Mathematical models of health focusing on diabetes: Delay differential equations and data mining

Jonathan Francis Easton

PhD

2014

Mathematical models of health focusing on diabetes: Delay differential equations and data mining

Jonathan Francis Easton

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of Doctor of Philosophy

Research carried out in the Faculty of Engineering and Environment

October 2014

Dedication

This PhD thesis is dedicated to my Gran for her endless love, support and guidance throughout my life.

Without the unconditional love and support you gave me, I would not be in the position I am now. You always taught me to give my best and were always bursting with pride in everything I did. I cannot tell you enough how lucky I am to have had you in my life. The bond we shared is something that will last forever and I know you are always with me, guiding me in everything I do.

I love you Gran

Abstract

Mathematical models have been applied to biology and health to gain a better understanding of physiological systems and disease, as well as to improve levels of treatment and care for certain conditions. This thesis will focus on two different methodologies to investigate models of health, namely delay differential equations and Bayesian based data mining.

The first approach uses delay differential equations to model the glucose-insulin regulation system. Many models exist in this area, typically including four exponential functions, and take a number of different forms. The model used here is a system of two delay differential equations with two time delays. The one delay form of this model has previously been widely studied, but less is known about the two delay system from an analytical view point.

This work improves upon the existing models by incorporating Hill functions instead of exponential functions. The new model presented is studied for its appropriateness and robustness to changing parameters such as glucose infusion rate and insulin degradation. A local and global stability of the two-delay system is presented both in general terms and explicitly using Lyapunov functionals and linear matrix inequalities.

The second method employs data mining techniques including a robust and transparent naïve Bayes classifier for classification and prediction of aspects of health. A study into prediction of post-stroke mortality is made on a data set of stroke patients. Interesting results are obtained for the classification of naturally arising mortality periods and an investigation into the role of age as a risk factor for post-stroke mortality. A wide range of risk factors are then investigated for significance which are used to build new predictive models.

These two approaches have the joint aim of improving the understanding of aspects of health through mathematical modelling techniques. A new model of the glucose-insulin regulatory system is developed and for the first time an analysis of the global stability of the two-delay model by use of a Lyapunov functional is provided. The second approach sees typical and robust data mining techniques used to analyse medical data. New models for stroke mortality and prediction of diabetes and obesity are created, which review risk factors and also illustrate the benefit of data mining techniques for analysing medical data.

Declaration

I, Jonathan Francis Easton, declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the University Ethics Committee on 06/01/2012.

I declare that the word count of this thesis is 41205 words.

Name: Jonathan Francis Easton

Signature:

Date:

Acknowledgements

I must first thank my Mum and Dad for everything they have done and continue to do for me. You have always strived to give me all I need to grow, learn and develop skills which have turned me into the man I am today. My entire life I have felt loved, supported and secure, knowing that in any situation you will be there for me and I will never be able to repay the love and faith you have in me. Together, you have given me all that you possibly could and you have taught me to always give of my best, in whatever I do. Your pride and belief in me in all that I have done is boundless, and I only hope I can continue to make you proud.

I would like to thank my family, Auntie Audrey and Uncle Charlie for your love and support. You have always been a big part of my life and I know you are both proud of me. To Dave, you are the closest thing I have to a brother and I appreciate our relationship. I thank you and Emma for your offers of support (and meals) this past year. I also include in this section family friends, people who have always had a part in my life and continue to ask after my welfare and show interest even though I do not see them very often.

I really must stress how grateful I am to my principal supervisor, Professor Maia Angelova. Your guidance and faith coupled with the opportunities and responsibilities you have given me during the last three years has brought me to this point now, it has been a wonderful experience thank you for your faith in me. I also must thank Dr Benoit Huard, who has been an absolute pleasure to work with during the last two years, your knowledge and willingness to help in all aspects has been fantastic.

Through the MATSIQEL project I have met many people, all of whom have been a pleasure to share experiences and work with in both Mexico and South Africa. I give a huge amount of thanks to Professor Christopher Stephens of the UNAM, Mexico. Following a brief meeting while on placement, you introduced me to and taught me an entirely new branch of mathematics which has become a large part of my thesis. Your enthusiasm for research and the desire to learn is something which I hope to take myself. Thank you for believing both in me and in our work. The time spent in Mexico and the people I have met, who have become a very important part of my life, made me feel so welcome that I now consider Mexico City a home away from home.

To my fellow PhD researchers of F7, the community we have in the office has made for a great experience. The sporting activities such as the football, badminton and gym will be remembered fondly, less so the ridiculous temperatures. I wish you all good luck in your research and for the future. I also thank my friends who have provided a distraction to PhD life and who have been there when needed.

Jonathan Easton

Thesis Conventions

The following conventions have been adopted in this thesis:

- **Spelling.** British English spelling conventions were used throughout this thesis, as defined in the Oxford English Dictionary.

- **Typesetting.** This document was written and compiled using LaTeX.

- **Mathematics.**
  - MATLAB codes were written using MATLAB 2010a/2013a.
    URL: http://www.mathworks.com
  - Calculations and graphs were produced using Mathematica 8.0. Wolfram Research, Inc., 2010.
    URL: http://www.wolfram.com/mathematica/
  - WEKA data mining software was used where stated using WEKA 3.6.11.
    URL: http://www.cs.waikato.ac.nz/ml/weka

- **Referencing.** The Harvard style has been used for referencing.

Publications

Elements of the research presented in this thesis have also been published in the following volumes:

- Huard, B., Easton, J.F., & Angelova, M. (2015) 'Investigation of stability in a two-delay model of the ultradian oscillations in glucose-insulin regulation.' Commun. Nonlinear Sci. Numer. Simulat. 26: 211-222.
  DOI: 10.1016/j.cnsns.2015.02.017

- Easton, J.F., Stephens, C.R. & Angelova, M. (2014) 'Risk factors and prediction of very short term versus short/intermediate term post-stroke mortality: A data mining approach.' Comput. Biol. Med. 54:199-210.
  DOI:10.1016/j.compbiomed.2014.09.003

- Angelova, M., Holloway, P.M., Gibson, H., Ben-Halim, A., Easton, J.F., St Clair Gibson, A., Cook, G., Rauch, L., Lambert, E. & Klein, B. (2012) *'MATSIQEL: Models for Ageing and Technological Solutions for Improving and Enhancing the Quality of Life.'* In Conf. Proc. Marie Curie People 2012 Conference, pp 92-100.

Presentations

- Easton, J.F. (2014, January 17th), Investigating nutrition in Mexico in relation to diabetes and obesity. MATSIQEL: Final Meeting, Northumbria University, UK.

- Easton, J.F. (2013, September 10th), Investigating mathematical models of health: a data mining approach to risk factors and prediction of stroke mortality. Colloquium at Department of Mathematical Sciences, Stellenbosch University, SA. URL: http://math.sun.ac.za/wp-content/uploads/2011/01/COLLEAST.pdf

- Easton, J.F. (2013, September 9th), Two-delay model of the ultradian oscillations in glucose-insulin regulation. Department of Mathematics and Applied Mathematics, University of Cape Town, SA.

- Easton, J.F. (2013, September 4th), Investigating mathematical models of health: a data mining approach to risk factors and prediction of stroke mortality. MATSIQEL Reporting Conference: Insights from Low and/or Middle Income Countries. Department of Human Biology, University of Cape Town, SA.

- Easton, J.F. (2013, February 20th), Investigating mathematical models of diabetes. CEIS research seminar. Northumbria University, UK.

- Easton, J.F. (2013, January 18th), Investigating mathematical models of health: Data mining and diabetes. MATSIQEL: End of year meeting. Sofia University, Bulgaria.
- Easton, J.F. (2012, September 26th), Investigating Mathematical Models of Diabetes. MML research seminar. Northumbria University, UK.

Posters

- Huard, B., Easton, J.F. & Angelova, M. (2013) Local and Global stability of a two-delay model of the ultradian oscillations in glucose-insulin regulation. BioDynamics 2013, Bristol, UK. Dates viewed: 11-13th September.
- Angelova, M., Holloway, P.M., Gibson, H., Ben-Halim, A., Easton, J.F., St Clair Gibson, A., Cook, G., Rauch, L., Lambert, E. & Klein, B. (2012) MATSIQEL: Models for Ageing and Technological Solutions for Improving and Enhancing the Quality of Life. Marie Curie People 2012 Conference. Date viewed: 6th November.
- Easton, J.F., Angelova, M. & Huard, B. (2012) Investigating mathematical models of diabetes. IoP conference: Physics in Healthcare. Date viewed: 31st October.
- Easton, J.F. (2012) Mathematical models of diabetes using delayed differential equations. LMS conference: Mathematics in Human Biology. Dates viewed: 6-8th June.

Funding

Travel for work and conferences as part of this thesis has been supported by the following grants:

- The MATSIQEL project, European FP7 research grant: FP7-PEOPLE-2009-IRSES-247541.
- The UNAM, Mexico City - DGAPA grant: IN 113414.
- Northumbria University PGR conference bursary scheme.

Contents

List of Figures

List of Tables

List of Abbreviations

AUC	Area Under the curve
BMI	Body Mass index
CFS	Correlation-based feature selection
DDE	Delay differential equation
ESS	European stroke scale
FN	False negatives
FP	False positives
GIST-UK	UK Glucose Insulin in Stroke Trial
GKI	Glucose, Potassium and Insulin
IVGTT	Intra-venous glucose tolerance test
LACS	Lacunar syndrome
LMI	Linear Matrix inequalities
PACS	Partial anterior circulation syndrome
POCS	Posterior circulation syndrome
ROC	Receiver operating characteristic
TACS	Total anterior circulation stroke
TN	True negatives
TP	True positives

Chapter 1

Introduction

"Few will have the greatness to bend history itself; but each of us can work to change a small portion of events, and in the total of all those acts will be written the history of this generation"

— Robert F. Kennedy (1925 - 1968)

1.1 Introduction

The application of mathematical methodology in relation to life sciences such as biology, physiology and medical studies is extremely beneficial, however there are many challenges faced. Before research begins, challenges in interdisciplinary work include a necessity to have a good knowledge base in two or more fields while also having the ability to convey results in the correct language to connect with research areas. For example, a mathematician, statistician, biologist and medical professional can each present one result in different ways.

This thesis will focus on mathematical models of health. Any aspect of health is a very complex system, which provides a large challenge in accurate mathematical modelling. The definition of complexity in this sense is given as having many different levels and different aspects. Any living system or disease is complex and can be studied via a number of different subsystems. It is possible to study a disease on a microscopic level in order to give understanding of chemical interactions and functionality for causes and effects within molecular subsystems. Alternatively, a disease can be studied on a macroscopic level to discover associations and causality of very different factors, for example, social, economic and behavioural influences. Due to this very complex nature, it is not possible to fully study a living system in one combined model, but, mathematically modelling different aspects on different levels helps to build a better knowledge and understanding. The area of multiscale modelling has begun to address this issue, applied to an increasing range of situations, for example see Leszczynski and Shukla (2009).

Progress in mathematical applications to biology have led to an increased level of interest in the area, creating great advances in the study of health and diseases. This thesis will present a study of two different mathematical modelling techniques applied to health and disease with a focus on diabetes, in order to highlight their complexity and the need for

different mathematical approaches on multiple scales.

1.1.1 Diabetes

The main health related focus of this thesis is the disease diabetes. Officially known as diabetes mellitus, it is a chronic disease which means it cannot be cured but only controlled through medication and treatment. Other examples of chronic diseases include asthma, Alzheimer's, heart disease and allergies. Further information can be found from groups such as The Center for Chronic Diseases (CMCD 2014).

Diabetes is a disorder of metabolism, causing problems in the way our food is broken down and used in the body. When we eat, almost all food is broken down into some element of glucose which comes from digesting carbohydrates commonly found in starchy and sweet foods. Glucose is a type of sugar which is primarily used by the cells in the body as fuel in order to function properly. A small amount of glucose is produced naturally in the liver; however, external glucose is the body's main source of energy. Once food has been broken down and glucose is present in the blood it requires insulin, which is a hormone secreted by the pancreas that allows the glucose to be used by the cells as fuel or stored in the liver and muscles as glucagon. However, in diabetic people there is either little or no insulin secreted by the pancreas or the cells do not respond to the presence of insulin, often referred to as insulin resistance. Type one diabetes is a genetic defect typically developed at a young age in which the insulin secreting cells of the pancreas are attacked by the immune system stopping production of insulin. The much more common type two diabetes accounts for approximately 90% of all diabetes sufferers in the United Kingdom and is a condition developed later in life where the body either stops producing insulin or insulin resistance is established (NHS 2014). There is also a third type known as gestational diabetes, which is developed during pregnancy and can occur with no prior

diabetic history or symptoms.

In terms of facts and figures, diabetes is not only a very serious and prominent problem worldwide, but it is also a growing problem. According to recent figures obtained from the UK based charity Diabetes UK released in February 2014, the incidence of diabetes across the UK is currently 3.2 million people, approximately 6% of the population for those actually diagnosed. Then there is an estimated further 630,000 people as yet unaware of their condition (Diabetes-UK 2014). By comparing this with the corresponding figures from the year before, it is clear to see that the incidence of diabetes is a growing problem (Diabetes-UK 2013a). The previous year's figures, released in April 2013 gave a total of 2.9 million people in the UK, an increase of approximately 300,000 diagnosed cases of diabetes in an approximate 12 month period. Further to this, it is expected that by the year 2025 there will be 5 million diagnosed cases of diabetes in the UK alone. This is a worldwide problem not just confined to the UK. A 2011 Lancet report gave figures that there are 347 million diabetic sufferers worldwide (Danaei et al 2011) and in addition, the World Health Organisation (WHO) predicts that diabetes will be the 7th highest leading cause of death in the world by the year 2030 (Mathers and Loncar 2006).

There are many different factors that can cause diabetes or increase the risk of the disease. Diabetes can be developed on a microscopic and genetic level, through interactions such as that of glucose and insulin in the metabolic system and up to macroscopic factors such as social and lifestyle elements. All of which contribute to the disease itself (Association et al 2014). The number one causal factor for type two diabetes is obesity (Chan et al 1994; Resnick et al 2000). Obesity is typically classified as having a body mass index (BMI) greater than 30. However, people with a lower BMI but who have a large percentage of body fat, particularly centred around the waist, are also highly at risk. By extension, the risk factors for obesity such as lack of exercise and unhealthy eating are also long term risk factors of diabetes due to the strong link between the two. Alongside

this, other causal factors of diabetes include a history of high blood pressure, heart attack or stroke. There is also evidence for mental illnesses controlled by medication, such as schizophrenia, to be related to diabetes (Dixon et al 2000; Sernyak et al 2002). Furthermore, if a close family member has had diabetes then risk is increased (Diabetes-UK 2013b).

Information from Diabetes UK shows that diabetes itself is a risk factor for many other diseases and conditions (Diabetes-UK 2014). An increased risk of cardiovascular diseases, including heart disease, stroke and angina amongst others can be related to diabetes. Due to an excess of glucose in the blood being passed through the kidneys to leave the body, there is an increased risk of kidney failure and related diseases. Kidney disease accounts for nearly 16% of all deaths in diabetics (Morrish et al 2001). Eye sight can be affected by diabetes and retinopathy is very common in long term diabetics, so much so that in the UK, diabetes is the leading cause of preventable loss of sight (Bunce and Wormald 2006). Other such effects include blood circulation issues leading to amputation and nerve damage as well as risk of mental illnesses such as depression.

Diabetes is a disease that is linked to many other diseases which can occur at the same time either dependently or independently, known as co-morbidity. It is clear that many conditions, such as stroke, are associated to cause (Diabetes-UK 2013b) and effect (Barrett-Connor and Khaw 1988) of diabetes. The significant complexity of diabetes and its relation with many causal factors and its presence as a risk factor for many other conditions, make it an extremely complex problem to study and treat.

The importance of studying diabetes to understand all elements of the disease will not only lead to better treatment and interventions to improve quality of life, but will also aid governments and hospitals in terms of health care costs. There are many studies from a wide range of different countries investigating the financial cost of healthcare related to

diabetes and associated factors. One particular study in 2010 suggested that global health care spending on diabetes stood at 376 billion USD. This accounts for 12% of global healthcare expenditure. It is also estimated that by the year 2030 the figure will be 490 billion USD (Zhang et al 2010). One specific and recent example of this comes from a BBC news story in July 2014 (Gallagher 2014) reporting that the UK is spending 10% of its health budget on diabetes and one extreme solution may be weight loss surgery for anybody with a body mass index in the very obese range, greater than 35. This would cost between £3,000 and £15,000 per operation with an estimated 460,000 people already meeting the criteria for assessment for surgery.

All of the health and financial reasons discussed make diabetes an extremely pressing matter in need of continued serious attention. The aforementioned complexity of the disease makes it very difficult to study and treat and as such there are many different investigations for different aspects of diabetes; however, there is much work to be done on a plethora of levels.

1.1.2 Approaches to mathematical modelling of health

A search of the literature for mathematical models related to diabetes produces many different results. These range from microscopic subsystems involving the affect of micro-receptor on β-cell secretion for release of insulin by the pancreas (Nesca et al 2013), up to very large scales such as economic models assessing health costs (Bagust et al 2001). Examples of a more physiological approach can be seen in models based on feedback loops and regulation of hormones as with glucose-insulin regulation (Bennett and Gourley 2004b; Li and Kuang 2007). The treatment of diabetes is modelled in a number of different areas with examples such as active insulin control (Lam et al 2002), and comparisons of techniques using statistical marginal structure modelling (Neugebauer et al

2012). Many models also exist to describe links to other factors such as stroke (Kothari et al 2002), or interventions such as exercise alongside diabetes and heart disease (Sacre et al 2014). These are very few examples of the mathematical modelling that exists to aid the understanding and care of diabetes and its related conditions.

1.2 Aims and motivation

This thesis will take the two approaches listed below in order to create very different models which investigate aspects of health on two different scales.

- Delay differential equations techniques used to model behaviour of glucose-insulin regulation within the metabolic system.
- Application of data mining techniques to a macroscopic medical data in order to study risk factors and build predictive models.

Further details of these two approaches and their application in this thesis are outlined below.

1.2.1 Methods of study

The method of study and type of techniques to be used in modelling depends largely on the initial hypothesis and the desired outcome of the investigation. Here, the focus will be on two particular styles of mathematical modelling used to study elements of health. When causality, the relationship between two or more events, is known then a small set of input variables can be used to give a small set of output variables. To model this a reductionist approach can be taken by modelling a subsystem and utilising

the factors that directly affect it. A technique involving a set of differential equations allows for greater understanding and gives a wide range of modelling capabilities for any system with continuously changing variables through time. Behavioural elements of a system based on certain factors through time as well as functionality of a system can be determined through use of differential equations. An example of application in diabetes modelling is in the glucose-regulation system which models the interaction between two substances whose quantities vary with time.

Where the causality is not known for a particular condition, it is desirable to study a large number of input variables to cover many possibilities in order to gain an understanding. A useful technique for this comes from a more statistical background in the area of data mining where many variables can be studied in one model for their relation to a specific condition. This type of scenario often occurs on a more macroscopic level, for example to investigate behavioral and social factors alongside medical aspects related to a disease. In these scenarios it is often difficult to give an initial hypothesis as there are many connections to be determined. Statistically, this approach is useful in shaping results related to risk factors and probabilistic models. In relation to health, a data mining based model can be created as a risk analysis to determine patients who are more likely to develop a given condition based on a set of wide ranging and unrelated variables. There is also the possibility to investigate one specific group of variables such as past medical histories.

1.2.2 Motivation

First, a system of delay differential equations will be used to model the interaction between glucose and insulin in the glucose-insulin regulation system. The system is known to have two time delays and the study of the oscillatory nature of this system with both one and two delays has been prominent in the field since the 1990s. Although the analysis of

the system with one time delay is extensive, the two delay system is much less explored, usually with a purely numerical approach. For the first time, a stability analysis both locally and globally of the two delay system of glucose and insulin regulation is provided. The general linearisation of the system will be used analytically for the local case with a linear matrix inequality (LMI) approach taken to simulate this numerically. A Lyapunov functional approach will be used to analytically give conditions for global stability, while providing numerical simulations to show the relation to model parameters. New aspects of the model are created by the introduction of Hill functions. This incorporates physiologically driven parameters into the model, given the original biological nature of the Hill function, as well as removing the need for auxiliary parameters with no physiological meaning. This gives an improved and clearer set of functions. Models for glucose-insulin regulation can be used for the advances in the treatment of diabetes through an artificial pancreas (Steil and Reifman 2009). However, the current models attempting to be applied to an artificial pancreas are restricted in design, such as the linear model with no delays presented by Huang et al (2012). The motivation is that a greater understanding of the glucose-insulin regulatory system with two delays will lead to a better application to the treatment of diabetes, as with the artificial pancreas.

Secondly, data mining techniques involving a naïve Bayes classifier will be used in order to produce classifying and predictive models based on risk factors and associations for aspects of health. There is a large increase in the amount of medical data being produced and stored which is often of very large volume and high dimension. Medical data is typically used to study one particular aspect based on a specific hypothesis, meaning that a number of potential insights could be being missed. The research area of data mining and knowledge discovery has been developed to investigate large data of many types, with a growing application to medical data. One aim of the thesis is to show how data mining techniques can be used more widely on the wealth of medical data being accumulated,

allowing for useful conclusions to be drawn. A data set of stroke patients will be studied to build new models for assessment of risk for post-stroke mortality based on rapidly accessible information. These models raise interesting points for risk factors and could be used to assess risk of mortality for different time periods to assist in the treatment of stroke patients in hospitals.

Individually, the aim of these two approaches related to health conditions and treatment is to create a useful investigation which will enhance the understanding of the systems and add to the knowledge base in the respective areas. Further to this, an appraisal of the two sides of these very different but very useful modelling techniques will be made. Both of which can be referred to as a mathematical model of health.

1.3 Open questions to address

By using the approaches of delay differential equations modelling and data mining, the work in this thesis attempts to address the following research questions:

- Can Hill functions be used as an appropriate way of representing aspects of the glucose-insulin regulatory system to produce the desired oscillatory behaviour?
- Is it possible to derive delay-dependent stability conditions in the local case for the two-delay system to define apparition of oscillations?
- Can a Lyapunov functional approach to global stability of the two-delay model of glucose-insulin regulation be used to find the optimal ranges of the Hill parameters by obtaining stability conditions dependent on these parameters?
- Will the use of data mining techniques applied to medical data provide useful insights into relations of variables and conditions?

- Can these data mining techniques be used to build successful predictive models for a specific condition in a medical scenario?

1.4 Statement of original contribution

The work in this thesis contains the following key contributions to knowledge:

- **Hill functions used to model the underlying regulatory mechanisms of the glucose-insulin regulatory system:** This thesis contributes to the models for glucose-insulin regulation by creating new functions to describe the regulatory mechanisms which are of Hill function form. Optimal ranges for the parameters of these functions are found.

- **Local stability analysis of the two-delay glucose-insulin system:** The two-delay model presented in this thesis is analysed for local stability for the first time. The relation between the two-time delays is calculated analytically to give the boundary for oscillations. This is then confirmed numerically using a method of linear matrix equalities.

- **Global stability analysis of the two-delay glucose-insulin system:** This thesis presents an analytical global stability analysis of the two-delay system using a Lyapunov functional approach. While providing conditions for stability, the results are used to numerically find the optimal ranges for the Hill coefficients.

- **A new classification of post-stroke mortality time scales:** An analysis into a data set of stroke patients shows that there are natural time scales arising for post-stroke mortality which suggest differences in risk factors for different time periods. A very short time scale as well as longer time scales, dependent on the analysis, are

proposed which differ from the typical one month time scale.

- **Risk factor analysis for post-stroke mortality over different time scales with a focus on age:** This thesis presents a focused study on age as a risk factor for post-stroke mortality for different time scales. Age is seen to not have any effect on post-stroke mortality in the very short period. Whereas, for the longer time period age plays a significant role. A wide range of other risk factors are analysed for two time periods, 1-7 days and 8-93 days.

- **Predictive models are successfully created for post-stroke mortality:** The analysis of risk factors is used to build predictive models over the very short, 1-7 days, and the short/intermediate, 8-93 days, time scales by using a naïve Bayes approach to assign variables a score function. Moreover, these models are compared to other techniques which show the benefit of a naïve Bayes approach.

1.5 Thesis structure

Following this introduction, Chapters 2 - 6 of this thesis will be related to the method of delay differential equations applied to the glucose-insulin regulation system, while Chapters 7 and 8 will look at the data mining approach to health to build models for post-stroke mortality.

In Chapter 1 a background to diabetes is given alongside the motivation for this work, while the questions to be addressed and the contributions of the thesis are discussed.

Chapter 2 gives an overview of the literature for models related to glucose-insulin regulation highlighting key studies in relation to the delay differential equation work in this thesis.

Chapter 3 introduces the concepts and methodologies that are used in the creation and analysis of the delay differential equation model of glucose-insulin regulation.

In Chapter 4 the model with the new Hill functions is introduced providing some preliminary numerical analysis of the parameters and their effect in relation to the fixed point. An analytical overview of local stability in the one-delay case is covered.

Chapter 5 presents analysis of local stability in order to determine conditions for the apparition of oscillations in the two-delay case. The characteristic equation is assessed for the general case when both delays are nonzero and the results are confirmed numerically using linear matrix inequalities. A Lyapunov functional is then created to give a set of sufficient conditions for local stability.

Chapter 6 assesses the global stability of the nonlinear two delay glucose-insulin regulations system by analytical derivation of a Lyapunov functional. The global analysis of the system allows for the new parameters in the model to be investigated further in relation to stability. Optimal ranges for the Hill parameters are given at the end of this chapter.

Chapter 7 marks the beginning of the data mining related work where an overview of some techniques and methodologies used in this thesis and from the literature for statistical based data mining and machine learning are introduced.

In Chapter 8 the study of risk factors in short and intermediate term post-stroke mortality with the creation of predictive models is presented. First, appropriate mortality time scales are determined before analysing age as a specific risk factor. A wide range of factors are studied and used to build new predictive models via a naïve Bayes approach.

Chapter 9 gives an overall conclusion summarising the work in this thesis and provides a discussion of possible avenues for future work.

Chapter 2

Literature review for glucose-insulin regulation

An overview of the literature for models related to the glucose-insulin regulation system. The focus will be on the key models in the field with their extensions and how they have inspired the work in this thesis.

2.1 Introduction

The glucose-insulin regulatory system is an important part of maintaining a healthy body. The pancreas and the liver regulate the production of insulin and glucose respectively, in order to keep the glucose level in check. Insulin, produced by the pancreas, is needed by the body to allow glucose to be used as fuel in the cells. Failure to regulate this system will result in high blood sugar leading to diabetes due to either a lack of insulin or insulin resistance and is the cause of many other long term problems. The system operates with ultradian oscillations, defined as cycles which are repeated throughout a 24 hour day, which were first discovered by Hansen (1923). These ultradian oscillations have been studied in a variety of ways since and can best be seen during constant glucose infusion in work originating from Shapiro et al (1988), an adaptation of which is seen in Figure 2.1.

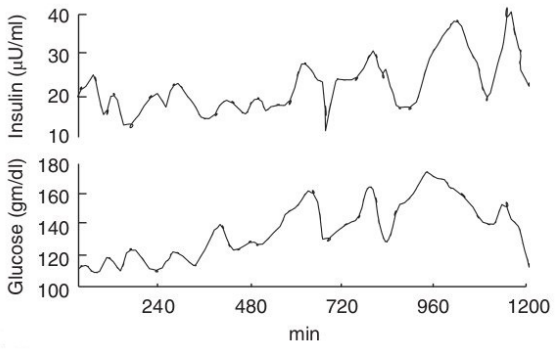

Figure 2.1: Graphical representation of the ultradian oscillations of insulin and glucose under conditions of constant glucose infusion. Adaption taken from Li et al (2006).

2.2 Original model and functions

The work in this thesis to model the ultradian oscillations of the glucose-insulin regulatory system originally stems from the work of Sturis et al (1991a,b, 1995). The paper (Sturis

et al 1991a) provides a model based on a negative feedback loop, typically found in most biological systems, and incorporates two insulin compartments, plasma and remote, and one glucose compartment. An important feature of the system is the inclusion of a time delay representing the delayed effect of insulin on inhibiting glucose production. The system includes five functions, $f_1, \ldots, f_5$, which represent the underlying regulatory mechanisms of the system. The full framework, as adopted by this thesis, is represented in the flow diagram of the model in Figure 2.2.

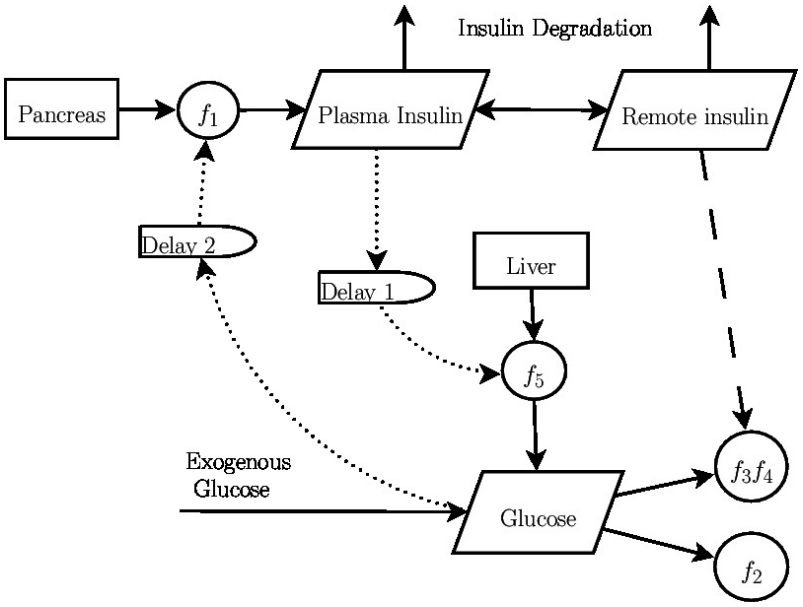

Figure 2.2: Modelling framework showing the regulatory mechanisms for glucose and insulin regulation. The functions f_i represent key features such that f_1 is the effect of glucose on insulin secretion, f_2 - insulin independent glucose utilisation, f_3 and f_4 represent insulin dependent glucose utilisation and f_5 - effect of insulin on glucose production.

The model is created as a system of six ordinary differential equations: three equations are used for the main variables of the model, they are plasma insulin, $x(mU)$, remote insulin, $y(mu)$ and plasma glucose, $z(mg)$ while the remaining three differential equations represent the time delay in a three tiered linear system introducing the auxiliary states,

$h_1,\ h_2$ and h_3,

$$\begin{aligned}
\frac{dx}{dt} &= f_1(z) - E\left(\frac{x}{V_p} - \frac{y}{V_i}\right) - \frac{x}{t_p},\\
\frac{dy}{dt} &= E\left(\frac{x}{V_p} - \frac{y}{V_i}\right) - \frac{y}{t_i},\\
\frac{dz}{dt} &= G_{in} + f_5(h_3) - f_2(z) - f_3(z)f_4(y),\\
\frac{dh_1}{dt} &= \frac{3(x-h_1)}{t_d}, \qquad \frac{dh_2}{dt} = \frac{3(h_1-h_2)}{t_d}, \qquad \frac{dh_3}{dt} = \frac{3(h_2-h_3)}{t_d}.
\end{aligned} \tag{2.1}$$

The three auxiliary variables model the delay via an expansion similar to that of Taylor's expansion. Taking the expression for the derivative of h_1 without the coefficient, where t_d is the total time delay, then it can be written such that,

$$x(t-t_d) = h_1(t-t_d) + h_1'(t-t_d)t_d. \tag{2.2}$$

This is equivalent to the Taylor expansion of $h_1(t)$ at time $t - t_d$. Therefore, the auxiliary variable $h_1(t) \approx x(t-t_d)$. When this is extended to the third order case including auxiliary variables h_2 and h_3 the expansion is not of Taylor expansion form and hence the delay is not third order but is purely a three-tiered system where each auxiliary variable is an approximation of the previous. This has been shown in Li et al (2006).

The functions $f_1, \ldots, f_5$ are defined as:

$$\begin{aligned}
f_1(z) &= \frac{R_m}{1+\exp\left(\frac{-z}{300V_g}+6.6\right)}, \qquad f_2(z) = U_b\left(1-\exp\left(\frac{-z}{144V_g}\right)\right),\\
f_3(z) &= \frac{0.01z}{V_g}, \qquad f_4(y) = \frac{U_m}{1+\exp(-1.772\log y\left(\frac{1}{V_i}+\frac{1}{(Et_i)}\right)+7.76)} + U_0,\\
f_5(h_3) &= \frac{R_g}{1+\exp\left(\frac{0.29h_3}{V_p}-7.5\right)}.
\end{aligned} \tag{2.3}$$

Parameter	Description
t_p	Time for plasma insulin degradation
t_i	Time for interstitial insulin degradation
t_d	Time delay on glucose production dependent on plasma insulin
E	Diffusion of insulin between plasma and interstitial space
G_{in}	Exogenous glucose infusion
V_p	Volume of insulin in plasma space
V_i	Volume of interstitial fluid
V_g	Volume of glucose space
U_b	Maximum velocity of insulin-independent glucose utilization
U_0	Minimum velocity of insulin-dependent glucose utilization
U_m	Maximum velocity of insulin-dependent glucose utilization
R_m	Maximum insulin infusion rate
R_g	Maximum glucose infusion rate

Table 2.1: Descriptions of the parameters in the model (Sturis et al 1991a).

The parameters and their descriptions can be seen in Table 2.1, the parameters not described in the table are auxiliary parameters and have no physiological meaning. It should also be noted that in the paper of Sturis et al (1991a), a second time delay for the slow effect of insulin on glucose utilization is mentioned but not explicitly modelled. This delay would represent a delay related to function f_4. However, this could also be represented as a delay in insulin production, f_1, as within the system the delay for these two functions would be equivalent. While it is not directly modelled by the system of ODEs (Sturis et al 1991a), it is an effect of the movement of insulin between the interstitial and remote compartments as the action on glucose is only seen in the interstitial fluid. Extensions of this model introduce this time delay and are discussed in more detail in Section 2.3.

The functions seen in Equation (2.3) are commonly used in relation to glucose-insulin regulation and the model of Sturis et al (1991a) has been used as a base for many models since its publication. One of the early developments of the model, seen in Tolic et al (2000), investigates the advantages of periodic versus constant insulin supply. A simplified version of the original model of Sturis et al (1991a) is used by replacing the functions, f_i, with constants and Taylor expansions. The five functions $f_1, \ldots, f_5$ are rewritten

slightly but remain inherently the same.

There are a number of ways in which the model of Sturis et al (1991a) can be developed. The main step is to incorporate explicitly the time delay for glucose production, into function f_5 creating a system of three delayed differential equations (Drozdov and Khanina 1995; Bennett and Gourley 2004a). This removes the reliance on three tiered system of differential equations for the delay model and amongst other modifications is applied in most papers since. The model of Drozdov and Khanina (1995) also reduces the number of functions to three by rewriting function f_2, noting that function f_3 is not written as an independent function but is included directly in the equation for glucose.

2.3 Systems of two delay differential equations

A further way to reduce the model is by combining the two insulin compartments into one by removing the diffusion term and giving an overall insulin degradation term with function f_1. This gives a model of two delayed differential equations which can be seen in many papers e.g. Engelborghs et al (2001a), Bennett and Gourley (2004b) and Li et al (2006). Models of this type will provide the base of analysis for the work presented in this thesis. A significant further development to models in this area is the introduction of a second time delay (Engelborghs et al 2001a; Li et al 2006). This second time delay relates to the slow effect of glucose utilisation dependent on insulin as mentioned in Sturis et al (1991a). The idea for approaching this delay is very different in Engelborghs et al (2001a) and Li et al (2006), their models will be discussed in further detail here.

The main analysis of these models is centred around two areas: (a) the effect on oscillations of altering parameters and (b) bifurcation analysis for stability. The work of Engelborghs et al (2001a) studies a model for two time delays by introducing a delayed form of

function f_1 alongside the non-delayed function f_1 in the equation for insulin levels such that the equation for $\dot{I}$ is,

$$\dot{I}(t) = \alpha f_1(G(t)) - \frac{I(t)}{t_1} + (1-\alpha) f_1(G(t-\tau_1)). \tag{2.4}$$

In this model α represents the degree of infliction of a diabetic person where a small α indicates a more significantly affected patient. The delay represents the slow movement of insulin supply through the cells in the body to interact with the internal system. This can be viewed, in effect, as a delay of the insulin response stimulated by glucose, i.e. the delay in insulin production. This is the reasoning in reusing the function, f_1. The work in Engelborghs et al (2001a) carries out a predominantly numerical analysis of the steady state for their two-equation model with two delays. This is done by finding branches of periodic solutions through investigations of Hopf bifurcations. The authors hypothesize that there is two distinct results representing diabetic and non-diabetic patients. The bifurcation analysis seen in Engelborghs et al (2001a) and further works by these authors are created using the MATLAB package DDE-BIFTOOL (Engelborghs et al 2000, 2001b, 2002).

Now, looking to the work of Li and his collaborators (Li et al 2006; Li and Kuang 2007), the second delay in their models is described more simply as the delayed effect of glucose on insulin response. The description of this delay stems form the β-cells in the pancreas which secrete insulin. The β-cells are stimulated when the glucose level is high, which causes secretion. The biological process for this involves the molecules of glucose entering the islets in the pancreas by transporter GLUT2, an increase of ATP:ADP protein transporters, the closure of K^+ channels and the opening of Ca^+ channels before the insulin is secreted in sacs via the β-cells (Gilon et al 2002; Li et al 2006). The insulin then moves into the remote space, modelled by the variable, y, (Sturis et al 1991a), to be used in the process of converting glucose to energy. Further details of this delay can

be seen for example in (Gilon et al 2002; Cherrington et al 2002). This gives a system of two delayed differential equations with two delays. The terms τ_1 and τ_2 represent the delay in insulin response to glucose levels and the delay in glucose production dependent on insulin. The model is presented here by Equation (2.5) and will be used as the base for the studies in this thesis. Further details of the model and functions in this thesis will be discussed in Chapter 4,

$$\begin{aligned} \dot{G}(t) &= G_{in} - f_2(G(t)) - f_3(G(t))f_4(I(t)) + f_5(I(t-\tau_2)), \\ \dot{I}(t) &= f_1(G(t-\tau_1)) - d_i I(t). \end{aligned} \tag{2.5}$$

Here, alongside the functions $f_1, \ldots, f_5$, the constant parameters d_i and G_{in} represent insulin degradation and external glucose infusion respectively. The analysis of this system in Li et al (2006) compares a range of models before numerically using a bifurcation analysis for the two time delays and two selected parameters, external glucose infusion, G_{in}, and insulin degradation, d_i. This gives threshold values for these parameters for the production of ultradian oscillations. The follow-up paper (Li and Kuang 2007), studies the model of Li et al (2006) with two delays by providing an analytical study of local stability for both the one-delay and the two-delay cases. The case where one of the delays is zero is thoroughly studied and for the two-delay case more general conditions are obtained. Across all of the literature the analysis of the two-delay model is significantly less present than the one-delay model. The two-delay model is more realistic in terms of the biological representation of the system however, it is more difficult to study than the one-delay model.

In terms of stability analysis, it is important to mention the work of Bennett and Gourley (2004a,b,c). The model (Bennett and Gourley 2004b) is a system of two delayed differential equations with one time delay and is investigated analytically. Importantly for this study, conditions for global stability are obtained by using the comparison principle

approach and secondly, the Lyapunov functional approach. Serving as part of the motivation for the work in this thesis, the Lyapunov functional approach to stability will be discussed in more detail in Chapter 3 and applied to the two-delay model in Chapter 6.

