. Armanet, P. Morel, N. Niclauss, A. Sgroi, Y. D. Muller, L. Giovannoni, G. Parnaud, and T. Berney, "Unique arrangement of alpha-and beta-cells in human islets of langerhans," *Diabetes*, 2010.

[96] M. Von Herrath and G. T. Nepom, "Animal models of human type 1 diabetes," *Nature immunology*, vol. 10, no. 2, p. 129, 2009.

[97] G. Kilimnik, B. Zhao, J. Jo, V. Periwal, P. Witkowski, R. Misawa, and M. Hara, "Altered islet composition and disproportionate loss of large islets in patients with type 2 diabetes," *PloS one*, vol. 6, no. 11, p. e27445, 2011.

[98] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors, "scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 6 2014.

[99] T. E. Oliphant, *A guide to NumPy*, vol. 1. 2006.

[100] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," 2001–. [Online; accessed <today>].

[101] scikit image, "scikit-image/uft." `https://github.com/scikit-image/scikit-image/blob/master/skimage/restoration/uft.py`.

[102] scikit image, "scikit-image/gauss." `https://github.com/scikit-image/scikit-image/blob/master/skimage/filters/_gaussian.py`.

[103] scikit image, "scikit-image/thresholding." `https://github.com/scikit-image/scikit-image/blob/master/skimage/filters/thresholding.py`.

[104] G. Zack, W. Rogers, and S. Latt, "Automatic measurement of sister chromatid exchange frequency.," *Journal of Histochemistry & Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977.

[105] scikit image, "scikit-image/misc." `https://github.com/scikit-image/scikit-image/blob/master/skimage/morphology/misc.py`.

[106] J. L. Fernández, L. Muriel, V. Goyanes, E. Segrelles, J. Gosálvez, M. Enciso, M. LaFromboise, and C. De Jonge, "Simple determination of human sperm dna fragmentation with an improved sperm chromatin dispersion test," *Fertility and sterility*, vol. 84, no. 4, pp. 833–842, 2005.

[107] S. Spatial, "Scipy/spatial/plotutils," *GitHub*.

[108] J.-R. Sack and J. Urrutia, *Handbook of computational geometry*. Elsevier, 1999.

[109] SciPy, "Scipy-stats." `https://github.com/scipy/scipy/blob/v1.1.0/scipy/stats/stats.py#L1420-L1488`.

[110] SciPy, "Scipy-stats." `https://github.com/scipy/scipy/blob/v1.1.0/scipy/stats/stats.py#L4980-L5094`.

[111] SciPy, "Scipy-stats." `https://github.com/scipy/scipy/blob/v1.1.0/scipy/stats/stats.py#L4836-L4925`.

[112] R. G. Newcombe, "Confidence intervals for an effect size measure based on the mann–whitney statistic. part 2: Asymptotic methods and evaluation," *Statistics in Medicine*, vol. 25, no. 4, pp. 559–573, 2006.

[113] J. D. Trudeau, J. P. Dutz, E. Arany, D. J. Hill, W. E. Fieldus, and D. T. Finegood, "Neonatal beta-cell apoptosis: a trigger for autoimmune diabetes?," *Diabetes*, vol. 49, no. 1, pp. 1–7, 2000.

[114] G. C. Weir and S. Bonner-Weir, "Five stages of evolving beta-cell dysfunction during progression to diabetes," *Diabetes*, vol. 53, no. suppl 3, pp. S16–S21, 2004.

[115] M. Abdul-Rasoul, H. Habib, and M. Al-Khouly, "'the honeymoon phase'in children with type 1 diabetes mellitus: frequency, duration, and influential factors," *Pediatric diabetes*, vol. 7, no. 2, pp. 101–107, 2006.

[116] R. A. Rizza, P. E. Cryer, and J. E. Gerich, "Role of glucagon, catecholamines, and growth hormone in human glucose counterregulation: effects of somatostatin and combined α-and β-adrenergic blockade on plasma glucose recovery and glucose flux rates after insulin-induced hypoglycemia," *The Journal of clinical investigation*, vol. 64, no. 1, pp. 62–71, 1979.

[117] J. M. Sharif, M. Miswan, M. Ngadi, M. S. H. Salam, and M. M. bin Abdul Jamil, "Red blood cell segmentation using masking and watershed algorithm: A preliminary study," in *Biomedical Engineering (ICoBE), 2012 International Conference on*, pp. 258–262, IEEE, 2012.

