Analysis of Non-Steady State Physiological and Pathological Processes

Volume I

Nathan R Hill

Harris-Manchester College

Thesis Submitted for the Degree of Doctor of Philosophy

University of Oxford

Michaelmas Term 2008

Nuffield Department of Medicine
John Radcliffe Hospital
Headington
Oxford OX3 9DU

Analysis of Non-Steady State Physiological and Pathological Processes

Nathan R Hill
Harris-Manchester College
Submitted for the Degree of Doctor of Philosophy
Michaelmas Term 2008

Abstract

The analysis of non steady state physiological and pathological processes concerns the abstraction, extraction, formalisation and analysis of information from physiological systems that is obscured, hidden or unable to be assessed using traditional methods.

Time Series Analysis (TSA) techniques were developed and built into a software program, Easy TSA, with the aim of examining the oscillations of hormonal concentrations in respect to their temporal aspects – periodicity, phase, pulsatility. The Easy TSA program was validated using constructed data sets and used in a clinical study to examine the relationship between insulin and obesity in people without diabetes. In this study fifty-six non-diabetic subjects (28M, 28F) were examined using data from a number of protocols. Fourier Transform and Autocorrelation techniques determined that there was a critical effect of the level of BMI on the frequency, amplitude and regularity of insulin oscillations.

Second, information systems formed the background to the development of an algorithm to examine glycaemic variability and a new methodology termed the Glycaemic Risk in Diabetes Equation (GRADE) was developed. The aim was to report an integrated glycaemic risk score from glucose profiles that would complement summary measures of glycaemia, such as the HbA1c. GRADE was applied retrospectively to blood glucose data sets to determine if it was clinically relevant. Subjects with type 1 and type 2 diabetes had higher GRADE scores than the non-diabetic population and the contribution of hypo- and hyperglycaemic episodes to risk was demonstrated. A prospective study was then designed with the aim to apply GRADE in a clinical context and to measure the statistical reproducibility of using GRADE. Fifty-three (Male 26, Female 27) subjects measured their blood glucose 4 times daily for twenty-one days. The results were that lower HbA1c's correlated with an increased risk of hypoglycaemia and higher HbA1c's correlated with an increased risk of hyperglycaemia. Some subjects had HbA1c of 7.0 but had median GRADE values ranging from 2.2 to 10.5. The GRADE score summarized diverse glycaemic profiles into a single assessment of risk. Well-controlled glucose profiles yielded GRADE scores <= 5 and higher GRADE scores represented increased clinical risk from hypo or hyperglycaemia.

Third, an information system was developed to analyse data-rich multi-variable retinal images using the concept of assessment of change rather than specific lesion recognition. A fully Automated Retinal Image Differencing (ARID) computer system was developed to highlight change between retinal images over time. ARID was validated using a study and then a retrospective study sought to determine if the use of the ARID software was an aid to the retinal screener. One hundred and sixty images (80 image

pairs) were obtained from Gloucestershire Diabetic Eye Screening Programme. Images pairs were graded manually and categorised according to how each type of lesion had progressed, regressed, or not changed between image A and image B. After a 30 day washout period image pairs were graded using ARID and the results compared. The comparison of manual grading to grading using ARID (Table 4.3) demonstrated an increased sensitivity and specificity. The mean sensitivity of ARID (87.9%) was increased significantly in comparison to manually grading sensitivity (84.1%) ($p<0.05$). The specificity of the automated analysis (87.5%) increased significantly from the specificity (56.3%) achieved by manually grading ($p<0.05$). The conclusion was that automatic display of an ARID differenced image where sequential photographs are available would allow rapid assessment and appropriate triage.

Forth, non-linear dynamic systems analysis methods were utilised to build a system to assess the extent of chaos characteristics within the insulin-glucose feedback domain. Biological systems exist that are deterministic yet are neither predictable nor repeatable. Instead they exhibit chaos, where a small change in the initial conditions produces a wholly different outcome. The glucose regulatory system is a dynamic system that maintains glucose homeostasis through the feedback mechanism of glucose, insulin, and contributory hormones and was ideally suited to chaos analysis. To investigate this system a new algorithm was created to assess the Normalised Area of Attraction (NAA). The NAA was calculated by defining an oval using the 95% CI of glucose & Insulin (the limit cycle) on a phasic plot. Thirty non-diabetic subjects and four subjects with type 2 diabetes were analysed. The NAA indicated a smaller range for glucose and insulin excursions with the non-diabetics subjects ($p<0.05$). The conclusion was that the evaluation of glucose metabolism in terms of homeostatic integrity and not in term of cut-off values may enable a more realistic approach to the effective treatment and prevention of diabetes and its complications.

Contents

List of Figures

List of Tables

List of Equations

Signal to noise ratio

EQ 1.1

$$SNR = \left(\frac{Asignal}{Anoise} \right)^2$$

Where A is the Root Mean Square of the amplitude.

Auto Regressive Moving Average

EQ 2.1

$$\frac{\sum_{-a}^{a} X_{(t-a)}}{m}$$

$$a = (m-1)/2$$

Where *X* is the dependent variable under observation (i.e. hormone concentration), *t* is time and *m* is an odd positive integer.

Three Point Auto Regressive Moving Average

EQ 2.2

$$\frac{\sum_{-1}^{1} X_{(10-1)}}{3}$$

$$a = (3-1)/2$$

Linear stationarisation

EQ 2.3

$$B = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sum X^2 - \frac{\left(\sum X\right)^2}{N}}$$

Where *B* is the slope of the line, *X* is time, *Y* is the observed value and *N* is number of data points.

End to end stationarisation

EQ 2.4

$$B = \frac{X_{Last} - X_{first}}{N}$$

$$X' = X - (B * t)$$

Where *B* = the slope, X_{Last} = last data point, X_{first} = first data point, N = total number of data points. X' is the new data point, X = current data point, t = current period.

Cubic polynomial

EQ 2.5

$$y = b_0 + b_1 x + b_2 x^2 + b_3 x^3$$

Fourier Transform

EQ 2.6

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j 2 \pi f t}\, \Delta t$$

Where $X(f)$ is the FT of $x(t)$. Frequency (f) is measured in Hertz and t is time. The code used to calculate this is in Chapter 1 of the code thesis that accompanies this thesis

OC5 from regression line for Observed Concentration example

EQ 2.7

$$y = 1.0573x + 4.2886$$

$$OC5 =$$

$$y = 3.3551 = 1.0573x + 4.2886 \therefore$$

$$3.3551 - 4.2886 = 1.0573x \therefore$$

$$\frac{3.3551 - 4.2886}{1.0573} = x \therefore$$

$$OC5 = 10^{x} \therefore$$

$$OC5 = 10^{-0.88} \therefore$$

$$OC5 = 0.13$$

Runs analysis

EQ 2.8

$$C_m^{m+n} = C_n^{m+n}$$

The sequence aabbbab, has 3 a's, 3 b's and forms 4 runs (aa, bbb, a, b). Therefore if there are m number of a's and n number of b's total number of possible groupings

Runs analysis groups

EQ2.9

$$P\{u \le u'\} = \frac{1}{C_n^{m+n}} \sum_{u=2}^{u'} \int u$$

If u is the number of groups (runs) in the current dataset then the probability of an arrangement of groups yielding u' or fewer groups can be represented by

Where $\int u = 2C_{k-1}^{m-1} \cdot C_{k-1}^{n-1}$, when $u = 2k$, i.e. u is even

in addition $\int u = C_{k-1}^{m-1} \cdot C_{k-2}^{n-1} + C_{k-2}^{m-1} \cdot C_{k-1}^{n-1}$,

when $u = 2k - 1$, i.e. u is odd

Null Hypothesis for Permutation analysis

EQ 2.10

$$H_0 : F = G$$

null hypothesis H_0 of no difference between the two groups F and ,

Random noise equation

EQ 2.11

If the random number is odd then $X_1 = X - n$

If it is even then $X_1 = X + n$

Where, $n = Rand * \left(\left[l / 100 \right] * X_{\max} \right)$

Where X is the data before the noise and X_1 is the data after the noise has been applied, l is equivalent to the percentage power of noise (i.e. 1%), n is equal to the absolute value of noise. The noise could either be added to or subtracted from the data profile depending on whether the random number was odd or even.

Observed Concentration Validation

EQ 2.11 OC5 **=** 10^((3.3551 – Intercept)/Gradient)
OC50 = 10^((5 – Intercept)/Gradient)
OC95 = 10^((6.6449 – Intercept)/Gradient)

Intercept and gradient from line of best fit

GRADE

EQ 3.1 Mmol/l GRADE value = 425 * (log (log(x)) + 0.16)

EQ 3.2 Mg/dl GRADE value = 425 * (log (log(x*18)) + 0.16)

Where x is the observed glucose value and the log is base 10.

GRADE

EQ 3.3 Equation for calculation of contribution to GRADE score as a percentage. BG measured in mmol/l

Hypoglycaemia (BS<3.9)

$$\frac{\sum \text{GRADE BS} < 3.9}{\sum \text{All GRADE Values}} \times 100$$

Euglycaemia (3.9≥BS≤7.8)

$$\frac{\sum \text{GRADE } 3.9 \geq \text{BS} \leq 7.8}{\sum \text{All GRADE Values}} \times 100$$

Hyperglycaemia (BS≥7.8)

$$\frac{\sum \text{GRADE BS} > 7.8}{\sum \text{All GRADE Values}} \times 100$$

Rotation

EQ4.1 x' = Cos (theta) * x – Sin (theta) * y
y' = Sin (theta) * x + Cos (theta) * y

Theta is the angle of rotation. Co-ordinates (x, y) are rotated to (x', y').The results of the equations are the new position of x and y

Fast Normalised Cross-Correlation

EQ4.2

$$C(u,v) = \frac{\sum_{xy}\left(I[u+x, v+y] - \bar{I}uv\right)\hat{T}(x,y)}{\sum_{xy}\left(I[u+x, v+y] - \bar{I}uv\right)^2 \sum_{xy}\hat{T}(x,y)^2}$$

Where I is image one, and $\hat{T}$ is image two. An increased correlation (C) the more exact the match.