2.4 Functional extensions to models

An alternative extension of the model of Li et al (2006) was presented in a paper which introduces an additional two functions, f_6 and f_7 (Chen and Tsai 2010). These two functions are included to account for the effect of hyperglycaemia which is the condition of having high blood sugar levels, just below diabetic classification. Function $f_6(G(t))$, is a function of glucose concentration and takes a very similar form to that of functions f_1 and f_5. While function seven, $f_7(G(t) - 330)$, again has a similar form but is a function of $G(t) - 330$, to accurately represent kidney glucose extraction which takes effect above the urine threshold, where the 330 represents the threshold of 330 mg/dl. Two additional parameters, α and β, are also included to estimate the dysfunctions of a system with diabetes such as hyperglycaemia. These additional functions and parameters are coupled with altering external glucose, G_{in}, to include a step function, $\mathcal{U}(t - t_m)$, in order to represent the effect of intermittent meal inputs at time t_m, rather than a continued glucose infusion. Overall, the model presented in Chen and Tsai (2010) is given by,

$$\begin{aligned}
\dot{G}(t) &= [G_{in} + f_5(I(t-\tau_2))f_6(G(t))] \\
&\qquad - [f_2(G(t)) + f_7(G(t) - 330) + \beta f_3(G(t))f_4(I(t))] \\
\dot{I}(t) &= \alpha f_1(G(t-\tau_1)) - d_i I(t) \\
G_{in}(t) &= \sum_{\forall m \in M} G_m(t - t_m)\mathcal{U}(t - t_m)
\end{aligned} \tag{2.6}$$

The developments in this model are useful as they show that small changes to the model of Li et al (2006) can produce a useful system for modelling glucose-insulin regulation under intermittent meal infusion, while also accounting for other diabetic related effects.

Another important work to discuss is that of Huang et al (2012). In this paper mathematical models are built for the glucose-insulin regulation system with impulsive injections of insulin to develop the algorithms for the insulin pump and an artificial pancreas. An artificial pancreas is a medical device designed to replace the function of the human pancreas in the metabolic system as a way of delivering a more normal method of treatment by automatically monitoring and controlling plasma glucose levels in diabetic patients. A body of work on the artificial pancreas can be seen by Steil and his collaborators (Steil and Reifman 2009; Steil et al 2010; Steil 2013). Currently the models available for the artificial pancreas are lacking in accuracy for predictability of effective control. The work in Huang et al (2012) investigates both type 1 and type 2 diabetes to improve the mathematical algorithms for use in the artificial pancreas. The model is again based on the work of Li (Li et al 2006; Li and Kuang 2007), but with simplifications related to the functions $f_1, \ldots, f_5$, such that,

$$\begin{aligned} f_1(x) &= \frac{\sigma_1 x^2}{\sigma_1^2 + x^2}, \qquad f_2(x) = \sigma_2 x, \qquad f_3(x) = ax, \\ f_4(x) &= c + \frac{mx}{n + x}, \qquad f_5(x) = b, \end{aligned} \tag{2.7}$$

where $\sigma_1,\ \sigma_2,\ a,\ b,\ c,\ m$ and n are all positive constants. The simplifications are made on the basis that the shape of the functions is more important than the function itself while the time delays are also removed. With these simplifications a successful model is produced to begin improving the algorithms for an artificial pancreas. Thus, the belief is that with a greater knowledge of the glucose-insulin system by extending the investigation of the two-delay system, an improvement to the models for the artificial pancreas can be achieved. Also, it is important to note in the model of Huang et al (2012), that the function

f_1 is represented by a Hill function of coefficient two. The Hill function will be used in the model proposed in this work for the glucose-insulin regulatory system and will be discussed in further detail in Section 3.1.

The model in Li et al (2006) continues to be used as a base for new research and developments are regularly published. One of the most recent is the model of Kissler et al (2014) which adds two main parameters in order to show the effect of severity of diabetes, with options for both type 1 and type 2. The modelling of type 1 and type 2 diabetes follows a similar idea to that of Chen and Tsai (2010), via simple additional parameters. The effect of exercise and medication is introduced in order to discuss treatment plans through the inclusion of external insulin, I_{in}, for the modelling of medication for diabetes. Once again the functions, $f_1, \ldots, f_5$, remain the same and hence can be traced back to the original model (Sturis et al 1991a). Mathematically, the analysis in the paper is split into type 1 and type 2 diabetes. It is entirely numerical with discussion of the apparition of oscillations and analysis of hypothetical patients for effects of parameter variation related to exercise and medication on the oscillatory behaviour of the system.

Alternatively, the model and functions in Sturis et al (1991a) can be used as a base for more complex models rather than condensing and simplification. The models of Kang et al (2012) and Han et al (2012) investigate movement of glucose on a molecular level from cell to cell using elements of the original functions to develop more complex models. These papers also introduce models to investigate the kinetics of glucose receptors and how they contribute to the processes involved in the regulation system, making use of the delays, functions and insulin equations in Sturis et al (1991a). Furthermore, the modelling of glucose can be split into two compartments in a similar way to the insulin creating a system of four delayed differential equations as seen in Wu et al (2011) where each of the original five functions are incorporated in the model.

2.5 Minimal model for IVGTT

The glucose-insulin regulation system can also be modelled in ways which do not stem directly from the model of Sturis et al (1991a), although similarities are clearly seen. There exists a large body of work for glucose-insulin regulation following an intra-venous glucose tolerance test (IVGTT) which is an effective method of testing for insulin resistance in humans to diagnose diabetes. In Pacini and Bergman (1986) the 'minimal model' was introduced to describe this with the lowest number of identifiable parameters, where extensions to include insulin dynamics (De Gaetano and Arino 2000) and one time delay (Li et al 2001) have been made. The paper of Panunzi et al (2007) introduces a model which uses predominantly linear terms for actions such as glucose utilisation, hepatic glucose production and insulin release and is shown in Equation (2.8). This model includes one time delay, τ_g for the response of insulin secretion dependent on glucose level and is given by,

$$\begin{aligned} \frac{dG(t)}{dt} &= -K_{xgI}I(t)G(t) + \frac{T_{gh}}{V_g} \\ \frac{dI(t)}{dt} &= -K_{xi}I(t) + \frac{T_{igmax}}{V_i}\frac{\left(\frac{G(t-\tau_g)}{G^*}\right)^\gamma}{1+\left(\frac{G(t-\tau_g)}{G^*}\right)^\gamma} \end{aligned} \tag{2.8}$$

The important thing to note is the use of a Hill function in the equation for insulin with the Hill coefficients γ and G^*, where in this case G^* represents glucose concentration at which the insulin is half its maximum secretion rate. This paper also mentions the possibility of a system with two delays by introducing a delay, τ_i, for insulin action on glucose uptake by replacing $I(t)$ with $I(t-\tau_i)$. This model has been analysed in following papers, for example, with respect to positive equilibriums and local stability (Palumbo et al 2007) and also the global stability has been analysed for delay-dependent oscillations (Giang et al 2008; Li et al 2012). The most recent work involved a development of the

model with respect to meal consumption (Palumbo et al 2013).

2.6 Concluding remarks on existing literature

What is seen from the literature focusing on models directly related to glucose-insulin regulation, i.e. without reference to more microscopic modelling of pancreatic $\beta-$cell functionality, is that there are two main directions. The analytical study of the behaviour of the model is very important in giving a detailed understanding of how the system works. The papers related to this focus on the one-delay case by setting either of the delays to zero. The work of Bennett and Gourley (2004b) provides good analysis of global stability in the one-delay case, where local stability is the focus of Li and Kuang (2007). There are some efforts to analyse the two-delay system but they tend to be numerical, as seen in Engelborghs et al (2001a). Therefore, there is much more to be understood about the system with two delays particularly from an analytical view point which will lead to better modelling and applications for the system. The other main direction is seen as applications of the model in a more real world scenario. One particular case is the use of these models for the improvement of diabetes treatment through an artificial pancreas. The paper of Huang et al (2012) addresses some issues of the algorithms for the artificial pancreas but is based on a very simplified linear model with no delays.

The belief is that a better understanding of the system of glucose-insulin regulation including two time delays both analytically and numerically could lead to an overall improvement in the treatment of diabetes.

Chapter 3

Concepts and methods for delay differential equations and stability

This chapter gives an introduction to some of the key concepts and methodologies to be used in relation to the delay differential equation modelling of the glucose-insulin regulation system. The focus will be on the Hill function and stability, while techniques for stability analysis including Lyapunov functionals and linear matrix inequalities are also discussed.

3.1 Hill function

All living organisms have a large number of biochemical reactions taking place within the body. This process involves catalysts, known as enzymes, reacting with compounds, called substrates, to cause a specific reaction. One mathematical example of this to represent such processes is the Hill function, which was first introduced in 1910 (Hill 1910) as a mechanism for describing the interaction between oxygen, the substrate, and haemoglobin, the enzyme, within the blood.

The work in this thesis introduces a new form of functions to the models of glucose-insulin regulation by using Hill functions in the four main model functions (Sturis et al 1991a; Li et al 2006). The original exponential form which has typically been used in the area to represent the regulatory mechanisms of the system was chosen to fit data and their sigmoidal shape. By using the Hill function, a stronger physiological reasoning for the parameters will be given, while removing the need for auxiliary variables. A discussion of the background and derivation of Hill function and its origins in reaction kinetics will be presented here, with its application to the model in this thesis presented in Chapter 4.

In order to replace the original exponential form of the main functions which are of logsitic function form, any funciton which takes as sigmoidal shape could be chosen. As well as the Hill function, another example is the Gompertz function which is used in modeling of cancer and tumor growth (Laird 1964; Ferreira Jr et al 2002). The Gompertz function contains two main parameters which determine the growth rate and the position on the x-axis. One key difference with the Gompertz function is that it is not a symmetrical sigmoidal shape as the right-hand asymptote is approached slower than the left-hand asymptote. Therefore, to replicate the regulatory mechanisms of the system the Hill function is the best choice.

As the idea for the Hill function stems from enzyme kinetics, it is applicable to models

involving the rate of product formation in a cooperative enzymatic reaction. In many standard textbooks the Hill function is often described alongside Michaelis-Menten enzyme kinetics (Murray 2002; Jones et al 2010). The base reaction can be seen in Equation (3.1) where an enzyme, E, reacts with a substrate, S to form a complex, ES, which is converted into a product, P and the enzyme,

$$E + nS \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES_n \xrightarrow{k_2} P + E. \tag{3.1}$$

The rates of reaction are represented by $k_1,\ k_{-1}$ and k_2 and n is a measure of negative ($n < 1$), zero ($n = 1$) or positive ($n > 1$) cooperativity. The derivation involves the law of mass action, stating that the rate of reaction and the product of the concentrations of reactants are proportional, to give a differential equation for each term in the system. This is followed by the conservation law for the enzyme, which is purely a catalyst, such that the total quantity of the enzyme remains the same. Thus,

$$[E] + [ES_n] = [E]_0, \tag{3.2}$$

where the brackets, [ ], denotes concentration and E_0 represents the initial condition. The final expression is known as the Hill function and takes the form,

$$[ES_n] = \frac{Q[S]^n}{K_m + [S]^n}, \tag{3.3}$$

where K_m is the Michaelis constant and $Q = k_2[E_0]$. This is the derivation of the basic Hill function from enzyme kinetics.

Since its initial use in enzyme kinetics for oxygen and haemoglobin, the Hill function has been widely applied in other areas for modelling biochemical properties and reactions. One main area where it has been used is clinical and pharmacological modelling for

drug trials and other such reactions. A review of the success of the Hill function in pharmacology can be seen for example in Goutelle et al (2008) and Gesztelyi et al (2012). Other instances include modeling of the *lac* operon using Hill function form for aspects of the model (Yildirim et al 2004; Yagil and Yagil 1971), and as mentioned in Chapter 2, it has uses in glucose-insulin regulation models. The general form of the Hill function is given in Equation (3.4) which shows a nonlinear function f, dependent on the variable x,

$$f(x) = \gamma \frac{x^h}{x^h + k^h}, \tag{3.4}$$

where h and k are the Hill coefficients and γ is a constant coefficient.

The Hill function is sigmoidal and the steepness of response is affected by the Hill coefficient, h. The parameter k denotes the concentration required for the function to be at half its maximal output and thus determines the position of the turning point. Figure 3.1 shows a graphical representation of the effect of altering h, moving from negative cooperativity at $h = 0.5$, through zero cooperativity at $h = 1$ and into positive cooperativity for $h = 2$ and $h = 4$. Increasing h to very large values would lead to a function similar in shape to a step function and because of this property it has been used as a switch in biological systems (Cherry and Adler 2000). Figure 3.2 gives graphical evidence of changing parameter k where a Hill function is plotted for values of coefficient $h = 1$ and $h = 2$ for two different values of k. Each Hill function with parameter h, will cross at the same turning point, specified at the half maximal output. Note, these graphs are plotted with Equation (3.4) where $\gamma = 1$.

There are some important properties of the Hill function which will be useful for the analysis in this thesis. The first derivative of a Hill function, as seen in Equation (3.4),

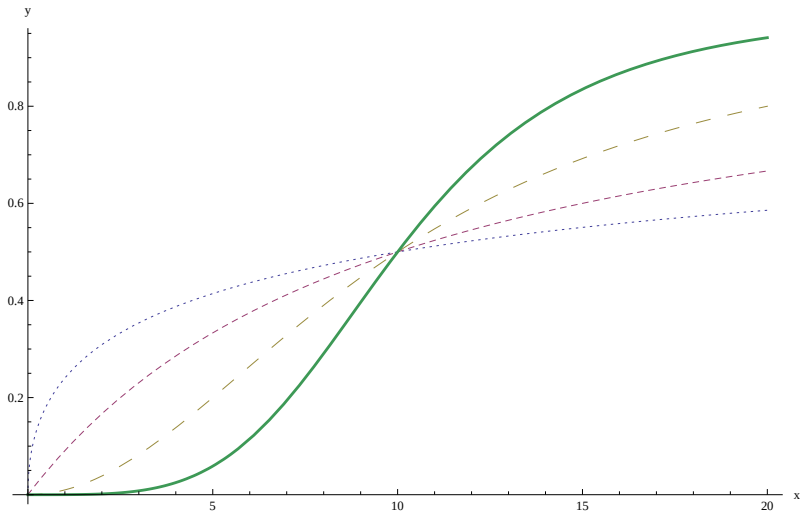

Figure 3.1: A plot of a Hill function with $\gamma = 1$ in range $0 < x < 20$ and $k = 10$ for four values of h. Dotted line represents $h = 0.5$, dashed line shows $h = 1$, wide dashed line is $h = 2$ and the solid line represents $h = 4$.

where $\gamma > 0,\ h,\ k$ are constant, is given by

$$f'(x) = \gamma h k^h \frac{x^{h-1}}{(x^h + k^h)^2}, \tag{3.5}$$

by using the quotient rule. The first derivative has an extremum when

$$x^h = \frac{h-1}{h+1} k^h. \tag{3.6}$$

The third derivative of $f(x)$ at this point is

$$f''' = -\gamma \frac{(h-1)^2(h+1)^2}{8hx^3}. \tag{3.7}$$

Thus, for $\gamma,\ x > 0$ the first derivative, f', has a maximum when $h > 0$, and respectively a minimum when $h < 0$. It is therefore possible to express the maximal value of the first derivative, denoted by f'^M , of each of these functions in terms of the Hill parameters.

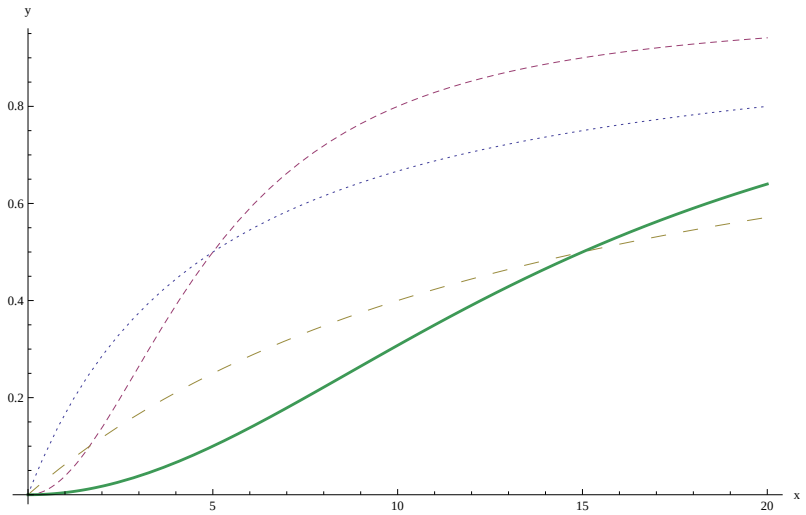

Figure 3.2: A plot of a Hill function with $\gamma = 1$ in range $0 < x < 20$ to show the effect of changing k. Four functions are plotted where the dotted line represents $h = 1$ with $k = 5$, the dashed line shows $h = 2$ with $k = 5$, the wide dashed line is $h = 1$ with $k = 15$ and the solid line represents $h = 2$ with $k = 15$.

This gives,

$$f'^{M} = \frac{\gamma}{hk}(h-1)^{\frac{h-1}{h}}(h+1)^{\frac{h+1}{h}}, \tag{3.8}$$

where the principal value has to be used whenever $\frac{h-1}{h}$ or $\frac{h+1}{h}$ is not an integer.

The Hill function has yet to be applied to all of the regulatory mechanisms, f_i, in the glucose-insulin system, but has been used in certain cases for single functions, typically f_1. As well as the example given in Chapter 2 related to the artificial pancreas (Huang et al 2012), a Hill function is used in Han et al (2012) to model additional action of the β-cells in the pancreas. The Hill function has also been used as a replacement for one of the original functions, f_i, as seen in Topp et al (2000), with coefficient $h = 2$ to represent insulin secretion, f_1. The model here places a focus on β-cell mass with the dynamics of glucose and insulin. These models use only one Hill function within a simplified model by way of linearising the remaining functions, leading to an incomplete understanding of the behaviour of the system and the related parameters. The Hill function will be applied

to four functions, $f_1,\ f_2,\ f_4$, and f_5, in the model for this thesis.

3.2 Stability of differential equations

The stability of a differential equation determines the behaviour of the solutions to the system under certain conditions and is important for all types of differential equations whether ordinary, partial or delay type. Here, some important features of stability will be discussed, including fixed points and the different types of stability.

In order to study a system of differential equations for stability it is necessary to have a fixed point for the system. The fixed point, sometimes known as an equilibrium point, represents a stationary condition for the dynamics of the system. A fixed point, x^*, is defined as the point at which the dynamical system,

$$\frac{dx}{dt} = F(x) \text{ is stationary, i.e. } F(x^*) = 0. \tag{3.9}$$

Therefore, if a system has an initial condition at the fixed point, $x(0) = x^*$, then the solution will remain at the fixed point for all time, t, $x(t) = x^*,\ t > 0$. A system can have any number of fixed points, including zero. Note that throughout this work a fixed point will be denoted by the $*$ symbol.

It is at the fixed point that the stability of the system is analysed. In general terms the stability of a fixed point determines whether the solution of a dynamical system moves towards, away from, or stays close to the fixed point as time increases. The trajectory of the solution can be defined as the path of the solution in time. All considerations in this section are related to the general definitions of stability, while Sections 3.4 and 3.5 will introduce some techniques for determining stability. This work will look at stability in the local and global senses and also discuss asymptotic stability.

Local is the term used for behaviour of the system with an initial condition sufficiently close to the fixed point. A locally stable fixed point can be described as a point for which any initial condition within a specified region, δ, close to the fixed point itself will remain within a region, ϵ, close to the fixed point. A typical definition for local stability for a fixed point x^* with initial condition $x(0)$ as summerised from Glendinning (1999) is given as,

$$||x(0) - x^*|| < \delta \Rightarrow ||x(t) - x^*|| < \epsilon. \tag{3.10}$$

Consequently there exists a real constant $\delta > 0$ where any initial conditions in this region will lead to a solution $x(t)$ to remain within the boundary of the region $\epsilon > 0$ for all $t > 0$. This condition requires the solution to stay close to the fixed point rather than to converge directly to it. A fixed point where the trajectories approach the fixed point as time, t, increases is known as asymptotically stable, $x(t) \to x^*$ when $t \to \infty$.

Alternatively if a system is not stable, it is unstable. In a similar style of definition, a system is unstable if there is always an initial condition $x(0)$ where $||x(0) - x^*|| < \epsilon$, such that the trajectory does not stay within the region and as $t \to \infty$. That is to say the solution moves away from the fixed point.

Global stability, sometimes referred to as stability in the large, is a stable system as defined in the local case, but without the restriction on initial conditions. A globally stable system has the property that any initial condition, $x(0)$, will result in the trajectories of the solution moving within a region, ϵ, close to the fixed point as time increases,

$$x(0) \in \mathbb{R} \Rightarrow ||x(t) - x^*|| < \epsilon, \text{ for sufficiently large } t. \tag{3.11}$$

Global asymptotic stability combines the definitions for global and asymptotic stability such that any initial conditions will result in trajectories converging exactly to the fixed point. It is also important to note that a system can only be globally stable if there is

only one fixed point. Having more than one fixed point invalidates the criteria for global stability as by definition, initial conditions at another fixed point will remain at that fixed point.

When determining the local stability of the system, the linearised system can be analysed allowing the use of linear analysis techniques, see Section 3.3. For global stability the original system must be investigated which is a more complicated process. A link between the linearisation of the system and the original system for stability comes via the Hartmann-Großman theorem which can be found in the texts for stability and differential equations (Glendinning 1999; Betounes 2010) and can be given as,

Hartman-Großman Theorem: If $x = 0$ is a hyperbolic stationary point of $\dot{x} = f(x)$ then there is a continuous invertible map, h, defined on some neighbourhood of $x = 0$ which takes orbits of the nonlinear flow to those of the linear flow. This map can be chosen so that the parameterisation of orbits by time is preserved (Grobman 1959; Hartman 1960).

The theory itself states that the behaviour of the linearised system close to a hyperbolic equilibrium point is qualitatively similar to that of the full system, dependent on there being no zero or purely imaginary eigenvalues of the linearised system. Similar results can also be found for delay equations. The book of Wu (1996) presents a Theorem which provides sufficient conditions for stability of the zero solution of the linearised system to imply local exponential stability of the corresponding equilibrium of the original nonlinear system with delays.

3.3 Linearisation

As mentioned in the discussion of stability, in order to study local stability it is necessary to study the linear system. When dealing with a nonlinear system, as is the case with the model presented in this thesis, a process of linearisation must be carried out. A linear approximation of the nonlinear system is made at a given point, which allows for the use of linear analysis techniques (Williamson 1997).

Linearisation of a system is made by taking the first order Taylor expansion at a given point. The description for linearisation of a nonlinear system with two time delays is seen here. Given a nonlinear system, where x is a vector, with two time delays such that,

$$\dot{x} = f(x,\ x(t-\tau_1),\ x(t-\tau_2)), \tag{3.12}$$

evaluated at its fixed point, x^*, which is defined as $f(x^*,\ x^*,\ x^*) = 0$. The Taylor expansion around that point at the first order is given by,

$$\dot{x} = \left.\frac{\partial f}{\partial x}\right|_{x^*} x + \left.\frac{\partial f}{\partial x(t-\tau_1)}\right|_{x^*} x(t-\tau_1) + \left.\frac{\partial f}{\partial x(t-\tau_2))}\right|_{x^*} x(t-\tau_2). \tag{3.13}$$

This can then be used to determine the characteristic equation in the study of local stabilty (Williamson 1997), see Chapter 4 for the application in this thesis.

3.4 Lyapunov functional method

Lyapunov functions are a very powerful method of determining the stability of a fixed point for a system of differential equations. Detailed information on Lyapunov functionals can be found in many texts related to differential equations and stability (Wiggins

2003; Glendinning 1999; Betounes 2010). A huge advantage of Lyapunov functionals is that they do not require the actual solution of the differential equation. This is a very useful property for more complex systems when it may not be easy or even possible to solve the system. In general terms a Lyapunov function, typically denoted by V, is a scalar energy-like function which is positive for all values of a function except at the fixed point, as described in Equation (3.14),

$$V(0) = 0 \text{ and } V(x) > 0. \tag{3.14}$$

Following this, the Lyapunov function is required to decrease on all trajectories of the system, that is to say that its time derivative is negative, $\dot{V}(x) < 0$. The definition of a Lyapunov function for local and global stability are covered below.

In the following definitions a neighborhood of the fixed point x^*, is denoted by C for simplicity. These definitions have been summarised from the texts Wiggins (2003); Glendinning (1999). For local stability it is required that for all x in a region, C, close to the fixed point, the time derivative of the Lyapunov functional is negative or equal to zero, also known as negative semidefinite,

$$\dot{V}(x) \leq 0, \ \forall x \in C \setminus \{x^*\}. \tag{3.15}$$

A Lyapunov function is locally asymptotically stable if the time derivative of V is strictly less than zero, this is known as negative definite,

$$\dot{V}(x) < 0, \ \forall x \in C \setminus \{x^*\}. \tag{3.16}$$

The definition for global stability and global asymptotic stability of a Lyapunov functional follows the same rules as the local stability but with the requirement that $\dot{V}$ is negative

for all x. Below is the definition for global asymptotic stability,

$$\dot{V}(x) < 0,\ \forall x \in \mathbb{R}^n \setminus \{x^*\}. \tag{3.17}$$

A Lyapunov function can take on different forms for different types of systems and thus different methods can be used in finding one and for any one system different Lyapunov functionals can be found (Glendinning 1999). One other key advantage of using Lyapunov functions is that it is known that any stable system will have a Lyapunov function. Hence, the construction of a Lyapunov function can be used to prove that the system is stable (Betounes 2010).

3.5 Linear matrix inequalities

One numerical method for solving the stability of a system of linear delayed differential equations is linear matrix inequalities (LMI). The LMI method provides delay-dependent criteria for stability of the linear system by using free weighting matrices. This concept for stability in delay systems has been developed in a number of papers for different types of systems in the control area (Wu et al 2004a,b; He et al 2005). The specific idea used here comes from He et al (2006) which provides the criteria for stability of a linear delayed differential equation system with multiple delays. For a linear system of dependent variables, represented by the vector, x, with two time delays, τ_1 and τ_2, of the general form,

$$\dot{x}(t) = A_0 x(t) + A_1 x(t-\tau_1) + A_2 x(t-\tau_2),\ \ x \in \mathbb{R}^p, \tag{3.18}$$

where p is the number of dependent variables and $A_0,\ A_1$ and A_2 are $p \times p$ matrices, a system of linear matrix inequalities is created where feasibility of the LMIs guarantees

local asymptotic stability. The methodology looks at the relationships for the two time delays stemming from the Newton-Leibniz formula, otherwise known as the fundamental theorem of calculus (Rudin 1986), for example

$$x(t-\tau_2) = x(t-\tau_1) - \int_{t-\tau_2}^{t-\tau_1} \dot{x}(s)ds, \tag{3.19}$$

which takes into account the direct relation between τ_1 and τ_2. The criteria for stability related to these conditions and the free weighting matrices are detailed here.

Considering the following Lyapunov functional,

$$\begin{aligned} V =& x^T(t)Px(t) + \int_{t-\tau_1}^{t} x^T(s)Q_1x(s)ds + \int_{t-\tau_2}^{t} x^T(s)Q_2x(s)ds \\ &+ \int_{-\tau_1}^{0}\int_{t+z}^{t} \dot{x}^T(s)W_1\dot{x}(s)dsdz + \int_{-\tau_2}^{0}\int_{t+z}^{t} \dot{x}^T(s)W_2\dot{x}(s)dsdz \\ &+ \int_{-\tau_1}^{-\tau_2}\int_{t+z}^{t} \dot{x}^T(s)W_3\dot{x}(s)dsdz, \end{aligned} \tag{3.20}$$

where $P,\ Q_i$ $(i = 1, 2)$ and W_j $(j = 1, 2, 3)$ are symmetric $p \times p$ matrices. It was shown in He et al (2006) that the requirement that V be a Lyapunov-Krasovskii functional for the linear system (3.18) is equivalent to the requirement that the following linear matrix inequalities are satisfied,

$$\begin{bmatrix} \Phi_{11} & \Phi_{12} & \Phi_{13} \\ \Phi_{12}^T & \Phi_{22} & \Phi_{23} \\ \Phi_{13}^T & \Phi_{23}^T & \Phi_{33} \end{bmatrix} < 0, \tag{3.21}$$

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} & N_1 \\ X_{12}^T & X_{22} & X_{23} & N_2 \\ X_{13}^T & X_{23}^T & X_{33} & N_3 \\ N_1^T & N_2^T & N_3^T & W_1 \end{bmatrix} \geq 0, \tag{3.22}$$

$$\begin{bmatrix} Y_{11} & Y_{12} & Y_{13} & S_1 \\ Y_{12}^T & Y_{22} & Y_{23} & S_2 \\ Y_{13}^T & Y_{23}^T & Y_{33} & S_3 \\ S_1^T & S_2^T & S_3^T & W_2 \end{bmatrix} \geq 0, \tag{3.23}$$

$$\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} & \kappa T_1 \\ Z_{12}^T & Z_{22} & Z_{23} & \kappa T_2 \\ Z_{13}^T & Z_{23}^T & Z_{33} & \kappa T_3 \\ \kappa T_1^T & \kappa T_2^T & \kappa T_3^T & W_3 \end{bmatrix} \geq 0, \tag{3.24}$$

where

$$\begin{aligned}
\Phi_{11} &= PA_0 + A_0^T + Q_1 + Q_2 + N_1 N_1^T + S_1 + S_1^T + A_0^T H A_0 \\
&\quad + \tau_1 X_{11} + \tau_2 Y_{11} + |\tau_1 - \tau_2| Z_{11}, \\
\Phi_{12} &= PA_1 - N_1 + N_2^T + S_2^T - T_1 + A_0^T H A_1 + \tau_1 X_{12} + \tau_2 Y_{12} + |\tau_1 - \tau_2| Z_{12}, \\
\Phi_{13} &= PA_2 + N_3^T + S_3^T - S_1 + T_1 + A_0^T H A_2 + \tau_1 X_{13} + \tau_2 Y_{13} + |\tau_1 - \tau_2| Z_{13}, \\
\Phi_{22} &= -Q_1 - N_2 - N_2^T - T_2 - T_2^T + A_1^T H A_1 + \tau_1 X_{22} + \tau_2 Y_{22} + |\tau_1 - \tau_2| Z_{22}, \\
\Phi_{23} &= -N_3^T - S_2 + T_2 - T_3^T + A_1^T H A_2 + \tau_1 X_{23} + \tau_2 Y_{23} + |\tau_1 - \tau_2| Z_{23}, \\
\Phi_{33} &= -Q_2 - S_3 - S_3^T + T_3 + T_3^T + A_2^T H A_2 + \tau_1 X_{33} + \tau_2 Y_{33} + |\tau_1 - \tau_2| Z_{33}, \\
H &= \tau_1 W_1 + \tau_2 W_2 + |\tau_1 - \tau_2| W_3
\end{aligned}$$

and

$$\kappa = \begin{cases} 1 & \text{if } \tau_1 \geq \tau_2 \\ -1 & \text{if } \tau_1 < \tau_2. \end{cases}$$

In this system the terms $P, Q, W, X, Y, Z, N, S, T$, are the free weighting matrices which are subject to the following conditions: The matrices $P = P^T > 0$ and $Q_i = Q_i^T > 0$ $(i = 1, 2)$ are symmetric definite positive, the matrices $W_i = W_i^T \geq 0$ $X_{ii} = X_{ii}^T \geq 0$, $Y_{ii} = Y_{ii}^T \geq 0$ and $Z_{ii} = Z_{ii}^T \geq 0$ where $i = 1, 2, 3$, are symmetric positive semi-definite, while the free matrices N_i, S_i, T_i, X_{ij}, Y_{ij} and Z_{ij} where $i, j = 1, 2, 3$, have no restrictions. Providing that these weighting matrices can be found to satisfy the conditions of the LMI above then the system being studied is locally asymptotically stable. A useful software package for the LMI method is found in the Robust Control Toolbox in MATLAB (MATLAB 2013).

Chapter 4

The two-delay differential equation model with Hill functions

In this chapter the modified model with the new regulatory mechanism functions, f_i, of Hill form is introduced. An investigation of the suitability of the model for glucose-insulin regulation and the effect of certain parameters on the oscillatory behaviour and the position of the fixed point will be presented. An analysis of the local stability for the one-delay case will be given. A condensed version of the results seen in the following three chapters has been published, see Huard et al (2015).

4.1 Mathematical model with Hill functions

The ultradian oscillations in the glucose-insulin regulatory system are studied in this thesis by considering the two-delay model as presented by Li et al (2006). The discussion of this model was given in Chapter 2 and the model will be summarised here for clarity. For glucose, G, and insulin, I, concentration, the rate of change is given by,

$$\begin{aligned} \dot{G}(t) &= G_{in} - f_2(G(t)) - f_3(G(t))f_4(I(t)) + f_5(I(t-\tau_2)), \\ \dot{I}(t) &= f_1(G(t-\tau_1)) - d_i I(t), \end{aligned} \tag{4.1}$$

where the two time delays, τ_1 and τ_2 are assumed to be positive and constant, d_i is the insulin degradation rate and G_{in} represents external glucose infusion. These two terms will be discussed in more detail later in this section. A key contribution of this thesis is the construction of regulatory mechanism functions, $f_1,\ f_2,\ f_4$ and f_5, by using Hill functions. This modification gives a more accurate account of the oscillatory behaviour of the system while allowing for the introduction of new meaningful parameters. Analytic conditions are obtained for these new parameters guaranteeing that the system exhibits sustained oscillations. This change will also negate the need for any auxiliary parameters that are not related to physiological aspects. As function f_3 describes a linear mechanism, it is not appropriate to use the Hill function form as this would be an unnecessary complication. Thus f_3 will remain linear. The precise form of these functions is given by,

$$\begin{aligned} f_1(G) &= \frac{R_m(G/V_g)^{h_1}}{(G/V_g)^{h_1} + k_1^{h_1}}, \quad f_2(G) = \frac{U_b(G/V_g)^{h_2}}{(G/V_g)^{h_2} + k_2^{h_2}}, \quad f_3(G) = \frac{G}{C_3 V_g}, \\ f_4(I) &= U_0 + \frac{(U_m - U_0)\left(I(1/V_i + 1/(Et_i))\right)^{h_4}}{\left(I(1/V_i + 1/(Et_i))\right)^{h_4} + k_4^{h_4}}, \quad f_5(I) = \frac{R_g(I/V_p)^{h_5}}{(I/V_p)^{h_5} + k_5^{h_5}}. \end{aligned} \tag{4.2}$$

Here, f_1 represents the effect of glucose on insulin secretion, f_2 is the insulin independent glucose utilisation, f_3 and f_4 represent insulin dependent glucose utilisation and f_5 is

effect of insulin on glucose production. Function f_5 is the only decreasing function, while $f_1, \ldots, f_4$ are all increasing. This dictates that the Hill coefficients $h_1, \ldots, h_4$ are positive and h_5 must be negative. The Hill coefficients could be evaluated by fitting to experimental data of glucose and insulin levels as well as the rates of glucose production and insulin secretion in relation to the functions. This work studies theoretical aspects of the Hill coefficients which could then be fitted if suitable data were available.

In order to determine the values of the Hill coefficients, $h_i,\ k_i,\ i = 1, 2, 4, 5$, two main criteria will be considered,

- The steady state of the system must be located in a physiologically relevant range.
- The model should be capable of rendering oscillations which match physiological behaviour.

The widely accepted physiological idea presented in Keener and Sneyd (1998) suggests that the shape of the functions is more important than the form. This adds further reasoning to the introduction of Hill functions as it is not necessary to have a specific function form provided the system is well represented. Furthermore, the idea in Keener and Sneyd (1998) motivates the work to obtain an initial estimate to the ranges of the Hill coefficients in the new functions by fitting them to the original functions. An optimisation is made dependent on physiological ranges in Section 4.3 and maximal stability region, see Chapter 6 where it is seen that parameter h_1 has the largest effect. For this reason function f_1 will not be fitted here. The process for fitting the Hill functions f_2 and f_5 is based on satisfying two specific conditions relating to the derivatives of the exponential form and the corresponding Hill function. Taking x as either glucose or insulin dependent on the function, the conditions are that the first derivatives with respect to x are equal, and that the second derivatives with respect to x equal zero.

The graphical comparison of the exponential function and the Hill function has been created with the values of Hill coefficients found using these methods and can be seen in Figures 4.1 and 4.2. In each figure the blue curve represents the original exponential function and the red curve represents the Hill function form. Thus, it can clearly be seen that the Hill functions can aptly display the desired shape for the regulatory mechanisms.

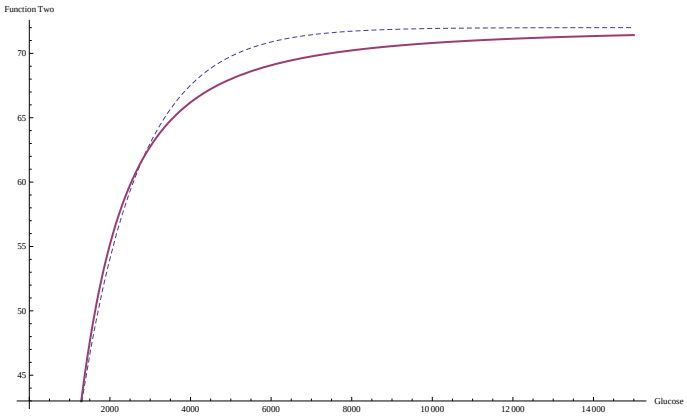

Figure 4.1: Comparison of function f_2 for the exponential form, represented by the dashed line, and the Hill form represented by the solid line.

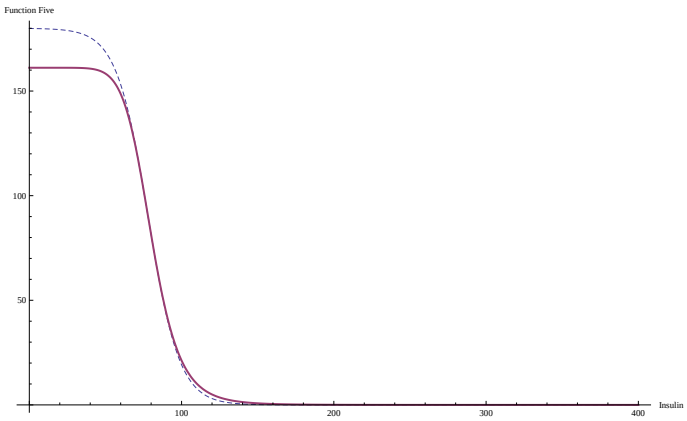

Figure 4.2: Comparison of function f_5 for the exponential form, represented by the dashed line, and the Hill form represented by the solid line.

The original form of function f_4 had been written specifically to contain an exponential, presumably for continuity. The Hill function form can be easily derived from this function and thus initial estimates for the Hill coefficients can be derived from the original form of the function, rather than through fitting to the form of the previous solution. The

derivation of the Hill function form of f_4 can be seen here, starting with the exponential form taken from Sturis et al (1991a) and the parameters as defined in Table 2.1,

$$f_4(y) = \frac{90}{1 + \exp\left(-\beta \ln\left(y\alpha\right) + 7.76\right)} + 4, \tag{4.3}$$

where in this section α is defined as,

$$\alpha = \frac{1}{V_i} + \frac{1}{Et_i}. \tag{4.4}$$

Raising the constant of the natural logarithm to a power and using the property of an exponential, $\exp\left(a + b\right) = \exp(a)\exp(b)$ gives,

$$f_4(y) = \frac{90}{1 + \exp\left(\ln\left(y^{-\beta}\alpha^{-\beta}\right)\right)\exp\left(7.76\right)} + 4. \tag{4.5}$$

Also, taking into account the cancellation of the exponential with the natural logarithm and reorganising produces,

$$f_4(y) = \frac{90}{1 + \frac{\exp(7.76)}{y^{\beta}\alpha^{\beta}}} + 4. \tag{4.6}$$

Multiplying through the fraction by y^{β} and further algebraic manipulation function f_4 is now in Hill function form,

$$f_4(y) = \frac{90y^{\beta}}{y^{\beta} + \left(\frac{\exp\left(\frac{7.76}{\beta}\right)}{\alpha}\right)^{\beta}} + 4. \tag{4.7}$$

From this the Hill coefficient, h_4, for function f_4, is equal to β from the original function

Function f_i	h_i	k_i
f_1	N/A	N/A
f_2	1.8	103.5
f_3	N/A	N/A
f_4	1.772	567.742
f_5	-8.54	26.726

Table 4.1: Initial estimates of the Hill coefficients from the process of fitting to the original functions.

and the parameter k_4 is defined as,

$$k_4 = \frac{\exp\left(\frac{7.76}{\beta}\right)}{\alpha} = \frac{\exp\left(\frac{7.76}{\beta}\right)}{\frac{1}{V_i} + \frac{1}{Et_i}}, \tag{4.8}$$

where $\beta = 1.77$ is a constant taken from Li et al (2006).