[118] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, "Methods for nuclei detection, segmentation, and classification in digital histopathology: a review of current status and future potential," *IEEE reviews in biomedical engineering*, vol. 7, pp. 97–114, 2014.

[119] C. Demir and B. Yener, "Automated cancer diagnosis based on histopathological images: a systematic survey," *Rensselaer Polytechnic Institute, Tech. Rep*, 2005.

[120] T.-C. Yang, "Modifiable areal unit problem," *GIS Resource Document*, vol. 5, no. 65, p. 135, 2005.

[121] P. Seiron, A. Wiberg, L. Krogvold, F. L. Jahnsen, K. Dahl-Jorgensen, O. Skog, and O. Korsgren, "Characterization of the endocrine pancreas in type 1 diabetes: islet size is maintained but islet number is markedly reduced," *bioRxiv*, p. 480509, 2018.

[122] E. Reaven, G. Gold, W. Walker, and G. Reaven, "Effect of variations in islet size and shape on glucose-stimulated insulin secretion," *Hormone and Metabolic Research*, vol. 13, no. 12, pp. 673–674, 1981.

[123] P. Soille, *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.

[124] K. Benkrid, D. Crookes, and A. Benkrid, "Design and fpga implementation of a perimeter estimator," in *Proceedings of the Irish Machine Vision and Image Processing Conference*, pp. 51–57, 2000.

[125] T. L. Heath *et al.*, *The thirteen books of Euclid's Elements*. Courier Corporation, 1956.

[126] D. Baetens, F. Malaisse-Lagae, A. Perrelet, and L. Orci, "Endocrine pancreas: three-dimensional reconstruction shows two types of islets of langerhans," *Science*, vol. 206, no. 4424, pp. 1323–1325, 1979.

[127] D. Chicheportiche and G. Reach, "In vitro kinetics of insulin release by microencapsulated rat islets: effect of the size of the microcapsules," *Diabetologia*, vol. 31, no. 1, pp. 54–57, 1988.

[128] M. T. Adams, J. M. Gilbert, J. H. Paiz, F. M. Bowman, and B. Blum, "Endocrine cell type sorting and mature architecture in the islets of langerhans require expression of roundabout receptors in β cells," *Scientific reports*, vol. 8, no. 1, pp. 1–12, 2018.

[129] C. M. Topaz, L. Ziegelmeier, and T. Halverson, "Topological data analysis of biological aggregation models," *PloS one*, vol. 10, no. 5, p. e0126383, 2015.

[130] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, "Novel type of phase transition in a system of self-driven particles," *Physical review letters*, vol. 75, no. 6, p. 1226, 1995.

[131] N. Atienza, L. M. Escudero, M. J. Jimenez, and M. Soriano-Trigueros, "Characterising epithelial tissues using persistent entropy," in *International Workshop on Computational Topology in Image Context*, pp. 179–190, Springer, 2019.

[132] K. Garside, R. Henderson, I. Makarenko, and C. Masoller, "Topological data analysis of high resolution diabetic retinopathy images," *PloS one*, vol. 14, no. 5, p. e0217413, 2019.

[133] A. Hatcher, *Algebraic topology*. Cambridge University Press, 2002.

[134] R. Ghrist, "Barcodes: the persistent topology of data," *Bulletin of the American Mathematical Society*, vol. 45, no. 1, pp. 61–75, 2008.

[135] P. M. Adler and J.-F. Thovert, *Fractures and fracture networks*, vol. 15. Springer Science & Business Media, 1999.

[136] A. Hagberg, P. Swart, and D. S Chult, "Exploring network structure, dynamics, and function using networkx," tech. rep., Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[137] P. Meda, "Gap junction proteins are key drivers of endocrine function," *Biochimica et Biophysica Acta (BBA)-Biomembranes*, vol. 1860, no. 1, pp. 124–140, 2018.

[138] K. Tsaneva-Atanasova, C. L. Zimliki, R. Bertram, and A. Sherman, "Diffusion of calcium and metabolites in pancreatic islets: killing oscillations with a pitchfork," *Biophysical journal*, vol. 90, no. 10, pp. 3434–3446, 2006.

[139] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, pp. 35–41, 1977.

[140] M. G. Everett and S. P. Borgatti, "The centrality of groups and classes," *The Journal of mathematical sociology*, vol. 23, no. 3, pp. 181–201, 1999.