Sensitivity & Specificity

EQ 4.1

$$sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

Sensitivity and specificity equations where TP = True positive, FP = False positive, TN = True negative , FN = False negative

Area of oval

EQ 5.1

$$Area = Height * Width * \frac{\pi}{4} \quad where,$$

$$Height = \log(\bar{y} + Y_{CI}) - \log(\bar{y} - Y_{CI}) \quad and$$

$$Radius = (X_{CI} * 0.5) \therefore$$

$$Width = 2 * (X_{CI} * 0.5) \therefore$$

$$Width = X_{CI}$$

Where $\bar{y}$ = mean insulin, y_{CI} =95%CI of insulin, x_{CI} = 95%CI of glucose

Abbreviations

AC Autocorrelation

ARID Automated Retinal Image Differencing

ARMA	Auto-Regressive Moving Average (see MPA)
ATP	Adenosine TriPhosphate
BG	Blood Glucose
BMI	Body Mass Index
CC	crosscorrelation
CCC	CrossCorrelation Correlogram
CD	Compact Disk
CGMS	Continuous Glucose Measuring System
CI	Confidence Interval
COREC	Central Oxford Research Ethics Committee
CRT	Cathode Ray Tube
CV	Coefficient of Variation
CWS	Cotton Wool Spot
DCCT	Diabetes Control and Complications Trial
DR	Discriminent Ratio
ECG	ElectroCardioGram
EDSPAG	English Diabetic Screening Project Advisory Group
ELISA	Enzyme-Linked Immunosorbent Assays
EQ	EQuation
ETDRS	Early Treatment of Diabetic Retinopathy study
FN	False Negative
FNCC	Fast Normalised cross-correlation
FP	False Positive
FPD	Fibrous Proliferations at Disc
FPE	Fibrous Proliferations Elsewhere
FPI	Fasting Plasma Insulin
FSH	Follicle-Stimulating Hormone
FT	Fourier Transform
GH	Growth Hormone
GnRH	Gonadotrophin Releasing Hormone
GRADE	Glycaemic Risk in Diabetes Equation
HBGM	Home blood glucose monitoring
HE	Hard Exudates
HOMA	Homeostasis Model Assessment
IRMA	IntraRetinal Microvascular Abnormalities
LH	Luteinising Hormone
LI	Lability Index
MA	Microaneurysm
MAGE	Mean Amplitude of Glycaemic Excursions
MPA	Moving Point Average (see ARMA)

NAA	Normalised Area of Attraction
NICE	National Institute for Clinical Excellence
NSF	National Service Framework
NVD	New Vessels at Disc
NVE	New Vessels Elsewhere
OC	Observed Concentration
OCDEM	Oxford Centre for Diabetes, Endocrinology & Metabolism
OGTT	Oral Glucose Tolerance Test
RGB	Red, Green, Blue
RGC	Retinopathy Grading Centre
ROI	Region Of Interest
SAA	Serial Array Averaging
SD	Standard Deviation
SEM	Standard Error of the Mean
SNR	Signal Noise Ratio
TN	True Negative
TP	True Positive
TSA	Time Series Analysis
TSH	Thyroid-Stimulating Hormone
UK	United Kingdom
UKPDS	United Kingdom Prospective Diabetes Study
USA	United States of America
VB	Venous Beading

Acknowledgements

I would like to thank the many people without whom the work presented in this thesis would not be possible.

Prof. David Matthews, my supervisor and boss, who provided a malnourished mind the opportunity to expand. Without the initial job offer to be a programmer I may never have been involved in the ongoing effort to cure diabetes. I would like to continue to thank Prof. Matthews for providing the support and guidance needed to complete this thesis and his excitement at the proof provided herein that a year is about 365 days long (further information in chapter 2).

I am grateful to Dr. Niki Meston and Dr. Richard Stevens for having the stamina to wade through a plethora of incarnations of my thesis with the ability of expressing their opinions without damaging my ego (too much). I am indebted to Ms. Irene Stratton for statistical help and acerbic wit on the current uses to which statistics are put in medical journals. I also thank Prof. Peter Hindmarsh for his collaborative energies and leadership on time series studies and Mr Steve Aldington for his collaborative work and instruction on retinal imaging and retinopathy. Thanks also go to Dr Jonathan Levy for his critical input into rectifying my lack of understanding of Beta-cells. I am grateful to Prof. Sir Mike Brady for his instruction and advice on computer vision techniques.

I would like to thank Mrs Carol Hill for administrative support throughout the years of my thesis development and for the unearthly ability to manage Prof Matthews diary so that I could expound on my latest (re-)discoveries with him.

Particular thanks are also owed to my wife Lisa for providing invaluable support over the last few years.

Thank You

1.0 Analysis of non-steady state physiological and pathological processes

‘Knowledge is power’ – Sir Francis Bacon[1]

1.1 Aim

To develop and implement methods that are able to manipulate megabyte data sets from subjects with metabolic disorders that are beyond the assessment of traditional methods, in order to produce clinically relevant, objective, user-independent information and outcomes.

1.1.1 Definitions

The definition of *'megabyte'* in the context of this thesis is data sets that contain multiple variables and/or multiple data points. The multi-variables are typically a substrate vs. time and the multi-data points are repeated measurements of the variables over a defined time period. The *'information'* is the clinically relevant, objective, user-independent output determined from the *'data'*.

For example, the assessment of home monitored blood glucose data contains 2 variables (glucose vs. time) and may contain multiple data points as the subject may measure their glucose throughout the day, over a period of weeks. The traditional assessment of the data is performed by a clinician's (subjective) assessment of hypo- and hyperglycaemic periods with the output as the advising information they provide. Such subjective analysis does not utilise the full information content of the data.

Other examples of low information return on data rich megabyte sets are often found in medicine. Dynamic endocrine tests, ECG's, retinal photographs, cytology and much clinical biochemistry are all generating vast data arrays at considerable expense with sparse information retrieval often based on single values and subjective assessments.

Trends, pulsatility and noise are typically ignored. The work within this thesis attempts to address this shortfall in analysis

1.2 Background to diabetes

Diabetes is a metabolic disorder resulting from an inability to produce, or utilise effectively, enough insulin to maintain glucose homeostasis[2]. Insulin is a hormone produced in the islets of Langerhans, within the pancreas, by a finite population of beta cells. Insulin functions are multiple and provide the main hormonal organisation of the post-prandial metabolism. Insulin co-ordinates a reduction in endogenous hepatic glucose production, stimulates glucose uptake into peripheral tissues, inhibits fatty acid release while stimulating triglyceride storage and increases glucose storage as glycogen within the liver. Insulin is the prime controlling mechanism of energy flux *in vivo [3]*. When there is insufficient insulin, peripheral tissues, such as skeletal muscle and adipose tissue, cannot access sufficient glucose from the blood due to a variety of intra-cellular mechanisms and reduced glucose receptor production/activation. An excess of glucose within plasma (hyperglycaemia) for an extended period results in the binding of glucose molecules to proteins (glycosylation) within vascular cells. Extended hyperglycaemia results in an increased potential for both microvascular and macrovascular pathology[4].

Diabetes is not a single disease. It occurs in several forms: type 1, type 2, gestational and a wide range of monogenic forms. Type 1 and type 2 are the two most common forms and represent about 9% and 90% respectively of the total diabetic population worldwide [5]. The distinction between what is now known as type 1 diabetes and type 2 diabetes was made by Sir Harold (Harry) Percival Himsworth in 1935 and the findings

were published in January 1936[6]. His work recognised that diabetes is actually a spectrum of diseases and not just a single pathological entity.

1.2.1 Type 1 diabetes

Type 1 diabetes is one of the most common non-communicable childhood diseases in developed nations. It is caused by an auto-immune process which results in a destruction of pancreatic islet beta-cells[7]. Presentation may occur at any age although there is a peak of increased incidence in the teenage years around puberty[8, 9]. It has a dramatic clinical presentation. Because of the rapid reduction in mass/function of beta cells, the pancreas produces little or no insulin. Even with careful insulin therapy, type 1 diabetes can cause a reduction in quality of life and shortening of an average life span by about 12 years [5]. Potential complications such as renal failure, blindness, neuropathy, amputations due to vascular complications, heart attacks and stroke are recognised aspects of the long-term disease process.

1.2.2 Type 2 diabetes

Type 2 diabetes is often associated with truncal obesity and hypertension (high blood pressure). It has a strong but complex genetic predisposition[2] in association with a significant environmental component, particularly related to a sedentary lifestyle and a diet inappropriately high in calories and saturated fat[10, 11]. Type 2 diabetes is a metabolic disorder caused by a relative insulin deficiency almost always found co-existing with an increase in insulin resistance associated with an elevated accumulation of fat or adipose tissue and can be accompanied by additional cardiovascular risk factors of dyslipidaemia (abnormal blood lipid levels) and hypertension. This combination is

sometimes characterised as the Metabolic Syndrome (Reavens syndrome, or syndrome X)[12]

The pathophysiology is not fully understood but the current consensus is that type 2 diabetes is involves a relative and progressive deficiency in insulin secretion, due to an unidentified underlying mechanism[13].