The analysis here proves that the Hill function can be derived directly from the exponential form of function f_4 as defined in Sturis et al (1991a). Each of the initial estimates of the Hill coefficients based on the model fitting are shown in Table 4.1.

Aside from the five functions in the model, $f_1, \ldots, f_5$, there are two other terms representing important elements of the regulation system. These are the insulin degradation rate for removal of insulin from the system, d_i, and the external glucose infusion rate, G_{in}. These two terms appear respectively in the insulin equation and the glucose equation of the model (4.1).

The insulin degradation rate has been studied in a number of papers particularly from the 1970s onwards, with early research such modelling insulin concentration with an insulin degrading compartment (Kitabchi et al 1971). Many developments exist and a comprehensive overview of insulin degradation and the methodology can be seen in Duckworth et al (1998). The cells and tissue of the liver and the kidneys are central to the degradation and clearance of inactive insulin to maintain a balance in the concentration within the sys-

Constant	Value	Units	Constant	Value	Units
R_m	210	min	V_i	11	l
V_g	10	l	E	0.2	l/min
U_b	72	mg/min	t_i	100	min
C_3	1000	mg/l	R_g	180	mg/min
U_0	40	mg/min	V_p	3	l
U_m	940	mg/min			

Table 4.2: Parameters values in model (4.1), from (Li et al 2006).

tem. This is an important aspect of glucose-insulin regulation and it has been modelled in different ways, but all models summarise results indicating that the insulin degradation rate is proportional to the level of insulin. The models of Sturis et al (1991a) and Bennett and Gourley (2004b) use the basis of a decaying half life whereas in Topp et al (2000) and Li et al (2006), a single positive parameter, d_i, represents rate of degradation.

4.2 Numerical simulations of oscillatory behaviour

One of the main requirements of the modified model is to reproduce the ultradian oscillations found in the glucose-insulin regulatory system. By modelling the full system with new Hill functions the desired ultradian oscillations can be produced. Figure 4.3 shows an example of the oscillatory behaviour under a constant glucose infusion rate of 1.35 mg/dl min, and an insulin degradation rate of 0.6 min^{-1} with the two time delays set at $\tau_1 = 6$ minutes and $\tau_2 = 36$ minutes. The Hill coefficients used to produce this graph are optimal values chosen to incorporate the results of Chapter 6, with ranges given in Table 6.1. The full set of parameter values for running this system can be seen in Table 4.2.

The value of d_i is not specifically set and can vary dependent on the patient. An analysis of varying this parameter by way of a phase plot can be seen in Figure 4.4. Results

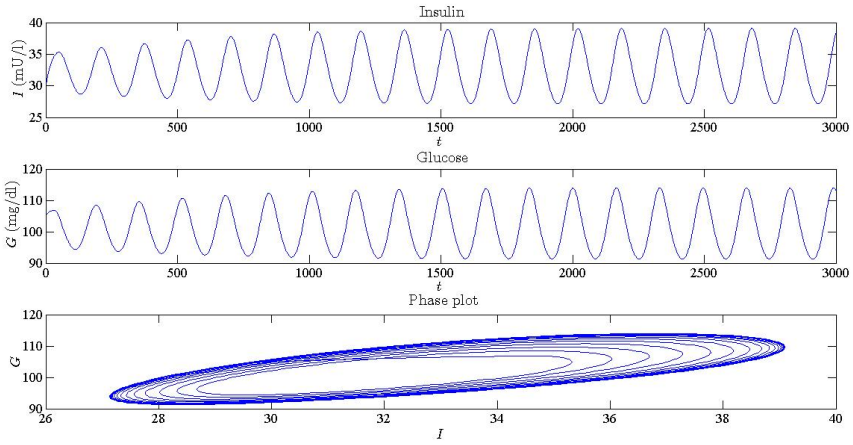

Figure 4.3: An example of ultradian oscillations of the glucose-insulin system using Hill functions. The top plot represents insulin concentration, the middle plot is glucose concentration and the bottom plot gives the phase plot for glucose and insulin.

in Li et al (2006) suggest a small value of insulin degradation rate of approximately $d_i = 0.06$ min^{-1}. Here, the rate is varied between $0.01 < d_i < 0.11$ in steps of 0.02.

The phase plots in Figure 4.4 show that a smaller insulin degradation rate gives a lower value for the glucose level. A lower degradation means there is more insulin in the system resulting in more utilisation of glucose, therefore, glucose does not reach higher levels. When the insulin degradation rate is high, the reverse of this is seen and an increased rate of degradation indicates a lower insulin level allowing the glucose to maintain higher concentrations. The system responds as expected to these conditions.

The parameter external glucose infusion, G_{in} represents glucose taken by the body from carbohydrates in equation $\dot{G}$ of model (4.1). To model the behaviour of the ultradian oscillations constant glucose infusion is used to allow the interaction of glucose and insulin to be shown clearly. Typical values for G_{in} range from 108 mg min^{-1} and 135 mg min^{-1} (Li et al 2006) to 216 mg min^{-1} (Sturis et al 1991a; Tolic et al 2000). Figure 4.5 shows the analysis of varying the glucose infusion rate such that, $1 < G_{in} < 251$, with a step size of 50.

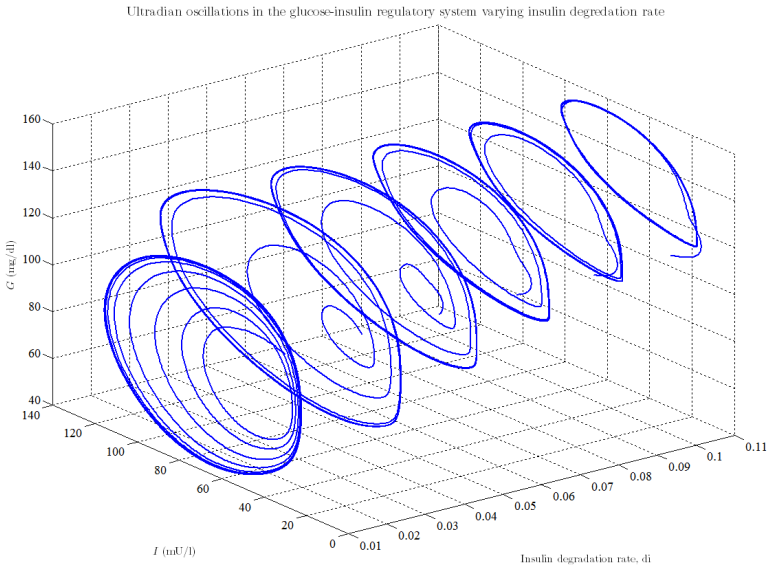

Figure 4.4: A graphical representation through the phase plot of glucose and insulin of the effect of altering the value of the parameter insulin degradation rate for $0.01 < d_i < 0.11$ in steps of 0.02. The values $\tau_1 = 6$ and $\tau_2 = 36$ were used for the delays.

It can be seen from the phase plots in Figure 4.5 that for lower values of external glucose infusion the concentrations of glucose and insulin drop very low in the system and as G_{in} increases, the system oscillates in a more uniform manner. A threshold value can be seen at which these oscillations stop and the solutions return to a single fixed point. This bifurcation point is in accordance with the analysis in Li et al (2006), showing that the modified model in this work can accurately reproduce expected behaviour. When studying the glucose-insulin regulation for an IVGTT, it is seen in published work that a high level of glucose infusion results in the system reducing to its basal level (Vicini et al 1997; Li et al 2012). Medically, a basal level is one which is the minimal necessary for life. This terminology is often used in diabetic studies for the minimum insulin needed to be supplied to diabetic patients.

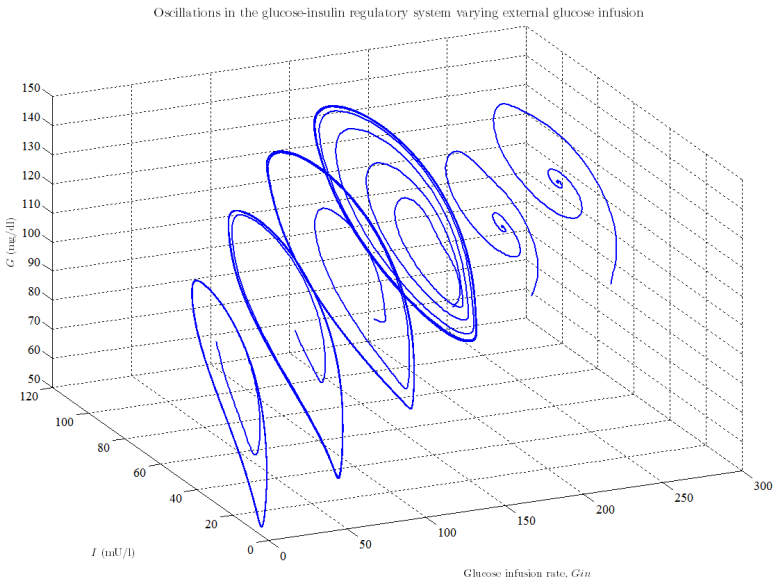

Figure 4.5: A graphical representation through the phase plot of glucose and insulin of the effect of altering the value of the parameter external glucose infusion rate for $1 < G_{in} < 251$ in steps of 50 measured in mg min^{-1}. The values $\tau_1 = 6$ and $\tau_2 = 36$ were used for the delays.

4.3 General effect of the Hill parameters

Although an explicit expression cannot be obtained, system (4.1) possesses a unique steady state (G^*, I^*), defined by the equations $\dot{G} = \dot{I} = 0$, which is always positive, as was shown in Bennett and Gourley (2004b). Since its value is influenced by the choice of the model parameters, the Hill coefficients h_i are constrained by the requirement that (G^*, I^*) be located in a physiologically feasible range. In order for the system to fit physiological values the fixed point for glucose, G^*, should be approximately within the range, $90 < G^* < 120$, and the fixed point for insulin, I^*, should approximately fit the range, $25 < I^* < 40$, (Sturis et al 1991a; Li et al 2006). Of the four Hill equations, the Hill coefficients h_1, h_4 and h_5 are investigated in this thesis for general effect on both fixed points. Parameters h_1 and h_4 are investigated because functions f_1 and f_4 are linked to the physiological mechanisms related to Type 1 and Type 2 diabetes respectively. Parameter h_5 is investigated as function f_5 contains the delay τ_2. As function f_2 represents

glucose utilisation independent of insulin, the parameters for this function should remain relatively constant and will not be directly investigated. General fixed values for the other parameters were taken from Table 4.1. Figure 4.6 shows the effect on the position of the fixed points as calculated for the model presented in this work for varying values of parameter h_1, while also altering h_4 which is represented by the vertical dots at each value of h_1. The corresponding graph for h_4 while altering h_1 is seen in Figure 4.7.

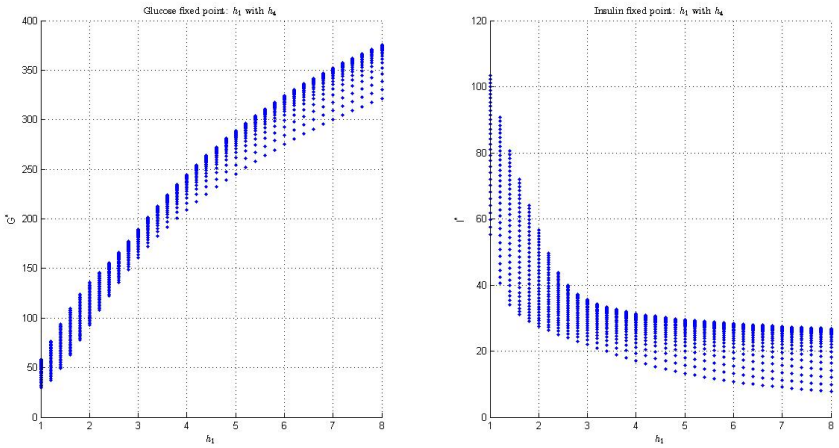

Figure 4.6: Effect on the fixed points of glucose (left) and insulin (right) of altering Hill coefficient h_1. The dots vertically for each value of h_1 represent the altering of h_4 such that a wider spread of points for one value of h_1 indicates a larger effect of h_4.

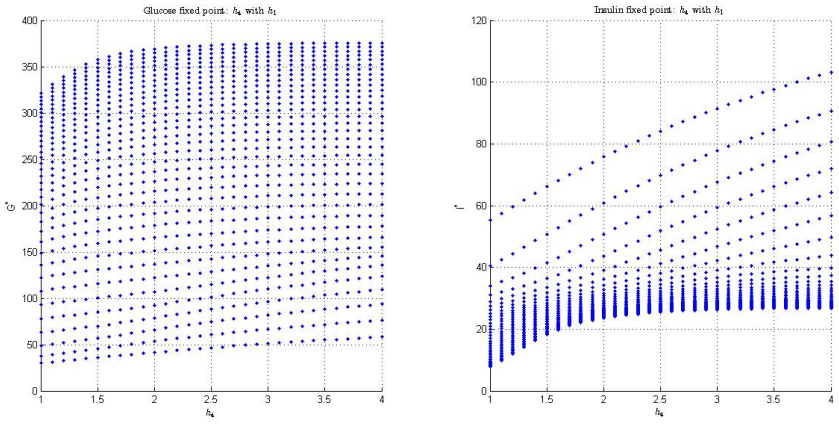

Figure 4.7: Effect on the fixed points of glucose (left) and insulin (right) of altering Hill coefficient h_4. The dots vertically for each value of h_4 represent the altering of h_1 such that a wider spread of points for one value of h_4 indicates a larger effect of h_1.

Figure 4.6 clearly displays a large effect of the value of h_1 on the fixed points, particularly for G^* where an increase in h_1 causes an increase in G^*. In this plot there is a relatively consistent and small range of values for h_4 for all values of h_1, suggesting h_1 has a more significant effect. This can be checked in the alternated graph in Figure 4.7 where h_4 is represented along the x-axis. For every value of h_4 there is a large effect by parameter h_1 on the fixed point. A similar set of results can be seen when comparing h_1 and h_5 as seen in Figures 4.8 and 4.9.

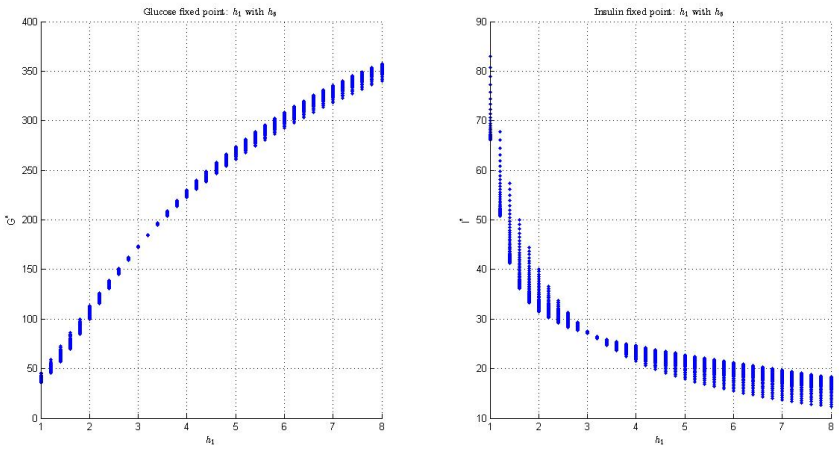

Figure 4.8: Effect on the fixed points of glucose (left) and insulin (right) of altering Hill coefficient h_1. The dots vertically for each value of h_1 represent the altering of h_5 such that a wider spread of points for one value of h_1 indicates a larger effect of h_5.

It is again clear that the parameter h_1 is dominant in affecting the position of the fixed point with a relatively small effect for h_5. Comparing Figures 4.6 and 4.8 it can be seen that the value of h_5 has a narrower range for in relation to h_1, particularly evident in the plots for I^*. The conclusion drawn from this is that the value of h_1 is the dominant Hill coefficient for the position of the fixed point, where the effect of h_5 is negligible.

This study shows that the position of the fixed points of the system are sensitive to the values of the Hill coefficients, in particular h_1. Therefore, this analysis can establish a range of possible values for the Hill parameters based on the fixed point. Further studies on the region of stability are made in Chapter 6.

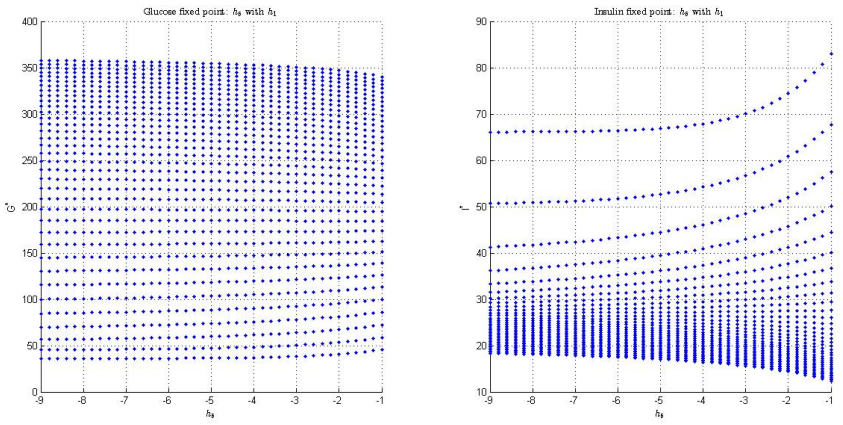

Figure 4.9: Effect on the fixed points of glucose (left) and insulin (right) of altering Hill coefficient h_5. The dots vertically for each value of h_5 represent the altering of h_1 such that a wider spread of points for one value of h_5 indicates a larger effect of h_1.

4.4 Linearisation and stability of the one-delay model

In order to study local stability it is necessary to obtain the linearised system, which will be introduced here, before analysing the model with one time delay. Both cases will be analysed by first setting $\tau_2 = 0$ and then separately $\tau_1 = 0$. This will provide conditions on local stability for the appearance of sustained ultradian oscillations in the one-delay cases. The local stability of model (4.1) was studied in Li et al (2006); Engelborghs et al (2001a) and Li and Kuang (2007) where several aspects were investigated with very generic assumptions on functions $f_1, \ldots, f_5$.

In matrix form, the linearisation of (4.1) at its unique steady state, (G^*, I^*), is given by

$$\begin{pmatrix} \dot{G}(t) \\ \dot{I}(t) \end{pmatrix} = \begin{pmatrix} -A & -B \\ 0 & -d_i \end{pmatrix} \begin{pmatrix} G(t) \\ I(t) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ D & 0 \end{pmatrix} \begin{pmatrix} G(t-\tau_1) \\ I(t-\tau_1) \end{pmatrix} + \begin{pmatrix} 0 & -C \\ 0 & 0 \end{pmatrix} \begin{pmatrix} G(t-\tau_2) \\ I(t-\tau_2) \end{pmatrix}, \tag{4.9}$$

where the following notation has been introduced,

$$\begin{aligned} &A = f_2'(G^*) + f_3'(G^*)f_4(I^*), \quad B = f_3(G^*)f_4'(I^*), \\ &C = -f_5'(I^*), \quad D = f_1'(G^*). \end{aligned} \tag{4.10}$$

To obtain the characteristic equation, one sets,

$$\begin{pmatrix} G(t) \\ I(t) \end{pmatrix} = \begin{pmatrix} G_0 \\ I_0 \end{pmatrix} e^{\lambda t}. \tag{4.11}$$

Rearranging by introduction of an identity matrix and dividing through by $e^{\lambda t}$, Equation (4.11) becomes

$$\begin{aligned} \left[\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -A & -B \\ 0 & -d_i \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ D & 0 \end{pmatrix} e^{-\lambda\tau_1} \right. \\ \left. - \begin{pmatrix} 0 & -C \\ 0 & 0 \end{pmatrix} e^{-\lambda\tau_2} \right] \begin{pmatrix} G_0 \\ I_0 \end{pmatrix} = 0. \end{aligned} \tag{4.12}$$

This gives,

$$\begin{pmatrix} \lambda + A & B + Ce^{-\lambda\tau_2} \\ -De^{-\lambda\tau_1} & \lambda + d_i \end{pmatrix} \begin{pmatrix} G_0 \\ I_0 \end{pmatrix} = 0. \tag{4.13}$$

For a non trivial solution to exist, it is required that the determinant vanishes, i.e.

$$(\lambda + A)(\lambda + d_i) - (B + Ce^{-\lambda\tau_2})(-De^{-\lambda\tau_1}) = 0. \tag{4.14}$$

This leads to the characteristic equation which is given as,

$$\lambda^2 + (A + d_i)\lambda + B\,D\,e^{-\lambda\tau_1} + C\,D\,e^{-\lambda(\tau_1+\tau_2)} + d_iA = 0. \tag{4.15}$$

The one-delay case, when $\tau_1\tau_2 = 0$, has been widely studied both for local and global stability. It has been shown in Engelborghs et al (2001a) and Li and Kuang (2007) for this situation that an explicit expression can be obtained for a critical value of the delay, above which the system produces sustained oscillations. Here, some of these results are recovered for the one-delay case by analysing Equation (4.15).

For the condition of local stability it must be seen that all λ's have negative real part. In the case of a single delay when $\tau_2 = 0$, the characteristic equation becomes,

$$\lambda^2 + (A + d_i)\lambda + (B\,D\, + C\,D)e^{-\lambda\tau_1} + d_iA = 0. \tag{4.16}$$

Introducing $\lambda = i\omega_1$, $\omega_1 > 0$, in order to find the crossing point on the real axis and using Euler's formula, $e^{-ix} = \cos(x) - i\sin(x)$, the following is obtained,

$$\begin{aligned} -\omega_1^2 + (BD + CD)\cos(\omega_1\tau_1) + d_iA \\ + i\big((A + d_i)\omega_1 - (BD + CD)\sin(\omega_1\tau_1)\big) = 0. \end{aligned} \tag{4.17}$$

The real and imaginary parts can be taken separately equal to zero such that,

$$\begin{aligned} -\omega_1^2 + (BD + CD)\cos(\omega_1\tau_1) + d_iA = 0, \\ (A + d_i)\omega_1 - (BD + CD)\sin(\omega_1\tau_1) = 0. \end{aligned} \tag{4.18}$$

Rearranging each equation for cosine and sine respectively gives,

$$\cos(\omega_1\tau_1) = \frac{\omega_1^2 - d_iA}{D(B + C)}, \quad \sin(\omega_1\tau_1) = \frac{(A + d_i)\omega_1}{D(B + C)}. \tag{4.19}$$

It is now possible to obtain an equation for τ_1 from the result in Equation (4.19), however the parameter ω_1 remains unknown. The two equations can be used to obtain a function for ω_1. First using the trigonometric identity, $\cos^2 + \sin^2 = 1$, and rearranging for ω_1^4, a

fourth order polynomial is obtained,

$$\omega_1^4 + (A^2 + d_i^2)\omega_1^2 + A^2 d_i^2 - (D(B+C))^2 = 0. \tag{4.20}$$

Equation (4.19) implies that ω_1 is a positive root for the fourth order polynomial seen above. As this is a quadratic equation in ω_1^2, the quadratic formula can be used to give an expression for ω_1^2,

$$\omega_1^2 = \frac{-(A^2 + d_i^2) \pm \sqrt{(A^2 + d_i^2)^2 - 4\left[A^2 d_i^2 - (D(B+C))^2\right]}}{2}. \tag{4.21}$$

Hence, ω_1 is defined by taking the positive branch of,

$$\omega_1 = \pm\frac{1}{\sqrt{2}}\sqrt{-(A^2 + d_i^2) \pm \sqrt{(A^2 - d_i^2)^2 + 4D^2(B+D)^2}}. \tag{4.22}$$

Therefore, from the first equation in (4.19) the following equation for the threshold value for oscillations for delay τ_1, denoted by τ_1^*, is obtained

$$\tau_1^* = \frac{1}{\omega_1}\arccos\left(\frac{\omega_1^2 - d_i A}{D(B+C)}\right), \tag{4.23}$$

where ω_1 is given in Equation (4.22) and A, B, C, D are defined in Equation (4.10). Thus, Equations (4.22) and (4.23) characterise the apparition of sustained oscillations for the system with one time delay, τ_1.

The same process can be carried out for the system when $\tau_1 = 0$ to obtain an expression for the threshold value to produce oscillations for the second time delay, τ_2^*, briefly explained here. The characteristic equation (4.15) with $\tau_1 = 0$, becomes,

$$\lambda^2 + (A + d_i)\lambda + C\,D\,e^{-\lambda\tau_2} + B\,D\, + d_i A = 0. \tag{4.24}$$

Again introducing, $\lambda = i\omega_2$, $\omega_2 > 0$, and using Euler's formula gives,

$$-\omega_2^2 + C\,D\,\cos(\omega_2\tau_2) + B\,D\, + d_i\,A + i((A + d_i)\omega_2 + C\,D\,\sin(\omega_2\tau_2)) = 0. \quad (4.25)$$

This leads to the equations in cosine and sine as seen below,

$$\cos(\omega_2\tau_2) = \frac{\omega_2^2 - (B\,D + d_iA)}{C\,D}, \quad \sin(\omega_2\tau_2) = \frac{\omega_2(A + d_i)}{C\,D}, \quad (4.26)$$

implying that ω_2 is a solution of the following quadratic,

$$\omega_2^4 + (A^2 + d_i^2 - 2B\,D)\omega_2^2 + (B\,D + A\,d_i)^2 - (C\,D)^2 = 0. \quad (4.27)$$

Applying the quadratic formula in ω_2^2 and rearranging produces the equation for ω_2,

$$\omega_2 = \pm\frac{1}{\sqrt{2}}\sqrt{2B\,D - (A^2 + d_i^2) \pm \sqrt{(A^2 - d_i^2)^2 + 4\,[(C\,D)^2 - B\,D(A + d_i)^2]}},$$

where again only the positive branch is considered. A value of τ_2^* is then obtained from (4.26),

$$\tau_2^* = \frac{1}{\omega_2}\arccos\left(\frac{\omega_2^2 - (B\,D + d_iA)}{C\,D}\right), \quad (4.28)$$

which gives the equations that characterise the emergence of sustained oscillations for the system with one delay, τ_2. These results for τ_1^* and τ_2^* recover and confirm the analysis for the one delay case seen in Li and Kuang (2007), which indicate that both delays are important in maintaining sustained oscillations in the glucose-insulin regulation system.

4.5 Discussion

In this chapter, the new modified model of glucose-insulin regulation has been introduced with the discussion of the new Hill functions and the manner of their application. The Hill functions have been shown to be an appropriate choice for the regulatory mechanisms of the system by producing the desired behaviour under variation of the glucose infusion rate and the insulin degradation. The modified model has also been shown to provide the desired oscillations of glucose and insulin in physiological ranges.

An analysis of the Hill parameters in relation to their effect on the fixed points of glucose and insulin has been made. This provides some initial ranges and conditions for the most appropriate values of the Hill coefficients. Further analysis based on stability region is made in Chapter 6. Further to this, a local stability analysis in the one delay cases has been made which recovers previously published work (Li and Kuang 2007), and leads to the following chapters for stability in the system with two delays.

Chapter 5

Local stability analysis of the two-delay system

This chapter investigates local stability of the system with two delays via two approaches. First, an extension of the one-delay analysis from Chapter 4 is made by assessing the characteristic equation for the general case when $\tau_1\tau_2 > 0$. This will produce conditions on both delays in order to find the separatrix for the region of asymptotic stability and the region of oscillations which is numerically confirmed using a LMI method. Secondly, a Lyapunov functional approach will be taken analytically for both the delay-independent and delay-dependent cases. The Lyapunov functional will be created from a general base function including a term to account for the interaction between the two delays.

5.1 Characteristic equation approach for two delays

Here, the analysis of stability for the one-delay model seen in Section 4.4 is extended for the two delay case. The characteristic equation of the linearised model (4.9), as seen in equation (4.15), will be investigated in the general case where both delays are nonzero. Previously it was shown in Li and Kuang (2007) that if $d_i A \geq D(B+C)$, then for any given $\tau_1 > 0$, there exists $\tau_2(\tau_1) > 0$ such that a Hopf bifurcation occurs at (τ_1, τ_2). By looking for pure imaginary solutions $\lambda = i\omega$, $\omega > 0$, the separatrix curve can be derived. Using the same methodology as for the one-delay case gives the following system,

$$\begin{aligned} \omega^2 - Ad_i + D\left[B\cos(\tau_1\omega) + C\cos((\tau_1+\tau_2)\omega)\right] = 0, \\ (A+d_i)\omega = D\left[B\sin(\tau_1\omega) + C\sin((\tau_1+\tau_2)\omega)\right]. \end{aligned} \tag{5.1}$$

By making use of the trigonometric identities for $\cos(x+y)$ and $\sin(x+y)$, while expanding in order to isolate the terms in τ_2, the subsequent terms are obtained,

$$\begin{aligned} \cos(\tau_2\omega) &= \frac{(\omega^2 - Ad_i)\cos(\tau_1\omega) + \omega(A+d_i)\sin(\tau_1\omega) - BD}{CD}, \\ \sin(\tau_2\omega) &= \frac{(Ad_i - \omega^2)\sin(\tau_1\omega) + \omega(A+d_i)\cos(\tau_1\omega)}{CD}. \end{aligned} \tag{5.2}$$

By applying the trigonometric identity $\cos^2 + \sin^2 = 1$, it is seen that ω therefore satisfies the following transcendental equation,

$$\begin{aligned} \omega^4 + \omega^2\left(A^2 - 2BD\cos(\tau_1\omega) + d_i^2\right) + A^2d_i^2 - 2BD\omega(A+d_i)\sin(\tau_1\omega) \\ +2ABDd_i\cos(\tau_1\omega) + D^2(B^2 - C^2) = 0. \end{aligned} \tag{5.3}$$

Solving (5.3) for $\sin(\tau_1\omega)$ gives the following expression,

$$\sin(\tau_1\omega) = \frac{\omega^4 + \omega^2\left(A^2 - 2BD\cos(\omega\tau_1) + d_i^2\right) + A^2d_i^2}{2BD\omega(A+d_i)}, \tag{5.4}$$

which when introduced into (5.2), gives the following expression,

$$\cos(\tau_2\omega) = \frac{(A^2+\omega^2)(d_i^2+\omega^2) - D^2(B^2+C^2)}{2BCD^2}. \tag{5.5}$$

This depends on τ_1 only implicitly through ω. Therefore, for a fixed value of τ_1, the transcendental equation (5.3) can be numerically solved for ω. Then by using Equation (5.5), the critical value for τ_2 is recovered from,

$$\tau_2^* = \frac{1}{\omega}\arccos\left[\frac{(A^2+\omega^2)(d_i^2+\omega^2) - D^2(B^2+C^2)}{2BCD^2}\right]. \tag{5.6}$$

Using this approach, a graphical representation of the curve delimiting zones of asymptotic stability and oscillatory behaviour can be obtained as seen in Figure 5.1. This shows for the first time that the boundary is not a straight line but is actually a curve. The boundary can be approximated by the line joining the two individual thresholds from the one-delay cases, τ_1^* and τ_2^* as was first suggested in the numerical experiments of Li and Kuang (2007). To produce Figures 5.1 and Figure 5.2, the following generic values are used,

$$A = 0.0101, \quad B = 0.8549, \quad C = 2.4681, \quad D = 0.0011, \tag{5.7}$$

which correspond to the choice of Hill parameters,

$$\begin{aligned} &h_1 = 2, \quad k_1 = 6000, \quad h_2 = 1.8, \quad k_2 = 103.5, \\ &h_4 = 1.5, \quad k_4 = 565, \quad h_5 = -8.54, \quad k_5 = 26.7. \end{aligned} \tag{5.8}$$

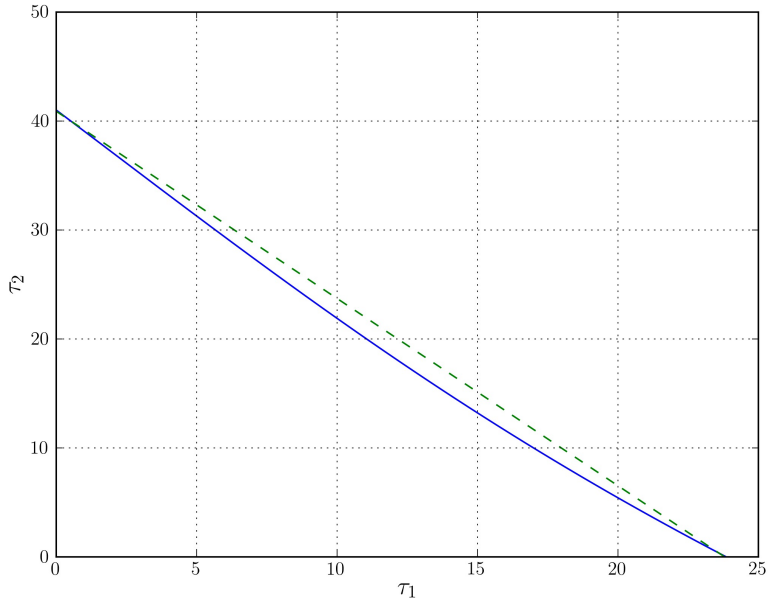

Figure 5.1: Threshold curve in the $\tau_1 - \tau_2$ plane with accompanying straight line approximation of the two individual delay thresholds.

In addition, the derivative of ω with respect to τ_1 is given by

$$\frac{d\omega}{d\tau_1} = \frac{\Delta}{\omega(A^2 + d_i^2 + 2\omega^2) - BD\left[(A + d_i)\sin(\tau_1\omega) + 2\omega\cos(\tau_1\omega)\right] - \tau_1\Delta}, \tag{5.9}$$

where

$$\Delta = BD\omega\left(\left(Ad_i - \omega^2\right)\sin(\tau_1\omega) + \omega(A + d_i)\cos(\tau_1\omega)\right), \tag{5.10}$$

which is equal to zero if and only if

$$\left(Ad_i - \omega^2\right)\sin(\tau_1\omega) + \omega(A + d_i)\cos(\tau_1\omega) = 0.$$

Comparing with (5.2), it can be seen that $\frac{d\omega}{d\tau_1} = 0$ if and only if $\sin(\tau_2\omega) = 0$. Further-

more, the second derivative at such points is

$$\frac{d^2\omega}{d{\tau_1}^2} = -\frac{B\omega\left(A^2+\omega^2\right)\left(d_i^2+\omega^2\right)}{C\left(A^2+d_i^2+2\omega^2\right)}. \tag{5.11}$$

As ω is assumed to be positive, this shows that ω attains locally maximal values at points where $\sin(\tau_2\omega) = 0$. The first positive value of τ_2 for which this occurs is $\tau_2 = 0$, this coincides with the one-delay case as seen in Section 4.4. As ω is maximal when $\tau_2 = 0$, this indicates that the delay in glucose production by the liver leads to a longer period in the regulation loop. Thus, these results further extend the knowledge mathematically while adhering to the expected behaviour from a biological view point.

5.1.1 Numerical investigation of asymptotic stability

The boundary of the asymptotic stability region can also be approached numerically by making use of Linear Matrix Inequalities (LMIs) as discussed in Chapter 3. The method of He et al (2006) numerically studies an appropriately designed Lyapunov functional for the linear system dependent on free weighting matrices.

This method was applied for the linearised system (4.9) using the parameter values as given in (5.7). The solution scheme, which relies on interior fixed point methods, was implemented using the Robust Control Toolbox in Matlab (MATLAB 2013). As a result, a plot of the region of asymptotic stability in the (τ_1, τ_2) plane has been obtained. The results are depicted in Fig. 5.2 where an almost linear pattern can be seen distinguishing between the area of asymptotic stability and the area of oscillations. It shows that the LMI optimisation based on the Lyapunov functional (3.20) allows for better results in the two-delay setting than in its one-delay limiting subcases. It appears to be remarkable that the numerical scheme approximates the stability region very accurately in the physiological range when $5 < \tau_1 < 20$.

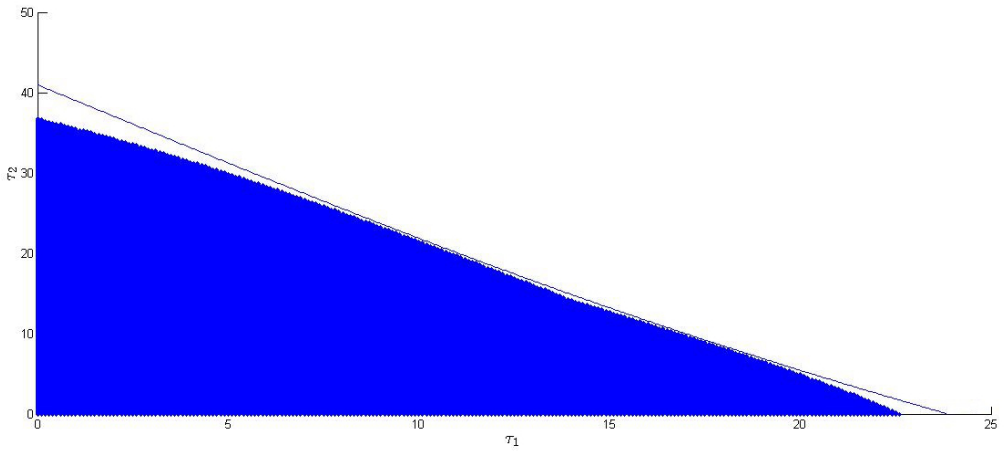

Figure 5.2: Plot of the region of asymptotic stability using the LMI method for $0 < \tau_1 < 30$ and $0 < \tau_2 < 50$. The solid line represents the theoretical curve, as obtained in Section 5.1.

5.2 Analytical Lyapunov functional approach

Following the LMI method for numerically studying stability based on a Lyapunov functional, an analytical investigation is made by constructing a general Lyapunov functional. It is possible to investigate the stability for both cases; delay-independent and delay-dependent. Studying the delay-independent case is much simpler, while still possible to obtain useful results regarding the system. However, it can be very restrictive and no conclusions can be drawn regarding the relationship of delays in a system. Both the delay-independent and the delay-dependent approaches will be taken to study local stability.

5.2.1 Delay-independent local stability

Here, the delay-independent local stability will be investigated by taking a Lyapunov functional approach using matrix form and the Schur complement for positive definiteness (Boyd and Vandenberghe 2009; Gallier 2010).

The Schur complement can be used for conditions of positive definiteness for a block

matrix, which in the case in this analysis, a 4×4 matrix can be written as a 2×2 matrix of four 2×2 blocks such that

$$M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}. \tag{5.12}$$

The Schur complement, S, of this matrix is denoted as $S = C - B^T AB$ (Gallier 2010). The conditions for positive definiteness using the Schur complement state that a symmetric block matrix, M, is positive definite if and only if $A > 0$ and $C - B^T AB > 0$. The conditions for semi positive definite allow $S \geq 0$. However, for the special case where C is a zero matrix the conditions become,

$$\text{If } C \geq 0 \text{ then } (I - CC^{\dagger})B^T = 0, \quad A - BC^{\dagger}B^T \geq 0, \tag{5.13}$$

where † represents the pseudo-inverse of a matrix, a generalised inverse of a singular matrix which can be used to solve systems of linear equations (Radhakrishna Rao and Mitra 1972). It is clear that in the case where $C = 0$, that it is required that $B^T = 0$ and hence $B = 0$ to have positive definiteness (Gallier 2010).