[141] S. Wasserman, K. Faust, *et al.*, *Social network analysis: Methods and applications*, vol. 8. Cambridge university press, 1994.

[142] B. Bollobás, *Modern graph theory*, vol. 184. Springer Science & Business Media, 2013.

[143] G. Drews, P. Krippeit-Drews, and M. Düfer, "Electrophysiology of islet cells," in *The islets of langerhans*, pp. 115–163, Springer, 2010.

[144] D. Korošak and M. Slak Rupnik, "Collective sensing of β-cells generates the metabolic code," *Frontiers in physiology*, vol. 9, p. 31, 2018.

[145] M. F. Brereton, E. Vergari, Q. Zhang, and A. Clark, "Alpha-, delta-and pp-cells: are they the architectural cornerstones of islet structure and co-ordination?," *Journal of Histochemistry & Cytochemistry*, vol. 63, no. 8, pp. 575–591, 2015.

[146] A. Kim, K. Miller, J. Jo, G. Kilimnik, P. Wojcik, and M. Hara, "Islet architecture: A comparative study," *Islets*, vol. 1, no. 2, pp. 129–136, 2009.

[147] M. A. Atkinson, M. Campbell-Thompson, I. Kusmartseva, and K. H. Kaestner, "Organisation of the human pancreas in health and in diabetes," *Diabetologia*, vol. 63, no. 10, pp. 1966–1973, 2020.

[148] R. A. e Drigo, Y. Ali, J. Diez, D. K. Srinivasan, P.-O. Berggren, and B. O. Boehm, "New insights into the architecture of the islet of langerhans: a focused cross-species assessment," *Diabetologia*, vol. 58, no. 10, pp. 2218–2228, 2015.

[149] C. Jopling, S. Boue, and J. C. I. Belmonte, "Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration," *Nature reviews Molecular cell biology*, vol. 12, no. 2, pp. 79–89, 2011.

[150] P. Leete, A. Willcox, L. Krogvold, K. Dahl-Jørgensen, A. K. Foulis, S. J. Richardson, and N. G. Morgan, "Differential insulitic profiles determine the extent of β-cell destruction and the age at onset of type 1 diabetes," *Diabetes*, vol. 65, no. 5, pp. 1362–1369, 2016.

[151] N. G. Morgan and S. J. Richardson, "Fifty years of pancreatic islet pathology in human type 1 diabetes: insights gained and progress made," *Diabetologia*, vol. 61, no. 12, pp. 2499–2506, 2018.

[152] P. In't Veld, "Insulitis in type 1 diabetes: a sticky problem," *Diabetes*, vol. 58, no. 6, pp. 1257–1258, 2009.

[153] Z. Wang, N. W. York, C. G. Nichols, and M. S. Remedi, "Pancreatic β cell dedifferentiation in diabetes and redifferentiation following insulin therapy," *Cell metabolism*, vol. 19, no. 5, pp. 872–882, 2014.

[154] M. Oshima, K.-P. Knoch, M. Diedisheim, A. Petzold, P. Cattan, M. Bugliani, P. Marchetti, P. Choudhary, G.-C. Huang, S. R. Bornstein, *et al.*, "Virus-like infection induces human β cell dedifferentiation," *JCI insight*, vol. 3, no. 3, 2018.

[155] C. Talchai, S. Xuan, H. V. Lin, L. Sussel, and D. Accili, "Pancreatic β cell dedifferentiation as a mechanism of diabetic β cell failure," *Cell*, vol. 150, no. 6, pp. 1223–1234, 2012.

[156] D. Jain, R. Jain, D. Eberhard, J. Eglinger, M. Bugliani, L. Piemonti, P. Marchetti, and E. Lammert, "Age-and diet-dependent requirement of dj-1 for glucose homeostasis in mice with implications for human type 2 diabetes," *Journal of molecular cell biology*, vol. 4, no. 4, pp. 221–230, 2012.

[157] H. Mizukami, K. Takahashi, W. Inaba, S. Osonoi, K. Kamata, K. Tsuboi, and S. Yagihashi, "Age-associated changes of islet endocrine cells and the effects of body mass index in j apanese," *Journal of diabetes investigation*, vol. 5, no. 1, pp. 38–47, 2014.

[158] Y. Matsuda, "Age-related morphological changes in the pancreas and their association with pancreatic carcinogenesis," *Pathology International*, vol. 69, no. 8, pp. 450–462, 2019.