1.2 Information systems

Information systems are methods for managing and processing information. The megabyte data sets require information systems to extract and formalise the relevant information from the data. Using information systems on large biological datasets can be problematic as they have common problems in several areas:

- An unclear hypothesis
- Data overload
- Noise
- Pulsatile information
- Hidden information
- Fractal information

1.2.1 An unclear hypothesis

It is essential to define clearly the hypothesis to be tested prior to data collection to avoid data overload. Aspects of this process include:

- What is the specific aim of this data collection?
- What questions can the data answer?

- What questions can the data not answer?
- How often should the data be collected?

For example, to establish if there may be a circadian rhythm of cortisol one could measure the hormone concentration at 8 am and again at midnight. Although it is a small dataset, the data are helpful in ascertaining whether there is or is not a diurnal rhythm present. However, such data cannot answer the questions about the quantity or detail of the cortisol signalling system *in vivo*. For such an analysis regular sampling and a time-series analysis are needed. Prior to investigation consideration of the frequency of the optimum sampling period is essential.

In the steady state the shortest advisable sample period can be judged from the half-life (t½) (clearance rate) of the hormone. Oscillations of period less than about 2(t½) are unlikely to be of physiological significance since any signalling based on such a period would be lost within the damping attributable to the half-life. For example, design an experiment to see if there were very fast oscillations of insulin secretion of period 20 seconds. Since insulin has a half-life of 3.8 minutes this would be a waste of time! The shortest oscillation of insulin that could be regarded as having physiological meaning (in terms of signalling, though not of release) would be a period of approximately 7 minutes.

1.2.2 Data overload

One of the problems faced in the clinic and in research is the ever-increasing amount of data available. As technology progresses additional ways are developed to record and measure physiological systems and/or reactions under investigation. For example, with the advent of the Continuous Glucose Measuring System (CGMS) it is now possible to have glucose concentration data generated at 3 minute intervals for periods of up to 72

hours or greater. It is difficult to decide how these data can be effectively used by clinicians, and it may even have a detrimental effect on the overall analysis of glycaemic control.

With the advent of fast computation, it has become possible to analyse large datasets in seconds, where previously this would have taken many days of calculation. Although computers excel at recursive analyses, as the data quantity increases the computation may take exponentially longer. For example, image processing involves vast datasets. A regular 8-mega pixel digital photo contains 8 million pixels of information, with each pixel being further subdivided in terms of visible colour in degrees of Red, Blue, or Green.

1.2.3 Noise

Physiological systems are generally complex and signalling is rarely unifactorial. Many inputs to a signal and complex output from that signal mean that information can be drowned out by noise or lost in a multiplicity of measurements. The aim is to take high volume data and convert this to tractable outcomes, which can be understood in physiological and pathological terms (this is examined further in Chapter 2). For example when measuring a hormone concentration in vivo, the information may be obscured by the action of other hormones, and the sub harmonics of other hormones. Likewise, the ability to grade retinopathy may be influenced by many associated factors such as fundal dilation, blood pressure, and ocular pressure.

It may be difficult to extract the ‘signal’ from the background data. The Oxford English dictionary defines noise as[14]:

1. *'In scientific and technical use: Random or irregular fluctuations or disturbances which are not part of a signal (whether the result is audible or not), or which interfere with or obscure a signal; oscillations with a randomly fluctuating amplitude over a usually continuous range of frequencies. Also (in extended use): distortions or additions which interfere with the transfer of information.'.*

2. *'In non-technical contexts: irrelevant or superfluous information or activity, esp. that which distracts from what is important.'*

Noise free data can never be realised in practice since some types of noise are the result of thermodynamic and quantum affects that cannot be avoided during measurement. Measurements derived from physiological systems are additionally subject to random data, feedback signals, assay errors, human error, secretion arrhythmia, and more.

There are methods to extract the signal from the noise, or the information from the raw data. These methods are known as 'Signal Processing'. The difficulty in extracting the information from the raw data depends both on the characteristics of the noise-free signal and the noise itself. The signal-to-noise-ratio (SNR) is the ratio of the power of the signal to the power of the noise. The higher the ratio, the easier it is to extract information and the more reliable the results. The SNR can be calculated by the equation in EQ 1.1 where A is the Root Mean Square of the amplitude.

EQ 1.1

$$SNR = \left(\frac{Asignal}{Anoise} \right)^2$$

1.2.4 Hidden information

If one were to take a normal pencil and look at it exactly head on, one could be excused for making the assumption that what was under observation was 2 concentric circles, one dark and one lighter in colour (Figure 1.1). However if one were to move the pencil by an angle of 90 degrees away from being head on one could then detect the previously ‘hidden’ information about the depth or length of the object. The analysis of some datasets closely resembles this analogy.

Figure 1.1 – Example of 'hidden' information. A pencil viewed from head-on contains no information about the pencil's length, when viewed at a 90-degree angle the length is now measurable.

1.2.4.1 Spectral domain

Joseph Fourier demonstrated in 1798 that a complex signal could be deconvoluted into terms of the sum of it frequencies[15]. When a complex signal is submitted to Fourier transformation, it yields data that can be displayed as the power spectrum of each component or as a range of frequencies. This deconvolutes the data to power/amplitude irrespective of phase, thus displaying the amount of power in the signal at each discrete frequency. This is analogous to white light passing through a prism, which splits it into its component parts (Figure 1.2) and is particularly useful in physiological analysis where datasets consist of complex waveforms that can be mistaken for indistinct background noise – further information in Chapter 2.2.3.1.

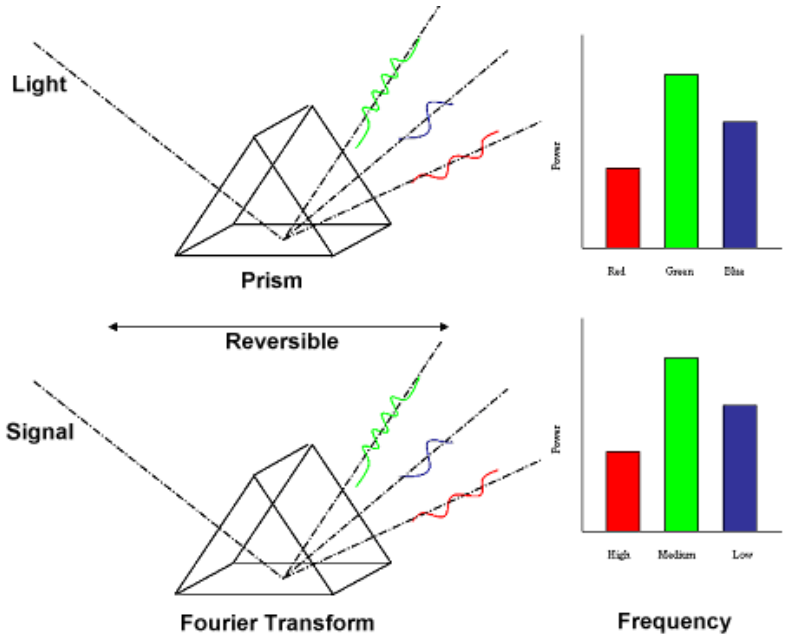


Figure 1.2 – Comparison of the Fourier Transform Time Series Analysis method to Sir Isaac Newton's Prism experiment. The methods split one incoming signal/beam into its component frequencies.

1.2.4.2 Phase domain

Phase space offers a method of turning numeric data into images. For example, a swinging pendulum is subject to friction. Every swing must end up at the centre, position 0 and eventually that is where the pendulum comes to rest. The central point "attracts" the swings. This can be described in a phase graph. As friction overcomes the energy in the system, a continual movement towards the centre is observed (Figure 1.3). This central point is known as the point attractor of the system.

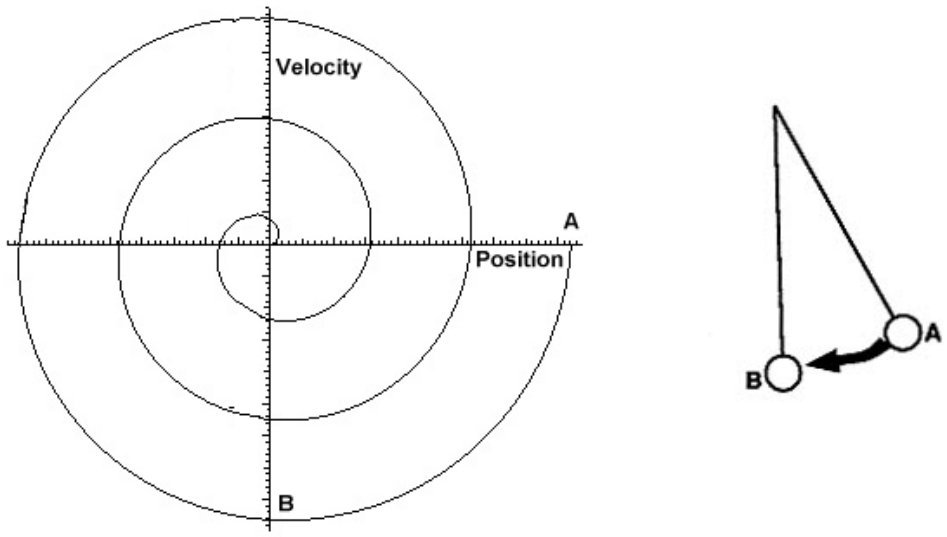


Figure 1.3 – Diagram of the movement of a pendulum oscillating from A to B and a graph of its diminishing power plotted as a phase space chart.