Throughout this analysis the notation for the variables with respect to time t, $u(t)$ and $v(t)$ will be written as u and v respectively. The equation to be used as the initial function for this analysis is,

$$V = \frac{1}{2}(\rho_1 u^2 + \rho_2 v^2 + 2\rho_3 uv), \qquad \rho_i > 0, \qquad i = 1, 2, 3. \tag{5.14}$$

The derivative of V is given by,

$$\frac{dV}{dt} = \rho_1 u\dot{u} + \rho_2 v\dot{v} + \rho_3(u\dot{v} + v\dot{u}) = (\rho_1 u + \rho_3 v)\dot{u} + (\rho_2 v + \rho_3 u)\dot{v}. \tag{5.15}$$

Introducing the linearised forms of $\dot{u}$ and $\dot{v}$ from (4.9) and expanding for terms in u^2, v^2 and uv gives,

$$\begin{aligned}\frac{dV}{dt} &= -u^2[\rho_1 A] - uv[\rho_1 B + \rho_3 A + \rho_3 d_i] - v^2[\rho_3 B + \rho_2 d_i] \\ &\quad - C(\rho_1 u + \rho_3 v)v(t-\tau_2) + D(\rho_2 v + \rho_3 u)u(t-\tau_1).\end{aligned} \tag{5.16}$$

When written in matrix form this is,

$$\frac{dV}{dt} \leq -\begin{bmatrix} u(t) \\ v(t) \\ u(t-\tau_1) \\ v(t-\tau_2) \end{bmatrix}^T \underbrace{\begin{bmatrix} \rho_1 A & \Phi_{12} & -\frac{1}{2}D\rho_3 & -\frac{1}{2}C\rho_1 \\ \Phi_{12} & \Phi_{22} & -\frac{1}{2}D\rho_2 & -\frac{1}{2}C\rho_3 \\ -\frac{1}{2}D\rho_3 & -\frac{1}{2}D\rho_2 & 0 & 0 \\ -\frac{1}{2}C\rho_1 & -\frac{1}{2}C\rho_3 & 0 & 0 \end{bmatrix}}_{\Phi_M} \begin{bmatrix} u(t) \\ v(t) \\ u(t-\tau_1) \\ v(t-\tau_2) \end{bmatrix}, \tag{5.17}$$

where $\Phi_{12} = \frac{1}{2}(\rho_1 B + \rho_3 A + \rho_3 d_i)$ and $\Phi_{22} = \rho_3 B + \rho_2 d_i$ are introduced. The system in (5.17) gives a linear matrix inequality which is required to be less than zero to have stability. Thus, the matrix Φ_M must be positive. The Schur complement can be used for conditions of positive definiteness for block matrices and thus writing matrix Φ_M in the form specified in (5.12) gives the following,

$$A = \begin{bmatrix} \rho_1 A & \Phi_{12} \\ \Phi_{12} & \Phi_{22} \end{bmatrix}, \quad B = \begin{bmatrix} -\frac{1}{2}D\rho_3 & -\frac{1}{2}C\rho_1 \\ -\frac{1}{2}D\rho_2 & -\frac{1}{2}C\rho_3 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

As the conditions in the special case when $C = 0$, as seen in Equation (5.13), require that $B = 0$, it is clear that it is not possible to obtain conditions for delay-independent stability by using this approach.

5.2.2 Delay-dependent local stability

The delay-dependent local stability will be studied analytically to obtain conditions related to the both delay terms, τ_1 and τ_2. To do this, the results of Bennett and Gourley (2004b) are generalised by extending their analysis for two delays. The form of the initial functional will be the same as that of the delay-independent case, including the term $\rho_3 uv$ to account for the interaction between the delay terms.

The calculation will begin with the initial Lyapunov candidate, working through for the linearised system (4.9) and explicitly introducing the delay terms. Some useful standard rules and propositions will be discussed here which will be useful in the calculations to obtain the Lyapunov functional.

Two different forms of Young's inequality will be used within this calculation (Young 1912; Beckner 1975). The standard Young's inequality with exponent two is seen in Lemma 1.

Lemma 1. *For any $\epsilon > 0$, it is given that*

$$\pm xy \leq \frac{1}{2}\left(\epsilon x^2 + \frac{1}{\epsilon}y^2\right). \tag{5.18}$$

Proof. Writing $\pm xy - \frac{1}{2}\left(\epsilon x^2 + \frac{1}{\epsilon}y^2\right) = -\frac{1}{2\epsilon}\left(\epsilon x \mp y\right)^2 \leq 0,$
the result is immediate. □

Also, a special case of Young's inequality with a free parameter a, as introduced in Park (1999), will be used. This is seen in Lemma 2.

Lemma 2. *For any $\epsilon > 0$, where constant a is free, it is given that*

$$\pm xy \leq \frac{1}{2}\left(\epsilon(x + ay)^2 + \frac{1}{\epsilon}y^2\right) \mp ay^2. \tag{5.19}$$

Proof. Writing Young's inequality as seen in Lemma 1 but replacing x with $x + ay$ produces $(x + ay)y \leq \frac{1}{2}\left(\epsilon(x + ay)^2 + \frac{1}{\epsilon}y^2\right)$. The rearrangement provides an immediate proof by following that of Lemma 1. □

Using the form of Young's inequality including the parameter a as in Lemma 2, allows for a greater level of freedom in finding the best conditions for stability. The restrictions on the other parameters introduced to the functional all require strictly positive constants. The inclusion of a free parameter which can be negative gives the opportunity to provide a balance to the restrictions of positive terms included, resulting in a better Lyapunov functional.

Also useful will be the double Leibniz rule (Abramowitz and Stegun 1972), given by Equation (5.21), detailed below,

Leibniz rule for the derivative of an integral with variable limits

$$\frac{d}{dt}\int_{f_1(t)}^{f_2(t)} g(t,z)dz = \int_{f_1(t)}^{f_2(t)} g_t(t,z)dz + g(t,f_2(t))f_2'(t) - g(t,f_1(t))f_1'(t) \tag{5.20}$$

This can be applied to a double integral to give

$$\frac{d}{dt}\int_{t-\tau}^{t}\int_{z}^{t} g(s)dsdz = \tau g(t) - \int_{t-\tau}^{t} g(s)ds.$$

Indeed, by denoting $H(t,z) = \int_z^t g(s)ds$, so that

$$\begin{aligned}\frac{d}{dt}\int_{t-\tau}^{t}\int_{z}^{t} g(s)dsdz = \frac{d}{dt}\int_{t-\tau}^{t} H(t,z)dz &= \int_{t-\tau}^{t}\frac{d}{dt}H(t,z)dz + H(t,t) - H(t,t-\tau)\\ &= \int_{t-\tau}^{t} g(t)dz + \int_{t}^{t} g(s)ds - \int_{t-\tau}^{t} g(s)ds\\ &= \tau g(t) - \int_{t-\tau}^{t} g(s)ds\end{aligned}$$

Therefore, it can be written that

$$\int_{t-\tau}^{t} g(s)ds = \tau g(t) - \frac{d}{dt}\int_{t-\tau}^{t}\int_{z}^{t} g(s)\,dsdz, \tag{5.21}$$

which defines the double integral Leibniz rule.

The following theorem is an original contribution of this thesis.

Theorem 1. *The functional*

$$\begin{aligned}
V_1 = {} & \frac{1}{2}(\rho_1 u^2 + \rho_2 v^2 + 2\rho_3 uv) + \frac{CD}{2}\epsilon_1 \int_{t-\tau_1}^{t}\int_{z}^{t} (\rho_1 u + \rho_3 v + a_1 u(s-\tau_1))^2 dsdz \\
& + \frac{CD}{2}\left(\frac{1}{\epsilon_1} - 2a_1\right)\int_{t-\tau_1}^{t}\int_{z}^{t} u^2(s-\tau_1)dsdz \\
& + \frac{Cd_i}{2}\left[\epsilon_2 \int_{t-\tau_1}^{t}\int_{z}^{t} (\rho_1 u + \rho_3 v + a_2 v(s))^2 dsdz + \left(\frac{1}{\epsilon_2} + 2a_2\right)\int_{t-\tau_1}^{t}\int_{z}^{t} v^2(s)dsdz\right] \\
& + \frac{DA}{2}\left[\epsilon_3 \int_{t-\tau_2}^{t}\int_{z}^{t} (\rho_2 v + \rho_3 u + a_3 u(s))^2 dsdz + \left(\frac{1}{\epsilon_3} - 2a_3\right)\int_{t-\tau_2}^{t}\int_{z}^{t} u^2(s)dsdz\right] \\
& + \frac{DB}{2}\left[\epsilon_4 \int_{t-\tau_2}^{t}\int_{z}^{t} (\rho_2 v + \rho_3 u + a_4 v(s))^2 dsdz + \left(\frac{1}{\epsilon_4} - 2a_4\right)\int_{t-\tau_2}^{t}\int_{z}^{t} v^2(s)dsdz\right] \\
& + \frac{DC}{2}\epsilon_5 \int_{t-\tau_2}^{t}\int_{z}^{t} (\rho_2 v + \rho_3 u + a_5 v(s-\tau_2))^2 dsdz \\
& + \frac{DC}{2}\left(\frac{1}{\epsilon_5} - 2a_5\right)\int_{t-\tau_2}^{t}\int_{z}^{t} v^2(s-\tau_2)dsdz \\
& + \frac{CD}{2}\tau_2\left(\frac{1}{\epsilon_1}(\epsilon_1 a_1 - 1)^2 + \frac{\epsilon_1 a_1}{\epsilon_6}\right)\int_{t-\tau_1}^{t} u^2(s)ds \\
& + \frac{DC}{2}\tau_1\left(\frac{1}{\epsilon_5}(\epsilon_5 a_5 - 1)^2 + \frac{\epsilon_5 a_5}{\epsilon_7}\right)\int_{t-\tau_2}^{t} v^2(s)ds,
\end{aligned} \tag{5.22}$$

is a Lyapunov functional for the linearised dynamical system seen in (4.9) if the parameter values are such that inequalities (5.40) - (5.46) are satisfied, and such that $\Psi_{11} > 0$ *and* $\Psi_{11}\Psi_{22} - \Psi_{12}^2 > 0$*, where the* Ψ_{ij} *are defined in Equations (5.35) - (5.37).*

The functional used as the base for this theorem will be the same as in the delay-independent

case, (5.14) which will be stated again for completeness and clarity here. Again, the notation $u = u(t)$ and $v = v(t)$ will be used and the dependence on t will only be stated when necessary.

Proof. Consider the functional V_0 as,

$$V_0 = \frac{1}{2}(\rho_1 u^2 + \rho_2 v^2 + 2\rho_3 uv), \qquad \rho_i > 0, \qquad i = 1, 2, 3. \tag{5.23}$$

and hence the derivative is given by,

$$\frac{dV_0}{dt} = (\rho_1 u + \rho_3 v)\dot{u} + (\rho_2 v + \rho_3 u)\dot{v}. \tag{5.24}$$

In this calculation, the terms $(\rho_1 u + \rho_3 v)$ and $(\rho_2 v + \rho_3 u)$ will be left in their bracketed form and treated as one term throughout until their expansion is required. Introducing the linearised form of $\dot{u}$ and $\dot{v}$ and separating the delayed variables from the non-delayed variables give,

$$\begin{aligned}\frac{dV_0}{dt} = &-(\rho_1 u + \rho_3 v)(Au + Bv) - d_i v(\rho_2 v + \rho_3 u) - C(\rho_1 u + \rho_3 v)v(t - \tau_2) \\ &+ D(\rho_2 v + \rho_3 u)u(t - \tau_1).\end{aligned} \tag{5.25}$$

Ghost terms can be introduced to the delayed terms as $[u(t - \tau_1) - u(t) + u(t)]$ and $[v(t - \tau_2) - v(t) + v(t)]$. This allows the inclusion of integrals as seen below,

$$\begin{aligned}\frac{dV_0}{dt} = &-(\rho_1 u + \rho_3 v)(Au + Bv + Cv) + (Du - d_i v)(\rho_2 v + \rho_3 u) \\ &+ C(\rho_1 u + \rho_3 v)\int_{t-\tau_2}^{t} \dot{v}(s)ds - D(\rho_2 v + \rho_3 u)\int_{t-\tau_1}^{t} \dot{u}(s)ds.\end{aligned} \tag{5.26}$$

The expressions for $\dot{u}$ and $\dot{v}$ need to be reintroduced again here, while bringing all possible

terms inside the integrals,

$$\begin{aligned}\frac{dV_0}{dt} &= -(\rho_1 u + \rho_3 v)(Au + Bv + Cv) + (Du - d_i v)(\rho_2 v + \rho_3 u) \\ &+ C\int_{t-\tau_2}^{t} [D(\rho_1 u + \rho_3 v)u(s-\tau_1) - d_i(\rho_1 u + \rho_3 v)v(s)]ds \\ &+ D\int_{t-\tau_1}^{t} [A(\rho_2 v + \rho_3 u)u(s) + B(\rho_2 v + \rho_3 u)v(s) + C(\rho_2 v + \rho_3 u)v(s-\tau_2)]ds.\end{aligned} \tag{5.27}$$

At this point the special case of Young's inequality, as seen in Lemma 2, is applied. It can be noted here that a simpler but more restrictive Lyapunov functional could be created by using the original form of Young's inequality where $a = 0$. However the inclusion of this parameter gives more freedom to the functional. Young's inequality is applied to the terms in the integrals to give,

$$\begin{aligned}\frac{dV_0}{dt} &= -(\rho_1 u + \rho_3 v)(Au + Bv + Cv) + (Du - d_i v)(\rho_2 v + \rho_3 u) \\ &+ \frac{CD}{2}\int_{t-\tau_2}^{t} \left[\epsilon_1 \left((\rho_1 u + \rho_3 v) + a_1 u(s-\tau_1)\right)^2 + \left(\frac{1}{\epsilon_1} - 2a_1\right) u^2(s-\tau_1)\right] ds \\ &+ \frac{Cd_i}{2}\int_{t-\tau_2}^{t} \left[\epsilon_2 \left((\rho_1 u + \rho_3 v) + a_2 v(s)\right)^2 + \left(\frac{1}{\epsilon_2} + 2a_2\right) v^2(s)\right] ds \\ &+ \frac{DA}{2}\int_{t-\tau_1}^{t} \left[\epsilon_3 \left((\rho_2 v + \rho_3 u) + a_3 u(s)\right)^2 + \left(\frac{1}{\epsilon_3} - 2a_3\right) u^2(s)\right] ds \\ &+ \frac{DB}{2}\int_{t-\tau_1}^{t} \left[\epsilon_4 \left((\rho_2 v + \rho_3 u) + a_4 v(s)\right)^2 + \left(\frac{1}{\epsilon_4} - 2a_4\right) v^2(s)\right] ds \\ &+ \frac{DC}{2}\int_{t-\tau_1}^{t} \left[\epsilon_5 \left((\rho_2 v + \rho_3 u) + a_5 v(s-\tau_2)\right)^2 + \left(\frac{1}{\epsilon_5} - 2a_5\right) v^2(s-\tau_2)\right] ds.\end{aligned} \tag{5.28}$$

Equation (5.28) can be rearranged by separating out the integral terms to allow the use of the identity of the double Leibniz rule as stated in Equation (5.21). By applying this

definite integral terms are created such that,

$$\begin{aligned}
\frac{dV_0}{dt} &= -(\rho_1 u + \rho_3 v)(Au + Bv + Cv) + (Du - d_i v)(\rho_2 v + \rho_3 u) \\
&+ \frac{CD}{2}\epsilon_1 \left[\tau_2 \left(\rho_1 u + \rho_3 v + a_1 u(t-\tau_1)\right)^2 \right. \\
&\left. - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} \left(\rho_1 u + \rho_3 v + a_1 u(s-\tau_1)\right)^2 dsdz\right] \\
&+ \frac{CD}{2}\left(\frac{1}{\epsilon_1} - 2a_1\right)\left[\tau_2 u^2(t-\tau_1) - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} u^2(s-\tau_1)dsdz\right] \\
&+ \frac{Cd_i}{2}\epsilon_2 \left[\tau_2 \left(\rho_1 u + \rho_3 v + a_2 v(t)\right)^2 - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} \left(\rho_1 u + \rho_3 v + a_2 v(s)\right)^2 dsdz\right] \\
&+ \frac{Cd_i}{2}\left(\frac{1}{\epsilon_2} + 2a_2\right)\left[\tau_2 v^2(t) - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} v^2(s)dsdz\right] \\
&+ \frac{DA}{2}\epsilon_3 \left[\tau_1 \left(\rho_1 v + \rho_3 u + a_3 u(t)\right)^2 - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} \left(\rho_2 v + \rho_3 u + a_3 u(s)\right)^2 dsdz\right] \\
&+ \frac{DA}{2}\left(\frac{1}{\epsilon_3} - 2a_3\right)\left[\tau_1 u^2(t) - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} u^2(s)dsdz\right] \\
&+ \frac{DB}{2}\epsilon_4 \left[\tau_1 \left(\rho_1 v + \rho_3 u + a_4 v(t)\right)^2 - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} \left(\rho_2 v + \rho_3 u + a_4 v(s)\right)^2 dsdz\right] \\
&+ \frac{DB}{2}\left(\frac{1}{\epsilon_4} - 2a_4\right)\left[\tau_1 v^2(t) - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} v^2(s)dsdz\right] \\
&+ \frac{DC}{2}\epsilon_5 \left[\tau_1 \left(\rho_1 v + \rho_3 u + a_5 v(t-\tau_2)\right)^2 \right. \\
&\left. - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} \left(\rho_2 v + \rho_3 u + a_5 v(s-\tau_2)\right)^2 dsdz\right] \\
&+ \frac{DC}{2}\left(\frac{1}{\epsilon_5} - 2a_5\right)\left[\tau_1 v^2(t-\tau_2) - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t} v^2(s-\tau_2)dsdz\right].
\end{aligned} \tag{5.29}$$

For clarity and in the interest of space saving in these rather long expressions, the definite double integral terms from Equation (5.29) will be denoted by Ω, as they remain unchanged. They are reintroduced in the final form of the expression for the $\dot{V}_0$. Further to that, the terms $u^2(t-\tau_1)$ and $v^2(t-\tau_2)$ can be written in integral form such as,

$$u^2(t-\tau_1) = u^2(t) - \frac{d}{dt}\int_{t-\tau_1}^{t} u^2(s)ds. \tag{5.30}$$

This gives two more definite integrals which will be included in the term, Ω. Thus, continuing here with the non-integral terms, expanded and rewriting for terms in $u^2(t)$, $v^2(t)$ and $u(t)v(t)$ gives,

$$\begin{aligned}\frac{dV_0}{dt} = \Omega + u^2 &\left[-\rho_1 A + \rho_3 D + \frac{CD}{2}\left(\frac{1}{\epsilon_1} - 2a_1\right)\tau_2 + \frac{Cd_i}{2}\epsilon_2\tau_2\rho_1^2\right. \\ &\left. + \frac{DA}{2}\epsilon_3\tau_1(\rho_3^2 + a_3^2 + 2\rho_3 a_3) + \frac{DA}{2}\left(\frac{1}{\epsilon_3} - 2a_3\right)\tau_1 + \frac{DB}{2}\epsilon_4\tau_1\rho_3^2\right] \\ + v^2&\Big[- \rho_2 d_i - \rho_3(B+C) + \frac{Cd_i}{2}\epsilon_2\tau_2(\rho_3^2 + a_2^2 + 2\rho_3 a_2) + \frac{Cd_i}{2}\left(\frac{1}{\epsilon_2} + 2a_2\right)\tau_2 \\ &+ \frac{DA}{2}\epsilon_3\tau_1\rho_2^2 + \frac{DB}{2}\epsilon_4\tau_1(\rho_2^2 + a_4^2 + 2\rho_2 a_4) + \frac{DB}{2}\left(\frac{1}{\epsilon_4} - 2a_4\right)\tau_1 \\ &+ \frac{DC}{2}\left(\frac{1}{\epsilon_5} - 2a_5\right)\tau_1\Big] + uv\left[-\rho_1(B+C) + \rho_2 D - \rho_3(A + d_i)\right. \\ &\left. + Cd_i\epsilon_2\tau_2(\rho_1\rho_3 + \rho_1 a_2) + DA\epsilon_3\tau_1(\rho_2\rho_3 + \rho_2 a_3) + DB\epsilon_4\tau_1(\rho_2\rho_3 + \rho_3 a_4)\right] \\ + \frac{CD}{2}&\epsilon_1\tau_2(\rho_1 u + \rho_3 v + a_1 u(t-\tau_1))^2 + \frac{DC}{2}\epsilon_5\tau_1\left(\rho_1 v + \rho_3 u + a_5 v(t-\tau_2)\right)^2.\end{aligned} \tag{5.31}$$

The final two terms in Equation (5.31) contain variables of delays which must be transformed into terms of $u(t)$ and $v(t)$ or integral terms to be taken into the functional. Focusing first on the term of $u(t-\tau_1)$ by expanding out the squared bracket produces a term involving $2a_1(\rho_1 u + \rho_3 v)u(t-\tau_1)$ for which the original form of Young's inequality in Lemma 1 can be applied. Rearranging and writing the $u^2(t-\tau_1)$ term in integral form produces,

$$\frac{CD}{2}\epsilon_1\tau_2\left[(a_1\epsilon_6 + 1)(\rho_1^2 u^2 + 2\rho_1\rho_3 uv + \rho_3^2 v^2) + \left(a_1^2 + \frac{a_1}{\epsilon_6}\right)\left[u^2 - \frac{d}{dt}\int_{t-\tau_1}^{2} u^2(s)\right]\right]. \tag{5.32}$$

The use of Young's inequality again provides a definite integral term which can be accounted for as part of the functional, while the remaining variables are functions of time,

t. Applying the same method for the term involving $v(t-\tau_2)$ in Equation (5.31) gives,

$$\frac{DC}{2}\epsilon_5\tau_1\left[(a_5\epsilon_7+1)(\rho_2^2v^2+2\rho_2\rho_3uv+\rho_3^2u^2)+\left(a_5^2+\frac{a_5}{\epsilon_7}\right)\left[v^2-\frac{d}{dt}\int_{t-\tau_2}^{2}v^2(s)\right]\right]. \tag{5.33}$$

Now, all of the elements from Equations (5.31, 5.32 and 5.33), as well as all of the integral terms can be brought together. The full expression for $\dot{V}_0$ is given by,

$$\begin{aligned}
\frac{dV_0}{dt} &= -u^2[\Psi_{11}]-v^2[\Psi_{22}]-uv[\Psi_{12}] \\
&-\frac{CD}{2}\epsilon_1\frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t}(\rho_1u+\rho_3v+a_1u(s-\tau_1))^2dsdz \\
&-\frac{CD}{2}\left(\frac{1}{\epsilon_1}-2a_1\right)\frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t}u^2(s-\tau_1)dsdz \\
&-\frac{Cd_i}{2}\epsilon_2\frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t}(\rho_1u+\rho_3v+a_2v(s))^2dsdz \\
&-\frac{Cd_i}{2}\left(\frac{1}{\epsilon_2}+2a_2\right)\frac{d}{dt}\int_{t-\tau_1}^{t}\int_{z}^{t}v^2(s)dsdz \\
&-\frac{DA}{2}\epsilon_3\frac{d}{dt}\int_{t-\tau_2}^{t}\int_{z}^{t}(\rho_2v+\rho_3u+a_3u(s))^2dsdz \\
&-\frac{DA}{2}\left(\frac{1}{\epsilon_3}-2a_3\right)\frac{d}{dt}\int_{t-\tau_2}^{t}\int_{z}^{t}u^2(s)dsdz \\
&-\frac{DB}{2}\epsilon_4\frac{d}{dt}\int_{t-\tau_2}^{t}\int_{z}^{t}(\rho_2v+\rho_3u+a_4v(s))^2dsdz \\
&-\frac{DB}{2}\left(\frac{1}{\epsilon_4}-2a_4\right)\frac{d}{dt}\int_{t-\tau_2}^{t}\int_{z}^{t}v^2(s)dsdz \\
&-\frac{DC}{2}\epsilon_5\frac{d}{dt}\int_{t-\tau_2}^{t}\int_{z}^{t}(\rho_2v+\rho_3u+a_5v(s-\tau_2))^2dsdz \\
&-\frac{DC}{2}\left(\frac{1}{\epsilon_5}-2a_5\right)\frac{d}{dt}\int_{t-\tau_2}^{t}\int_{z}^{t}v^2(s-\tau_2)dsdz \\
&-\frac{CD}{2}\tau_2\left(\frac{1}{\epsilon_1}(\epsilon_1a_1-1)^2+\frac{\epsilon_1a_1}{\epsilon_6}\right)\frac{d}{dt}\int_{t-\tau_1}^{t}u^2(s)ds \\
&-\frac{DC}{2}\tau_1\left(\frac{1}{\epsilon_5}(\epsilon_5a_5-1)^2+\frac{\epsilon_5a_5}{\epsilon_7}\right)\frac{d}{dt}\int_{t-\tau_2}^{t}v^2(s)ds.
\end{aligned} \tag{5.34}$$

Where the following functions are defined as,

$$\begin{aligned}\Psi_{11} &= \rho_1 A - \rho_3 D - \frac{CD}{2}\tau_2\left(\frac{1}{\epsilon_1}(\epsilon_1 a_1 - 1)^2 + \frac{\epsilon_1 a_1}{\epsilon_6}\right) - \frac{CD}{2}\tau_2\epsilon_1\rho_1^2(a_1\epsilon_6 + 1) \\ &\quad -\frac{Cd_i}{2}\tau_2\epsilon_2\rho_1^2 - \frac{AD}{2}\tau_1\epsilon_3(\rho_3 + a_3)^2 - \frac{AD}{2}\tau_1\left(\frac{1}{\epsilon_3} - 2a_3\right) - \frac{BD}{2}\tau_1\epsilon_4\rho_3^2 \\ &\quad -\frac{CD}{2}\tau_1\epsilon_5\rho_3^2(a_5\epsilon_7 + 1), \qquad (5.35) \\ \Psi_{22} &= \rho_3(B + C) + \rho_2 d_i - \frac{Cd_i}{2}\tau_2\epsilon_2(\rho_3 + a_2)^2 - \frac{Cd_i}{2}\tau_2\left(\frac{1}{\epsilon_2} + 2a_2\right) \\ &\quad -\frac{CD}{2}\tau_2\epsilon_1\rho_3^2(a_1\epsilon_6 + 1) - \frac{AD}{2}\tau_1\epsilon_3\rho_2^2 - \frac{BD}{2}\tau_1\epsilon_4(\rho_2 + a_4)^2 \\ &\quad -\frac{BD}{2}\tau_1\left(\frac{1}{\epsilon_4} + 2a_4\right) - \frac{CD}{2}\tau_1\epsilon_5\rho_2^2(a_5\epsilon_7 + 1) \\ &\quad -\frac{CD}{2}\tau_1\left(\frac{1}{\epsilon_5}(\epsilon_5 a_5 - 1)^2 + \frac{\epsilon_5 a_5}{\epsilon_7}\right), \qquad (5.36)\end{aligned}$$

and,

$$\begin{aligned}\Psi_{12} &= \rho_1(B + C) + \rho_3(A + d_i) - \rho_2 D - Cd_i\tau_2\epsilon_2(\rho_1\rho_3 + \rho_1 a_2) \\ &\quad -CD\tau_2\epsilon_1\rho_1\rho_3(a_1\epsilon_6 + 1) - AD\tau_1\epsilon_3(\rho_2\rho_3 + \rho_2 a_3) \\ &\quad -BD\tau_1\epsilon_4(\rho_2\rho_3 + \rho_2 a_4) - CD\tau_1\epsilon_5\rho_2\rho_3(a_5\epsilon_7 + 1). \qquad (5.37)\end{aligned}$$

Therefore, by defining the functional V_1 as,

$$
\begin{aligned}
V_1 = V_0 &+ \frac{CD}{2}\epsilon_1 \int_{t-\tau_1}^{t}\int_{z}^{t}(\rho_1 u + \rho_3 v + a_1 u(s-\tau_1))^2 dsdz \\
&+ \frac{CD}{2}\left(\frac{1}{\epsilon_1} - 2a_1\right)\int_{t-\tau_1}^{t}\int_{z}^{t} u^2(s-\tau_1)dsdz \\
&+ \frac{Cd_i}{2}\epsilon_2 \int_{t-\tau_1}^{t}\int_{z}^{t}(\rho_1 u + \rho_3 v + a_2 v(s))^2 dsdz + \frac{Cd_i}{2}\left(\frac{1}{\epsilon_2} + 2a_2\right)\int_{t-\tau_1}^{t}\int_{z}^{t} v^2(s)dsdz \\
&+ \frac{DA}{2}\epsilon_3 \int_{t-\tau_2}^{t}\int_{z}^{t}(\rho_2 v + \rho_3 u + a_3 u(s))^2 dsdz + \frac{DA}{2}\left(\frac{1}{\epsilon_3} - 2a_3\right)\int_{t-\tau_2}^{t}\int_{z}^{t} u^2(s)dsdz \\
&+ \frac{DB}{2}\epsilon_4 \int_{t-\tau_2}^{t}\int_{z}^{t}(\rho_2 v + \rho_3 u + a_4 v(s))^2 dsdz + \frac{DB}{2}\left(\frac{1}{\epsilon_4} - 2a_4\right)\int_{t-\tau_2}^{t}\int_{z}^{t} v^2(s)dsdz \\
&+ \frac{DC}{2}\epsilon_5 \int_{t-\tau_2}^{t}\int_{z}^{t}(\rho_2 v + \rho_3 u + a_5 v(s-\tau_2))^2 dsdz \\
&+ \frac{DC}{2}\left(\frac{1}{\epsilon_5} - 2a_5\right)\int_{t-\tau_2}^{t}\int_{z}^{t} v^2(s-\tau_2)dsdz \\
&+ \frac{CD}{2}\tau_2\left(\frac{1}{\epsilon_1}(\epsilon_1 a_1 - 1)^2 + \frac{\epsilon_1 a_1}{\epsilon_6}\right)\int_{t-\tau_1}^{t} u^2(s)ds \\
&+ \frac{DC}{2}\tau_1\left(\frac{1}{\epsilon_5}(\epsilon_5 a_5 - 1)^2 + \frac{\epsilon_5 a_5}{\epsilon_7}\right)\int_{t-\tau_2}^{t} v^2(s)ds,
\end{aligned}
\tag{5.38}
$$

then the derivative is given by,

$$
\frac{dV_1}{dt} = -u^2[\Psi_{11}] - v^2[\Psi_{22}] - uv[\Psi_{12}]. \tag{5.39}
$$

A Lyapunov functional must be positive-definite for all cases, and thus for a functional to be considered a Lyapunov functional it must also be positive definite. Therefore, it is required from Equation (5.38) that $V_1 > 0$. The free parameters in this functional are subject to certain constraints as defined when introduced. The ρ and the ϵ terms are strictly positive such that $\rho_i > 0,\ i = 1, 2, 3$ and $\epsilon_j > 0,\ j = 1, \ldots, 7$, while the $a_k,\ k = 1, \ldots, 5$ terms have no restrictions. Further constraints are developed through combinations of terms involving the a_k, which may be negative. In order for the functional V_1 to be

positive, the following constraints must be satisfied,

$$\frac{1}{\epsilon_1} - 2a_1 \geq 0 \quad \Rightarrow \quad a_1 - \frac{1}{2\epsilon_1} \leq 0 \tag{5.40}$$

$$\frac{1}{\epsilon_2} + 2a_2 \geq 0 \quad \Rightarrow \quad -\left(a_2 + \frac{1}{2\epsilon_2}\right) \leq 0 \tag{5.41}$$

$$\frac{1}{\epsilon_3} - 2a_3 \geq 0 \quad \Rightarrow \quad a_3 - \frac{1}{2\epsilon_3} \leq 0 \tag{5.42}$$

$$\frac{1}{\epsilon_4} - 2a_4 \geq 0 \quad \Rightarrow \quad a_4 - \frac{1}{2\epsilon_4} \leq 0 \tag{5.43}$$

$$\frac{1}{\epsilon_5} - 2a_5 \geq 0 \quad \Rightarrow \quad a_5 - \frac{1}{2\epsilon_5} \leq 0 \tag{5.44}$$

$$\frac{1}{\epsilon_1}(\epsilon_1 a_1 - 1)^2 + \frac{\epsilon_1 a_1}{\epsilon_6} \geq 0 \quad \Rightarrow -\left[\frac{1}{\epsilon_1}(\epsilon_1 a_1 - 1)^2 + \frac{\epsilon_1 a_1}{\epsilon_6}\right] \leq 0 \tag{5.45}$$

$$\frac{1}{\epsilon_5}(\epsilon_5 a_5 - 1)^2 + \frac{\epsilon_5 a_5}{\epsilon_7} \geq 0 \quad \Rightarrow -\left[\frac{1}{\epsilon_5}(\epsilon_5 a_5 - 1)^2 + \frac{\epsilon_5 a_5}{\epsilon_7}\right] \leq 0 \tag{5.46}$$

It is also necessary to define the conditions for the derivative of the Lyapunov functional to be stable. At this stage, the derivative, $\dot{V}_1$, will be written in matrix form with the symmetric matrix Ψ, such that,

$$-\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}^T \underbrace{\begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix}}_{\Psi} \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} \tag{5.47}$$

Where $\Psi_{11},\ \Psi_{22}$ and Ψ_{12} are given by Equations (5.35) to (5.37), and $\Psi_{12} = \Psi_{21}$. Therefore, provided the constraints (5.40) to (5.46) are met and that the matrix Ψ is positive definite then the functional V_1 as defined in Theorem 1 will be clearly positive and hence provides a Lyapunov functional for the linearised dynamical system (4.9). □

For stability, the derivative of the Lyapunov functional must be negative, $\dot{V}_1 < 0$ and therefore the matrix Ψ, defined in Equation (5.47), must be positive definite. Writing the expression for $\dot{V}_1$ in matrix form allows for the use of Sylvester's criterion for positive

definiteness (Gilbert 1991). It states that a $n \times n$ Hermitian matrix is positive definite if all of the principal minors are positive. This is generalised in the case here such that, a Hermitian matrix is a symmetric matrix if there are no imaginary components. For a 2×2 matrix the principal minors are the term Ψ_{11} and the determinant of the matrix Ψ itself. By meeting these conditions it is equivalent to say that the matrix is positive definite.

Therefore, in order to have stability of the Lyapunov functional, V_1, two conditions must be met: $\Psi_{11} > 0$ and also the determinant of matrix Ψ is positive, i.e. $\Psi_{11}\Psi_{22} - \Psi_{12}^2 > 0$. Thus, when these conditions are met the system is stable.

5.3 Discussion

The local stability analysis, both numerically and analytically, has been investigated for the first time in relation to the two-delay glucose-insulin regulation system. An extension of the characteristic equation approach as seen in the one-delay case in Chapter 4 is made for two delays which results in an expression for τ_2 which only implicitly depends on τ_1 which can be solved for conditions of stability. The graphical plot of these calculations leads to the discovery that the boundary for the region of stability is a curve and not a straight line. This is confirmed by further numerical simulation using the LMI method which gives an accurate representation for the range $5 < \tau_1 < 20$.

The local stability is further analysed using a Lyapunov functional approach to provide sufficient conditions for stability, where the delay-independent and the delay-dependent cases are investigated. The delay-dependent case yields a series of conditions on the parameters introduced in order to give stability of the system. In order to allow more freedom in the choice of parameters, a special case of Young's inequality was repeatedly used.

Chapter 6

Global stability analysis of the two-delay system

An analysis of global stability of the nonlinear two-delay glucose-insulin regulation system is presented in this Chapter. Sufficient conditions for stability are derived analytically by taking a Lyapunov functional approach for the original nonlinear form of the model. While more complex than using the linearised system as in the local case, use of the nonlinear model allows for a series of stability conditions to be obtained that are related to the biological parameters in the model. The effect of altering the Hill parameters will be investigated to find optimal ranges dependent on maximising the stability region.

6.1 Lyapunov functional approach

In order to study the global stability of the system, it is necessary to work with the non-linear model in its original form. Here, the aim is to investigate the relationship between the global stability and the two time delays in the system, τ_1 and τ_2. Note that throughout this section the notation $u = u(t)$ and $v = v(t)$ is used once again in order to make the equations clearer, the dependency of the variables will be explicitly stated when necessary. Following the work of Bennett and Gourley (2004b), the steady state of the system, (G^*, I^*), is studied. The shifted variables,

$$G(t) = u(t) + G^*, \quad I(t) = v(t) + I^*, \tag{6.1}$$

are introduced such that $(u, v) = (0, 0)$ represents the steady state of the translated system,

$$\begin{aligned} \dot{u} &= G_{in} - f_2(u + G^*) - f_3(u + G^*)f_4(v + I^*) + f_5(v(t - \tau_2) + I^*), \\ \dot{v} &= f_1(u(t - \tau_1) + G^*) - d_i(v + I^*). \end{aligned} \tag{6.2}$$

The Taylor expansion approximated at the first order of a given function $f(x)$ is given as,

$$f(x + x_0) = f(x_0) + xf'(c),$$

where c is some number between x_0 and x. By writing $c = x_0 + \theta x$, with $0 \leq \theta \leq 1$, the Taylor expansion can then be written as,

$$f(x + x_0) = f(x_0) + xf'(x_0 + \theta x), \quad 0 \leq \theta \leq 1.$$

Therefore, there exists $0 \leq \theta_i \leq 1,\ i = 1, \ldots, 5$, such that when applied to $\dot{u}$ and $\dot{v}$ from

Equation (6.2), the following are true,

$$\begin{aligned}\dot{u} &= \underbrace{G_{in} - f_2(G^*) - f_3(G^*)f_4(I^*) + f_5(I^*)}_{=0} - uf_2'(G^* + \theta_2 u) \\ &\quad - uf_4(I^*)f_3'(G^* + \theta_3 u) - vf_3(G^*)f_4'(I^* + \theta_4 v) \\ &\quad + uvf_3'(G^* + \theta_3 u)f_4'(I^* + \theta_4 v) + v(t-\tau_2)f_5'(I^* + \theta_5 v(t-\tau_2)) \\ &= -uf_2'(G^* + \theta_2 u) - uf_4(I^*)f_3'(G^* + \theta_3 u) - vf_3(G^*)f_4'(I^* + \theta_4 v) \\ &\quad + uvf_3'(G^* + \theta_3 u)f_4'(I^* + \theta_4 v) + v(t-\tau_2)f_5'(I^* + \theta_5 v(t-\tau_2)), \end{aligned} \tag{6.3}$$

and,

$$\begin{aligned}\dot{v} &= \underbrace{f_1(G^*) - d_i I^*}_{0} + u(t-\tau_1)f_1'(G^* + \theta_1 u(t-\tau_1)) - d_i v \\ &= u(t-\tau_1)f_1'(G^* + \theta_1 u(t-\tau_1)) - d_i v. \end{aligned} \tag{6.4}$$

A feature of this system is that it is bounded, as proved in Bennett and Gourley (2004b). This leads to the upper bound on $G(t)$ which can be seen in the proposition below:

Proposition 1. *Writing $f_3(G) = qG$ for some $q \in \mathbb{R}$,*

$$G(t) \leq \max\left\{G(0), \frac{G_{in} + f_5(0)}{qf_4(0)}\right\}.$$

This is the upper bound for maximal glucose as used in Bennett and Gourley (2004b) and will be used in the calculation of the Lyapunov functional.

By taking $V_0 = \frac{1}{2}(u^2 + \omega v^2)$, with $\omega > 0$ for the nonlinear dynamical system given by (4.1) a Lyapunov functional and stability condtions can be obtained as follows. In deriving this Lyapunov functional Young's inequality with exponent two as given by Lemma 1 and the double integral Leibniz rule as given by Equation 5.21, as detailed in Chapter 5 will be repeatedly used.