[159] H. Hani, Z. N. Allaudin, T. A. T. Ibrahim, M.-A. Mohd-Lila, K. Sarsaifi, S. N. Camalxaman, and A. M. Othman, "Morphological changes of post-isolation of caprine pancreatic islet," *In Vitro Cellular & Developmental Biology-Animal*, vol. 51, no. 2, pp. 113–120, 2015.

[160] A. Höög, A. Sandberg-Nordqvist, S. M. Abdel-Halim, C. Carlsson-Skwirut, A. Guenifi, M. Tally, C.-G. Ostenson, S. Falkmer, V. R. Sara, S. Efendić, *et al.*, "Increased amounts of a high molecular weight insulin-like growth factor

ii (igf-ii) peptide and igf-ii messenger ribonucleic acid in pancreatic islets of diabetic goto-kakizaki rats," *Endocrinology*, vol. 137, no. 6, pp. 2415–2423, 1996.

[161] J. S. Canzano, L. H. Nasif, E. A. Butterworth, D. A. Fu, M. A. Atkinson, and M. Campbell-Thompson, "Islet microvasculature alterations with loss of beta-cells in patients with type 1 diabetes," *Journal of Histochemistry & Cytochemistry*, vol. 67, no. 1, pp. 41–52, 2019.

[162] R. B. Reinert, M. Brissova, A. Shostak, F. C. Pan, G. Poffenberger, Q. Cai, G. L. Hundemer, J. Kantz, C. S. Thompson, C. Dai, *et al.*, "Vascular endothelial growth factor-a and islet vascularization are necessary in developing, but not adult, pancreatic islets," *Diabetes*, vol. 62, no. 12, pp. 4154–4164, 2013.

[163] M. Brissova, A. Shostak, C. L. Fligner, F. L. Revetta, M. K. Washington, A. C. Powers, and R. L. Hull, "Human islets have fewer blood vessels than mouse islets and the density of islet vascular structures is increased in type 2 diabetes," *Journal of Histochemistry & Cytochemistry*, vol. 63, no. 8, pp. 637–645, 2015.

[164] D.-S. Koh, J.-H. Cho, and L. Chen, "Paracrine interactions within islets of langerhans," *Journal of Molecular Neuroscience*, vol. 48, no. 2, pp. 429–440, 2012.

[165] N. R. Johnston, R. K. Mitchell, E. Haythorne, M. P. Pessoa, F. Semplici, J. Ferrer, L. Piemonti, P. Marchetti, M. Bugliani, D. Bosco, *et al.*, "Beta cell hubs dictate pancreatic islet responses to glucose," *Cell metabolism*, vol. 24, no. 3, pp. 389–401, 2016.

[166] L. Briant, T. Reinbothe, I. Spiliotis, C. Miranda, B. Rodriguez, and P. Rorsman, "δ-cells and β-cells are electrically coupled and regulate α-cell activity via somatostatin," *The Journal of physiology*, vol. 596, no. 2, pp. 197–215, 2018.

[167] A. Foulis and J. Stewart, "The pancreas in recent-onset type 1 (insulin-dependent) diabetes mellitus: insulin content of islets, insulitis and associated

changes in the exocrine acinar tissue," *Diabetologia*, vol. 26, no. 6, pp. 456–461, 1984.

[168] M. Campbell-Thompson, M. Atkinson, A. Butler, N. Chapman, G. Frisk, R. Gianani, B. Giepmans, M. Von Herrath, H. Hyöty, T. Kay, *et al.*, "The diagnosis of insulitis in human type 1 diabetes," *Diabetologia*, vol. 56, no. 11, pp. 2541–2543, 2013.

[169] J. Meier, A. Bhushan, A. Butler, R. Rizza, and P. Butler, "Sustained beta cell apoptosis in patients with long-standing type 1 diabetes: indirect evidence for islet regeneration?," *Diabetologia*, vol. 48, no. 11, pp. 2221–2228, 2005.

[170] A. Xin, H. Mizukami, W. Inaba, T. Yoshida, Y.-k. Takeuchi, and S. Yagihashi, "Pancreas atrophy and islet amyloid deposition in patients with elderly-onset type 2 diabetes," *The Journal of Clinical Endocrinology & Metabolism*, vol. 102, no. 9, pp. 3162–3171, 2017.

[171] Y. Katsuda, T. Ohta, K. Miyajima, Y. Kemmochi, T. Sasase, B. Tong, M. Shinohara, and T. Yamada, "Diabetic complications in obese type 2 diabetic rat models," *Experimental Animals*, vol. 63, no. 2, pp. 121–132, 2014.