Diabetes is associated with a disruption in glucose-insulin homeostasis[3] and it is possible to analyse glycaemic-insulin data in the normal Cartesian way (line of best fit). However the glucose-insulin relationship comprises a dynamically complex and noisy system [16]. This system reflects rapid and continuous homeostatic modifications within normal physiology in response to alterations such as changes in activity, hormonal milieu, or dietary intake and should be analysed with this is mind. It is possible to describe such a system at a single instant in time as a point. However, the next instant the system will have subtly changed and so the coordinates of the point have changed. If one were to plot these points, a graph observing the history of the glucose insulin relationship would be produced. This is known as describing the system in 'phase space' and could provide additional information about the system (This is discussed in more detail in Chapter 5 Chaos).

1.2.5 Pulsatile information

The examination of the pulsatility information in a megabyte dataset may lead to insightful clinical outcomes. Many hormones are secreted in pulses and the examination of the pulsatility may lead to an objective result. Research into endocrine systems has revealed the significance and relevance of Circadian rhythms (temporal biological rhythms in approximate 24h cycles) of a variety of hormones. Cushing's syndrome may be diagnosed by analysis of the changes in diurnal cortisol release[17]. More rapid cycles of hormonal release (e.g. Growth hormone) have been detected recently[18-20] and the secretion of insulin from β-cells is pulsatile and the absence or suppression of this pulsatility has been demonstrated as BMI increases [21].

The difficulties in measuring hormone pulsatility are being overcome by increasingly sensitive and specific analytical assays such as the use of enzyme-linked immunosorbent assays (ELISA). The detection of pulsatility requires a specialised analysis such as Time Series Analysis (More in Chapter 2). Time series analysis techniques can be used to examine the phase, frequency & amplitude of hormonal concentrations and processes.

1.2.5.1 Waveforms

Waveforms or the shape of the dataset once plotted may be of use clinically when assessing the data objectively. Plotting a megabyte dataset as a graph enables a subjective assessment to be made on the dataset as a whole. The examination of waveforms within the datasets can be made objectively using Time Series Analysis techniques such as Autocorrelation (See Chapter 2.2.3.2) to determine if waveforms

reoccur and what waveforms are dominant within the dataset. It is further possible to compare waveforms using Cartesian methods such as area under the curve.

1.2.6 Fractal information

Over the last forty five years, since Edward N. Lorenz seminal publication[22] on chaos theory, there has been a marked development of mathematical thought relating to this area. Fractal analysis is a central area of chaos theory[23]. One of the most remarkable aspects of this has been the understanding that many phenomena in nature are fractal (entities that are scale invariant). Further scrutiny of detail in any observed phenomenon cannot yield an answer that regresses to a solution, but may diverge into a huge complexity of detail. One of the best examples of this is the Mandelbrot fractal[24] which is the plot of the real and imaginary domain of a regressive series (Figure 1.4).

This type of approach has been applied to a number of aspects of biology e.g. blood flow[25], heart rate analysis[26] and growth[27]. The heart rate analysis performed by Yeragani et al[26] determined that the fractal dimension of a heart rate time series was an effective measure of autonomic function and a more applicable analysis than traditional time series analysis.

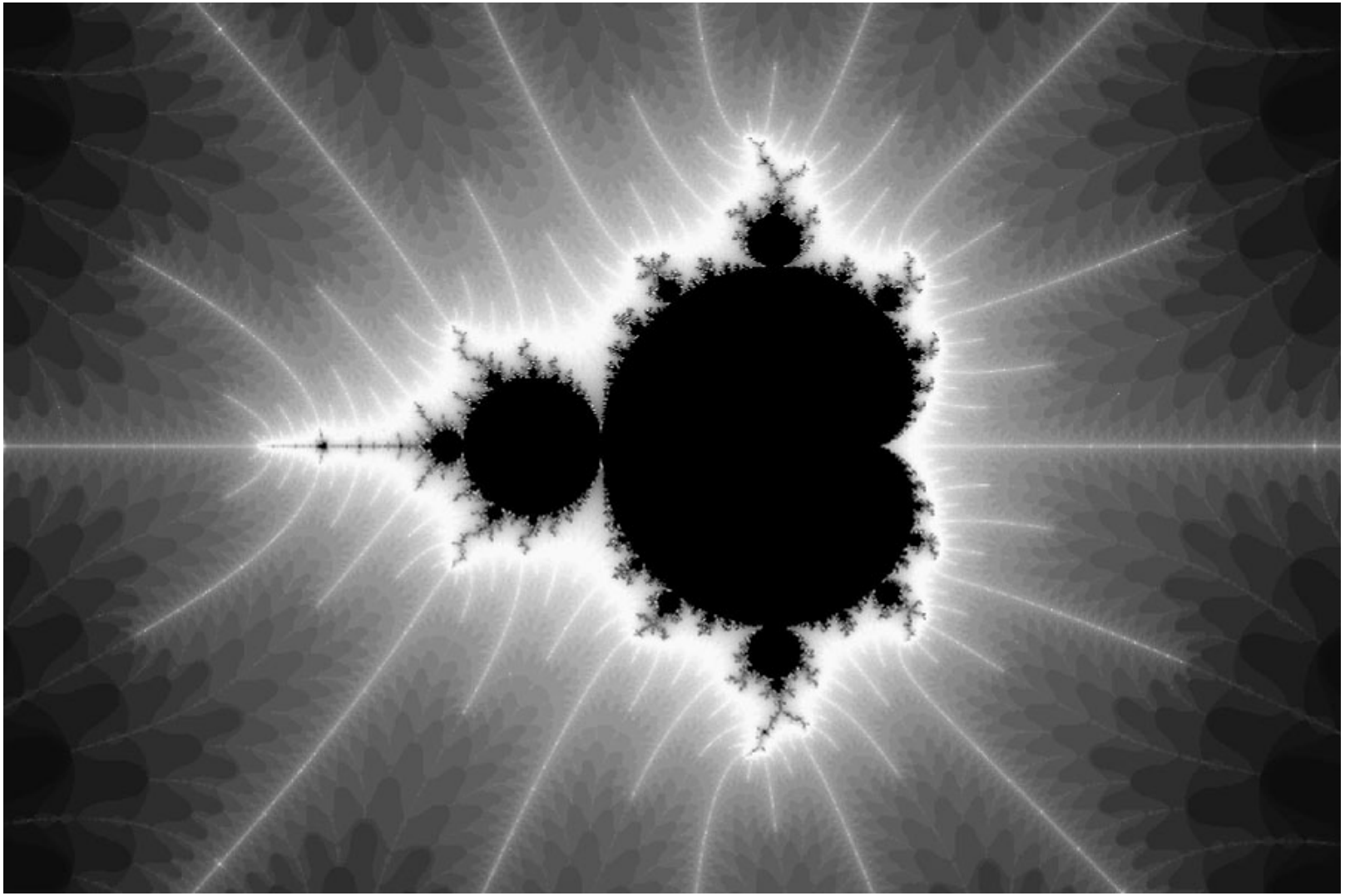


Figure 1.4- Mandelbrot fractal shown above is the perfect exemplification of the Strange Attractor.

1.5 Information systems in diabetes

The analysis of non steady state physiological and pathological processes concerns the abstraction, extraction, formalisation and analysis of information from physiological systems that is obscured, hidden or unable to be assessed using traditional methods. Photographic images of retinopathy require an abstraction of the images before grading (especially if analysed by computer) to reduce the amount of data needing processing, and extraction of the retinal architecture, and assessment of any ischaemia (further information in Chapter 4). Detailed glycaemic assessment in someone with diabetes might involve thousands of data items from a continuous glucose monitoring system that requires formalisation and abstraction before a clinically relevant judgement can be reached (further information in chapter 3). Hormonal analysis in those with diabetes can yield insights into the causes and effects of hormonal imbalances and feedback mechanisms (further information in chapter 5).

Information systems in diabetes range from a single measurement of a single variable, to determine a disease, to megabytes of information of multi-variables (Table 1.1). As the number of variables and the number of data points increase the assessment of the system reaches a threshold where the assessment changes from an objective observer independent assessment to a subjective observer dependent assessment.

Number of variables (excluding time)	Number of independent data points	Assay	Measurement Frequency	Process/Disease diagnosis	'Normal' ranges	
1	1	Thyroid-Stimulating Hormone (TSH)	Once	Hypo- and Hyperthyroidism	0.5-6 mU/L	
1	2	Cortisol	Am, Midnight	Hypercortisolism	280-700 nmol/L	Objective
1	4	Oral Glucose Tolerance Test (OGTT)	T1, T60, T90, T120	Impaired fasting glucose, Diabetes	<7.8 mmol/l post-prandial[28]	
2	10 to 30	Minimal model of Insulin/Glucose	Regular period of testing	secretion & kinetics	Varies[29]	
1	100's	Home blood glucose monitoring	1 to 10 times daily	Hypo- and Hyperglycaemia	4 – 6mmol/l pre-prandial	
1	1000's	Continuous glucose monitoring systems	T1,T3, …., Tn	Hypo- and Hyperglycaemia	3.5 to 11mmo/l	Subjective
3	Megabyte's	Retinal screening images	12 months	Retinopathy	No retinopathy present	

Table 1.1 Assessment of in vivo processes: How does an increase in the number of variables measured correlate with the subjectivity of the assessment.