Consider the function $V_0 = \frac{1}{2}(u^2 + \omega v^2)$, with $\omega > 0$. Then

$$\begin{aligned}\frac{dV_0}{dt} &= u\dot{u} + \omega v\dot{v} \\ &= u\left[-uf_2'(G^* + \theta_2 u) - uf_4(I^*)f_3'(G^* + \theta_3 u) - vf_3(G^*)f_4'(I^* + \theta_4 v)\right. \\ &\quad \left. + uvf_3'(G^* + \theta_3 u)f_4'(I^* + \theta_4 v) + v(t-\tau_2)f_5'(I^* + \theta_5 v(t-\tau_2))\right] \\ &\quad + \omega v\left[u(t-\tau_1)f_1'(G^* + \theta_1 u(t-\tau_1)) - d_i v\right].\end{aligned} \tag{6.5}$$

For simplicity, function $f_3(G)$ will be written as a linear term such that $f_3(G) = qG$, where $f_3(G^*) = qG^*$ and $f_3'(G^*) = q$. Grouping the u^2 and v^2 terms gives,

$$\begin{aligned}\frac{dV_0}{dt} &= u^2\left[-f_2'(G^* + \theta_2 u) - qf_4(I^*)\right] - \omega d_i v^2 - uvq(G^* + u)f_4'(I^* + \theta_4 v) \\ &\quad + uv(t-\tau_2)f_5'(I^* + \theta_5 v(t-\tau_2)) + \omega vu(t-\tau_1)f_1'(G^* + \theta_1 u(t-\tau_1)).\end{aligned} \tag{6.6}$$

Here, ghost terms are introduced to the delayed variables to give terms in the form $[v(t-\tau_1) - v]$, which can be written in integral form to give,

$$\begin{aligned}\frac{dV_0}{dt} &= u^2\left[-f_2'(G^* + \theta_2 u) - qf_4(I^*)\right] - \omega d_i v^2 - uvq(G^* + u)f_4'(I^* + \theta_4 v) \\ &\quad - uf_5'(I^* + \theta_5 v(t-\tau_2))\int_{t-\tau_2}^{t}\dot{v}(s)ds + uvf_5'(I^* + \theta_5 v(t-\tau_2)) \\ &\quad - \omega vf_1'(G^* + \theta_1 u(t-\tau_1))\int_{t-\tau_1}^{t}\dot{u}(s)ds + \omega vuf_1'(G^* + \theta_1 u(t-\tau_1)).\end{aligned} \tag{6.7}$$

The expressions for $\dot{u}(s)$ and $\dot{v}(s)$ are introduced into the integrals where function $f_3(G) = qG$. A series of rearrangements are made to give terms which combine variables u and v and clearly show the positive terms. The modulus of the terms related to function f_5 will be taken as function f_5 is a decreasing function allowing the negative f_5 terms to be written as positives. Also, the terms in $v(t)v(s)f_4'(I^* + \theta_4 v(s))$ are grouped giving the

expression,

$$\begin{aligned}\frac{dV_0}{dt} &= u^2\left[-f_2'(G^*+\theta_2 u) - qf_4(I^*)\right] - \omega d_i v^2 - uvq(G^*+u)f_4'(I^*+\theta_4 v)\\ &\quad - uv|f_5'(I^*+\theta_5 v(t-\tau_2))| + \omega v u f_1'(G^*+\theta_1 u(t-\tau_1))\\ &\quad + |f_5'(I^*+\theta_5 v(t-\tau_2))|\int_{t-\tau_2}^{t}\Big[u(t)u(s-\tau_1)f_1'(G^*+\theta_1 u(s-\tau_1))\\ &\quad - d_i v(s)u(t)\Big]ds + \omega f_1'(G^*+\theta_1 u(t-\tau_1))\int_{t-\tau_1}^{t}\Big[u(s)v(t)f_2'(G^*+\theta_2 u)\\ &\quad + qv(t)u(s)f_4(I^*) + qv(s)v(t)\left(G^*+u(s)\right)f_4'(I^*+\theta_4 v(s))\\ &\quad + v(t)v(s-\tau_2)|f_5'(I^*+\theta_5 v(s-\tau_2))|\Big]ds.\end{aligned} \tag{6.8}$$

The term $(G^* + u(s))$ is the addition of two glucose values and as such is less than the maximal glucose level, M_G. By introducing M_G as a replacement for this, the dependence on a combination of three variables is removed. At this stage, Young's inequality Lemma 1, will be applied where possible to split the variables. Further to that, the maximum quantities of the functions will be introduced and are denoted by f_i^M and $f_i'^M$. Due to the sigmoidal nature of the Hill function, each function is bounded and can be expressed in terms of the Hill parameters, as seen in Equation (3.8), in order to give a single value which still satisfies the inequality. This gives,

$$\begin{aligned}\frac{dV_0}{dt} &\leq u^2\left[-f_2'(G^*+\theta_2 u) - qf_4(I^*)\right] - \omega d_i v^2 + q\frac{1}{2}\left(\epsilon_1 u^2 + \frac{1}{\epsilon_1}v^2\right)M_G f_4'^M\\ &\quad + \frac{1}{2}\left(\epsilon_2 u^2 + \frac{1}{\epsilon_2}v^2\right)|f_5'^M| + \omega\frac{1}{2}\left(\epsilon_3 u^2 + \frac{1}{\epsilon_3}v^2\right)f_1'^M\\ &\quad + |f_5'^M|\int_{t-\tau_2}^{t}\Big[\frac{1}{2}\left(\epsilon_4 u^2(s-\tau_1) + \frac{1}{\epsilon_4}u^2(t)\right)f_1'^M + \frac{d_i}{2}\left(\epsilon_5 v^2(s) + \frac{1}{\epsilon_5}u^2(t)\right)\Big]ds\\ &\quad + \omega f_1'^M\int_{t-\tau_1}^{t}\Big[\frac{1}{2}\left(\epsilon_6 u^2(s) + \frac{1}{\epsilon_6}v^2(t)\right)f_2'^M + q\frac{1}{2}\left(\epsilon_7 u^2(s) + \frac{1}{\epsilon_7}v^2(t)\right)f_4(I^*)\\ &\quad + q\frac{1}{2}\left(\epsilon_8 v^2(s) + \frac{1}{\epsilon_8}v^2(t)\right)M_G f_4'^M + \frac{1}{2}\left(\epsilon_9 v^2(s-\tau_2) + \frac{1}{\epsilon_9}v^2(t)\right)|f_5'^M|\Big]ds.\end{aligned} \tag{6.9}$$

As each of the variables u and v, as functions of s and t and the delays are now separated, the integrals can be evaluated for the variables dependent on t. Each of the integrals are with respect to s, therefore those variables not dependent on s can be treated as constants within the integral. Evaluating these terms and then grouping for u^2 and v^2 produces the following expression,

$$\begin{aligned}\frac{dV_0}{dt} \leq & -u^2\left[f_2'(G^*+\theta_2 u)+qf_4(I^*)-\frac{1}{2}\Big[q\epsilon_1 M_G f_4'^M+\epsilon_2|f_5'^M|+\omega\epsilon_3 f_1'^M\right.\\ & \left.+|f_5'^M|\tau_2\left(\frac{f_1'^M}{\epsilon_4}+\frac{d_i}{\epsilon_5}\right)\Big]\right]-v^2\left[\omega d_i-\frac{1}{2}\left[q\frac{1}{\epsilon_1}M_G f_4'^M\right.\right.\\ & +\frac{1}{\epsilon_2}|f_5'^M|+\omega\frac{1}{\epsilon_3}f_1'^M+\omega f_1'^M\tau_1\Big[\frac{1}{\epsilon_6}f_2'^M+\frac{1}{\epsilon_7}qf_4(I^*)\\ & \left.\left.+\frac{1}{\epsilon_8}qM_G f_4'^M+\frac{1}{\epsilon_9}|f_5'^M|\Big]\right]\right]+|f_5'^M|\frac{\epsilon_4}{2}f_1'^M\int_{t-\tau_2}^{t}u^2(s-\tau_1)ds\\ & +|f_5'^M|\frac{\epsilon_5}{2}d_i\int_{t-\tau_2}^{t}v^2(s)ds+\omega f_1'^M[\frac{\epsilon_6}{2}f_2'^M+\frac{\epsilon_7}{2}qf_4(I^*)]\int_{t-\tau_1}^{t}u^2(s)ds\\ & +\omega f_1'^M\frac{\epsilon_8}{2}qM_G f_4'^M\int_{t-\tau_1}^{t}v^2(s)ds+\omega f_1'^M\frac{\epsilon_9}{2}|f_5'^M|\int_{t-\tau_1}^{t}v^2(s-\tau_2)ds.\end{aligned} \tag{6.10}$$

The Leibniz rule (Abramowitz and Stegun 1972), used for a double integral as shown in Equation (5.21), can be applied to each of the integral terms. This gives derivative of integral terms which can be accounted for as part of the Lyapunov functional definition, V. Using the Leibniz double integral rule produces,

$$\begin{aligned}\frac{dV_0}{dt} \leq -u^2 \Bigg[& f_2'(G^* + \theta_2 u) + q f_4(I^*) - \frac{1}{2}\Big[q\epsilon_1 M_G f_4'^M + \epsilon_2 |f_5'^M| + \omega\epsilon_3 f_1'^M \\ & + |f_5'^M|\tau_2 \left(\frac{f_1'^M}{\epsilon_4} + \frac{d_i}{\epsilon_5}\right)\Big]\Bigg] - v^2 \Bigg[\omega d_i - \frac{1}{2}\Bigg[q\frac{1}{\epsilon_1} M_G f_4'^M \\ & + \frac{1}{\epsilon_2}|f_5'^M| + \omega\frac{1}{\epsilon_3} f_1'^M + \omega f_1'^M \tau_1 \Big[\frac{1}{\epsilon_6} f_2'^M + \frac{1}{\epsilon_7} q f_4(I^*) \\ & + \frac{1}{\epsilon_8} q M_G f_4'^M + \frac{1}{\epsilon_9}|f_5'^M|\Big]\Bigg]\Bigg] + |f_5'^M|\frac{\epsilon_4}{2} f_1'^M \Big[\tau_2 u^2(t-\tau_1) \\ & - \frac{d}{dt}\int_{t-\tau_2}^{t}\int_z^t u^2(s-\tau_1)dsdz\Big] + |f_5'^M|\frac{\epsilon_5}{2} d_i \left[\tau_2 v^2 - \frac{d}{dt}\int_{t-\tau_2}^{t}\int_z^t v^2(s)dsdz\right] \\ & + \omega f_1'^M [\frac{\epsilon_6}{2} f_2'^M + \frac{\epsilon_7}{2} q f_4(I^*)] \left[\tau_1 u^2 - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_z^t u^2(s)dsdz\right] \\ & + \omega f_1'^M \frac{\epsilon_8}{2} q M_G f_4'^M \left[\tau_1 v^2 - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_z^t v^2(s)dsdz\right] \\ & + \omega f_1'^M \frac{\epsilon_9}{2}|f_5'^M| \left[\tau_1 v^2(t-\tau_2) - \frac{d}{dt}\int_{t-\tau_1}^{t}\int_z^t v^2(s-\tau_2)dsdz\right].\end{aligned} \tag{6.11}$$

This then motivates the definition of the functional, V_1, to account for the integral terms such that,

$$\begin{aligned}V_1 = V_0 & + |f_5'^M|\frac{\epsilon_4}{2} f_1'^M \int_{t-\tau_2}^{t}\int_z^t u^2(s-\tau_1)dsdz + |f_5'^M|\frac{\epsilon_5}{2} d_i \int_{t-\tau_2}^{t}\int_z^t v^2(s)dsdz \\ & + \omega f_1'^M [\frac{\epsilon_6}{2} f_2'^M + \frac{\epsilon_7}{2} q f_4(I^*)] \int_{t-\tau_1}^{t}\int_z^t u^2(s)dsdz \\ & + \omega f_1'^M \frac{\epsilon_8}{2} q M_G f_4'^M \int_{t-\tau_1}^{t}\int_z^t v^2(s)dsdz + \omega f_1'^M \frac{\epsilon_9}{2}|f_5'^M| \int_{t-\tau_1}^{t}\int_z^t v^2(s-\tau_2)dsdz,\end{aligned} \tag{6.12}$$

where the derivative of V_1 is given by,

$$\begin{aligned}\frac{dV_1}{dt} \leq -u^2 \Bigg[& f_2'(G^* + \theta_2 u) + q f_4(I^*) - \frac{1}{2}\Big[q\epsilon_1 M_G f_4'^M + \epsilon_2 |f_5'^M| + \omega\epsilon_3 f_1'^M \\ & + |f_5'^M|\tau_2 \left(\frac{f_1'^M}{\epsilon_4} + \frac{d_i}{\epsilon_5}\right)\Big] - \omega f_1'^M [\frac{\epsilon_6}{2} f_2'^M + \frac{\epsilon_7}{2} q f_4(I^*)]\tau_1 \Bigg] \\ & - v^2 \Bigg[\omega d_i - \frac{1}{2}\Bigg[q\frac{1}{\epsilon_1} M_G f_4'^M + \frac{1}{\epsilon_2}|f_5'^M| + \omega\frac{1}{\epsilon_3} f_1'^M \\ & + \omega f_1'^M \tau_1 \Big[\frac{1}{\epsilon_6} f_2'^M + \frac{1}{\epsilon_7} q f_4(I^*) + \left(\epsilon_8 + \frac{1}{\epsilon_8}\right) q M_G f_4'^M + \frac{1}{\epsilon_9}|f_5'^M|\Big]\Bigg] \\ & - |f_5'^M|\frac{\epsilon_5}{2} d_i \tau_2\Bigg] + |f_5'^M|\frac{\epsilon_4}{2} f_1'^M \tau_2 u^2(t-\tau_1) + \omega f_1'^M \frac{\epsilon_9}{2}|f_5'^M|\tau_1 v^2(t-\tau_2).\end{aligned} \tag{6.13}$$

The last two terms in Equation (6.13) are functions of the delays and they must be removed in this final step to give the desired functional. By ensuring that there are only two distinct variables in the expression, $u^2(t)$ or $v^2(t)$, then there will only be conditions based on two expressions in order to have stability.

With this in mind, the delayed variables are directly replaced by their integral form such that,

$$u^2(t-\tau_1) = \frac{d}{dt}\int_t^{t-\tau_1} u^2(s)ds + u^2(t), \tag{6.14}$$

and likewise for $v^2(t-\tau_2)$, so that no further inequalities are used. This change produces two further definite integral terms which can be accounted for in the functional itself, while the u^2 and v^2 terms will be included in their respective brackets. Therefore the functional V_G is defined as,

$$V_G = V_1 + |f_5'^M|\frac{\epsilon_4}{2} f_1'^M \tau_2 \int_t^{t-\tau_1} u^2(s)ds + \omega f_1'^M \frac{\epsilon_9}{2}|f_5'^M|\tau_1 \int_t^{t-\tau_2} v^2(s)ds, \tag{6.15}$$

which is clearly positive for all functions u, v. The derivation of functional V_G shows that

the following inequality relating to the derivative holds,

$$\begin{aligned}\frac{dV_G}{dt} \leq -u^2 \Bigg[& f_2'(G^* + \theta_2 u) + q f_4(I^*) - \frac{1}{2} \Big[q\epsilon_1 M_G f_4'^M + \epsilon_2 |f_5'^M| + \omega \epsilon_3 f_1'^M \\ & + |f_5'^M| \tau_2 \left(\left(\epsilon_4 + \frac{1}{\epsilon_4} \right) f_1'^M + \frac{d_i}{\epsilon_5} \right) \Big] - \omega f_1'^M [\frac{\epsilon_6}{2} f_2'^M + \frac{\epsilon_7}{2} q f_4(I^*)] \tau_1 \Bigg] \\ & - v^2 \Bigg[\omega d_i - \frac{1}{2} \Bigg[q \frac{1}{\epsilon_1} M_G f_4'^M + \frac{1}{\epsilon_2} |f_5'^M| + \omega \frac{1}{\epsilon_3} f_1'^M + \omega f_1'^M \tau_1 \Big[\frac{1}{\epsilon_6} f_2'^M \\ & + \frac{1}{\epsilon_7} q f_4(I^*) + \left(\epsilon_8 + \frac{1}{\epsilon_8} \right) q M_G f_4'^M + \left(\epsilon_9 + \frac{1}{\epsilon_9} \right) |f_5'^M| \Big] \Bigg] - |f_5'^M| \frac{\epsilon_5 d_i}{2} \tau_2 \Bigg] . \end{aligned} \tag{6.16}$$

For the trivial solution to be stable, the brackets multiplied by u^2 and v^2 must be positive. Taking this into consideration a final change can be made by introducing the minimal value of the derivative for function f_2' in the u^2 bracket, such that minimum $f_2'(G^* + \theta_2 u) = f_2'^m$. Thus, the following Theorem 2 for stability has been formulated:

Theorem 2. *The nonlinear system (4.1) is globally asymptotically stable if there exist strictly positive constants* ω, ϵ_i, $i = 1, \ldots, 9$ *such that the inequalities,*

$$\begin{aligned} & f_2'^m + q f_4(I^*) - \frac{1}{2} \left[q\epsilon_1 M_G f_4'^M + \epsilon_2 |f_5'^M| + \omega \epsilon_3 f_1'^M \right. \\ & \left. + |f_5'^M| \tau_2 \left(\left(\epsilon_4 + \frac{1}{\epsilon_4} \right) f_1'^M + \frac{d_i}{\epsilon_5} \right) \right] - \omega f_1'^M [\frac{\epsilon_6}{2} f_2'^M + \frac{\epsilon_7}{2} q f_4(I^*)] \tau_1 > 0, \end{aligned} \tag{6.17}$$

$$\begin{aligned} & \omega d_i - \frac{1}{2} \Bigg[q \frac{1}{\epsilon_1} M_G f_4'^M + \frac{1}{\epsilon_2} |f_5'^M| + \omega \frac{1}{\epsilon_3} f_1'^M + \omega f_1'^M \tau_1 \Big[\frac{1}{\epsilon_6} f_2'^M + \frac{1}{\epsilon_7} q f_4(I^*) \\ & + \left(\epsilon_8 + \frac{1}{\epsilon_8} \right) q M_G f_4'^M + \left(\epsilon_9 + \frac{1}{\epsilon_9} \right) |f_5'^M| \Big] \Bigg] - |f_5'^M| \frac{\epsilon_5 d_i}{2} \tau_2 > 0, \end{aligned} \tag{6.18}$$

are satisfied.

6.2 Conditions for stability based on parameters

To obtain conditions for global stability, the values of the epsilon terms, ϵ_i, must be chosen to satisfy the two inequalities in Theorem 2. There are many possible choices for the ϵ_i however it is important to try and find a set of conditions which will aid the understanding of the model and the relationship of the two time delays τ_1 and τ_2.

One method of selection which was first attempted was to extend the approach of Bennett and Gourley (2004b) for global stability conditions which was successful in the one-delay case. Their choice of the ϵ_i terms was made to remove the time delay from one of the two inequalities leaving a squared delay term in the other. For the model with two time delays the idea is to separate the delays such that there is an individual inequality for each delay, which results in a pair of disjoint quadratic inequalities. This is obtained by first setting all epsilon terms not related to the delays to one, $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon_4 = \epsilon_8 = \epsilon_9 = 1$, and letting $\epsilon_5 = \frac{1}{\tau_2}, \epsilon_6 = \epsilon_7 = \frac{1}{\tau_1}$ giving the inequalities,

$$\begin{aligned}
&f_2'^m + qf_4(I^*) - \frac{1}{2}\left[qM_G f_4'^M + |f_5'^M| + \omega f_1'^M + 2|f_5'^M| f_1'^M \tau_2\right. \\
&\left. + d_i |f_5'^M| \tau_2^2\right] - \frac{\omega}{2} f_1'^M [f_2'^M + qf_4(I^*)] > 0, \\
&\omega d_i - \frac{1}{2}\left[qM_G f_4'^M + |f_5'^M| + \omega f_1'^M + \omega f_1'^M \left[f_2'^M + qf_4(I^*)\right]\tau_1^2\right. \\
&\left. + \left(2qM_G f_4'^M + |f_5'^M|\right)\tau_1\right] - |f_5'^M| \frac{d_i}{2} > 0.
\end{aligned} \tag{6.19}$$

Through the numerical experimentation seen in Chapter 5, it is clear that the relationship between the two delays is almost linear and thus these inequalities in the quadratic delay terms are not felt to be appropriate for the two-delay model.

The motivation for the choice of the ϵ_i values to satisfy the inequalities in Theorem 2 is to maximise the area of stability. From this, further conditions on appropriate ranges of the Hill parameters can be determined. In all cases, it is first appropriate to set $\epsilon_4 = \epsilon_8 =$

$\epsilon_9 = 1$, as this leads to a minimal value for the terms in the form $\left(\epsilon_i + \frac{1}{\epsilon_i}\right)$.

Consider the first inequality as written in the following form,

$$z\left(\frac{\tau_1}{\tau_1^*} + \frac{\tau_2}{\tau_2^*} - 1\right) \leq 0, \tag{6.20}$$

where $z > 0$ and τ_1^*, τ_2^* represent the crossing points with the corresponding axes. By comparing (6.17) and (6.20), one can write the expressions

$$\begin{gathered}\frac{1}{2}f_1'^M\omega\left(\epsilon_6 f_2'^M + \epsilon_7 q f_4(I^*)\right) = \frac{z}{\tau_1^*}, \quad |f_5'^M|\left(f_1'^M + \frac{d_i}{2\epsilon_5}\right) = \frac{z}{\tau_2^*}, \\ f_2'^m + q f_4(I^*) - \frac{1}{2}\left[f_4'^M M_G q\epsilon_1 + \epsilon_2|f_5'^M| + \epsilon_3\omega f_1'^M\right] = z\end{gathered} \tag{6.21}$$

which can be solved for ϵ_2, ϵ_5 and ϵ_6 to give

$$\begin{gathered}\epsilon_2 = \frac{2(f_2'^m + q f_4(I^*) - z) - q\epsilon_1 f_4'^M M_G - \epsilon_3\omega f_1'^M}{|f_5'^M|}, \\ \epsilon_5 = \frac{d_i|f_5'^M|\tau_2^*}{2\left(z - f_1'^M|f_5'^M|\tau_2^*\right)}, \qquad \epsilon_6 = \frac{2z - f_1'^M f_4(I^*)q\epsilon_7\omega\tau_1^*}{f_1'^M f_2'^M\tau_1^*\omega}.\end{gathered} \tag{6.22}$$

The requirement that these define positive quantities leads to the following inequality in relation to τ_2^*,

$$\tau_2^* < \frac{z}{f_1'^M|f_5'^M|}, \quad \text{and} \quad z < \frac{1}{2}\left(2(f_2'^m + q f_4(I^*)) - q\epsilon_1 f_4'^M M_G - \epsilon_3\omega f_1'^M\right) \tag{6.23}$$

and therefore,

$$\tau_2^* < \frac{2(f_2'^m + q f_4(I^*)) - q\epsilon_1 f_4'^M M_G - \epsilon_3\omega f_1'^M}{2f_1'^M|f_5'^M|} < \frac{2(f_2'^m + q f_4(I^*))}{2f_1'^M|f_5'^M|}. \tag{6.24}$$

In order to maximise the value of τ_2^*, the variation of this quantity is studied with respect to the Hill parameter h_1 and also h_4. This follows from the analysis of the position of the fixed point seen in Chapter 4 where a larger effect was seen for parameter h_1, with a

smaller effect for parameter h_4. Since $f_4(I^*)$ and $f_1'^M$ are the only variables affected by changes of h_1 and h_4, Figure 6.1 presents the graphs of the variation of $f_4(I^*)/f_1'^M$ with respect to these parameters.

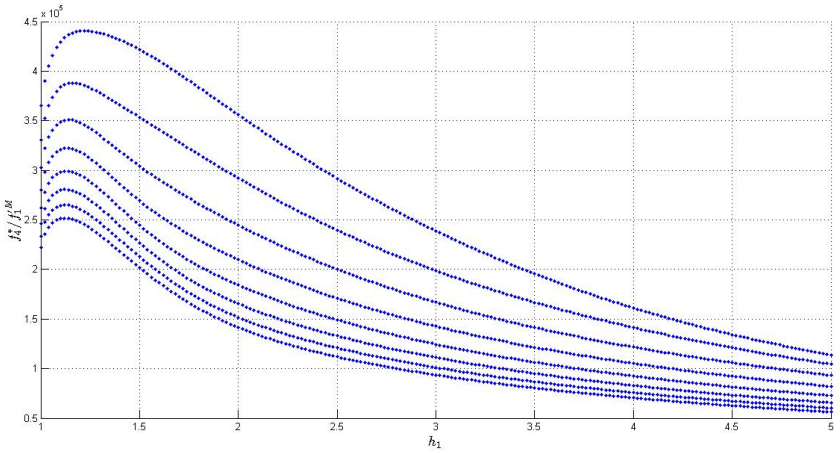

Figure 6.1: Graph for the value of $f_4(I^*)/f_1'^M$ with respect to h_1. The dots for each value of h_1 represent values of h_4.

Figure 6.1 shows that higher values of h_1 will produce a smaller region of asymptotic stability. Testing the oscillatory behaviour of the system for values close to the peak of stability causes the system to reach asymptotic stability almost immediately. In accordance with recent analytical and experimental results (Huang et al 2012), the analysis presented here indicates that values of h_1 in the vicinity of 2 will provide a model capable of describing both asymptotically stable and the desired oscillatory behaviour of the system. Further to this, it can be seen that while there is an effect of the parameter h_4, it is less significant than the effect of parameter h_1. The same pattern is observed in the relationship between h_1 and h_4 as in the case for fixed points, Section 4.3.

As the analysis of h_1 indicates that function f_1 is significantly important in the appearance of oscillations and the position of the fixed point, it is necessary to investigate parameter k_1. By analysing the maximum threshold value for τ_2^* dependent on varying k_1 with h_1 produces Figure 6.2. These graphs confirm that the Hill parameter h_1 is the dominant

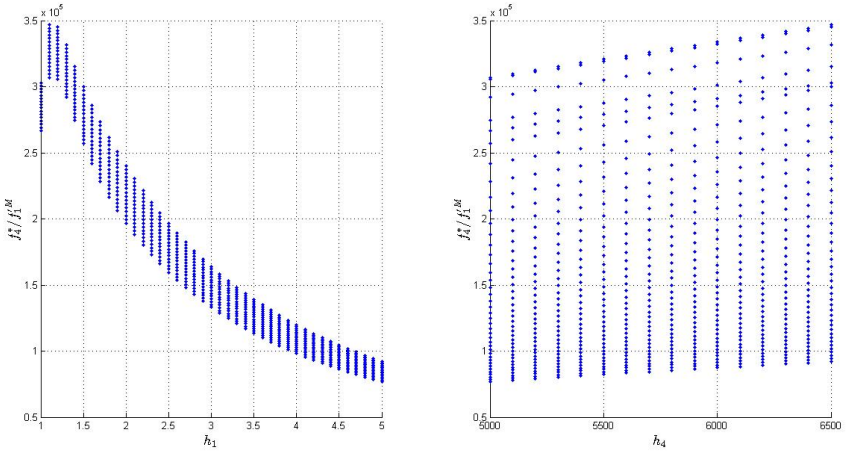

Figure 6.2: Graphs for the value of $f_4(I^*)/f_1'^M$ with respect to: (left) h_1 with dots representing k_1, and (right) k_1 with dots representing h_1.

parameter, while the effect of k_1 on the size of the stability region through τ_2^* is limited. Thus, when selecting an appropriate value for h_1, the value of k_1 can be allocated to give the best oscillatory behaviour without concern for stability regions.

By the same manner, it is possible to get an expression for τ_1^* given by,

$$\tau_1^* < \frac{2(f_2'^m + qf_4(I^*))}{\epsilon_7 q\omega f_4(I^*) f_1'^M}. \tag{6.25}$$

The variation in h_1 is also studied for maximising the stability region in terms of τ_1^* to see if there is a difference with respect to the maximal delay thresholds. The expression for τ_1^* in Equation (6.25) is simplified to $1/f_1'^M$ when varying h_1, due to the cancellation of the $f_4(I^*)$ terms. The analysis for this can be seen in Figure 6.3 which shows a similar result to the analysis of τ_2^*, confirming the optimal region for the value of h_1.

Following the analysis in this thesis, a set of ranges on Hill parameters can be suggested which allow an accurate modelling of the ultradian oscillations in the glucose-insulin regulatory system. These are reported in Table 6.1. The analysis here is focused on the h_i coefficients for which it was determined that h_1 is the key parameter. The value of h_4 also

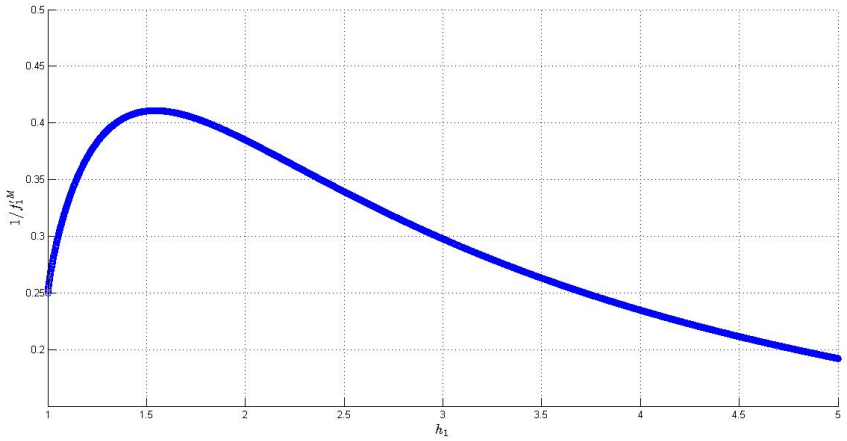

Figure 6.3: Value of $1/f_1'^M$ with respect to h_1.

Hill coefficient	Parameter range
h_1	$[1.5, 2.2]$
h_2	$[1.7, 1.8]$
h_4	$[1.4, 1.6]$
h_5	$[-9, -7]$

Table 6.1: Parameter ranges for the Hill coefficients for the modelling of ultradian oscillations.

has an effect whereas the effect of h_5 is small. The ranges for the parameters of function f_2, which was not investigated, and f_5 are based on the fitting of the Hill function to the original exponential functions, see Chapter 4. The range for parameters of function f_4 also remain close to the fitted values but are varied slightly based on obtaining a better fit. The range of values for the parameters of function f_1 has been optimised, based on the effect it has on both the position of the fixed points and maximisation of the stability region.

6.3 Discussion

This chapter has given for the first time an analysis of global stability of the nonlinear two-delay model of glucose-insulin regulation by the construction of a Lyapunov functional. A set of sufficient conditions for stability, which are in the form of two inequalities, are obtained via this new Lyapunov functional and the analysis of the nonlinear system allowed for optimal parameter ranges to be investigated in relation to the stability region.

A key contribution of this work is the analysis of meaningful parameters, which also serves as one motivation of using the Hill function. This analysis of the global stability of the two-delay system has allowed for the ranges of the Hill parameters to be optimised in relation to the stability region. The numerical analysis complements the investigation of the position of the fixed point in Chapter 4 to give a set of optimal ranges for the values of the Hill parameters, h_i.

Chapter 7

Data mining concepts and methodology

This chapter gives an introduction to the second modelling approach taken in this thesis, the application of data mining techniques to medical data. Here, the data set and some of the data mining techniques that are used to produce the results in Chapter 8 are introduced. The ideas discussed in this section will cover coarse graining of variables, obtaining a statistical significance parameter and the naïve Bayes approach to classification and prediction for building predictive models. Some comparative techniques for model building will also be discussed, along with methods of model testing. In order to illustrate the techniques, some examples from the study of risk factors for post-stroke mortality in Chapter 8 will be used where necessary.

7.1 Introduction to data mining approach

The focus of this part of the thesis will now be on the application of data mining techniques as a method of investigation of health and onset of disease for a complex system. This approach to mathematical modelling is used on a medical data set to study variables on a macroscopic level in order to build mathematical models, in this case for prediction. Diabetes, as discussed in Chapter 1, has strong links to many different conditions, one such condition is Stroke, see for example Barrett-Connor and Khaw (1988); Jorgensen et al (1994) or Diabetes-UK (2014). The data mining approach presented in this thesis will be applied to a data set of stroke patients to analyse risk factors and use them to build new predictive models of post-stroke mortality over different time horizons. This work will highlight the benefits of a data mining inspired approach to a medical data set.

This chapter will provide a background to the study in Section 7.2, before introducing the data used for analysis in this thesis followed by some of the techniques and concepts in this area which are used to create the results presented in Chapter 8.

7.2 Background

Healthcare and medicine are two important areas where the revolution in data storage and analysis has permitted access to large quantities of data which can in principle be analysed for multiple purposes. Two important scenarios in which data are collected are: (i) data associated with the clinical histories of patients - this is of relevance, for instance, to healthcare providers, insurers and healthcare authorities; (ii) data associated with clinical trials, where an intervention, such as a new drug or treatment regime, is tested on a particular group and outcomes measured. In case (i) the focus is on the individual, or a cohort of similar individuals, whereas in (ii) it is on the cohort itself or, rather, the intervention.

The existence of large electronic data bases is of course not unique to medicine, as it is common to many different fields. The data is typically of large volume and of very high dimension. This presents several complicating factors for traditional statistical modelling, and so, data mining, or knowledge discovery in data bases, as a general research area (Hand et al 2001), has developed as a useful form of statistical analysis alongside the exponential explosion of data and is now being used extensively. However, its impact in medicine and healthcare (Bellazzia and Zupan 2008) has been less, and slower, than one might expect or hope. Part of the reason for this is associated with privacy issues and also the large data fragmentation often inherent in medical data. Another potential reason, as hinted at in Mullins et al (2006), is that medicine is largely hypothesis driven, such as in testing an intervention or treatment. Indeed, this is the case with the data examined in this chapter. Whereas much of data mining is associated with knowledge discovery, where a specified hypothesis to be examined does not necessarily exist allowing for exploration of the data to provide interesting results.

7.3 Data

The data used in this case study, analysed using the techniques presented in this chapter to produce the results seen in Chapter 8, is associated with the academic multicentre randomised control trial, UK Glucose Insulin in Stroke Trial (GIST-UK) (Gray et al 2007). Patient confidentiality makes any simple, large scale analysis of clinical records difficult and so most studies involve analysis of controlled groups where there is a specific goal to the analysis. A cohort of patients are studied who have suffered an acute stroke under the conditions specified by GIST-UK protocol.

The inclusion and exclusion criteria for patients in this study is highlighted here. Eligible patients had a fixed neurological deficit and no evidence of rapid improvement during a 60

minute period after stroke onset. Patients were excluded from the trial with subarachnoid haemorrhage, isolated posterior circulation syndromes with no physical disability, pure language disorders, renal failure (urea > 20 mmol/L or creatinine > 200 mol/L), anaemia (haemoglobin < 9 g/dL), or coma (motor response ≤ 3 on the Glasgow coma scale). Also excluded were patients with an established history of insulin-treated diabetes mellitus, previous disabling stroke (modified Rankin scale [mRS] score > 3), established history of dementia or abbreviated mental test score $< 7/10$, or symptomatic cardiac failure (New York Heart Association grade III or IV) (Gray et al 2007). There were 933 patients in the study where mortality was determined over a three month period in which 267 patients died. The main types of variables used in the analysis in this thesis relate to social demographics, admission medication and readings and past medical histories. The main focus of the analysis in Chapter 8 will be post-stroke mortality.

7.4 Coarse graining for data

In the case of classification and prediction where statistical significance is important, it is necessary to have well proportioned variable groups and thus when taking a data mining approach it is often appropriate to coarse grain the variables. Throughout this work the term used for a coarse grained group will be 'bins' in order to make it clear where coarse graining has been used. Coarse graining is a method of grouping a variable by partitioning into a certain number, n, of bins. The division is dependent on the hypothesis to be tested or the desired level of data in each bin. In this work, continuous variables are coarse grained into bins each with an approximately equal amount of data. Taking an example variable of age, basic grouping of equal range, e.g. 20-30, 30-40, ..., 90-100 years may leave some groups with low quantities. When dealing with medical data certain conditions may have a highly skewed age distribution, as is the case with stroke which is

studied in Chapter 8, thus grouping in equal ranges would cause a loss of significance in the groups with small populations. Therefore, it is appropriate to create a number of bins such that the number of items in each bin is equal, this ensures that the significance of each analysis using a coarse grained variable has similar statistical weight.

The choice of number of bins is influenced by finding a balance between two important factors. Firstly, a larger number of bins potentially allows for a finer resolution in terms of time scale and the possibility of identifying the relative changes over smaller time intervals. However, by increasing the number of bins one also reduces the statistical sample in each bin thereby potentially losing statistical significance further into the analysis.

When coarse graining the dependent variable to assess risk factors for a classification problem, it may not be appropriate to choose bins of equal size. The hypothesis may require specific ranges, however it is still important to consider the statistical significance in the selection of bins for classification. From a data mining perspective, the establishment of n bins for classification of a dependent variable corresponds to n potential classes, $C_1, \ldots, C_n$. The goal is to estimate the probability of the dependent class bin, as a function of the set of risk factors $\mathbf{X}$, this is given by $P(C_i|\mathbf{X})$. This can determine which risk factors, if any, are distinct for different classes of the dependent variable. In other words, is there a change in the importance of a given risk factor?

In relation to the analysis of post-stroke mortality in Chapter 8, two choices of n bins are made for different scenarios. First, four time bins are considered, $n = 4$, so that the number of days since stroke onset which correspond to 25%, 50%, 75% and 100% of the total observed cases. The second division will be mortality over 1-7 days and 8-93 days which is to be used for building predictive models. The choices of $n = 2$ and $n = 4$ bins are a compromise between a meaningful scale for classification and a large enough sample for good statistical significance. More details can be found in Chapter 8.

7.5 Statistical significance

When creating predictive models based on a number of variables, it is important to study which variables are actual risk factors for the dependent variable. This research tests the statistical significance of potential risk factors in the data set by using a binomial test for the class of interest. This involves having a null hypothesis such that the individual variable being studied, X, does not affect the class of the dependent variable, C. This allows for the creation of a significance parameter, ε. Equation (7.1) gives the diagnostic used in this work (Grover and Vriens 2006),

$$\varepsilon = \frac{N_x(P(C|X) - P(C))}{\sqrt{N_x P(C)(1 - P(C))}}, \tag{7.1}$$

where N_x represents the size of the variable population, $P(C)$ gives the proportion of cases of the dependent variable in the total population and $P(C|X)$ gives the proportion of cases of the dependent variable given the presence of variable X. This approximation works well where there is a large N_x and a small $P(C)$. In the data introduced in Section 7.3 the probability of mortality for the full dataset is 28.6%, $P(C) = 0.286$. The values of N_x vary depending on the variable being tested, where each continuous variable is coarse-grained to ensure equally distributed 'bins' with large quantities. Some typical examples of binary variables include; a past medical history of diabetes (Yes), $N = 154$, past medical history of angina (Yes), $N = 123$, and aspirin on admission (Yes), $N = 384$.

In the case of analysing the variable age as a risk factor of post-stroke mortality as in Chapter 8, the null hypothesis is that age does not affect mortality in a given time period and thus equation (7.1) is explained such that $P(C) = P(C_T) =$ proportion of deaths in the time period T (1-5 days, 6-12 days etc.) and X = age group (33-70, 70-76 etc.). The numerator of equation (7.1) can be explained as the difference between the actual number of cases of C_T where variable X is present in relation to the expected number of

cases from an overall population. This is then measured in terms of standard deviations of the binomial distribution which is given by the denominator of the equation (Grover and Vriens 2006). The statistical significance, ε, is a measure of the extent to which the null hypothesis is verified by the data. In the circumstance, which is valid here, where the binomial distribution can be approximated by a normal distribution, $|\varepsilon| > 2$ corresponds to the standard 95% confidence interval.

The statistical significance, ε, will be used as a filter selection algorithm to determine which variables to include in the model. A cut off value for $|\varepsilon|$ is taken such that any variables with a statistical significance greater than the cut off will be included in the model. This cut off value is chosen in each case to be low enough to include an ample mix of attributes, while still maintaining a good degree of statistical significance given the size of the data set.