[172] C. Laing and G. J. Lord, *Stochastic methods in neuroscience*. Oxford University Press, 2010.

[173] H. Peiris, C. S. Bonder, P. T. H. Coates, D. J. Keating, and C. F. Jessup, "The β-cell/ec axis: how do islet cells talk to each other?," *Diabetes*, vol. 63, no. 1, pp. 3–11, 2014.

[174] C. Ionescu-Tirgoviste, P. A. Gagniuc, E. Gubceac, L. Mardare, I. Popescu, S. Dima, and M. Militaru, "A 3d map of the islet routes throughout the healthy human pancreas," *Scientific reports*, vol. 5, no. 1, pp. 1–14, 2015.

[175] E. Papadopoulou and D.-T. Lee, "A new approach for the geodesic voronoi diagram of points in a simple polygon and other restricted polygonal domains," *Algorithmica*, vol. 20, no. 4, pp. 319–352, 1998.

[176] J. M. Osborne, A. G. Fletcher, J. M. Pitt-Francis, P. K. Maini, and D. J. Gavaghan, "Comparing individual-based approaches to modelling the self-organization of multicellular tissues," *PLoS computational biology*, vol. 13, no. 2, p. e1005387, 2017.

[177] P. Pathmanathan, J. Cooper, A. Fletcher, G. Mirams, P. Murray, J. Osborne, J. Pitt-Francis, A. Walter, and S. Chapman, "A computational study of discrete mechanical tissue models," *Physical biology*, vol. 6, no. 3, p. 036001, 2009.

[178] A. G. Fletcher, J. M. Osborne, P. K. Maini, and D. J. Gavaghan, "Implementing vertex dynamics models of cell populations in biology within a consistent computational framework," *Progress in biophysics and molecular biology*, vol. 113, no. 2, pp. 299–326, 2013.

[179] J. Nava-Sedeño, H. Hatzikirou, F. Peruani, and A. Deutsch, "Extracting cellular automaton rules from physical langevin equation models for single and collective cell migration," *Journal of mathematical biology*, vol. 75, no. 5, pp. 1075–1100, 2017.

[180] M. C. Ozturk, Q. Xu, and A. Cinar, "Agent-based modeling of the interaction between cd8+ t cells and beta cells in type 1 diabetes," *PloS one*, vol. 13, no. 1, p. e0190349, 2018.

[181] K. C. Wedgwood, S. J. Richardson, N. G. Morgan, and K. Tsaneva-Atanasova, "Spatiotemporal dynamics of insulitis in human type 1 diabetes," *Frontiers in Physiology*, vol. 7, 2016.

[182] M. Scianna and L. Preziosi, "Multiscale developments of the cellular potts model," *Multiscale Modeling & Simulation*, vol. 10, no. 2, pp. 342–382, 2012.

[183] N. Chen, J. A. Glazier, J. A. Izaguirre, and M. S. Alber, "A parallel implementation of the cellular potts model for simulation of cell-based morphogenesis," *Computer physics communications*, vol. 176, no. 11-12, pp. 670–681, 2007.

[184] U. Christen and M. G. von Herrath, "Ip-10 and type 1 diabetes: a question of time and location," *Autoimmunity*, vol. 37, no. 4, pp. 273–282, 2004.

[185] M. Scianna, L. Preziosi, and K. Wolf, "A cellular potts model simulating cell migration on and in matrix environments," *Mathematical Biosciences & Engineering*, vol. 10, no. 1, pp. 235–261, 2013.

[186] M. Kragl and E. Lammert, "Basement membrane in pancreatic islet function," in *The Islets of Langerhans*, pp. 217–234, Springer, 2010.

[187] T. Otonkoski, M. Banerjee, O. Korsgren, L.-E. Thornell, and I. Virtanen, "Unique basement membrane structure of human pancreatic islets: implications for β-cell growth and differentiation," *Diabetes, obesity and metabolism*, vol. 10, pp. 119–127, 2008.

[188] G. R. Mirams, C. J. Arthurs, M. O. Bernabeu, R. Bordas, J. Cooper, A. Corrias, Y. Davit, S.-J. Dunn, A. G. Fletcher, D. G. Harvey, *et al.*, "Chaste: an open source c++ library for computational physiology and biology," *PLoS computational biology*, vol. 9, no. 3, p. e1002970, 2013.