The change from objective to subjective assessment appears to be paradoxical. With more data, one would expect the objectivity to increase. The reason it does not do so is that the processing of such complex data has been delegated to the human brain and an 'opinion' sought. This we term subjective. This does not mean the assessment is incorrect, but in quantifiable terms it becomes unrepeatable. In brief, opinions vary! However, if a study is conducted multiple times the more objective (and precise) the assessment can become (Figure 1.5) as the data points under consideration are no longer independent. The variables observed/measured at the same time period in the additional studies can be compared to the variable already obtained.

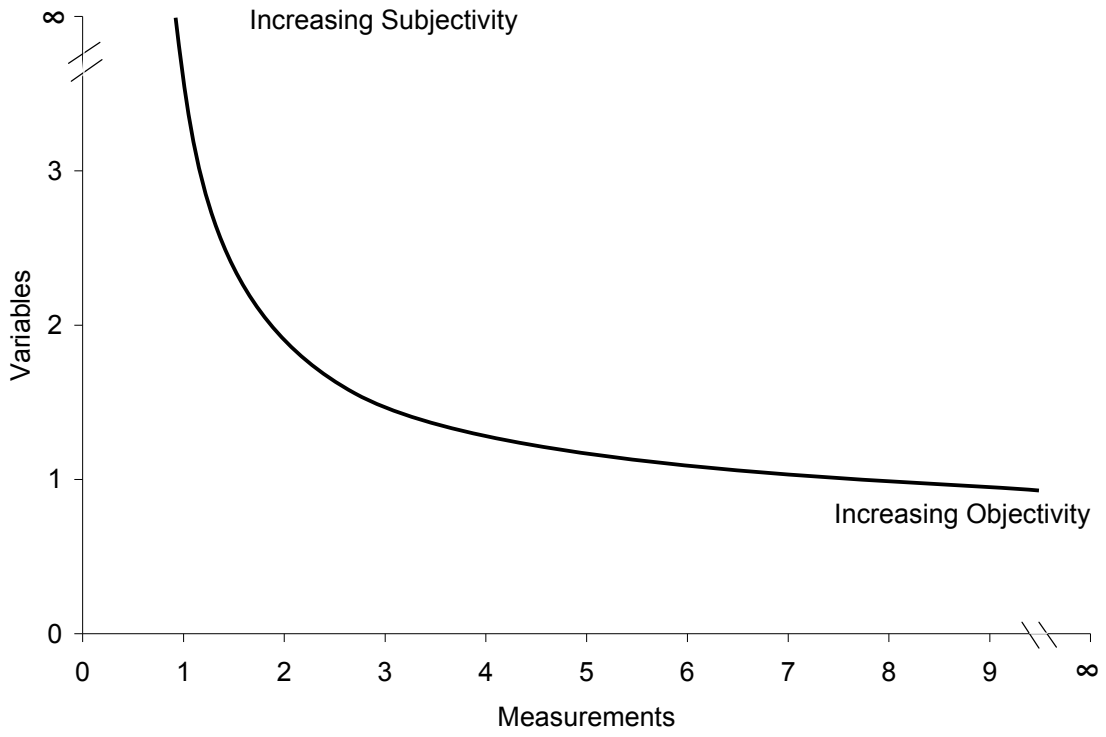


Figure 1.5 – An example of the subjectivity/objectivity of an assessment of a dataset

To illustrate further a single input variable (1 dimensional) dataset from a study of the height of 15 year old children. The objectivity would increase as the number of children measured was increased (Table 1.2). The number of measurements of the same variable is increased.

Number of children measured in a population	Objectivity of the mean assessment	Measure of dispersion
1	'Poor' estimate	None
100	'Good' estimate	'Fair' estimate of CI
1000	'Excellent' estimate	'Good' estimate of CI
Every 15 year old	'Truth'	'Accurate' estimate of CI

Table 1.2 – Repeated measures of a single variable (children's height)

A 2 dimensional dataset (2 input variables) in clinical practice is typically the measurement of an analyte vs. time. The traditional method of assessing 2 dimensional data in clinical practice is by manual inspection. For example, continuous glucose monitoring systems measure glucose concentrations at fixed time periods over 48-72 hours. A clinician will note the highs and lows but ignore the majority of the data. This analytical method is wholly observer dependent. An observer-independent method involves the assessment of the 2 dimensional data by the use of software programs developed for such measurements. It is typical practice with these programs, when analysing a continuous glucose profile, to perform data analysis on a single variable (the analyte) and ignore the other (time). The relationship between the two variables is not assessed, nor is the accuracy of time generally considered, though it may be critically important. The currently available software for analysing continuous glucose monitoring calculates the mean glucose, delineates the high values vs. the low values of glucose, and examines the dispersion of the data. These calculations are performed irrespective of the 2nd variable (time). To analyse both variables Time Series Analysis (TSA) techniques (further information in Chapter 2) have been developed that allow an objective assessment of both the input variables in two dimensional data. These techniques are more complex than standard one dimensional data analysis.

As the number of input variables (dimensions) increases objective assessment of data becomes increasingly difficult and specialised tools are required. For example, images have at least 3 variables for each pixel - the 3 components of colour (red, green, blue). The examination of images longitudinally introduces another dimension time. The traditional approach of image examination is opinion based (which is subjective and observer dependent) and not mathematical. The brain can perform complex cerebral integration of all the colours in an image to make an assessment of what is seen, but this

assessment will vary depending on the person. A method is required that can assess independently and objectively the images according to their input variables.

The number of input variables and the number of measurements of those input variables are related to the independence and objectivity of an assessment. It is this type of assessment that is the aim of this thesis. A mathematical calculation of the mean of a data set is observer independent as no matter who calculates the mean it should always be the same. As the number of related input dimensions increases and the number of measurements increases the complexity of the assessment increases. Methods are needed to analyse megabyte, multi-dimension, multi-measured data independently and objectively. The techniques available for the examination of physiological information systems are numerous. However, they depend to a certain extent on the hypothesis under test. To analyse certain conditions in vivo one has had to establish new methods of analysis or adapt/collate techniques from differing fields of science.

The use of Information system may thus aid diagnosis of complex pathologies in those with diabetes. All these aspects are further examined in this Volume I of the thesis. The detailed software can be found referenced in Volume II.

2.0 Time Series Analysis

‘Sometime a cigar is just a cigar’ - Sigmund Freud (attrib.)

2.1 Introduction to Time Series Analysis

During the last 25 years Time Series Analysis (TSA) has become one of the most widely used branches of Mathematical Statistics[15]. Its fields of application range from neurophysiology[30] to astrophysics[31] and it covers such well-known areas as economic forecasting, study of physiological data, control systems, signal processing, communications and vibrations engineering.

A time series is a collection of observations made sequentially in time [15]. The analyses of these observations occur in a variety of different ways dependent upon what aspects are considered to be of most interest. The time series that have been used here are of a discrete format (readings are taken at set intervals in time rather than continuously) and the datasets are equally spaced. TSA functions may be applied to a dataset in order to describe the main properties. They describe pulsatility, amplitude, phase, frequency, probabilities, and pattern reoccurrences. TSA can be used to explain how the variations in one profile may affect or be related to variations in another profile (e.g. the relationship between glucose and insulin whereby insulin concentrations are directly affected by glucose concentrations). TSA can be used as a method to understand basic mechanisms such as whether insulin oscillations differ between normal subjects and those with type 2 diabetes. Biological Time series data sets may contain few data points but are multi-variable as they contain a time variable and normally a substrate concentration variable. These megabyte data sets can be rapidly assessed by time series analysis; the techniques employed allow a evaluation of a profiles' properties. The properties that may be of relevance to an investigation are changes in hormonal oscillatory secretion or changes in the periodicity of these oscillations.

A time series comprises several key components that may require modification or extraction before the information can be analysed. These components are:

- Trends in the data profile
- Periodic signals in the data profile
- Irregular components of the data profile
- The signal of interest in the data profile

2.1.1 Trends in the data profile

A trend is a long-term change in the mean level or a smooth regular component having a period longer than the length of the time series. It is the general direction in which a variable changes (increasing or decreasing). Trends may be either deterministic and proceed in a fixed predictable fashion. For example, the age of a cohort in a trial will increase progressively, though, because of drop-outs will not be exactly deterministic. A further example at a physiological level would be the clearance rate of insulin from the bloodstream which has a measurable half-life. The clearance rate can be considered deterministic. It proceeds in a fixed and predictable way. Alternatively, trends can be stochastic and contain processes that have indeterminate or random characteristics. For example in examining hormone concentration the measurements observed will contain stochastic components from the assay used, additional processes occurring in vivo, and feedback mechanisms. Because of this, care must be taken when analysing such data.

2.1.2 Periodic signals in the data profile

Periodic signals are clearly marked cycles that recur at regular intervals. For example, the seasonal component (annual cycle) of weather and inter-annual phenomena such as El Niño are periodic signals. When periodicities approach the length of the time series, it becomes difficult to discriminate these from trends. If these periods are recognised they can be factored into the analysis. An example in meteorology would be the seasonal change in temperature obscuring an analysis of the differences between maximum and minimum temperatures on a daily basis (Figure 2.0).