As a value of $|\varepsilon| > 2$ represents two standard deviations away from the mean, this is a typical value selected to be the cut off for significant variables. However the value itself can be optimised for any particular problem as $|\varepsilon| > 2$ may be too restrictive or not restrictive enough in finding the most appropriate model. The optimisation can be made in relation to a certain performance metric. For instance, the performance metric may be the model sensitivity or model specificity, a desired number of true positives or a particular number of variables in the model itself, see for example Steyerberg et al (2010). One scenario where preferable to have a higher sensitivity would be a medical screening test where it is necessary to correctly identify patients with the condition, those incorrectly classified as having the condition when they do not would be removed at the next stage. The reverse situation would lead to a patient with a condition not receiving treatment. By analysing models through varying the value of the cut off for ε, an optimal value can be found to give the most appropriate model given the goal of the investigation. An important outcome of this particular feature selection procedure is that it is highly transparent. Any

significant differences between features that appeared in different models can be easily identified. Hence, by using this methodology over an automatic feature selection which may be obtained without many details, relations between variables on a more baseline level can be uncovered and thus investigated further should this be desired.

7.5.1 Correlation-based feature selection

Other automated methods exist for selecting subsets of values, which, while they are less transparent than the epsilon method described here, are often much quicker with the aid of ready built programs in software packages such as WEKA for machine learning (Hall et al 2009). An alternative method for selecting which variables are most appropriate to be used in a model is that of a correlation-based feature selection (CFS) technique. This technique will be used for comparison in Chapter 8 and will be introduced here.

Developed by Hall (1999), CFS is a method for obtaining an optimum subset of variables based on the idea that the variables selected are correlated to, (i.e. predictive of), the dependent class variable but they are not correlated to other variables. The space of variables is searched heuristically and each subset is evaluated based on a parameter known as merit. The heuristic search allows the method to quickly find the subset with the highest merit, although with any heuristic method it may not be the very best subset. The merit parameter is a standardised version of Pearsons correlation coefficient, as seen in Equation (7.2), where $\overline{r_{cf}}$ is the average of the class to feature correlation and $\overline{r_{ff}}$ is the mean feature to feature correlation within the subset,

$$M_s = \frac{k\overline{r_{cf}}}{k + K(k-1)\overline{r_{ff}}}. \tag{7.2}$$

The three correlation methods which can be used in CFS are symmetrical uncertainty (CFS-UC), relief (CFS-Relief) and minimum description length (CFS-MDL). Symmet-

rical uncertainty, sometimes known as mutual information (Duncan 1970), is a measure based on information gain. Related to Shannon entropy (Shannon 1948), information gain gives the decrease in entropy of a variable X, given variable Y, see for example Dunham (2003). Information gain is a symmetrical measure, which is a useful property in correlation assessment, as it allows two way comparison between variables. Symmetrical uncertainty normalises the range of information gain and overrides the problem of bias towards variables with a larger number of cases. The second measure, relief, is an algorithm for feature weighting based on difference of probabilities for class interactions (Kira and Rendell 1992). Relief is not a symmetrical measure but for two variables the average can be taken from calculating, variables X to Y and Y to X. The third measure is minimum description length which looks to build the simplest possible algorithm, thus minimising the complexity (Rissanen 1978). Minimum description length is a balance of finding the simplest model which accounts for the structure of the data with each variable having a cost of been included in the data.

It is also important to point out that the CFS method assumes conditional independence of variables as is the case with the naïve Bayes analysis, (see the following subsection for details of the naïve Bayes). This is deemed to be appropriate and acceptable due to the fact that naïve Bayes produces good results even when the assumptions are violated (Zhang 2005; Domingos and Pazzani 1997). The CFS method can also deal with continuous variables by discretising the training data before the process is started by using the supervised discretisation method (Fayyad and Irani 1993). This method of discretisation extends the approach of binary discretisation via entropy, as is used in decision tress with the C4.5 algorithm (Quinlan 1993), which recursively applies the binary discretisation to a continuous variable until a criterion is met to stop making divisions. This criterion is a cost type function where a cost is applied to a correct decision and to a wrong decision. This in itself is a binary problem which uses the Bayes decision criterion to minimise the

expected cost of a making a division.

Using the CFS feature selection through the CfsSubsetEval function in WEKA allows a quick method of feature selection producing a good subset of values for a specified class variable. The default correlation method is symmetrical uncertainty, this shall be used for comparison in Chapter 8. A downside to using a method such as CFS in WEKA is that it is not possible to retrieve the information for how significant each of the independent variables selected for the subset are separately. A set of values are returned with an overall merit but without an indication of individual significance or individual correlation to the class variable. Thus, it is not possible to evaluate the variables individually. It is also not possible to view the discretisation of the continuous variables. The CFS method may return a subset of variables including the continuous variable age, but without the number of groups or the ranges of groupings used. Thus, this method is less transparent on a base level than using a statistical significance parameter such as ε.

7.6 Naïve Bayes and the score function

This research will use a common and robust classification algorithm - the naïve Bayes classifier. The naïve Bayes approximation is highly efficient in computational terms, due to the assumption that all variables are independent. It is also a very adaptable method which can perform well in many different scenarios, even if the assumption of independence is violated (Zhang 2005; Domingos and Pazzani 1997). The naïve Bayes is a common choice of prediction algorithm in many initial investigations for data mining which can set a useful baseline for more specific and sophisticated techniques.

The goal is to calculate the posterior probabilities $P(C|\mathbf{X})$, where C represents the class of interest and $\mathbf{X}$ is a vector of conditioning attributes used as predictors of the class C.

The posterior probability is the conditional probability calculated after the evidence has been taken into account. Using Bayes theorem, as given here in Equation (7.3),

$$P(C|\mathbf{X}) = \frac{P(\mathbf{X}|C)P(C)}{P(\mathbf{X})}, \tag{7.3}$$

the posterior probabilities can be related to the likelihood functions $P(\mathbf{X}|C)$. The likelihood functions can be simplified by assuming independence of the conditioning factors, $\mathbf{X} = (X_1, X_2, , X_N)$, so that $P(\mathbf{X}|C) = \prod_{(i=1)}^{N} P(X_i|C)$ – this is the naïve Bayes approximation (Hand et al 2001; Grover and Vriens 2006), as seen in Equation (7.4),

$$P(C|\mathbf{X}) = \frac{\prod_{i=1}^{N} P(X_i|C)P(C)}{P(\mathbf{X})}. \tag{7.4}$$

Equation (7.3) shows that an estimate of the posterior probabilities requires a knowledge of $P(\mathbf{X})$. However, if the overall goal is classification, where it is sufficient to know if $P(C|\mathbf{X}) > P(\bar{C}|\mathbf{X})$, with $\bar{C}$ being the complement of C, then we can consider the ratio $P(C|\mathbf{X})/P(\bar{C}|\mathbf{X})$ and assume the naïve Bayes approximation for the likelihood function for $\bar{C}$, $P(\mathbf{X}|\bar{C}) = \prod_{(i=1)}^{N} P(X_i|\bar{C})$. Rather than the ratio $P(C|\mathbf{X})/P(\bar{C}|\mathbf{X})$, it is equivalent to consider the logarithm of the ratio, defining a score function, S_{total},

$$S_{total} = \sum_{i=1}^{N} \ln \frac{P(X_i|C)}{P(X_i|\bar{C})} + \ln \frac{P(C)}{P(\bar{C})}. \tag{7.5}$$

The score function for an individual variable is denoted by S_i. This score function is the log odds ratio of the probability of having the condition, C, given X_i versus the probability of not having the condition, $\bar{C}$, given X_i. The odds ratio itself can be very sensitive to some relations and thus taking the logarithm of the odds ratio stabilises the variance and provides a symmetric distribution giving a confidence interval for the odds ratio that lies in a narrower range. Another benefit of the naïve Bayes approximation is

its transparency and interpretability. This can be seen in Equation (7.5), noting that the total risk is a sum of risk scores from the individual variables thus allowing for a simple comparison of the relative contributions of the different risk factors. As the total risk score is a monotonic function of the posterior probability $P(C|\mathbf{X})$, it is equivalent for the purposes of classification.

For classification, a cut off value can be taken to classify positive and negative results. Here, a positive result indicates an association between the variable and the class variable, C, whereas a negative result indicates that variable is not associated with the class being tested. This cut off value can be optimised based on a desired criteria. A typical method for selecting a criteria is by using a receiver operating characteristic (ROC) curve and finding the optimal balance between model sensitivity and specificity, see Section 7.8 for more details. A general guideline before optimisation would be to take $S_{total} > 0$ although optimal results are obtained through varying this cut-off, this is applied in Section 8.5. A data set with a low number of cases for the dependent variable will typically need a much lower cut off value to provide good results.

7.7 Alternative algorithms in data mining

The research presented in this thesis uses a classification approach to the problems being investigated as the principal aim is to determine risk probabilities based on the naïve Bayes algorithm. However, there are many alternative data mining techniques available for classification, as well as also viewing problems from a regression point of view. Here, some basic methods are introduced which are used for comparison in Chapter 8. Due to the nature of the naïve Bayes method, two other commonly used methods for classification will be used, these are logistic regression and C4.5 decision trees. Then from a regression view point, linear regression will also be discussed. For more detailed descrip-

tions of these methods, see for example Dunham (2003); Russell and Norvig (2010) and Wu et al (2008).

7.7.1 Classification techniques

One common, robust and applicable method for predicting a binary variable for classification is logistic regression (Russell and Norvig 2010). This technique uses a logistic function to classify an outcome as a function of the independent variables. In this thesis a simple additive logistic regression will be used assuming no interactions, as with the naïve Bayes. An advantage of the logistic regression is that it does not make assumptions of normality, linearity and homogeneity of the independent variables as is the case with linear regression, see Section 7.7.2. Logistic regression assumes that the dependent variable is binary, this is the case in Chapter 8 where post-stroke mortality is modelled. There is also the same assumption on independence of variables as with the naïve Bayes classifier, see Section 7.6. Again, similarly to the naïve Bayes, logistic regression is a relatively robust technique which can produce a good model even if the assumption of independence is not met. However, one drawback of the logistic regression is that it is recommended to have a large number of cases per independent variable, sometimes referred to as events per variable (EPV), and so may not prove successful for smaller data sets with many variables (Peduzzi et al 1996). Logistic regression gives a probability of success for the independent variable, y, as,

$$p(y) = \frac{\exp^{a+b_1x_1+b_2x_2+...}}{1+\exp^{a+b_1x_1+b_2x_2+...}}, \tag{7.6}$$

where a, is a constant term, x_i represents the independent variables and b_i are their corresponding coefficients. Models for classification can be built with feature selection dependent on different criteria such as the Akaike information criteria (Akiake 1974). When

tested, the probability cut off for classification can be optimised in a similar way as discussed for the score total in Section 7.6.

Another common methodology used in data mining is the decision tree (Russell and Norvig 2010). This technique involves splitting the problem into subsets which systematically lead to a prediction based on information at a number of nodes. Many different classification algorithms exist for building decision trees including ID3, C4.5, CART and CHAID (Wu et al 2008). A typical choice for classification is the C4.5 algorithm which will be the focus here (Quinlan 1993). C4.5 is a developed version of ID3 that can deal with missing values and both continuous and discrete data. The continuous variables are discretised using a binary discretisation method to find the optimal cut off point to split the variable. A restriction of this method is that it only gives two possible groups for a continuous variable at any one node, when it may be more appropriate to split the continuous variable into more groups. This can also lead to problems when continuous and discrete variables are combined.

The C4.5 algorithm uses information gain to create a decision tree (Quinlan 1993). Information gain typically has bias towards variables with a higher number of values, giving them higher worth. Thus, when a mix of variables are used, including discrete variables with a large number of options, they are more likely to be picked out as significant by the algorithm. To build a decision tree, the variable with highest information gain is chosen as its root node, this is the first node at the top of the decision tree. The data is then partitioned along the values of the selected variables. This process is repeated to select the following nodes one by one until all cases in the data being used to build the tree are correctly classified. The information gain is based on information entropy, otherwise known as Shannon entropy (Shannon 1948), and is given as,

$$Entropy = -\sum_{i=1}^{n} P(x_i) \log_2 P(x_i), \tag{7.7}$$

where n is the number of possible outcomes for the discrete variable. The information entropy calculates the amount of information needed on a scale from one to zero to correctly classify an event. Once the full tree is complete, post-pruning occurs aiming to reduce overfitting. This is the problem of the tree being too well trained to the training set, making it a poor model for different data. Pruning removes unnecessary branches dependent on a statistical significance. C4.5 models the error rates as a Bernoulli random variable and approximates the error as a normal distribution (Quinlan 1993). A confidence level is set and the error estimate for each node is tested, if possible the tree is pruned. Similarly to the logistic regression, this works best for large data sets. With a small data set containing an imbalance of cases it can be difficult to build a good decision tree to classify the cases.

7.7.2 Linear Regression

Aside from classification, a prediction approach could be taken by using a continuous target function. This allows direct consideration of regression techniques, such as linear regression. Multivariate linear regression looks for a simple linear relationship between the dependent variable and the independent variables in order to give an expression for prediction of the dependent variable, see for example Dunham (2003). Linear regression can only take place if certain assumptions are met, failure of the data to meet with these assumptions results in a very poor model. These include a linear relationship between the dependent variable and all independent variables, normal distribution of the residual errors, constant variance across the variables known as homoscedasticity, and there must be no multicollinearity as is the case for all multiple regression models. Linear regression normally uses a continuous dependent variable. For the purpose of classification, the continuous variable can then be split into the respective classes following the analysis and hence categorisation could be made during testing of the model.

7.8 Methods of testing binary classification models

7.8.1 Score decile testing

The first method used for testing the performance of the naïve Bayes predictive classification models in this research is decile testing. After building the model and testing on the hold out one third test data set, each individual will have a resultant total score, S_{total} as seen in equation (7.5). The decile method of testing involves ranking the resultant scores from highest to lowest score and dividing into ten equally sized groups, hence the name deciles. The decile containing the highest score values, either score decile 1 or 10 depending on how they have been specified, will contain the top 10% of scores, the second decile will contain the next 10% and so on. Within each decile the number of identified positive cases can be counted and the proportion of the full decile can be calculated. The pattern for a good model is to see the decile of highest scores contain the highest proportion of positive cases and this decreasing through the deciles. This can also be checked against the average proportion of cases in each decile. An example of this is that for a particular model, decile 2 may have a range of score from $2.4 - 3.1$ with a result for the decile performing 85% better than the average. Therefore a new case with a score in that range for decile 2 would be correctly categorised with an 85% higher probability than the average.

7.8.2 Confusion matrices and related measures

Presented here is a short discussion of confusion matrices and their relation to sensitivity and specificity as a measure of model performance. A confusion matrix, sometimes referred to as a contingency table, is a common technique in statistical based data mining for representing the performance of a classification model (Dunham 2003). Any classifi-

cation model for classification with a clearly defined positive and negative result can be tested via a confusion matrix. There are four possible outcomes which are used to build a confusion matrix, true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN). For example, taking a dependent variable where the response 'Yes' is a positive and 'No' is a negative such that,

- TP - Correctly identified as 'Yes' result
- FP - A 'No' response incorrectly identified as 'Yes' class
- TN - Correctly identified as 'No' class
- FN - A 'Yes' response incorrectly identified as 'No' class.

A confusion matrix visualises these four outcomes in the following form,

		Actual diagnosis	
		Positive	Negative
Prediction	Positive	TP	FP
	Negative	FN	TN

From the confusion matrix it is easy to calculate a number of different statistical measures and error rates to test the performance of the model. Here, the focus is placed on sensitivity and specificity.

Sensitivity is a measure of the number of correctly identified positive cases from the classification. This is calculated as the number of correctly identified positive results out of the total number of actually positive results in the data. Similarly the specificity relates to the negative cases, namely, the number of correctly identified negative results out of the total number of actually negative results (Lalkhen and McCluskey 2008). The equations

for sensitivity and specificity are shown below,

$$Sensitivity = \frac{TP}{TP + FN} \qquad Specificity = \frac{TN}{TN + FP}. \tag{7.8}$$

These measures are a simple but effective base guideline to the performance of a model. A perfect model would have 100% sensitivity and specificity, however this is virtually impossible, especially in a real world scenario where there is usually a balance between the two. A model with high sensitivity means it correctly identifies the positive results, therefore a negative result is a very strong indicator that it is not in the class. The specificity is again the reverse of this where high values means it is highly successful at finding the negative results for those not in the class, and hence a positive result for a model with high specificity indicates that there is a very high chance of it being a correct positive classification. For analysis of medical data, it is important to have models with a higher sensitivity to a lower specificity, rather than the reverse so that a higher proportion of the patients with a particular disease/illness (a positive result) are identified and the appropriate treatment can be given.

The relation between sensitivity and specificity can be graphically represented by a ROC curve which plots the sensitivity as a function of one minus the specificity. This can also be referred to as the true positive rate plotted against the false positive rate. When testing models using a ROC curve, the preferred model can be selected based on the area under the curve (AUC). A larger AUC indicates a better model, where a perfect model, having 100% sensitivity and specificity would be a right angle with maximum area under the curve. Viewing a ROC curve with the naked eye, the best model is that which runs closest to the top left corner of the graph, the point (0,1) where both axes run from zero to one.

Chapter 8

Investigation of risk factors and prediction of post-stroke mortality for different time scales

This chapter presents the results of the data mining based approach, discussed in Chapter 7, used to assess the risk factors of post-stroke mortality with the aim of building usable predictive models. Initially, an investigation into the natural time scales of post-stroke mortality is made in order to provide reasonable time scales for the application of risk factor analysis and prediction. A study on the role of age in post-stroke mortality is presented before giving an analysis of all risk factors. These risk factors are then used in the building of predictive models using a naïve Bayes approach for very short term mortality and short/intermediate term mortality. The naïve Bayes based approach proposed in this work is then compared to other common algorithms in the area.

8.1 Introduction

The work in this chapter will provide an investigation into the risk factors of very short and short/intermediate term post-stroke mortality in order to build predictive models to correctly classify those who are at risk of death in the specified time scales. A naïve Bayes approach will be taken to build these models and the investigation will use the techniques discussed in the previous chapter to deliver interesting and useful results. All results in this chapter have been published, see Easton et al (2014).

In principle, any one of a host of different statistical modelling techniques could be used, each of which have their own advantages and disadvantages dependent on their assumptions, along with the type of data being used and the hypothesis to be tested. Comparisons of such techniques can be found for a different number of scenarios (Long et al 2009; Caruana and Niculescu-Mizil 2006). Searching through the literature shows that there are not many predictive models for post-stroke mortality and those which do exist use a range of different techniques and are often linked to specific aspects such as tomography for intracerebral hemorrhage (Zahuranec et al 2012), intensive care unit therapeutic intervention (Riachy et al 2008) or coronary artery bypass (Kluck et al 2007). Other prediction models related to stroke also exist but related to a particular condition, examples include post-stroke infection (Fluri et al 2012) and recurrence of vascular disease (Castillo et al 2009).

An original contribution of this work is to build predictive models for post-stroke mortality over different time horizons, to do this a naïve Bayes approach will be taken due to the benefits discussed in Chapter 7. Due to the lack of general predictive models for post-stroke mortality the naïve Bayes is an ideal method to produce initial models.

This chapter will first see a discussion of the time scales to be studied for post-stroke mortality. A detailed investigation specifically into age as a risk factor will be made

before an overview of risk factors will be discussed. These risk factors will then be used to build models to classify those patients at risk of very short and short/intermediate term post-stroke mortality with finally testing of these models and a comparison to other techniques.

8.2 Classification of stroke mortality timescales

As there is an elevated risk of mortality for an extended period after a stroke, it is of interest to ask if the risk factors for mortality are constant across time. For example, are the risk factors for mortality over different time horizons the same? A relevant question is then: what are the appropriate time horizons for considering mortality? What might be considered short, intermediate or long term?

Many analyses of stroke mortality, as a function of days from stroke onset or days since admission, have considered short term to be associated with periods of less than a month, with typical categorisations being short term < 1 month, intermediate term 1-3 months and long term 3-12 months (Collins et al 2003; Bronnum-Hansen et al 2001; Palnum et al 2009; Bamford et al 1990). However, typical mortality curves, as a function of days since admission, show that between approximately 50% (ischaemic) and 70% (haemorragic) of stroke patients who will die in the intermediate term (up to 93 days) are already dead within 30 days (Collins et al 2003).

With this in mind, the motivation here is that a more appropriate characterisation of short, intermediate and long term comes from an examination of the cumulative mortality curve itself. The curve for this data set can be seen in Figure 8.1, where the cumulative proportion of deaths as a function of all deaths through the 3 month period in which mortality was measured. The form of the curve suggests that immediacy of death should be more

appropriately categorised by dividing the cumulative curve for death related to days from stroke onset into n coarse grained time bins.

First, four time bins, $n = 4$, are considered so that from Figure 8.1 the number of days since stroke onset which correspond to 25%, 50%, 75% and 100% of the total number of deaths will be observed. As can be deduced from the data of Figure 8.1, the corresponding time frames for the four bins are: 5 days (25%), 12 days (50%) and 33 days (75%). Four bins are used in Section 8.3 as a specific variable is being investigated and it is of interest to see the change in risk over smaller time scales. The second division will be mortality over 1-7 days and 8-93 days to be used for building predictive models. In this case, the motivation for choosing two mortality periods was so that the statistical significance of the results would be higher when investigating a wide range of variables and represents an approximately 30% to 70% split. Also, one week was chosen as it is better interpretability as a time period than 1-5 days. Given that short term is usually defined in the literature as < 1 month, here the period $1-5$ days is defined as *very short* term, $6-33$ days *short* term and $34-93$ days *intermediate* term for the analysis of age alone. For the predictive model analysis 1-7 days will be termed *very short* term and 8-93 days as *short/intermediate* term.

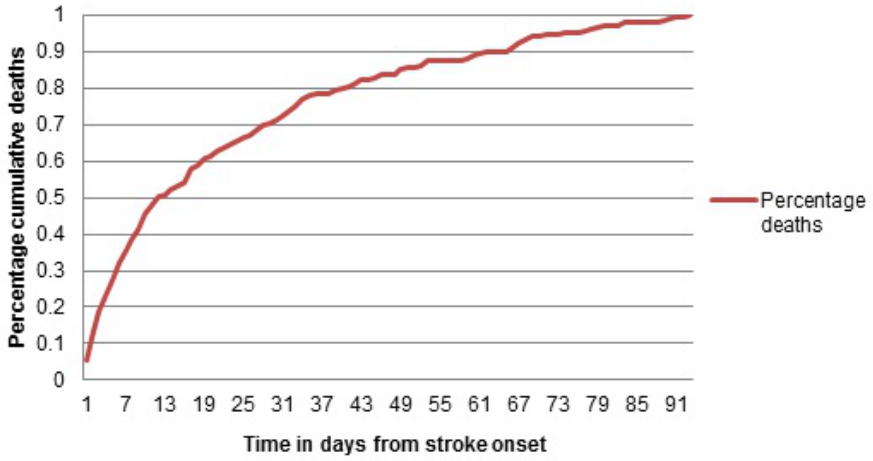

Figure 8.1: Cumulative proportion of post-stroke mortality as a function of days from stroke onset.

8.3 Investigation into age as a risk factor for post-stroke mortality

With the goal of determining the variation of risk factors over time, a particularly relevant variable is that of age, where the phenomenon of frailty might be expected to play a significant role. Age has been chosen due to its strong relationship to stroke and also type two diabetes (Sacco et al 1997), while being relatively interpretable. Figure 8.2 gives the proportion of patients of a given age who have died in different mortality periods in number of days from stroke onset. The age ranges were determined by a process of coarse graining based on $n = 4$ bins such that roughly equal numbers of participants were included in each bin. Thus, each age bin has the same statistical weight. It can clearly be seen that the oldest age group has the highest proportion of deaths, corresponding to the fact that age itself, as is well known (Mustacchi 1985; Sacco et al 1997), is an important risk factor for stroke mortality.

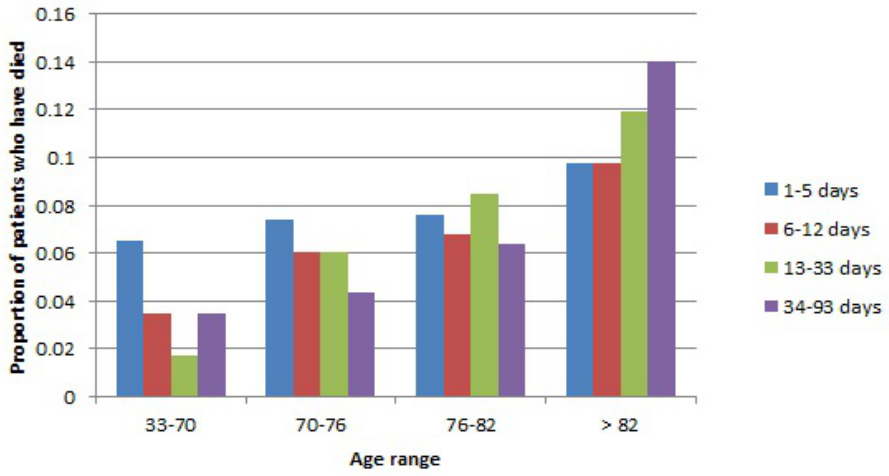

Figure 8.2: Proportion of patients from the full population who have died in a certain number of days since admission as a function of the patients age.

What is clear from the graph is that the importance of age as a risk factor for mortality is quite different for shorter or longer term mortality, with a relatively weak dependence

for very short term death, 1-5 days, versus a very strong dependence for death in the period 34-93 days. Younger stroke victims (33-70) are clearly more robust with respect to death as a function of time, with 1-5 days being a much more significant timescale than 6-93 days. Whereas, for the oldest stroke victims it is suggested here that the strongly increasing mortality as a function of time from stroke onset is an indication of increased frailty in this group. In this context, frailty is given as the higher susceptibility to adverse outcomes which is common in elderly people as a general deterioration of health. Thus, over a longer time period there is more opportunity for the detrimental effects of suffering a stroke to cause mortality for an already frail group (Fulop et al 2010; Koller and Rockwood 2013; Pendergast et al 1993).

As the different time bins for days since admission are not uniform in terms of number of days, Figure 8.3 shows the proportion of deaths per day as a function of age and days since admission. The relative mortality risk as a function of age as seen in Figure 8.2 remains clearly visible, with the death risk per day being only about 40% higher for age group > 82 than age group $33 - 70$ in the very short term $1 - 5$ day period, but more than 400% higher for the intermediate $34 - 93$ days period. Here, it can also be seen that the true risk as a function of days since admission is a strongly decreasing function of time for all age groups.

The statistical significance of these results can be tested with a binomial test for the class of interest using the approach discussed in Chapter 7. The null hypothesis is that age does not affect mortality in a given time period. In the circumstance, which is valid here, where the binomial distribution can be approximated by a normal distribution, $|\varepsilon| > 2$ corresponds to the standard 95% confidence interval. The results are given in Figure 8.4, which shows that the enhanced mortality of the oldest age group as a function of time from stroke onset is statistically significant at a greater than 95% confidence level for intervals 6-12, 13-33 and 34-93 days. For the age group 33-70, mortality in the intervals

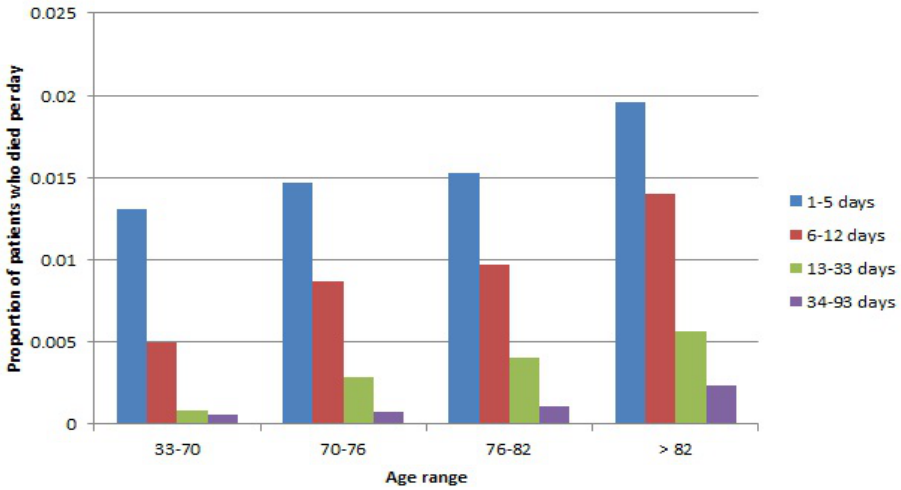

Figure 8.3: Proportion of deaths per day as a function of age and days from admission.

6-12, 13-33 and 34-93 days are statistically significantly smaller than average. There are no statistically significant differences between mortality rates of age groups 70-76 and 76-82 for any of the time intervals. From the perspective of time interval as opposed to age group, for 1-5 days - very short term mortality - we see that there is no statistically significant difference between age groups.

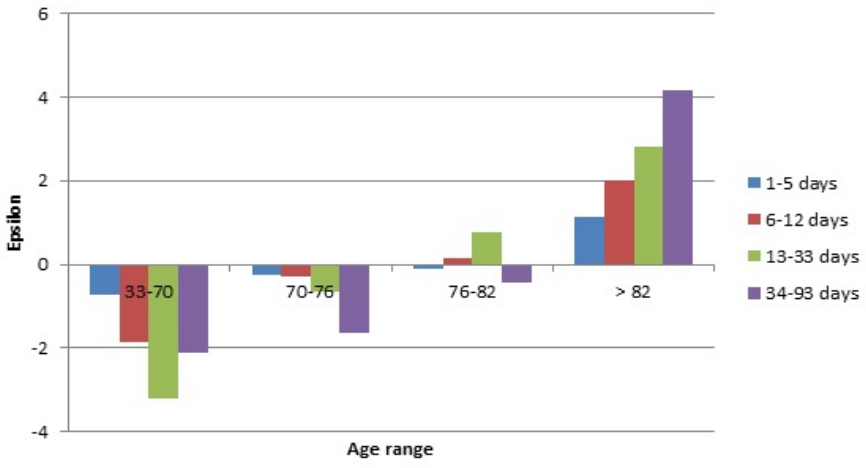

Figure 8.4: Binomial tests for the statistical significance, ε, of the results shown in Figure 8.2, for proportion of patients who have died in a specified time range with age.

In understanding whether this pronounced age effect is a consequence of frailty in the oldest age group, it could instead be suggested that the effect is a result of a higher incidence

of more severe strokes in the older versus younger population. There are four levels of stroke severity contained in the data set as classified by the Bamford stroke classification, also known as the Oxford stroke classification (Bamford et al 1991). These are known as TACS, PACS, POCS and LACS, listed here in decreasing order of severity. Within the data each stroke type is well represented except for POCS where there are 16 cases. The acronym TACS refers to a total anterior circulation stroke, and is the most severe. The patients who suffered a TACS type of stroke is used here. Figure 8.5 gives the proportion of TACS sufferers who have died over a certain time period as a function of age, and in Figure 8.6 the statistical significance of the results is given. It can be seen that qualitatively the results are as in Figures 8.2 and 8.4. In other words, the effect of advanced age is such as to lead to increased mortality over longer time intervals from admission independently of the stroke type. Due to the strong effect of age, the same results would be expected for all stroke types.

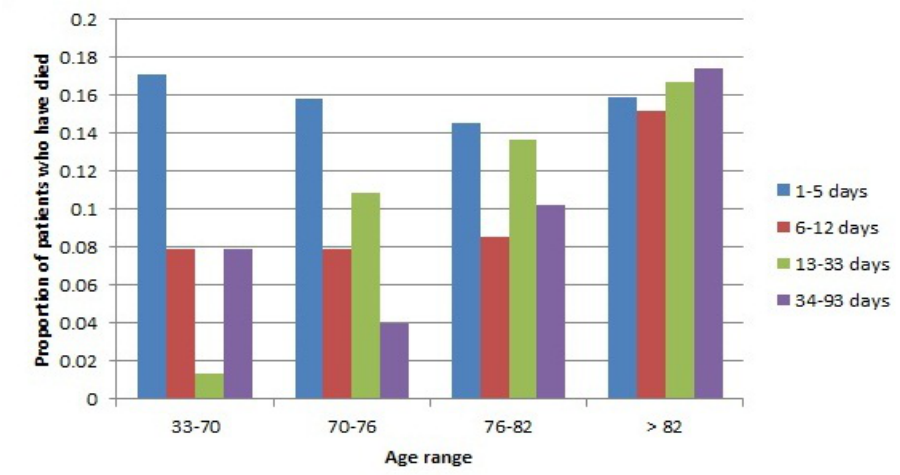

Figure 8.5: Proportion of patients who suffered a TACS stroke type who have died in a certain time period since admission as a function of the patients age.

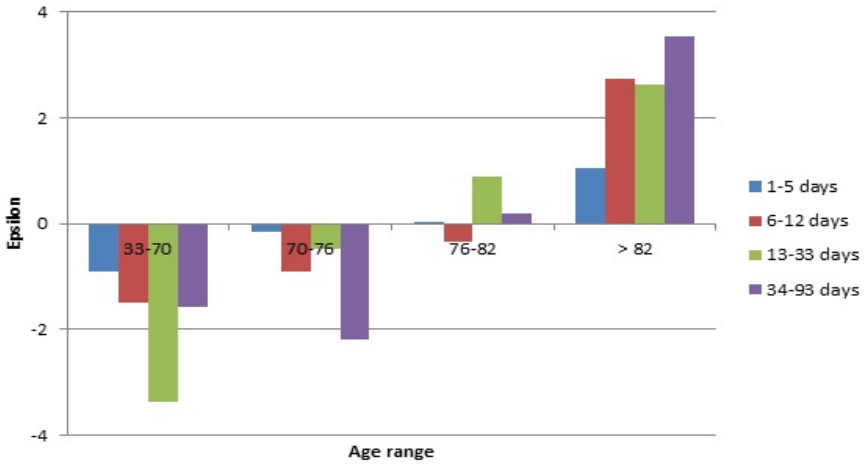

Figure 8.6: Binomial tests for statistical significance, ε, of the results shown in Figure 8.5, for proportion of TACS patients who have died in a specified time range with age.

8.4 Analysis of risk factors for stroke mortality as a function of time

As well as age, any other factor can be examined as a potential discriminator between shorter and longer term mortality. The null hypothesis is that for the mortality period of interest, the mortality rate is not affected by the factor, X. The alternative hypothesis is the two-sided general case where the factor, X, does affect the mortality rate. Table 8.1, which has been included here as an example, and Tables A.1 - A.4 of Appendix A display the list of all variables selected for initial investigation from the data set. The type of variables selected were loosely based on the criteria that the information about the patients could be obtained within the first 24 hours of admission, this includes socio-demographics, past medical history and drugs on admission. The tables give the corresponding probabilities, P, and statistical significance, ε, values for two different ranges of death, very short term death in 1-7 days and short/intermediate term death, 8-93 days. The restriction to two time intervals rather than four, as seen with the age analysis, is made to enhance the statistical significance of the results. Also included is the relative

difference in the probabilities of death between very short and intermediate term - the relative change in the odds ratios, D. This is calculated for death 1-7 days versus 8-93 days for each potential predictive driver, X, against the overall odds ratio for the two classes. Thus, $D > 0$, respectively < 0, means that the driver is associated with a higher/lower mortality rate 1-7 days versus 8-93 days for that risk factor than would be expected, the reference point being the average ratio over all risk factors.

Table 8.1 shows the results for stroke type, age and solution administered. Assessing stroke type, it is logical that the most severe TACS stroke, used in the age analysis in Section 8.3, has the highest risk for both short and medium term death. Also, for stroke type, both partial anterior circulation syndrome, PACS, and lacunar syndrome, LACS, have an increased mortality rate for the longer time period than the initial 1-7 days. Interestingly, the usage of the GKI (glucose, potassium and insulin) intervention, assessed as the original purpose of the GIST-UK study, is associated with an enhanced survival rate in the period 1-7 days and an enhanced mortality risk over the period 8-93 days. As the intervention was only given for a short period of time, this could indicate that the intervention is efficacious only if continued. Table 8.1 also shows the effect of age, particularly evident in the 8-93 day period.

Considering drugs on admission, seen in Table A.1 of Appendix A, it is evident that warfarin and digoxin both correlate with an enhanced mortality over the period 1-7 days when compared to the 8-93 day period. As both drugs can be used to treat heart problems, namely atrial fibrillation for digoxin (NHS 2012b) and, in the case of warfarin, blood clots (NHS 2012a), this enhanced risk may be due to other health conditions the patient is suffering from or, alternatively, could be a direct affect for stroke patients. In contrast, the opposite is seen when administering lipid lowerers or diuretics, as this is associated with a reduced mortality rate over the interval 1-7 days. For variables associated with past medical history, having either had a previous stroke or hypocholesterolemia is associated

Variable	Group	P(Dead 1-7|X)	P(Dead 8-93|X)	P(Survived|X)	ε Dead 1-7	ε Dead 8-93	ε Survived	Relative change in odds 1-7 v 8-93
Stroke Type	TACS Stroke	0.19	0.29	0.52	4.56	4.24	- 6.74	0.16
	PACS Stroke	0.04	0.13	0.82	- 2.79	- 1.93	3.55	- 0.42
	LACS Stroke	0.01	0.06	0.93	- 3.64	- 4.05	5.95	- 0.57
	POCS Stroke	0.33	0.33	0.33	1.24	0.63	- 1.39	0.74
Stroke Side	Left side	0.10	0.20	0.70	- 0.56	0.54	- 0.09	- 0.15
	Right side	0.11	0.18	0.71	0.28	- 0.54	0.27	0.12
	Bilateral	1.00	-	-	2.85	- 0.48	- 1.53	
Arm (Infusion Type)	GKI	0.08	0.23	0.69	- 1.55	1.74	- 0.44	- 0.38
	Saline control	0.14	0.15	0.71	1.52	- 1.71	0.43	0.55
Age	0< 1 <63.5	0.12	0.08	0.81	0.17	- 2.54	2.06	1.60
	63.50<= 2 <70.16	0.08	0.10	0.82	- 0.92	- 1.96	2.31	0.30
	70.16<= 3 <73.34	0.12	0.17	0.72	0.17	- 0.52	0.33	0.20
	73.34<= 4 <75.86	0.09	0.15	0.76	- 0.55	- 0.81	1.07	0.01
	75.86<= 5 <79.48	0.09	0.15	0.77	- 0.55	- 0.81	1.32	0.01
	79.48<= 6 <82.20	0.10	0.22	0.67	- 0.19	0.64	- 0.66	- 0.18
	82.20<= 7 <85.95	0.18	0.25	0.57	2.04	1.28	- 2.48	0.28
	8: 85.95 +	0.10	0.40	0.49	- 0.15	4.76	- 3.98	- 0.55

Table 8.1: Probabilities and significance levels for stroke type variables. Shading in the ε columns correspond to variables that were included in the predictive corresponding model. In the relative change column highlighted drivers correspond to variables that discriminate between 1-7 and 8-93 day mortality.

with a reduced mortality rate for 1-7 days, whereas atrial fibrillation and heart failure are associated with an enhanced mortality in the 8-93 day period. Interestingly, despite being well known causal risk factors having a stroke (Diabetes-UK 2014), neither the presence of diabetes mellitus or hypertension are statistically significant risk factors for this group in terms of post-stroke mortality.

The variables related to readings on admission are seen in Tables A.3 and A.4. Due to the coarse grained nature of these continuous variables it is possible to see clear patterns. These patterns often have association to reduced or enhanced mortality in the extremes. This is the case with glucose, temperature and plasma creatine. Certain readings do not differentiate between the two mortality periods at lower values but become significant for one mortality period as the level increases. This occurs in white cell count and systolic blood pressure where the highest group is a significant factor for the 8-93 day mortality period. It can also be seen that the different stroke scales which currently exist, Euro-

pean Stroke Scale (Hantson et al 1994), Barthel index (Mahoney and Barthel 1965) and modified Rankin (van Swieten et al 1979), in and of themselves offer little systematic discrimination between short term and intermediate term mortality although they are predictive of mortality in general.