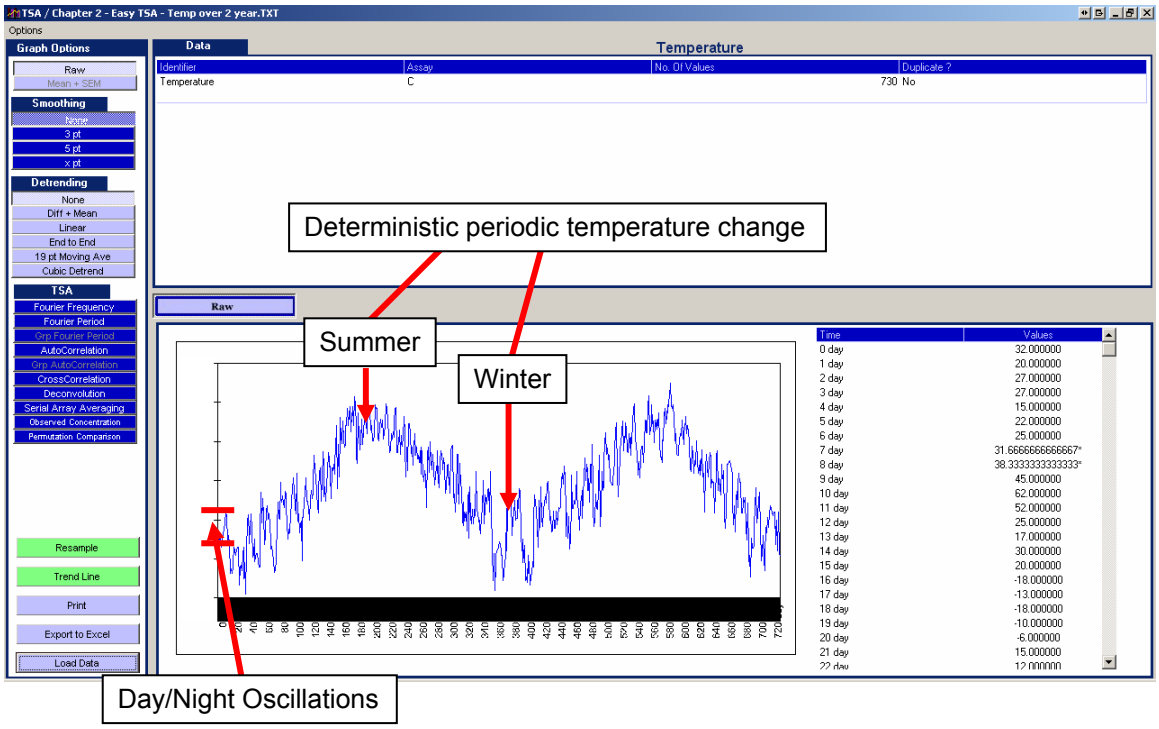


Figure 2.0 – Temperature oscillations over a 2 year period

2.1.3 Irregular components of the data profile

Physiological systems are generally complex and signalling is rarely unifactorial. There may be many inputs to a signal, and complex effects of that signal. Information may be obscured by noise or lost in a multiplicity of measurements. So noise reduction is a key feature of any time series analysis and is achieved by iterative computing processes and especially by identifying average changes and their periodicities. By removing phase information(information about the time at which oscillations begin) data from different sources can be pooled without loss of signal, but with reduction in the noise component.

2.1.4 The signal of interest in data

The underlying pattern is normally the signal that is being sought from the data. It is the component under investigation, the positive aspect. In a time variant data series the deviations from the mean may be entirely noise or trend, and there may be no information containing a periodic component, the negative aspect. In the investigation of megabyte data sets, it is beneficial to enlarge or enhance the positive aspect under consideration whilst eliminating or minimising the negative aspects of the data. This can be achieved by any or all of the following methods:

1. The signal being sought within defined limits – e.g. analysis of the start (4-20 mins) oscillations/pulsations in insulin secretion.
2. The data being measured at appropriate time intervals to extract the required information
3. The noise is reduced by assaying in duplicate
4. The data are smoothed, and outlier minimised by 3 or 5pt moving average (more in Chapter 2.2.2)

5. The data is stationarised (more in Chapter 2.2.2)
6. Time series analysis techniques are used that can extract the periodicity and amplitude of components (more in Chapter 2.2.3)
7. If N series of data are utilised to improve the signal to noise ratio

2.1.5 The use of Time Series Analysis in endocrinology

Wide spectrums of biological rhythms exist. The most evident environmental rhythm is the alternation between day and night. The photoperiodism responses of organisms are those many biological processes oscillating with a frequency similar to or identical with that of the solar day, i.e. a circadian rhythm.

The underlying aetiology of circadian rhythms remains unclear, but attributes of this process are understood, such as the influence of light/dark cycles intrinsic to sleep[32, 33]. The oscillations are not simple and may contain background interference (noise) from other physiological processes such as sub-harmonics of other pulsatile phenomena.

Since almost all endocrine systems utilise pulsatile hormone secretion rather than continuous release to signal end-organ receptor systems it is worth considering how such a system has evolved. There is growing evidence now to suggest that the pulsatile presentation of the information is a more efficient method of signalling[34-37]. Many hormones are secreted in pulses, for example insulin[38-40] glucagon[41, 42], growth hormone (GH)[43, 44], cortisol[45] and luteinising hormone (LH)[46].

Endocrinology is not unique in biology where a pulsed pattern of data presentation is suggested to be the most efficient method. For example, in the regulation of blood pressure by the baroreceptor reflex arc, a greater response results when the signal is pulsed than when it is applied in a continuous mode[47].

In endocrinology, the control of puberty and reproductive function by gonadotrophin releasing hormone (GnRH) is optimum when GnRH is administered in a pulsatile fashion whereas rapid down regulation of the gonadotroph may be achieved by continuous administration of gonadotrophin releasing hormone[48, 49]. This latter effect has been employed in the development of gonadotrophin releasing hormone analogues for treatment of patients with prostatic cancer. Different patterns of GnRH pulsatility regulate LH and follicle-stimulating hormone (FSH) secretion[50]. Insulin administered in a pulsatile mode has a greater hypoglycaemic effect than the same dose administered as a continuous infusion[51]. In the rodent, the secretory pattern of GH is an central determinant of body growth and the activator of a number of liver enzymes[52, 53, 54].

2.2 The development of the Easy TSA program

Existing TSA packages do not allow a wide enough interrogation of biological datasets. There was a need to examine the phase, amplitude, pulsatility and regularity of hormonal concentrations. The analysis of time series required a varied and specialised range of techniques and methods that were not present in any one package. It was decided specifically to build a software program that contained the necessary tools needed for the analysis of biological rhythms. A software package was developed for the

purpose of analysis, Easy TSA. The software accompanies the thesis on a CD and the code is available in Thesis Volume II – Chapter 1.

The aim was to integrate the judicially chosen methods and techniques of TSA into one comprehensive program that automated the processing of the mathematics. The program was thoroughly validated mathematically against a constructed dataset consisting of sine waves and artificial data. The program, following validation, was used as a research tool on disparate datasets concerned with the endocrine system and evaluated in a clinical environment.

The program was specifically tailored for clinicians and scientists and was tuned to accept very short time series. This was in contrast to many TSA analytical tools that require several thousand observations to allow robust analysis or operate on continuous data sets. Many hormones have specific oscillatory activity and Easy TSA may be used to provide both descriptive and predictive information.

2.2.1 An overview of the development of Easy TSA

The software was developed in a Visual Studio environment (Microsoft). It has eight distinct time series analysis functions, three smoothing techniques and five stationarisation methods. This amount of functionality enables a user to explore complex waveforms from insulin pulsatility to growth hormone levels. The methods incorporated into the Easy TSA program (Figure 2.1) were selected for their applicability to clinical research as the program has been developed specifically for clinicians and physiologists analysing time series data comprising relatively series of discrete observations from data-poor 50 point megabyte datasets to data-rich 5000 point megabyte datasets.

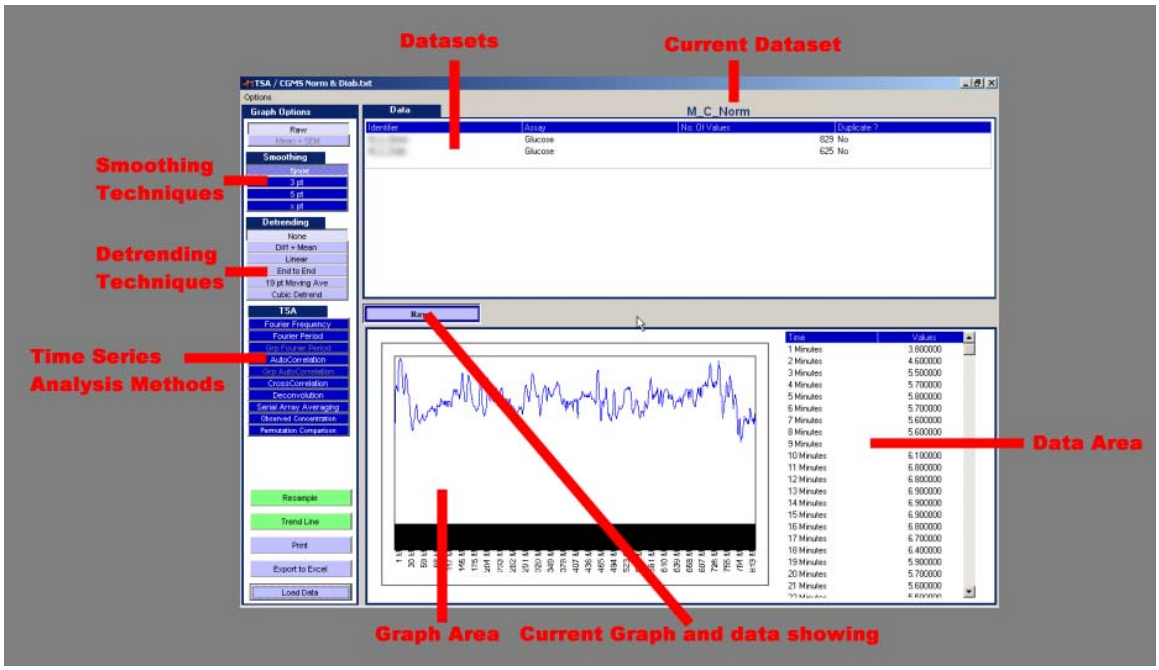


Figure 2.1 – A screen capture detailing an overview of the aspects in the Easy TSA Program.

2.2.2 Pre-processing functions in Easy TSA

The pre-processing of a dataset before analysis is a critical part of the investigative procedure. Biological data tends to be noisy and selecting the appropriate smoothing and detrending algorithm enables a robust analysis.

2.2.2.1 Moving averages and smoothing of data

It is necessary to prepare the data by compensating for spurious values and lessening the impact of outliers on the data analysis. When analysing hormonal pulsations it would be preferable to estimate using the precision associated with the detection of each pulse. In theory, the optimum method of achieving an estimation of the dataset would be to assay each point many times and plot both the mean and standard error. 'Noisy' or imprecise datasets would yield larger standard errors, and the standard error would become smaller as the sampling increased. It is not always practical to increase sampling, since when sampling increases both the costs and complexity involved in the testing increase.

Physiological data are regularly assayed in duplicate, but duplicate values do not always improve the estimation, if one is an outlier. There are a number of smoothing techniques applicable to TSA. For Example, exponential smoothing can be used to reduce random fluctuations in data and provide a more distinct view of the underlying behaviour of the series. It is often used to predict future values (forecasting). The most common technique used termed Auto-Regressive Moving Average (ARMA) was developed by Box and Jenkins[55] and is particular suited to the difficulty of outliers and noise when used as a pre-processing technique in Easy TSA. The aim of smoothing is to make a signal more distinct by reducing the effect of other fluctuations.

2.2.2.1.1 Autoregressive moving average

This technique is sometimes known as a 'moving average' and is used to improve precision. The ARMA process (EQ 2.1) is analogous to a mathematical integration of sampling (it averages the data depending on the values of its immediate neighbouring data points).

EQ 2.1

$$\frac{\sum_{-a}^{a} X_{(t-a)}}{m}$$
$$a = (m-1)/2$$

Where *X* is the dependent variable under observation (i.e. hormone concentration), *t* is time and *m* is an odd positive integer.

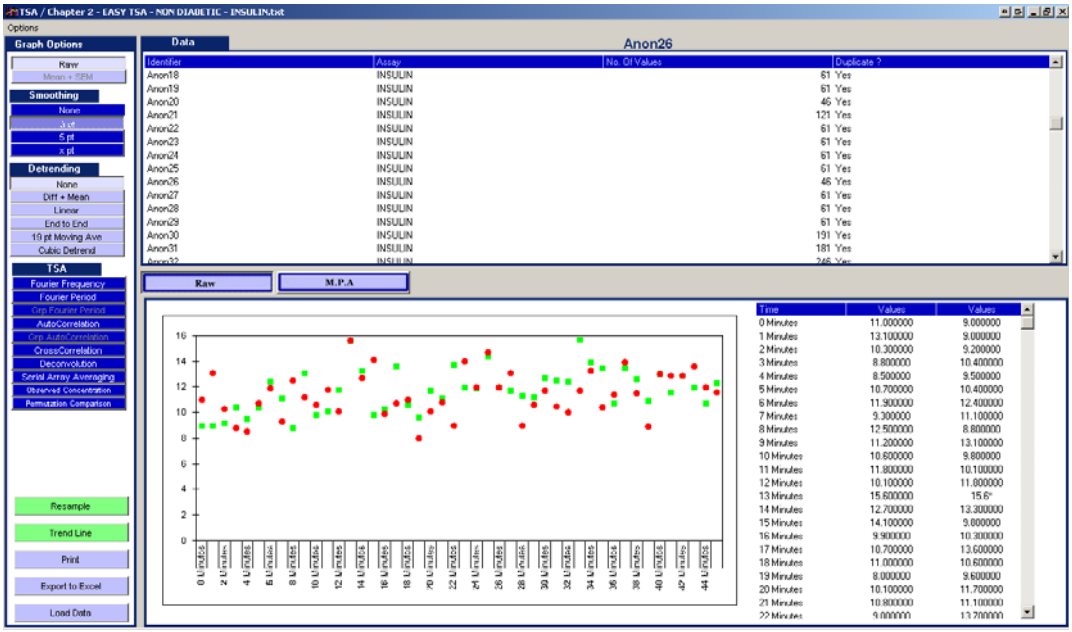
For example, consider the insulin profile from a person without diabetes (Figure 2.2a) where *X* was the insulin concentration measured and blood samples were taken every minute. To smooth these data a 3-point moving average may be used. Such that at the ten minute mark (t=10) the 3-point moving average equation (EQ 2.2) would average X_9, X_{10} and X_{11} where m = 3 and samples of blood were taken each minute (t=1) to assess an oscillatory phenomenon.

EQ 2.2

$$\frac{\sum_{-1}^{1} X_{(10-1)}}{3}$$
$$a = (3-1)/2$$

Each one-minute sampling point in the example may be regarded as a static point on the curve of the dataset and one may regard each three minute period as being measurable by using the assays from that sample together with the previous and subsequent samples. This would provide 3 data points to estimate each 1 minute period. If performed in duplicate 6 data points would estimate each 1 minute period (Figure 2.2).

a)



b)

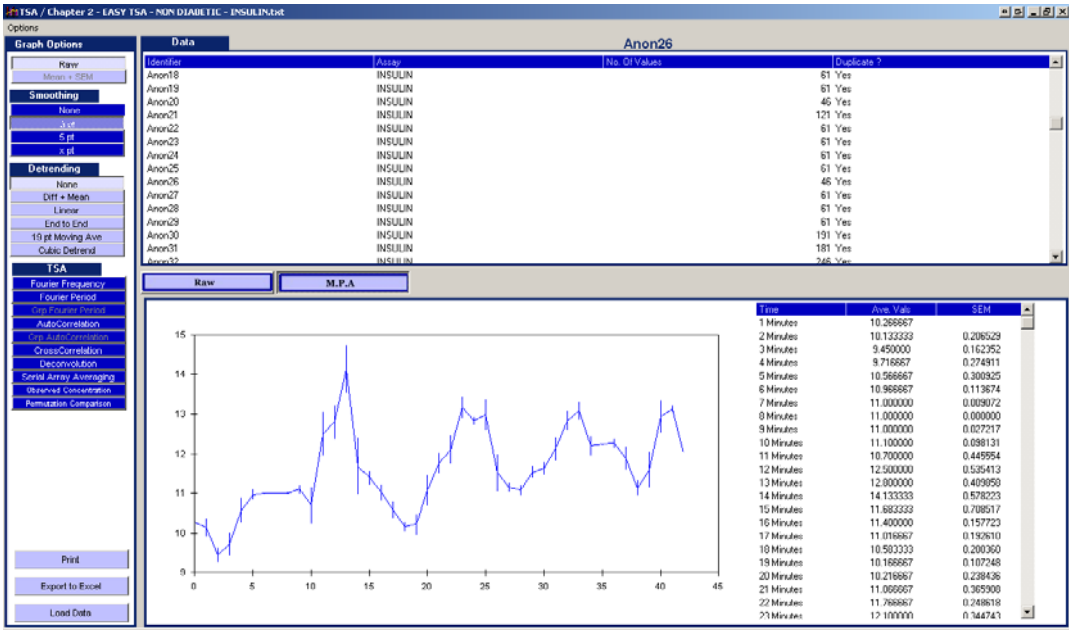


Figure 2.2 a) Insulin sampled in duplicate every minute b) 3-Point Moving Average of duplicate insulin data in a)

The 3-point moving average technique is not the same as taking samples every three minutes. 3-point ARMA estimates are generated for every data point in the dataset except the first and last data points. The first and last data points do not have neighbouring data points against which to average. If n = 'total number of samples taken'

then the new dataset length = n – 2. The resulting smoothed profile has a mean and a standard error derived from six assays for each 1-minute point.

The method achieves a result that is needed for signal detection, namely that one outlier does not cause an apparent pulse. The calculation of the standard error of the mean for each new data point allows a moving estimate of precision to be plotted on the data (Figure 2.2b). Noise is shown by large error bars.

Moving averages are one of the simplest forms of smoothing algorithms. They produce data that show less of a peak when one or two points deviate from the mean. However in a three point moving average when three or more data are available to demonstrate a peak the smoothing does not influence the assessed amplitude when analysed used a TSA method (For example, Fourier Transform) Equal sample intervals are important for TSA. Clinical tests, such as the intravenous glucose tolerance test, have often used short sample intervals initially followed by longer sample intervals. This gives variable precision at different times and is only analysable by graphs or by mathematical methods of interpolation.

Smoothing using larger intervals (e.g. 5, 7 or 9-point averages) eliminate high frequencies and produce data more indicative of trends and have been used for prediction[56] by other groups. Autoregressive moving average smoothing is a particular instance of the averaging method it is possible to perform smoothing with an even number of points (2, 4 etc) but would then refer to the mid point of the central two times. With data assayed in duplicate a 3 point moving average would encompass 6 data points so that 6 readings would contribute to each single value used. Equal time intervals for sampling is a pre-requisite for the all methods used in Easy TSA.

Accuracy of a dataset is not a necessity for pulse detection, but is relevant when comparisons between patients as a group are being made. The precision of a dataset as well as the number of samples taken are critical for pulse detection. Weak amplitude signals may be detected with high-precision assays and/or duplicate assays at given sample-points. One might compromise on assay repeats if small-amplitude cycles were regarded as not relevant to the hypothesis under test. To apply smoothing algorithms effectively it is advantageous that they be applied with this in mind.