8.5 Building predictive models for post-stroke mortality with analysis

Now that significant risk factors which are distinct for very short term versus short/intermediate term mortality have been identified, two predictive models are constructed. These are of binary form, where a death in the specified time scale is defined as the positive set, against all other patients defined as the negative set. Thus, there is one model to predict mortality in the first week after the stroke, with a second model to predict death between 8 and 93 days after the stroke. By using the naïve Bayes approach, as discussed in Chapter 7, each variable can be assigned a score. These individual scores are used to give each patient a total score, S_{total}. Models can be built for a range of variables dependent on feature selection by an optimisation of the significance parameter, ε. In order to select which variables are included in the model a simple filter feature selection algorithm was used by ranking the features based on the statistical significance, ε. Only those features with $|\varepsilon| > 1.5$ were included in the predictive model. This value was chosen by an optimisation procedure based on the performance metric of model sensitivity. The value of ε is low enough so as to include an ample mix of attributes, while still maintaining a good degree of statistical significance given the relatively small size of the data set. For comparison, the well-known correlation-based feature selection (CFS) technique (Hall 1999), as discussed in Section 7.5.1, was used. CFS produced a similar set of attributes to the $|\varepsilon| > 1.5$ criterion, including age, stroke type and digoxin. It appears that the difference

in results between feature selection methods is minimal.

This approach was used for both mortality periods to give two predictive models. The models were built using features with a statistical significance, $|\varepsilon| > 1.5$, and the naïve Bayes algorithm to assign a score. These models can now be tested on the test data set.

To evaluate the performance of the predictive models, the patients in the test set were ranked according to their risk score and then divided into score deciles. The number of deaths associated with each time range were counted for each decile. Both score models were tested for both time scales and the survivors in each decile were also counted. The results can be seen in Figures 8.7 and 8.8.

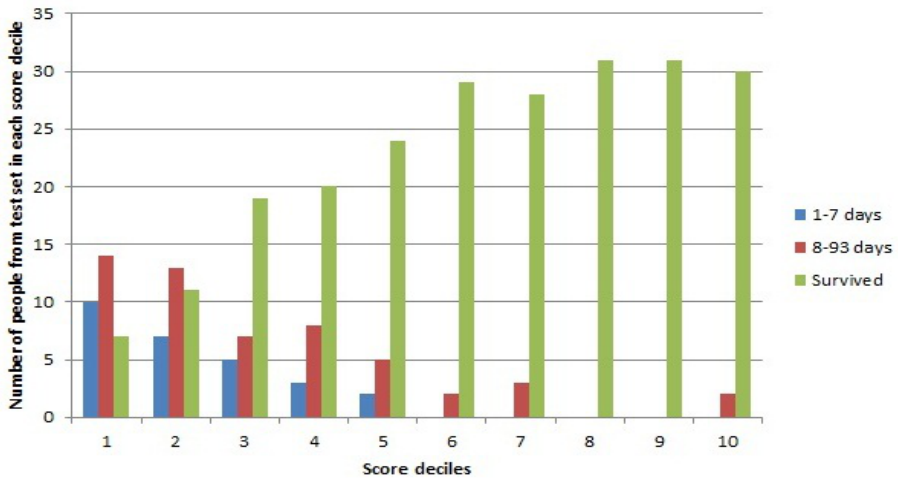

Figure 8.7: Performance of the predictive model for short term mortality risk (1-7 days), including risk factors with $|\varepsilon| > 1.5$.

Figure 8.7 shows the results of testing for the model built on mortality in the first week. By assessing the columns for death in 1-7 days, it can be seen that decile one has the most deaths in one to seven days and the number of deaths follows a decreasing pattern towards decile five. No patients who were categorised in deciles six to ten died within the first week. This shows that the model proposed in this work is a good predictor of those patients who are unlikely to die within the first week. Therefore, if a patient enters a clinic and receives a score which corresponds to the range from deciles six to ten, it is

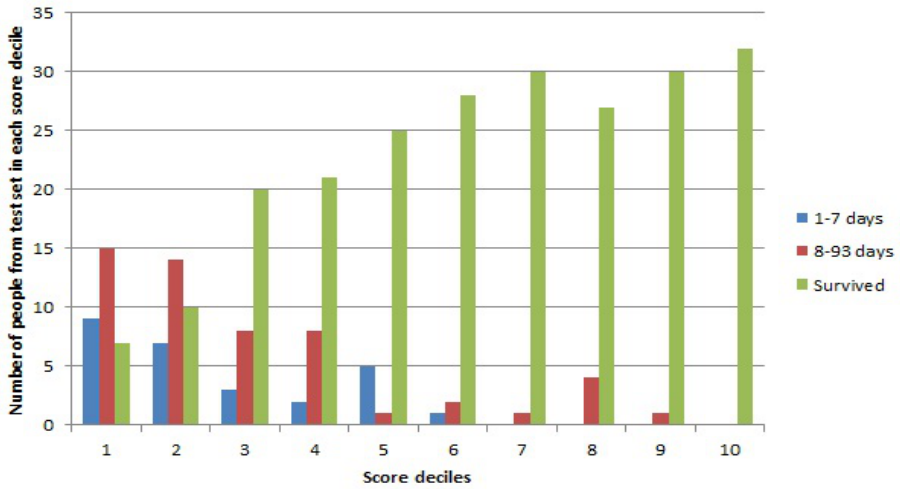

Figure 8.8: Performance of the predictive model for intermediate term mortality risk (8-93 days), including risk factors with $|\varepsilon| > 1.5$.

very unlikely they will suffer short term death and hence are not at immediate risk.

The model testing also counts the patients in the 8-93 day category and survival. The model for 1-7 days is, albeit to a lesser extent, a good indicator of patients who are also unlikely to die within 8-93 days. The deaths in this category are more spread out across the deciles, but a strong tendency towards the first five deciles still exists. Even for the surviving patients this model shows the expected inverse pattern of the mortality periods.

Figure 8.8 shows the same analysis carried out for the 8-93 day model. Focusing on the columns representing death in 8-93 days, there is somewhat better discrimination for this dedicated 8-93 day model as opposed to using the 1-7 days model to predict over the interval 8-93 days. The pattern is as expected, the difference being it is much more spread out than the one week model. The decreasing trend is displayed from decile one to decile nine but only the final score decile, 10, has no deaths. The results of the 8-93 days model may be less conclusive due to the much wider time period involved compared to the one week model.

The corresponding confusion matrices can also be considered for model performance on

our test set. A classification criteria was selected through optimisation for sensitivity in order to give a well fitting model. This value can be obtained through a ROC curve as those seen in Figures 8.9 and 8.10. The selected criteria were $S_{total}^{1-7} > -2.7$ and $S_{total}^{8-93} > -1.0$. Below the confusion matrix for the 1-7 day model can be seen and in (8.1) the corresponding sensitivity and specificity. The analogous results for the 8-93 day model are shown in (8.2).

		Actual diagnosis		
1-7 day		Positive	Negative	Total
Prediction	Positive	25	87	112
	Negative	2	197	199
	Total	27	284	311

$$\text{Sensitivity} = \frac{25}{25+2} = 92.6\%, \qquad \text{Specificity} = \frac{197}{197+87} = 69.4\% \tag{8.1}$$

		Actual diagnosis		
8-93 day		Positive	Negative	Total
Prediction	Positive	45	75	120
	Negative	9	182	191
	Total	54	257	311

$$\text{8 - 93 day sensitivity} = 83.3\%, \qquad \text{8 - 93 day specificity} = 70.8\% \tag{8.2}$$

Prospective models for the prognosis of death over different time scales could in principle be used for predicting mortality risk so as to delineate high-risk patients and facilitate prevention strategies. An important aspect of the analysis here is that it emphasises that the degree of risk depends on the time scale of interest and so intimates at the fact that prevention strategies can and should change over time. The profile of a very high risk

patient over the very short term is not necessarily the same as one who is high risk over the intermediate term.

Various different stroke risk scores have been used for predicting different outcomes (Hantson et al 1994; Mahoney and Barthel 1965; van Swieten et al 1979; Muir et al 1996; Nedeltchev et al 2010; Hénon et al 1995). Here, a comparison is made between the predictive models proposed in this work and the ESS score for both the 1-7 day and 8-93 days intervals. Figures 8.9 and 8.10 show the ROC curves for each model. For short term 1-7 day interval our predictive model does not outperform the ESS score model with the area under the curve (AUC) given as 0.858 for the seven day model compared to 0.867 for the ESS. However, for the 8-93 day interval there is a significant improvement from the bespoke model with AUC of 0.807 compared to 0.757 for the ESS, thereby showing the potential utility of different predictive models for different time intervals after stroke onset.

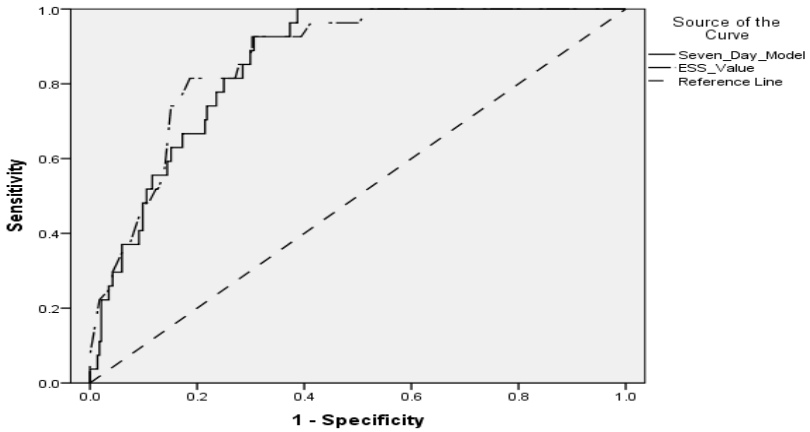

Figure 8.9: ROC curves for the 1-7 day prediction model and ESS score.

Rather than just give a score that measures relative risk only, the risk score can be used to actually determine the mortality probability. This is done by taking the test set divided into ranked deciles, as before, and calculating $P(death|\bar{S}_i)$ for each decile, where $\bar{S}_i$ is the average score for the i decile. The results can be seen for the 1-7 day model in Figure 8.11. The advantage of this is that, in principle, once a risk score has been determined for

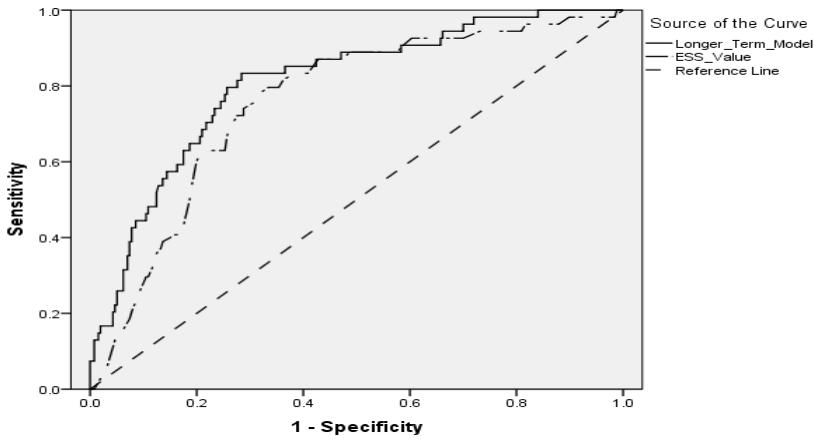

Figure 8.10: ROC curves for the 8-93 day prediction model and ESS score.

any individual it can be directly related to a mortality probability. For instance, a patient with a risk score of 4.0 would have a probability of dying in the first 7 days post-stroke of 30%, whereas a patient with risk score 1.0 would have a probability of 15%.

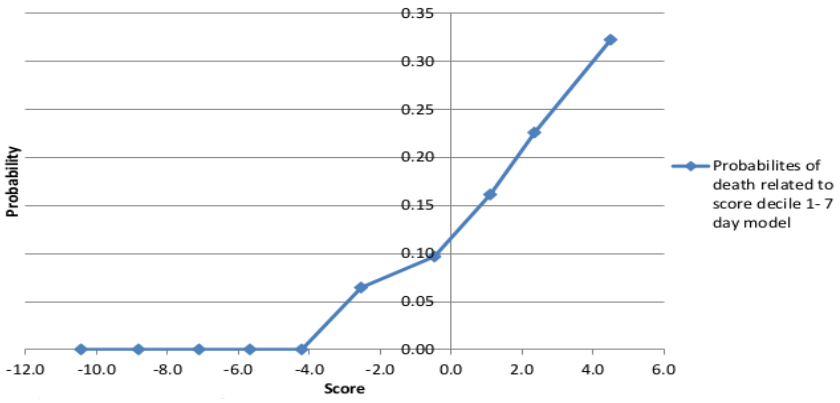

Figure 8.11: Relationship between risk score and mortality probability for 1-7 day model.

8.6 Comparison to other algorithms and techniques

Here the alternative techniques discussed in Chapter 7 are used to compare with the naïve Bayes models.

Variable	Coefficient	Std.Error	Variable	Coefficient	Std.Error
Constant	-3.846	0.844	PMH Diabetes	0.898	0.599
GKI solution	-0.785	0.302	PACS Stroke	-1.161	0.421
Warfarin	1.683	0.655	LACS Stroke	-1.771	0.768
Digoxin	0.721	0.435	POCS Stroke	2.971	1.513
Lipid Lowerer	-1.225	0.761	Glucose	0.205	0.072
Hypoglycaemic	-1.904	0.783	Creatine	0.015	0.005
PMH CVA	-1.004	0.487	ESS motor	-0.067	0.020

Table 8.2: Summary of 1-7 day logistic regression model variables and coefficients.

Variable	Coefficient	Std.Error	Variable	Coefficient	Std.Error
Constant	-11.922	5.955	Diastolic BP	0.015	0.008
Saline solution	-0.555	0.228	Temperature	0.233	0.158
Age	0.054	0.014	White Cells	-1.771	0.025
PMH CCF	0.917	0.344	ESS non-motor	2.971	0.010
Systolic BP	-0.015	0.005			

Table 8.3: Summary of 8-93 day logistic regression model variables and coefficients.

Logistic regression, as discussed in Section 7.7.1, is the first method used for comparison. The logistic regression model was created in a stepwise fashion and was implemented using the statistical software R. For classification of 1-7 day mortality the model can be seen in Table 8.2. The model contains 11 different variables, the majority of which were also selected by the naïve Bayes model. The variables digoxin, warfarin, lipid lowerers are highlighted from the admission medication alongside previous history of stroke and stroke type itself. The analysis was also carried out for classification of death in 8-93 days which produced the model in Table 8.3.

A decision tree method using the C4.5 algorithm, as described in Chapter 7, implemented through the WEKA data mining software and its J48 decision tree C4.5 tool was also tested for comparison. This was applied to create two models, one for each mortality period. A graphical representation of the best fitting pruned decision tree can be seen in Figures B.1 and B.2 of Appendix B.

From the regression view point rather than binary classification, multivariate linear re-

gression is used as another comparison. In testing the assumptions specified in Section 7.7.2, not all variables met the required conditions, meaning any model created would be very poor and invalid. This is not unusual for such a data set, and is a major limitation of linear regression. For reference, a stepwise linear regression model was attempted which produced a maximum R^2 value of 39.69%, where R^2 is the percentage of variation accounted for by the model. This confirms that linear regression is not suitable for this data.

All of the models for naïve Bayes, logistic regression and C4.5 decision trees were tested and their respective sensitivity and specificity have been calculated and can be seen in Table 8.4. Testing by way of sensitivity and specificity gives an intrinsic measure which is easily calculated, yet very useful for binary classification models. Comparing the three methodologies across both mortality periods shows that the naïve Bayes model has the highest sensitivity for both models, thus correctly classifying more of the patients who died in the given period. The logistic regression model was optimised in a similar manner to the naïve Bayes by selecting a cut-off probability according to the ROC curve. The cut off probabilities were selected as $p_{1-7} > 0.05$ and $p_{8-93} > 0.1$. The model performs well for sensitivity, although it comes at a much higher cost of specificity. This may be a result of having a small data set and thus a small number of cases per variable (Peduzzi et al 1996), implying logistic regression cannot reach its asymptotic error rate which could be lower than the naïve Bayes for larger data sets. The pruned decision tree also performs poorly, particularly for 8-93 days. The unpruned tree was also tested which produced a sensitivity of 29.6% for the 8-93 day period. While still poor it is significantly better than the pruned tree. This may be a result of the low number of events within the data set and thus pruning certain leaves, while it is necessary according to significance, can cause a greater loss in sensitivity. To test this further the feature selection method presented in this thesis was applied by taking only the top 14 variables according to ε, which were

Model	1 - 7 Day		8 - 93 Day	
	Sensitivity	Specificity	Sensitivity	Specificity
Naïve Bayes	92.6%	69.4%	83.3%	70.8%
Logistic regression	88.9%	59.9%	81.5%	51.0%
Decision tree	33.3%	94.6%	9.3%	95.0%

Table 8.4: Comparison table for sensitivity and specificity of all models

used to build an unpruned tree. The results for sensitivity were reduced to 18.5% for the 1-7 days model but were greatly increased to 37% for 8-93 days. Thus, for this particular data set decision tree building using C4.5 is volatile which again, can be attributed to the low number of cases in the data causing problems for a C4.5 decision tree.

Overall, while the different algorithms are qualitatively similar through feature selection, in terms of classification the naïve Bayes produces the best results for this data.

8.7 Discussion

The overall aim of this chapter has been to produce predictive models for post-stroke mortality. This has been successful, while providing other interesting results. Initially, a new classification of the time scales for which post-stroke mortality is classified has been proposed as clear patterns are seen in the mortality curve. This is confirmed in the analysis of risk factors for different time scales. A specific investigation into age concludes that age is not a risk factor over very short time scales, but is highly significant in the longer term. The significant risk factors were used to build new predictive models for post-stroke mortality for two different time scales, very short term for 1-7 days, and short/intermediate term for 8-93 days. Upon testing and comparison to other techniques, the naïve Bayes approach used in this work proved to be the most successful for this application to medical data.

As well as the specific results related to stroke, this work highlights the benefits of using a data mining inspired approach for a medical data set. Typically a medical data set is created to study one particular case or intervention however the techniques used in this work show that interesting results and conclusions can be made when a widespread investigation into a specific data set is made.

Chapter 9

Thesis conclusions, critical evaluations and further work

"If we knew what it was we were doing it wouldn't be called research, would it?"

— Attributed to Albert Einstein (1879 - 1955)

9.1 Introduction

This thesis has adopted two different mathematical modelling techniques in order to study diseases, namely diabetes and stroke, on very different scales. The first approach taken was on a microscopic level to model the metabolic subsystem of glucose-insulin regulation by using delay differential equations. The second approach was on a more macroscopic scale with a naïve Bayes data mining approach to the study of risk factors for post-stroke mortality to build new predictive models. A range of mathematical modelling techniques is necessary to gain a full understanding of living systems due to their complexity.

This chapter provides a summary and evaluation of the findings presented in this thesis, while giving a discussion of possible avenues for future work.

9.2 Thesis conclusions

The work in this thesis initially focused on the delay differential equations approach. Chapters 2 and 3 provided a review of the current literature for models of glucose-insulin regulation and an introduction to some of the key techniques used to obtain the results in the related results chapters.

Chapter 4 of this thesis presented the key contribution to the model of glucose-insulin regulation by creating a new Hill form of the functions describing the regulatory mechanisms of the system. The creation of the new Hill form to replace the exponential form allows for a significant biological background, given the base of the Hill function in modelling biochemical reactions in the blood (Hill 1910), while the need for auxiliary parameters is avoided. The new form of the model was shown to provide an accurate representation

of the desired ultradian oscillations in the glucose-insulin regulation system. The ranges obtained for the Hill coefficients were optimised in relation to position of the fixed point and the size of the global stability region (see Chapter 6). These calculations were all based on the model with two time delays as presented by Li et al (2006).

Following the investigation of the newly proposed application of Hill functions, Chapters 5 and 6 presented the stability analysis of the system in order to understand the behaviour of the two-delay model. The stability analysis of the glucose-insulin system in previously published work focused either on the global stability of one-delay systems (Bennett and Gourley 2004b; Giang et al 2008), or numerical bifurcation analysis of the two-delay model (Li et al 2006; Li and Kuang 2007), whereas in this thesis, the two-delay case was studied analytically to describe the onset of periodic behaviour and devise ranges for physiological parameters for the first time.

Chapter 5 was concerned with the local stability. In analysing the linearised system by the direct method using the characteristic equation, conditions for the apparition of oscillations were obtained in relation to the time delays. It was possible to obtain an expression for the critical value of delay τ_2, which was only implicitly dependent on the delay τ_1. This allowed for the graphical representation to be plotted revealing that the boundary for oscillatory behaviour is a curve in the delay domain. This new result extends the existing analysis where a linear relation is suggested but not investigated (Li and Kuang 2007). In addition to this, a Lyapunov functional was created for the linearised two-delay case to provide a set of sufficient conditions for stability.

In Chapter 6 the stability analysis was centred on the global case investigating the nonlinear system. This work, extending the approach of Bennett and Gourley (2004b), created a Lyapunov functional to provide new sufficient conditions for global stability in the two-delay case. The analysis of the nonlinear system allows for conditions of stability to be

related to the new Hill parameters of the model. A numerical analysis to maximise the stability region, based on maximising the threshold for the two delays, provided further conditions on the ranges of the Hill parameters. A set of optimal ranges for these parameters were given.

Chapters 4 to 6 present an introduction of a new form of the regulatory mechanism functions in the nonlinear model for glucose-insulin regulation with two time delays. An analytical and numerical analysis of both local stability and global stability is made for the first time for the two-delay system.

Chapters 7 and 8 investigated the naïve Bayes data mining approach applied to a medical data set of stroke patients. In Chapter 7 an introduction to the problem and a background in the methods used for the analysis in the following chapter was presented.

The results in Chapter 8 stem from the main goal of creating predictive models for post-stroke mortality, where the investigation uncovered further interesting results. The first part of the analysis was to determine the most appropriate time scales for assessing risk factors of post-stroke mortality. The analysis of the cumulative mortality curve over time showed that the distribution of deaths is not uniform and the choice was made to classify different sets of mortality periods. For investigations into specific variables, as presented with age, four periods were chosen to represent quartiles of the mortality curve. The second set of time periods were chosen as very short term, 1-7 day mortality, and short/intermediate term mortality for 8-93 days. Typically, previous studies considered one month as the point of investigation. However, it was expected that the risk factors would differ over shorter time scales. This was initially tested with a concentrated investigation into age over the four mortality periods, results here found that age was not a factor in 1-5 day mortality, but was a very significant risk factor in the longer term 34-93 days. This was confirmed by a joint analysis with the most severe stroke type.

In the second half of Chapter 8 predictive models were built. Based on the very short and short/intermediate mortality periods, all variables were analysed for their significance based on a parameter, ε and were then assigned a score stemming from the naïve Bayes approximation. A patients total score was then used for analysis in terms of prediction. It was seen that, particularly in the 1-7 day case, the model built was highly successful in determining post-stroke mortality with good results seen in terms of score decile testing and a sensitivity and specificity analysis. The new models built in this thesis were tested against some other common techniques including logistic regression and decision trees, where the naïve Bayes approach was shown to be the most appropriate for this data.

9.3 Evaluation of modelling approaches

This thesis highlights the ability and necessity to study living systems, in this case disease, on different scales via mathematical modelling. The two approaches presented in this thesis represent two scales of mathematical models of health with the aim of improving the knowledge and understanding of a disease. The complexity of diseases, as highlighted in Chapter 1 for diabetes, in terms of cause, effect, treatment and implications is very wide ranging. Therefore, to understand a disease it is necessary to take many multi-scale approaches to build an overall picture.

The first approach in this thesis, the application of delay differential equations, shows the ability to investigate a subsystem on a microscopic level. The study of the glucose-insulin regulation system gives the opportunity to view the core causes for diabetes with regard to the key physiological relations involving the liver and pancreas. For a system with a small number of inputs and outputs, this type of modelling allows for the intricate behaviour of systems to be seen and can be investigated under different conditions for different levels of the disease. This approach is extremely useful for such a system and is a classical

mathematical modelling technique.

Mathematical modelling via the second approach in this thesis, naïve Bayes based data mining, shows the ability to discover relations and new results for a wide range of variables where the causality is not known. The work in this thesis has shown that a more data mining inspired approach can provide new and interesting insights when analysing a clinical data set, which would otherwise be used for one specific purpose. This type of mathematical modelling is of a much more exploratory, data driven nature in order to discover unknown relations between variables.

It is clear that these two forms of mathematical modelling of health or disease can provide new and useful results extending the existing knowledge and understanding. It is also clear that an attempt to swap these two techniques with their two applications would not be possible. Therefore, in order to continue to build an increased understanding of diseases, to cover all aspects of their complexity, it is necessary to use a plethora of different techniques applied to a wide range of aspects.

Multiscale modelling is a relatively new technique which can be used to link models on different scales. Although initially created for solving physical and chemical problems, there are examples of mathematical and biological applications, including tumour growth (Macklin et al 2009) and in relation to diabetes, insulin secretion (Pedersen et al 2011). This type of modelling is a possible area of investigation for two appropriately related models.

9.4 Thesis appraisal and further work

There are many possible avenues for future work arising from this thesis through development of the ideas presented. Here, some of these options which would benefit from

future research will be discussed.

9.4.1 DDE: Inclusion of type 1 and type 2 diabetic parameters

The delay differential equation model presented in this thesis represents a functioning glucose-insulin regulatory system which has been successfully analysed in the two-delay case. One avenue for further study is to apply a parameter to represent the effect of type one diabetes and a second parameter to represent type two diabetes. The parameter for type one diabetes, α, would represent a deficiency in insulin production and would thus be included with the function f_1. The parameter for type two diabetes, β, would be a measure of the lack of glucose utilisation and would be included as a coefficient with the functions for glucose utilisation, $f_3 f_4$. The model would then take the following form,

$$\begin{aligned} \dot{G}(t) &= G_{in} - f_2(G(t)) - \beta f_3(G(t)) f_4(I(t)) + f_5(I(t-\tau_2)), \\ \dot{I}(t) &= \alpha f_1(G(t-\tau_1)) - d_i I(t). \end{aligned} \tag{9.1}$$

This model would then allow for investigation of theoretical patients with different levels of severity for both types of diabetes. This model can then be used to assess the behaviour of the system under these circumstances. Work of a similar nature can be seen in the literature for different models (Chen and Tsai 2010; Huang et al 2012; Kissler et al 2014), but an analysis as presented in this thesis has not been made. Initially, these new parameters would be set as constants to investigate general effect on behaviour.

9.4.2 DDE: Inclusion of insulin infusion for treatment

Another option for further study in relation to the model for glucose-insulin regulation is to include a parameter for insulin infusion, I_{in}, into the insulin equation. The inclusion

of insulin infusion can be seen in the artificial pancreas paper where the insulin level is increased by a set amount at regular intervals (Huang et al 2012), where in Kissler et al (2014) an I_{in} parameter is included alongside exercise. The inclusion of a direct parameter would give,

$$\begin{aligned} \dot{G}(t) &= G_{in} - f_2(G(t)) - \beta f_3(G(t)) f_4(I(t)) + f_5(I(t-\tau_2)), \\ \dot{I}(t) &= I_{in} + \alpha f_1(G(t-\tau_1)) - d_i I(t). \end{aligned} \tag{9.2}$$

The inclusion of insulin infusion allows for the analysis of diabetic treatments combating a lack of insulin within the system and hence is related to the addition of the parameters α and β, for type 1 and type 2 diabetes. An investigation for the two-delay model can be made for the insulin level required to return to normal dependent on severity of diabetes. The level of insulin infusion can be adjusted as a constant, which would represent the behaviour of an insulin pump, or alternatively it can be administered at intervals as insulin therapy via injection. Further developments include designing the insulin infusion as a function of both glucose and insulin, $I_{in}(G, I)$ to develop a relation with the aim of real time automatic treatment.

9.4.3 DDE: Alteration of glucose infusion for different scenarios

The study of the oscillatory behaviour in the glucose-insulin regulation system is undertaken in a number of different scenarios. As well as during constant glucose infusion as presented in this thesis, the cases of three meal infusion (Polonsky et al 1988), and oral glucose infusion (Kraegen et al 1972), can also be studied. Therefore, an area for further study for the model in this thesis is the alteration to glucose infusion to represent these scenarios. An example of an alteration to glucose infusion can be seen in Chen and Tsai (2010), where G_{in} represents overall effectiveness of food uptake.

This analysis can be extended to give new models for these cases in relation to the two time delay behaviour of the system. As well as individually, this analysis could be made in conjunction with the ideas suggested above to give a more complete idea of the behaviour of the system under typical daily conditions for a diabetic patient.

9.4.4 Data Mining: Extension to other diseases

The data mining work in this thesis showed how useful the application of such techniques to medical based data sets can be. Due to the nature of the data used in this study, which were subject to certain conditions, it would be useful to obtain a more general data set of post-stroke mortality. The belief is that the results in this thesis are applicable to the wider stroke population but a study of more general data would confirm the results further.

The analysis was carried out with post-stroke mortality as the dependent variable, however, any variable contained in the data set could be chosen as the dependent. The nature of data mining based techniques is that the causality is unknown and so the point of interest determines the dependent variable. In relation to this thesis it would be appropriate to investigate the relations to diabetes for a similar set of variables as used in the post-stroke mortality study. Alternatively, for any other data set which is available, the same transparent naïve Bayes data mining can be applied to investigate a variable of interest.

9.5 In conclusion

This thesis has presented original work containing new results for two different approaches to mathematical modelling of health. Both of these approaches have highlighted the necessity in developing models for all aspects of health due to the complexity of living systems. While individually, each approach has contributed to the knowledge in their re-

spective areas, developing a new model and new analysis of the two-delay glucose-insulin regulation system and creating new models for assessing stroke mortality on different time scales. The work in this thesis has provided a platform for further study in both areas.

Appendix A :

Variable	Group	P(Dead 1-7|X)	P(Dead 8-93|X)	P(Survived|X)	ε Dead 1-7	ε Dead 8-93	ε Survived	Relative change in odds 1-7 v 8-93
Drugs on Admission	Aspirin 1	0.10	0.21	0.69	- 0.38	0.74	- 0.38	- 0.15
1 = Yes	Aspirin 2	0.11	0.18	0.71	0.33	- 0.59	0.28	0.12
2 = No	Anti-Platelet 1	0.07	0.20	0.73	- 0.75	0.14	0.39	- 0.42
	Anti-Platelet 2	0.10	0.19	0.71	- 0.46	- 0.09	0.39	- 0.05
	Warfarin 1	0.24	0.20	0.56	2.09	0.13	- 1.54	1.08
	Warfarin 2	0.10	0.19	0.71	- 0.43	- 0.03	0.32	- 0.05
	Digoxin 1	0.22	0.29	0.49	2.59	1.91	- 3.40	0.30
	Digoxin 2	0.10	0.18	0.72	- 0.81	- 0.60	1.06	- 0.05
	Beta-Blockers 1	0.11	0.19	0.70	0.18	- 0.02	- 0.11	0.04
	Beta-Blockers 2	0.11	0.19	0.70	- 0.10	0.03	0.04	- 0.02
	Calcium Antagon	0.11	0.20	0.69	0.16	0.27	- 0.34	- 0.01
	Calcium Antagon	0.11	0.19	0.70	- 0.06	- 0.10	0.13	0.00
	Anti-arrythmic 1	0.11	0.21	0.68	- 0.06	0.23	- 0.16	- 0.13
	Anti-arrythmic 2	0.11	0.19	0.70	0.02	- 0.02	0.00	0.00
	ACE inhibitor 1	0.12	0.21	0.67	0.44	0.47	- 0.70	0.03
	ACE inhibitor 2	0.11	0.19	0.71	- 0.18	- 0.19	0.29	- 0.01
	Nitrate 1	0.11	0.21	0.68	- 0.07	0.52	- 0.40	- 0.13
	Nitrate 2	0.11	0.19	0.70	0.06	- 0.26	0.18	0.03
	Diuretic 1	0.12	0.26	0.62	0.60	2.72	- 2.74	- 0.19
	Diuretic 2	0.10	0.15	0.75	- 0.41	- 1.94	1.95	0.18
	Lipid lowerer 1	0.03	0.18	0.79	- 1.92	- 0.19	1.47	- 0.68
	Lipid lowerer 2	0.12	0.19	0.69	0.63	0.06	- 0.48	0.07
	Hypoglyceamic 1	0.05	0.15	0.79	- 1.49	- 0.85	1.75	- 0.37
	Hypoglyceamic 2	0.12	0.19	0.69	0.54	0.31	- 0.64	0.04
	Other Drugs 1	0.08	0.19	0.72	- 1.56	0.12	0.96	- 0.25
	Other drugs 2	0.14	0.18	0.68	1.24	- 0.16	- 0.71	0.30
ESS stroke scales	1: 0 - <14	0.23	0.40	0.37	3.97	5.57	- 7.47	- 0.01
ESS non-motor score	2: 14 - <22	0.19	0.26	0.55	2.71	1.82	- 3.41	0.29
	3: 22 - <28	0.12	0.16	0.72	0.20	- 0.68	0.45	0.22
	4: 28 - <38	0.09	0.16	0.75	- 0.74	- 0.68	1.09	- 0.08
	5: 38 - <42	0.01	0.13	0.87	- 3.26	- 1.68	3.66	- 0.87
	6: 42 +	0.02	0.02	0.96	- 2.90	- 4.38	5.73	0.74
ESS motor score	1: 0 - <3	0.17	0.35	0.48	2.08	4.07	- 4.90	- 0.13
	2: 3 - <5	0.16	0.22	0.62	1.77	0.82	- 1.91	0.28
	3: 5 - <10	0.21	0.21	0.58	3.34	0.57	- 2.76	0.74
	4: 10 - <20	0.06	0.17	0.77	- 1.69	- 0.43	1.52	- 0.42
	5: 20 - <31	0.04	0.13	0.84	- 2.32	- 1.68	3.02	- 0.47
	6: 31 +	0.01	0.06	0.93	- 3.22	- 3.37	5.08	- 0.71
ESS Total score	1: 0 - <18	0.23	0.39	0.38	3.97	5.32	- 7.26	0.02
	2: 18 - <30	0.22	0.26	0.52	3.65	1.82	- 4.05	0.48
	3: 30 - <42	0.13	0.20	0.66	0.83	0.32	- 0.84	0.16
	4: 42 - <53	0.04	0.13	0.83	- 2.32	- 1.43	2.81	- 0.50
	5: 53 - <69	0.03	0.08	0.89	- 2.63	- 2.93	4.30	- 0.35
	6: 69 +	-	0.07	0.93	- 3.54	- 3.12	5.08	- 1.00

Table A.1: Probabilities and significance levels for drugs on admission variables and ESS scales. Shading in the ε columns correspond to variables that were included in the corresponding predictive model. In the relative change column highlighted drivers correspond to variables that discriminate between 1-7 and 8-93 day mortality.

Variable	Group	P(Dead 1-7|X)	P(Dead 8-93|X)	P(Survived|X)	ε Dead 1-7	ε Dead 8-93	ε Survived	Relative change in odds 1-7 v 8-93
Past Medical History								
1 = Yes	CVA 1 Stroke	0.06	0.21	0.73	- 1.61	0.61	0.57	- 0.49
2 = No	CVA 2	0.12	0.18	0.70	0.67	- 0.38	- 0.14	0.13
Diabetes Mellitus	DM 1	0.10	0.14	0.77	- 0.42	- 1.27	1.38	0.20
	DM 2	0.11	0.20	0.69	0.18	0.54	- 0.58	- 0.02
Hypertension	BP 1	0.10	0.19	0.71	- 0.75	0.13	0.40	- 0.13
	BP 2	0.12	0.19	0.69	0.78	- 0.13	- 0.41	0.15
Atrial Fibrillation	MAF 1	0.15	0.28	0.57	1.32	2.27	- 2.84	- 0.07
	MAF 2	0.10	0.17	0.72	- 0.55	- 0.95	1.18	0.02
Myocardial Infarction	MI 1	0.14	0.25	0.60	1.03	1.47	- 1.96	- 0.01
	MI 2	0.10	0.18	0.72	- 0.40	- 0.58	0.77	0.00
Angina	ANG 1	0.10	0.23	0.67	- 0.42	1.07	- 0.63	- 0.26
	ANG 2	0.11	0.18	0.71	0.21	- 0.53	0.31	0.08
Heart Failure	CCF 1	0.16	0.37	0.47	1.09	3.33	- 3.59	- 0.27
	CCF 2	0.11	0.17	0.72	- 0.33	- 1.00	1.07	0.05
Valvular Heart Disea	VHD 1	0.15	0.30	0.56	0.65	1.41	- 1.65	- 0.13
	VHD 2	0.11	0.18	0.71	- 0.14	- 0.30	0.35	0.01
Carotid Stenosis	CSTEN 1	-	0.15	0.85	- 1.26	- 0.33	1.14	- 1.00
	CSTEN 2	0.11	0.19	0.70	0.24	- 0.18	- 0.01	0.04
Hypercholesterolaem	CHOL 1	0.04	0.14	0.82	- 1.81	- 1.05	2.13	- 0.48
	CHOL 2	0.12	0.20	0.68	0.76	0.42	- 0.87	0.05
Smoking	SMOK 1	0.11	0.17	0.73	- 0.14	- 0.63	0.63	0.10
	SMOK 2	0.09	0.19	0.71	- 1.03	0.23	0.51	- 0.17

Table A.2: Probabilities and significance levels for past medical history variables. Shading in the ε columns correspond to variables that were included in the corresponding predictive model. In the relative change column highlighted drivers correspond to variables that discriminate between 1-7 and 8-93 day mortality.

Variable	Group	P(Dead 1-7|X)	P(Dead 8-93|X)	P(Survived|X)	ε Dead 1-7	ε Dead 8-93	ε Survived	Relative change in odds 1-7 v 8-93
Stroke Scales								
Pre-stroke Barthel	Barthel 1: 0-19	0.15	0.24	0.61	1.71	1.63	- 2.56	0.10
	Barthel 2: 20	0.08	0.17	0.75	- 1.69	- 1.12	2.11	- 0.15
Modified Rankin	MRS 0	0.07	0.15	0.77	- 1.93	- 1.52	2.62	- 0.17
Score (MRS)	MRS 1	0.11	0.18	0.72	- 0.11	- 0.34	0.37	0.04
	MRS 2	0.15	0.22	0.63	1.05	0.71	- 1.32	0.16
	MRS 3	0.16	0.30	0.54	1.28	2.35	- 2.89	- 0.09
	MRS 4	-	1.00	-	- 0.35	2.07	- 1.53	- 1.00
Admission Variables								
Pulse on admission	40<= 1 <65	0.14	0.17	0.68	1.14	- 0.43	- 0.41	0.45
	65<= 2 <75	0.10	0.18	0.72	- 0.43	- 0.18	0.45	- 0.09
	75<= 3 <80	0.09	0.14	0.77	- 0.74	- 1.18	1.52	0.04
	80<= 4 <89	0.11	0.18	0.71	- 0.12	- 0.18	0.24	0.00
	89<= 5 <98	0.07	0.20	0.73	- 1.37	0.32	0.66	- 0.42
	6: 98 +	0.15	0.24	0.61	1.23	1.36	- 2.00	0.06
Systolic BP	94<= 1 <136	0.13	0.23	0.63	0.83	1.07	- 1.48	0.01
on admission	136<= 2 <150	0.13	0.22	0.64	0.83	0.82	- 1.26	0.06
	150<= 3 <163	0.09	0.23	0.68	- 0.74	1.07	- 0.41	- 0.35
	163<= 4 <177	0.11	0.15	0.74	- 0.12	- 0.93	0.88	0.19
	177<= 5 <194	0.09	0.17	0.74	- 0.74	- 0.43	0.88	- 0.13
	6: 194 +	0.09	0.12	0.78	- 0.52	- 1.66	1.78	0.30
Diastolic BP	48<= 1 <71	0.14	0.19	0.66	1.14	0.07	- 0.84	0.30
on admission	71<= 2 <80	0.11	0.19	0.70	- 0.12	0.07	0.02	- 0.05
	80<= 3 <87	0.09	0.20	0.71	- 0.74	0.32	0.24	- 0.26
	87<= 4 <95	0.12	0.16	0.72	0.20	- 0.68	0.45	0.22
	95<= 5 <106	0.08	0.21	0.71	- 1.06	0.57	0.24	- 0.37
	6: 106 +	0.11	0.18	0.71	0.13	- 0.36	0.22	0.12
Tempreture	33.7<= 1 <36.0	0.14	0.15	0.70	1.14	- 0.93	0.02	0.63
on admission	36.0<= 2 <36.3	0.09	0.19	0.72	- 0.74	0.07	0.45	- 0.22
	36.3<= 3 <36.5	0.09	0.19	0.72	- 0.74	0.07	0.45	- 0.22
	36.5<= 4 <36.8	0.13	0.15	0.71	0.83	- 0.93	0.24	0.52
	36.8<= 5 <37.2	0.12	0.21	0.67	0.20	0.57	- 0.62	- 0.05
	6: 37.2 +	0.09	0.29	0.63	- 0.63	2.05	- 1.32	- 0.48
ECG	Sinus Rhythm	0.09	0.15	0.76	- 1.48	- 2.04	2.76	- 0.00
on admission	Atrial fibrillatior	0.17	0.28	0.56	2.21	2.70	- 3.81	0.04
	Other Rhythm	0.12	0.24	0.64	0.22	0.77	- 0.81	- 0.13

Table A.3: Probabilities and significance levels for stroke scales and first section of admission variables. Shading in the ε columns correspond to variables that were included in the corresponding predictive model. In the relative change column highlighted drivers correspond to variables that discriminate between 1-7 and 8-93 day mortality.

Variable	Group	P(Dead 1-7|X)	P(Dead 8-93|X)	P(Survived|X)	ε Dead 1-7	ε Dead 8-93	ε Survived	Relative change in odds 1-7 v 8-93
Admission Variables								
Plasma sodium	126<= 1 <136	0.12	0.23	0.65	0.20	1.07	- 1.05	- 0.13
on admission	136<= 2 <138	0.09	0.22	0.69	- 0.74	0.82	- 0.19	- 0.32
	138<= 3 <139	0.13	0.13	0.73	0.83	- 1.43	0.66	0.74
	139<= 4 <140	0.09	0.13	0.78	- 0.74	- 1.43	1.73	0.12
	140<= 5 <142	0.10	0.20	0.70	- 0.43	0.32	0.02	- 0.17
	6: 142 +	0.13	0.22	0.65	0.66	0.77	- 1.11	0.03
Plasma potassium	2.4<= 1 <3.5	0.14	0.23	0.63	1.14	1.07	- 1.69	0.08
on admission	3.5<= 2 <3.8	0.08	0.17	0.75	- 1.06	- 0.43	1.09	- 0.23
	3.8<= 3 <3.94	0.06	0.20	0.74	- 1.69	0.32	0.88	- 0.50
	3.94<= 4 <4.12	0.16	0.19	0.64	1.77	0.07	- 1.26	0.48
	4.12<= 5 <4.5	0.09	0.14	0.77	- 0.74	- 1.18	1.52	0.04
	6: 4.5 +	0.12	0.20	0.68	0.41	0.18	- 0.43	0.08
Plasma Urea	1.3<= 1 <4.4	0.10	0.19	0.71	- 0.43	0.07	0.24	- 0.13
on admission	4.4<= 2 <5.4	0.04	0.17	0.79	- 2.32	- 0.43	1.95	- 0.61
	5.4<= 3 <6.3	0.11	0.15	0.74	- 0.12	- 0.93	0.88	0.19
	6.3<= 4 <7.3	0.14	0.18	0.67	1.14	- 0.18	- 0.62	0.37
	7.3<= 5 <9.0	0.13	0.18	0.69	0.51	- 0.18	- 0.19	0.19
	6: 9.0 +	0.14	0.26	0.60	0.94	1.74	- 2.13	- 0.07
Plasma creatine	37<= 1 <79	0.08	0.15	0.77	- 1.06	- 0.93	1.52	- 0.13
on admission	79<= 2 <89	0.06	0.23	0.71	- 1.69	1.07	0.24	- 0.57
	89<= 3 <99	0.13	0.16	0.71	0.51	- 0.68	0.24	0.33
	99<= 4 <107	0.13	0.20	0.67	0.51	0.32	- 0.62	0.07
	107<= 5 <123	0.11	0.17	0.72	- 0.12	- 0.43	0.45	0.06
	6: 123 +	0.16	0.22	0.62	1.58	0.72	- 1.69	0.26
Haemoglobin	4.9<= 1 <12.5	0.13	0.23	0.64	0.51	1.07	- 1.26	- 0.06
on admission	12.5<= 2 <13.2	0.07	0.17	0.76	- 1.37	- 0.43	1.31	- 0.33
	13.2<= 3 <13.9	0.08	0.16	0.76	- 1.06	- 0.68	1.31	- 0.18
	13.9<= 4 <14.6	0.09	0.21	0.70	- 0.74	0.57	0.02	- 0.29
	14.6<= 5 <15.5	0.15	0.20	0.64	1.46	0.32	- 1.26	0.32
	6: 15.5 +	0.13	0.16	0.70	0.78	- 0.62	0.00	0.41
White Cell count	3.0<= 1 <7.0	0.10	0.17	0.73	- 0.43	- 0.43	0.66	- 0.04
on admission	7.0<= 2 <8.3	0.06	0.13	0.82	- 1.69	- 1.68	2.59	- 0.20
	8.3<= 3 <9.5	0.12	0.16	0.72	0.20	- 0.68	0.45	0.22
	9.5<= 4 <11.2	0.13	0.23	0.64	0.51	1.07	- 1.26	- 0.06
	11.2<= 5 <13.5	0.13	0.17	0.69	0.83	- 0.43	- 0.19	0.35
	6: 13.5 +	0.11	0.28	0.61	0.02	2.30	- 1.99	- 0.32
Platelet Count	53<= 1 <187	0.10	0.25	0.65	- 0.43	1.57	- 1.05	- 0.33
on admission	187<= 2 <218	0.12	0.16	0.72	0.20	- 0.68	0.45	0.22
	218<= 3 <248	0.08	0.17	0.75	- 1.06	- 0.43	1.09	- 0.23
	248<= 4 <279	0.09	0.15	0.76	- 0.74	- 0.93	1.31	- 0.02
	249<= 5 <336	0.12	0.21	0.67	0.20	0.57	- 0.62	- 0.05
	6: 336 +	0.15	0.19	0.67	1.35	0.06	- 0.75	0.37
Glucose	2.3<= 1 <6.5	0.06	0.13	0.81	- 1.69	- 1.43	2.38	- 0.26
on admission	6.5<= 2 <7.0	0.05	0.16	0.79	- 2.00	- 0.68	1.95	- 0.49
	7.0<= 3 <7.6	0.11	0.22	0.67	- 0.12	0.82	- 0.62	- 0.17
	7.6<= 4 <8.4	0.13	0.22	0.64	0.83	0.82	- 1.26	0.06
	8.4<= 5 <10.0	0.16	0.19	0.64	1.77	0.07	- 1.26	0.48
	6: 10.0 +	0.14	0.21	0.65	0.94	0.47	- 1.04	0.16

Table A.4: Probabilities and significance levels for second section of admission variables. Shading in the ε columns correspond to variables that were included in the corresponding predictive model. In the relative change column highlighted drivers correspond to variables that discriminate between 1-7 and 8-93 day mortality.

Appendix B :

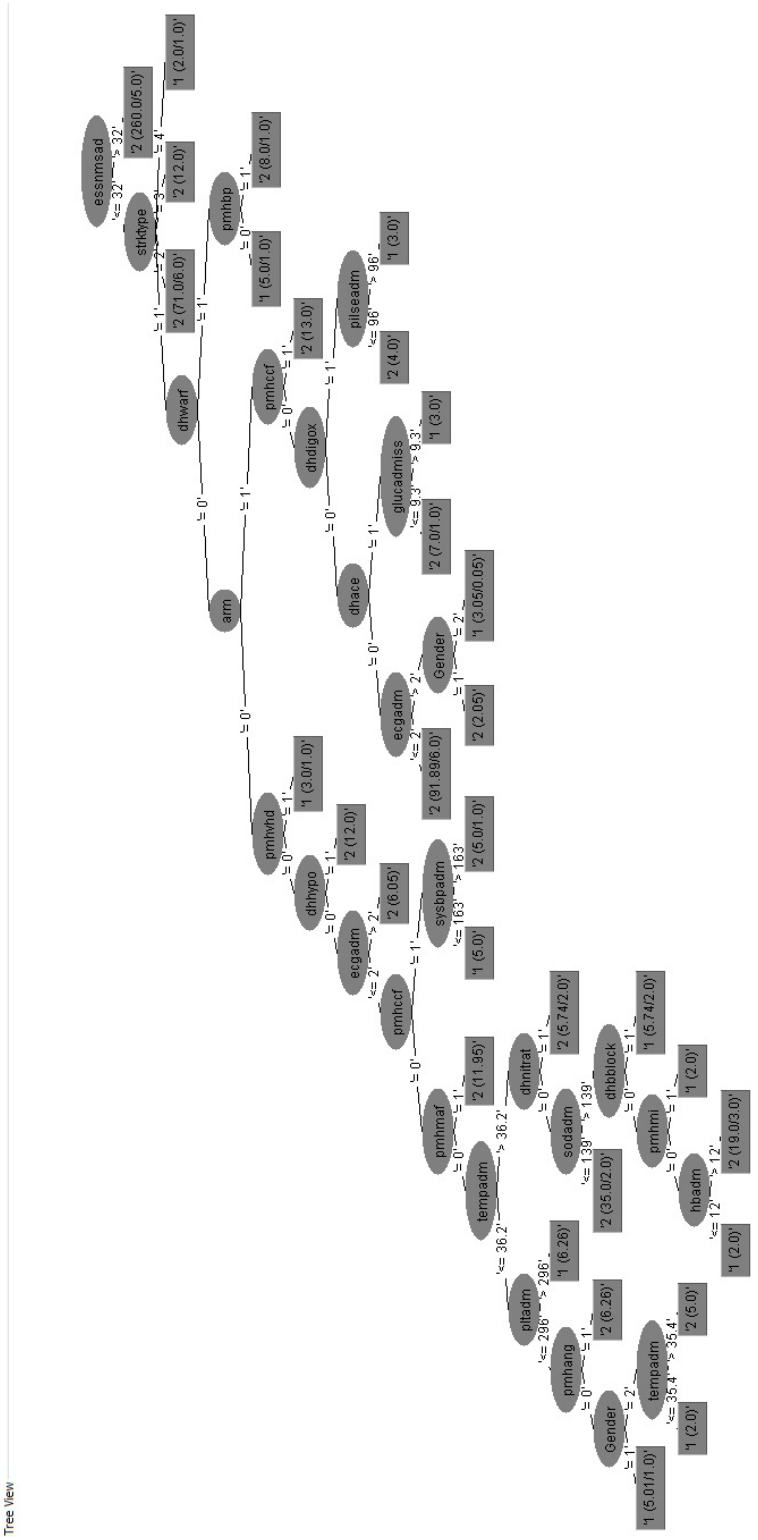

Figure B.1: The visualisation of the pruned decision tree built using the C4.5 algorithm for the classification of mortality in 1-7 days.

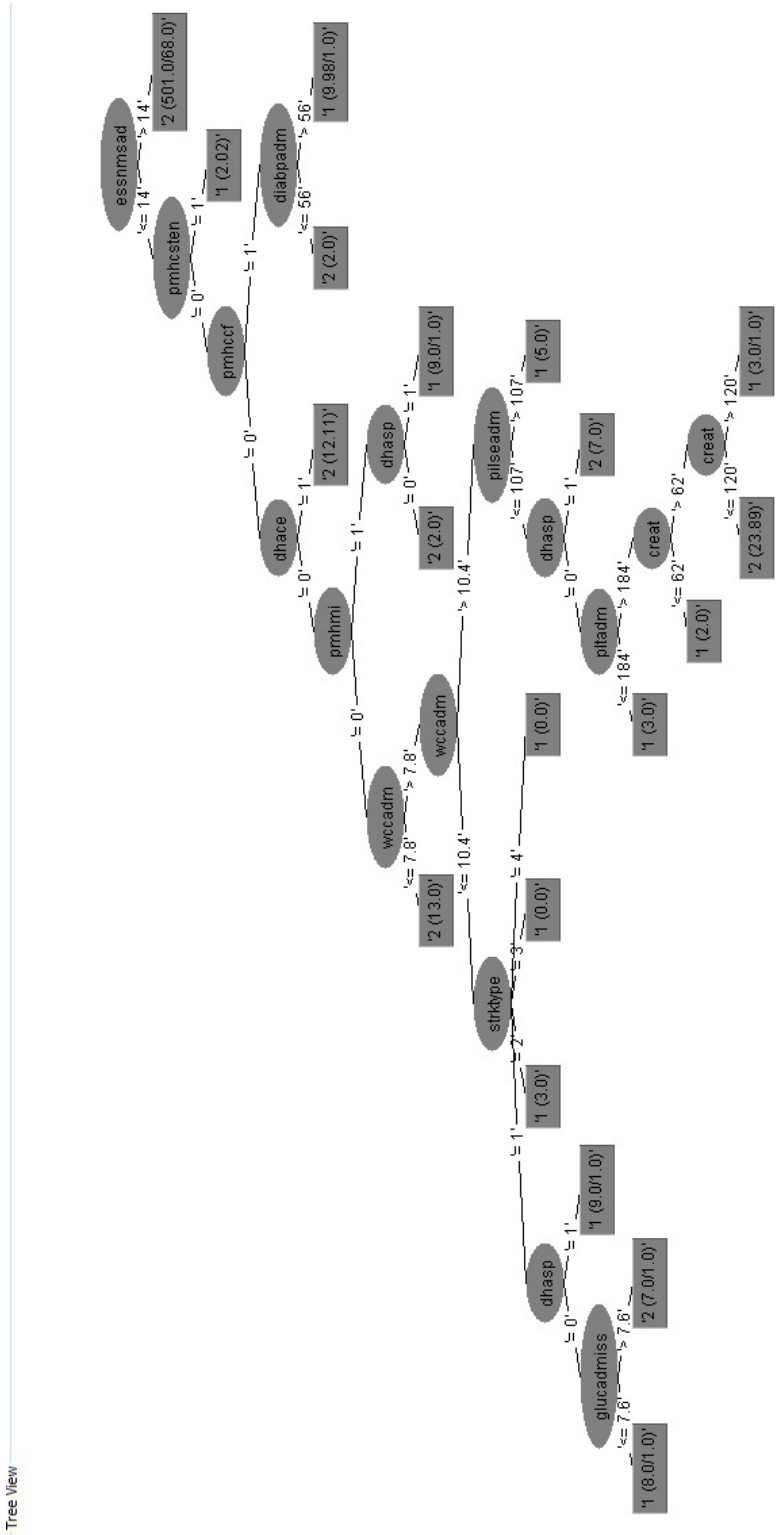

Figure B.2: The visualisation of the pruned decision tree built using the C4.5 algorithm for the classification of mortality in 8-93 days.

Bibliography

Abramowitz M, Stegun I (1972) Handbook of mathematical functions: with formulas, graphs, and mathematical tables. 55, Courier Dover Publications

Akiake H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Association AD, et al (2014) Diagnosis and classification of diabetes mellitus. Diabetes Care 37(Supplement 1):S81–S90

Bagust A, Hopkinson P, Maier W, Currie C (2001) An economic model of the long-term health care burden of type ii diabetes. Diabetologia 44(12):2140–2155

Bamford J, Dennis M, Sandercock P, Burn J, Warlow C (1990) The frequency, causes and timing of death within 30 days of a first stroke: the Oxfordshire community stroke project. J Neurol Neurosurg Psychiatry 53:824–829

Bamford J, Sandercock P, Dennis M, Warlow C, Burn J (1991) Classification and natural history of clinically identifiable subtypes of cerebral infarction. Lancet 337:1521–1526

Barrett-Connor E, Khaw KT (1988) Diabetes mellitus: an independent risk factor for stroke? Am J Epidemiol 128(1):116–123

Beckner W (1975) Inequalities in Fourier analysis. Ann Math pp 159–182

Bellazzia R, Zupan B (2008) Predictive data mining in clinical medicine: Current issues and guidelines. Int J Med Inform 77:81–97

Bennett DL, Gourley SA (2004a) Asymptotic properties of a delay differential equation

model for the interaction of glucose with plasma and interstitial insulin. Appl Math Comput 151(1):189–207

Bennett DL, Gourley SA (2004b) Global stability in a model of the glucose-insulin interaction with time delay. Eur J Appl Math 15(2):203–221

Bennett DL, Gourley SA (2004c) Periodic oscillations in a model of the glucose-insulin interaction with delay and periodic forcing. Dyn Syst 19(2):109–125

Betounes D (2010) Differential equations theory and applications, 2nd edn, Springer

Boyd S, Vandenberghe L (2009) Convex optimization, Cambridge university press

Bronnum-Hansen H, Davidsen M, Thorvaldsen P (2001) Long-term survival and causes of death after stroke. Stroke 32(1):2131–2136

Bunce C, Wormald R (2006) Leading causes of certification for blindness and partial sight in england & wales. BMC Public Health 6(1):58

Caruana R, Niculescu-Mizil A (2006) An empirical comparison of supervised learning algorithms. Proceedings of the 23rd International Conference on Machine Learning (ICML2006) pp 161–168

Castillo J, Alvarez-Sabin J, Martinez-Vila E, Montaner J, Sobrino T, Vivancos J (2009) Inflammation markers and prediction of post-stroke vascular disease recurrence: The mitico study. J Neurol 256:217–224

Chan J, Rimm E, Colditz G, Stampfer M, Willett W (1994) Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. Diabetes Care 17(9):961–969

Chen CL, Tsai HW (2010) Modeling the physiological glucose–insulin system on normal and diabetic subjects. Comput Meth Prog Bio 97(2):130–140

Cherrington AD, Sindelar D, Edgerton D, Steiner K, McGuinness OP (2002) Physiological consequences of phasic insulin release in the normal animal. Diabetes 51(suppl 1):S103–S108

Cherry J, Adler F (2000) How to make a biological switch. J Theor Biol 203:117–133

CMCD (2014) The centre for managing chronic diseases. URL http://cmcd.sph.umich.edu/

Collins T, Petersen N, Menke T, Souchek J, Foster W, Ashton C (2003) Short-term, intermediate-term, and long-term mortality in patients hospitalized for stroke. J Clin Epidemiol 56(1):81–87

Danaei G, Finucane M, Lu Y, Singh G, Cowan M, Paciorek C, Lin J, Farzadfar F, Khang Y, Stevens G, et al (2011) Global burden of metabolic risk factors of chronic diseases collaborating group (blood glucose) national, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2.7 million participants. Lancet 378(9785):31–40

De Gaetano A, Arino O (2000) Mathematical modelling of the intravenous glucose tolerance test. J Math Biol 40(2):136–168

Diabetes-UK (2013a) Diabetes in the UK 2012: Key statistics on diabetes. URL http://www.diabetes.org.uk/Documents/Reports/Diabetes-in-the-UK-2012.pdf

Diabetes-UK (2013b) Diabetes risk factors. URL http://www.diabetes.org.uk/Guide-to-diabetes/What-is-diabetes/Diabetes-risk-factors/

Diabetes-UK (2014) Diabetes in the UK 2013: Key statistics on diabetes. URL http://www.diabetes.org.uk/Documents/About%20Us/Statistics/Diabetes-key-stats-guidelines-April2014.pdf

Dixon L, Weiden P, Delahanty J, Goldberg R, Postrado L, Lucksted A, Lehman A (2000) Prevalence and correlates of diabetes in national schizophrenia samples. Schizophrenia Bull 26(4):903

Domingos P, Pazzani M (1997) On the optimality of the simple bayesian classifier under

zero-one loss. Mach Learn 29(2-3):103–130

Drozdov A, Khanina H (1995) A model for ultradian oscillations of insulin and glucose. Math Comput Model 22(2):23–38

Duckworth W, Bennett R, Hamel F (1998) Insulin degradation: Process and potential. Endocr Rev 19(5):608–624

Duncan T (1970) On the calculation of mutual information. SIAM J Appl Math 19(1):215–220

Dunham M (2003) Data Mining Introductory and Advanced Topics, 1st edn, New Jersey: Prentice Hall

Easton JF, Stephens CR, Angelova M (2014) Risk factors and prediction of very short term versus short/intermediate term post-stroke mortality: A data mining approach. Comput Biol Med 54:199–210

Engelborghs K, Luzyanina T, Samaey G (2000) DDE-BIFTOOL: a Matlab package for bifurcation analysis of delay differential equations. TW Reports, Department of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium (305)

Engelborghs K, Lemaire V, Bélair J, Roose D (2001a) Numerical bifurcation analysis of delay differential equations arising from physiological modeling. J Math Biol 42:361–385

Engelborghs K, Luzyanina T, Samaey G (2001b) DDE-BIFTOOL v. 2.00: a matlab package for bifurcation analysis of delay differential equations. TW Reports, Department of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium (330):61

Engelborghs K, Luzyanina T, Roose D (2002) Numerical bifurcation analysis of delay differential equations using dde-biftool. ACM T Math Software 28:1–21

Fayyad U, Irani K (1993) Multi-interval discretisation of continuousvalued attributes for classification learning. In: Proceedings of the Thirteenth International Join Conference on Artificial Intelligence, Morgan Kaufmann

Ferreira Jr S, Martins M, Vilela M (2002) Reaction-diffusion model for the growth of avascular tumor. Phys Rev E 65(2):021,907–1021,907–8

Fluri F, Morgenthaler N, Mueller B, Christ-Crain M, Katan M (2012) Copeptin, procalcitonin and routine inflammatory markers-predictors of infection after stroke. PLoS One 7(10)

Fulop T, Larbi A, Witkowski J, McElhaney J, Loeb M, Mitnitski A, Pawelec G (2010) Aging, frailty and age-related diseases. Biogerontology 11:547–563

Gallagher J (2014) More weight loss operations for diabetes. URL http://www.bbc.co.uk/news/health-28246641

Gallier J (2010) The Schur complement and symmetric positive semidefinite (and definite) matrices. December 44:1–12

Gesztelyi R, Zsuga J, Kemeny-Beke A, Varga B, Juhasz B, Tosaki A (2012) The Hill equation and the origin of quantitative pharmacology. Arch Hist Exact Sci 66(4):427–438

Giang D, Lenbury Y, De Gaetano A, Palumbo P (2008) Delay model of glucose-insulin systems: Global stability and oscillated solutions conditional on delays. J Math Anal Appl 343:996–1006

Gilbert G (1991) Positive definite matrices and Sylvester's criterion. Amer Math Monthly pp 44–46

Gilon P, Ravier M, Jonas JC, Henquin JC (2002) Control measures of the oscillations of insulin secretion in vitro and in vivo. Diabetes 51:S144–S151

Glendinning P (1999) Stability, instability and chaos: an introduction to the theory of nonlinear differential equations, 2nd edn, Cambridge University Press

Goutelle S, Maurin M, Rougier F, Barbaut X, Bourguignon L, Ducher M, Maire P (2008) The Hill equation: a review of its capabilities in pharmacological modelling. Fundam Clin Pharm 22:633–648

Gray C, Hildreth A, Sandercock P, OConnell J, Johnston D, Cartlidge N, Bamford J, James O, Alberti K (2007) Glucose-potassium-insulin infusions in the management of post-stroke hyperglycaemia: the UK glucose insulin in stroke trial (GIST-UK). Lancet Neurol 6(1):397–406

Grobman DM (1959) Homeomorphism of systems of differential equations. Dokl Akad Nauk SSSR 128(5):880–881

Grover R, Vriens M (2006) The Handbook of Marketing Research, 1st edn, London: Sage, chap 22

Hall M (1999) Correlation-based feature selection for machine learning. PhD thesis, The University of Waikato

Hall M, Frank, E Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: An update. SIGKDD explorations 11(1)

Han K, Kang H, Choi M, Kim J, Lee MS (2012) Mathematical model of the glucose-insulin regulatory system: From the bursting electrical activity in pancreatic beta-cells to the glucose dynamics in the whole body. Phys Lett A 376:3150–3157

Hand D, Mannila H, Smyth P (2001) Principles of Data Mining, 1st edn, Massachusetts: The MIT Press.

Hansen K (1923) Oscillations in the blood sugar in fasting normal persons. Acta Med Scand Suppl 4:27–58

Hantson L, De Weerdt W, De Keyser J, Diener H, Franke C, Palm R, Van Orshoven M, Schoonderwalt H, De Klippel N, Herroelen L (1994) The european stroke scale. Stroke 11:2215–2219

Hartman P (1960) A lemma in the theory of structural stability of differential equations. Proc Amer Math Soc 11(4):610–620

He Y, Wu M, She JH, Liu GP (2005) Robust stability for delay lur'e control systems with multiple nonlinearities. J Comput Appl Math 176(2):371–380

He Y, Wu M, She J (2006) Delay-dependent stability criteria for linear systems with multiple time delays. IEE Proc-Control Theory Appl 153(4):447–452

Hénon N, Godefroy O, Leys D, Mounier-Vehier F, Lucas C, Rondepierre P, Duhamel A, Pruvo JP (1995) Early predictors of death and disability after acute cerebral ischemic event. Stroke 26:392–398

Hill A (1910) The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J Physiol 40:iv–vii

Huang M, Li J, Song X, Guo H (2012) Modeling impulsive injections of insulin: towards artificial pancreas. SIAM J Appl Math 72:1524–1548

Huard B, Easton J, Angelova M (2015) Investigation of stability in a two-delay model of the ultradian oscillations in glucose–insulin regulation. Commun Nonlinear Sci Numer Simul 26(1):211–222

Jones D, Plank M, Sleeman B (2010) Differential equations and Mathematical Biology, 2nd edn, CRC Press

Jorgensen H, Nakayama H, Olsen T, Raaschou H (1994) Effect of blood pressure and diabetes on stroke in progression. Lancet 344(8916):156–159

Kang H, Han K, Choi M (2012) Mathematical model for glucose regulation in the whole-body system. Islets 4(2):84–93

Keener J, Sneyd J (1998) Mathematical physiology, vol 8. Springer

Kira K, Rendell L (1992) The feature selection problem: Traditional methods and a new algorithm. In: Proceedings of the Tenth National Conference on Artificial Intelligence, pp 129–134

Kissler S, Cichowitz C, Sankaranarayanan S, Bortz D (2014) Determination of personalized diabetes treatment plans using a two-delay model. J Theor Biol 359:101–111

Kitabchi A, Duckworth W, Brush J, Heinemann M (1971) Direct measurement of proinsulin in human plasma by the use of an insulin-degrading enzyme. J Clin Invest

50(9):1792–1799

Kluck O, Berman M, Stamler A, Sahar G, Kogan A, Porat E, Sagie A (2007) Value of echocardiography for stroke and mortality prediction following coronary artery bypass grafting. Interact Cardiovasc Thorac Surg 6:30–34

Koller K, Rockwood K (2013) Frailty in older adults: Implications for end-of-life care. Clev Clin J Med 80:168–174

Kothari V, Stevens R, Adler A, Stratton I, Manley S, Neil H, Holman R, et al (2002) UKPDS 60 risk of stroke in type 2 diabetes estimated by the UK prospective diabetes study risk engine. Stroke 33(7):1776–1781

Kraegen E, Young J, George E, Lazarus L (1972) Oscillations in blood glucose and insulin after oral glucose. Horm Metab Res 4:409–413

Laird A (1964) Dynamics of tumour growth. Brit J Cancer 18(3):490–502

Lalkhen A, McCluskey A (2008) Clinical tests: sensitivity and specificity. Contin Educ Anaesth Crit Care Pain 8(6):221–223

Lam ZH, Hwang KS, Lee JY, Chase J, Wake G (2002) Active insulin infusion using optimal and derivative-weighted control. Med Eng Phys 24(10):663–672

Leszczynski J, Shukla M (2009) Practical Aspects of Computational Chemistry: Methods, Concepts and Applications. Springer

Li J, Kuang Y (2007) Analysis of a model of the glucose-insulin regulatory system with two delays. SIAM J Appl Math 63(3):757–776

Li J, Kuang Y, Li B (2001) Analysis of IVGTT glucose-insulin interaction models with time delay. Discrete Contin Dyn Syst Ser B 1(1):103–124

Li J, Kuang Y, Mason C (2006) Modeling the glucose-insulin regulatory system and ultradian insulin secretory oscillations with two explicit time delays. J Theor Biol 242(3):722–735

Li J, Wang M, De Gaetano A, Palumbo P, Panunzi S (2012) The range of time delay and

the global stability of the equilibrium for an ivgtt model. Math Biosci 235:128–137

Long N, Gianola D, Rosa G, Weigel K, Avendano S (2009) Comparison of classification methods for detecting associations between snps and chick mortality. Genet Sel Evol 48:4031–4039

Macklin P, McDougall S, Anderson A, Chaplain M, Cristini V, Lowengrub J (2009) Multiscale modelling and nonlinear simulation of vascular tumour growth. J Math Biol 58(4-5):765–798

Mahoney F, Barthel D (1965) Functional evaluation: the Barthel index. Md State Med J 14:61–65

Mathers C, Loncar D (2006) Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 3(11):e442

MATLAB (2013) version 8.1.0 (R2013a). The MathWorks Inc., Natick, Massachusetts

Morrish N, Wang SL, Stevens L, Fuller J, Keen H (2001) Mortality and causes of death in the who multinational study of vascular disease in diabetes. Diabetologia 44(2):S14–S21

Muir K, Weir C, Murray G, Povey C, Lees K (1996) Comparison of neurological scales and scoring systems for acute stroke prognosis. Stroke 27:1817–1820

Mullins I, Siadatya M, Lymana J, Scully K, Garrett C, Miller W, Muller R, Robson B, Aptec C, Weiss W, Rigoutsos I, Platt D, Cohen S, Knaus W (2006) Data mining and clinical data repositories: Insights from a 667,000 patient data set. Comput Biol Med 36:1351–1377

Murray J (2002) Mathematical Biology 1: An Introduction, 3rd edn, New York: Springer

Mustacchi P (1985) Risk factors in stroke. Western J Med 143(2):186–192

Nedeltchev K, Renz N, Karameshev A, Haefeli T, Brekenfeld C, Meier N, Remond L, Schroth G, Arnold M, Mattle H (2010) Predictors of early mortality after acute ischaemic stroke. Swiss Med Wkly 140:254–259

Nesca V, Guay C, Jacovetti C, Menoud V, Peyot ML, Laybutt D, Prentki M, Regazzi R (2013) Identification of particular groups of microRNAs that positively or negatively impact on beta cell function in obese models of type 2 diabetes. Diabetologia 56(10):2203–2212

Neugebauer R, Fireman B, Roy J, O'Connor P, Selby J (2012) Dynamic marginal structural modeling to evaluate the comparative effectiveness of more or less aggressive treatment intensification strategies in adults with type 2 diabetes. Pharmacoepidem Dr S 21(S2):99–113

NHS (2012a) Anticoagulants, warfarin - NHS choices. URL http://www.nhs.uk/Conditions/Anticoagulants-warfarin-/Pages/Introduction.aspx

NHS (2012b) Digoxin - heart rhythm disorders medicines and drugs - NHS choices. URL http://www.nhs.uk/medicine-guides/pages/MedicineOverview.aspx?condition=Heart%20rhythm%20disorders&medicine=digoxin

NHS (2014) NHS choices: Diabetes. URL http://www.nhs.uk/Conditions/Diabetes/Pages/Diabetes.aspx

Pacini G, Bergman R (1986) Minmod: a computer program to calculate insulin sensitivity and pancreatic responsivity from the frequently sampled intravenous glucose tolerance test. Comput Meth Prog Bio 23(2):113–122

Palnum K, Andersen G, Ingeman A, Krog B, Bartels P, Johnsen S (2009) With acute stroke in Denmark : A nationwide follow-up study. sex-related differences in quality of care and short-term mortality among patients. Stroke 40:1134–1139

Palumbo P, Panunzi S, De Gaetano A (2007) Qualitative behavior of a family of delay-differential models of the glucose-insulin system. Discrete Contin Dyn Syst Ser B 7:399–424

Palumbo P, Pepe P, Panunzi S, De Gaetano A (2013) Observer-based closed-loop control for the glucose-insulin system: local input-to-state stability with respect to unknown meal disturbances. In: Proceedings of the 2013 American Control Conference (ACC 2013), pp 1751–1756

Panunzi S, Palumbo P, De Gaetano A (2007) A discrete single time delay model for the intra-venous glucose tolerance test. Theor Biol Med Model 4(35):1–16

Park P (1999) A delay-dependent stability criterion for systems with uncertain time-invariant delays. IEEE Trans Autom Control 44(4):876–877

Pedersen M, Man C, Cobelli C (2011) Multiscale modeling of insulin secretion. IEEE Trans Bio-Med Eng 58(10):3020–3023

Peduzzi P, Concato J, Kemper E, Holford T, Feinstein A (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49(12):1373–1379

Pendergast D, Fisher N, Calkins E (1993) 9 cardiovascular, neuromuscular, and metabolic alterations with age leading to frailty. J Gerontol 48:61–67

Polonsky K, Given B, Van Cauter E (1988) Twenty-four-hour profiles and pulsatile patterns of insulin secretion in normal and obese subjects. J Clin Invest 81(2):442–448

Quinlan JR (1993) C4.5: Programs for Machine Learning, 1st edn, California: Morgan Kaufmann Publishers

Radhakrishna Rao C, Mitra S (1972) Generalized inverse of a matrix and its applications. In: Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, vol 1, pp 601–620

Resnick H, Valsania P, Halter J, Lin X (2000) Relation of weight gain and weight loss on subsequent diabetes risk in overweight adults. J Epidemiol Commun H 54(8):596–602

Riachy M, Sfeir F, Sleilaty G, Hage-Chahine S, Dabar G, Bazerbachi T, Aoun-Bacha Z, Khayat G, Koussa S (2008) Prediction of the survival and functional ability of severe

stroke patients after icu therapeutic intervention. BMC Neurol 8(24)

Rissanen J (1978) Modeling by shortest data description. Automatica 14(5):465–471

Rudin W (1986) Real and complex analysis, 3rd edn. New York: McGraw-Hill Inc

Russell S, Norvig P (2010) Artificial Intelligence A Modern Approach, 3rd edn, New Jersey: Pearson

Sacco R, Benjamin E, Broderick J, Dyken M, Easton J, Feinberg W, Goldstein L, Gorelick P, Howard G, Kittner S, Manolio T, Whisnant J, Wolf P (1997) Risk factors. Stroke 28(1):1507–1517

Sacre J, Jellis C, Jenkins C, Haluska B, Baumert M, Coombes J, Marwick T (2014) A six-month exercise intervention in subclinical diabetic heart disease: effects on exercise capacity, autonomic and myocardial function. Metabolism

Sernyak M, Leslie D, Alarcon R, Losonczy M, Rosenheck R (2002) Association of diabetes mellitus with use of atypical neuroleptics in the treatment of schizophrenia. Am J Psychiat 159(4):561–566

Shannon C (1948) A mathematical theory of communication. Bell Syst Tech J 27(3):379–423

Shapiro E, Tillil H, Polonsky K, Fang V, Rubenstein A, Cauter E (1988) Oscillations in insulin secretion during constant glucose infusion in normal man: Relationship to changes in plasma glucose. J Clin Endocrinol Metab 67(2):307–314

Steil G (2013) Algorithms for a closed-loop artificial pancreas: the case for proportional-integral-derivative control. J Diabetes Sci Technol 7(6):1621–1631

Steil G, Reifman J (2009) Mathematical modeling research to support the development of automated insulin-delivery systems. J Diabetes Sci Technol 3(2):388–395

Steil G, Hipszer B, Reifman J (2010) Update on mathematical modeling research to support the development of automated insulin delivery systems. J Diabetes Sci Technol 4(3):759–769

Steyerberg E, Vickers A, Cook N, Gerds T, Gonen M, Obuchowski N, Pencina M, Kattan M (2010) Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology 21(1):128

Sturis J, Polonsky K, Mosekilde E, Van Cauter E (1991a) Computer model for mechanisms underlying ultradian oscillations of insulin and glucose. Am J Physiol 260(5):E801–E809

Sturis J, Van Cauter E, Blackman J, Polonsky K (1991b) Entrainment of pulsatile insulin secretion by oscillatory glucose infusion. J Clin Invest 87(2):439–445

Sturis J, Knudsen C, O'Meara N, Thomsen J, Mosekilde E, Cauter E, Polonsky K (1995) Phase-locking regions in a forced model of slow insulin and glucose oscillations. Chaos 5(1):193–199

van Swieten J, Koudstaal P, Visser M, Schouten H, van Gijn J (1979) Interobserver agreement for the assessment of handicap in stroke patients. Stroke 19:604–607

Tolic I, Mosekilde E, Sturis J (2000) Modeling the insulin-glucose feedback system: The significance of pulsatile insulin secretion. J Theor Biol 207(3):361–375

Topp B, Promislow K, DeVries G, Miura R, Finegood D (2000) A model of β-cell mass, insulin, and glucose kinetics: Pathways to diabetes. J Theor Biol 206:605–619

Vicini P, Caumo A, Cobelli C (1997) The hot IVGTT two-compartment minimal model: indexes of glucose effectiveness and insulin sensitivity. Am J Physiol-Endoc M 273(5):E1024–E1032

Wiggins S (2003) Introduction to applied nonlinear dynamical systems and chaos, 2nd edn, Springer

Williamson R (1997) Introduction to differential equations and dynamical systems, 2nd edn, McGraw Hill

Wu J (1996) Theory and applications of partial functional differential equations, vol 119. Springer

Wu M, He Y, She JH (2004a) New delay-dependent stability criteria and stabilizing method for neutral systems. IEEE Trans Autom Control 49(12):2266–2271

Wu M, He Y, She JH, Liu GP (2004b) Delay-dependent criteria for robust stability of time-varying delay systems. Automatica 40(8):1435–1439

Wu X, Kumar V, Quinlan J, Ghosh J, Yang Q, Motoda H, McLachlan G, Ng A, Liu B, Yu P, Zhou Z, Steinbach M, Hand D, Steinberg D (2008) Top 10 algorithms in data mining. Knowl Inf Syst 14:1–37

Wu Z, Chui CK, Hong GS, Chang S (2011) Physiological analysis on oscillatory behavior of glucose–insulin regulation by model with delays. J Theor Biol 280(1):1–9

Yagil G, Yagil E (1971) On the relation between effector concentration and the rate of induced enzyme synthesis. Biophys J 11:11–27

Yildirim N, Santillian M, Horike D, Mackey M (2004) Dynamics and bistability in a reduced model of the lac operon. Chaos 14(2):279–292

Young W (1912) On classes of summable functions and their Fourier series. P Roy Soc Lond A Mat 87(594):225–229

Zahuranec D, Sanchez B, Brown D, Wing J, Smith M, Garcia N, Meurer W, Morgenstern L, Lisabeth L (2012) Computed tomography findings for intracerebral hemorrhage have little incremental impact on post-stroke mortality prediction model performance. Cerebrovasc Dis 34:86–92

Zhang H (2005) Exploring conditions for the optimality of naive bayes. Int J Pattern Recogn 19(02):183–198

Zhang P, Zhang X, Brown J, Vistisen D, Sicree R, Shaw J, Nichols G (2010) Global healthcare expenditure on diabetes for 2010 and 2030. Diabetes Res Clin Pr 87(3):293–301