Further work is planned on the development of a guidance wizard for the users of Easy TSA to aid in the selection of the appropriate smoothing technique depending on the data and the information sought.

2.2.2.2 Trends and stationarisation of data sets

The Easy TSA program incorporates another set of pre-processing techniques and these are designed for removing trends in datasets. Trends are defined as present in a time series when there is a 'long-term change in the mean'[15] and it may be advantageous to remove this trend. e.g. In the UK the temperature is generally higher in summer and lower in the winter but this oscillation in temperature may occur on a background trend of systematic global warming (see previous Figure 2.0). A trend need not be a function of a season but may exhibit variation at a fixed period due to other physical cause. Irregular fluctuations are those that remain after trend and cyclic changes are removed.

A time series is said to be stationary if there is no change with time in the mean and variance. Easy TSA uses methods that describe the data in terms of a trend. A decision has to be made whether the trend is of interest. The trend may be the information from the data that is being sought, or it may be more desirable to remove the trend in order to analyse the shorter fluctuations. For example, again using the global warming example, one may want to establish that the trend is increasing or remove this aspect. Easy TSA includes functions that can remove this trend and effectively stationarise the data. Stationarisation techniques for achieving this are alternatively known as detrending[57].

It is possible to determine visually that the insulin values are increasing over time. Insulin concentrations start lower and end higher. To analyse the short term pulsatility of the secretion rates it is necessary to remove this upward trend. A typical resulting dataset will begin and end at approximately the same value after a detrending has been performed. The oscillations of the data are preserved.

An approach such as detrending is particularly applicable for hormone analysis where, because of limitations in long-term sampling, the effects of long-term trends cannot be excluded. The presence of trends may interfere with the effectiveness of TSA because they may dominate the other features. What defines 'long-term' is a matter for debate. If an oscillation occurred every 50 hours and the data available was for 20 hours the long-term cycle would appear as a trend. If several hundred hours' data were available, the long-term oscillation may be extracted.

Granger[58] has suggested that a trend comprises the totality of frequency components whose wavelengths exceed the length of the observed time series. For example if observations of 12 hour periods in GH rhythms were recorded it would mean that stationarisation of the data was essential as GH has a 24 hour diurnal rhythm that would appear as a trend in a 12 hour dataset[44].

If datasets are not correctly stationarised before time-series analysis then the results may reflect predominantly (or entirely) the trend. Trends may be of interest to the investigator, but the data may not have been collected to analyse this attribute. Stationarisation is therefore an absolute requirement in TSA. The selection of the correct method to stationarise needs to be related to the nature of the data being analysed. The methods and the techniques included in Easy TSA were:

- Linear
- End to End
- Difference + Mean
- 19 Point ARMA
- Cubic Polynomial

The choice of stationarisation will affect the analysis in terms of its discrimination of certain attributes. There is no uniquely correct stationarisation to apply to all datasets.

2.2.2.2.1 Linear stationarisation

Linear stationarisation[59] is a commonly used stationarisation technique. The detrending is achieved by calculating a regression line through the data and then subtracting the slope from the dataset. The equation is:

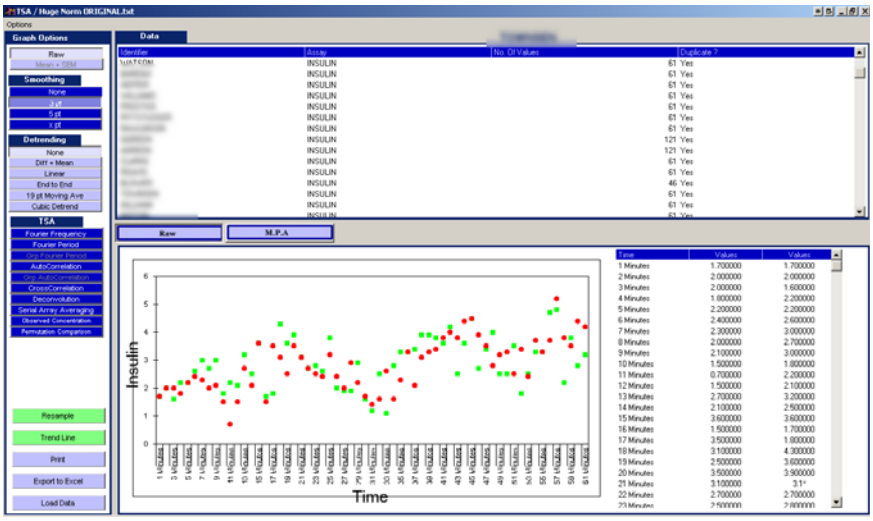
EQ 2.3

$$B = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sum X^2 - \frac{\left(\sum X\right)^2}{N}}$$

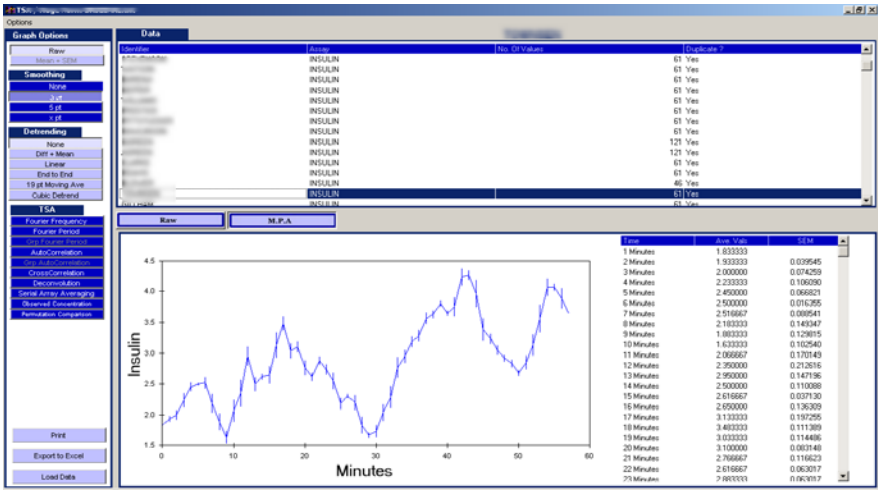
Where *B* is the slope of the line, *X* is time, *Y* is the observed value and *N* is number of data points.

The slope of the line is then used to stationarise each point (Y) by EQ 2.3 ($y_{stat} = y_{observed} - B_x + c$). Linear stationarisation is most appropriately applied to data with a predominant and single trend. The application of linear detrending to an insulin dataset (Figure 2.4a) where Y_{stat} = 'stationarised insulin concentration', Y_{obs} = 'observed insulin concentration', *X* = 'time of measurement', c is the intercept, N is the total number of observations and *B* is the trend(slope). The result is a dataset with the slope (trend) *B* removed (Figure 2.4b).

a)



b)



c)

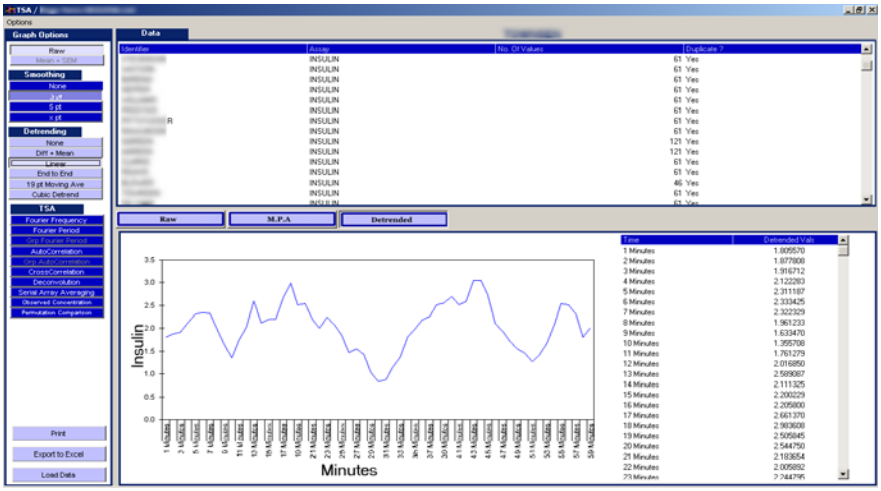


Figure 2.4 a) Insulin sampled in duplicate every minute. b) Three point moving average of duplicate data) c) Graph of insulin concentrations after the 3-point ARMA (graph b) and linear detrending has been applied

2.2.2.2.2 End-to-End stationarisation

The investigator makes an assumption that deviations are 'from' and 'back to' a baseline and the data start at a baseline and finish on it, the data may have to be truncated to achieve this. End-to-End detrending (EQ 2.4) will plot a line of best fit between the first and last data points in the profile and calculate the slope (*B*). The slope is then used to stationarise each data point in the profile (Figure 2.5b).

EQ 2.4

$$B = \frac{X_{Last} - X_{first}}{N}$$
$$X' = X - (B * t)$$

Where *B* = the slope, X_{Last} = last data point, X_{first} = first data point, N = total number of data points. X' is the new data point, X = current data point, t = current period.

For example, insulin oscillations are pulsatile and End to End detrending can be used to analyse a profile with multiple peaks, by splitting the data set into ‘pulses’ (Figure 2.5a,b). The resulting detrended profiles can then be aggregated (Figure 2.5c,d). The aggregated profile could then be compared to an End – End detrended and aggregated Insulin profile to assess amplitude.

a)

b)

c)

d)

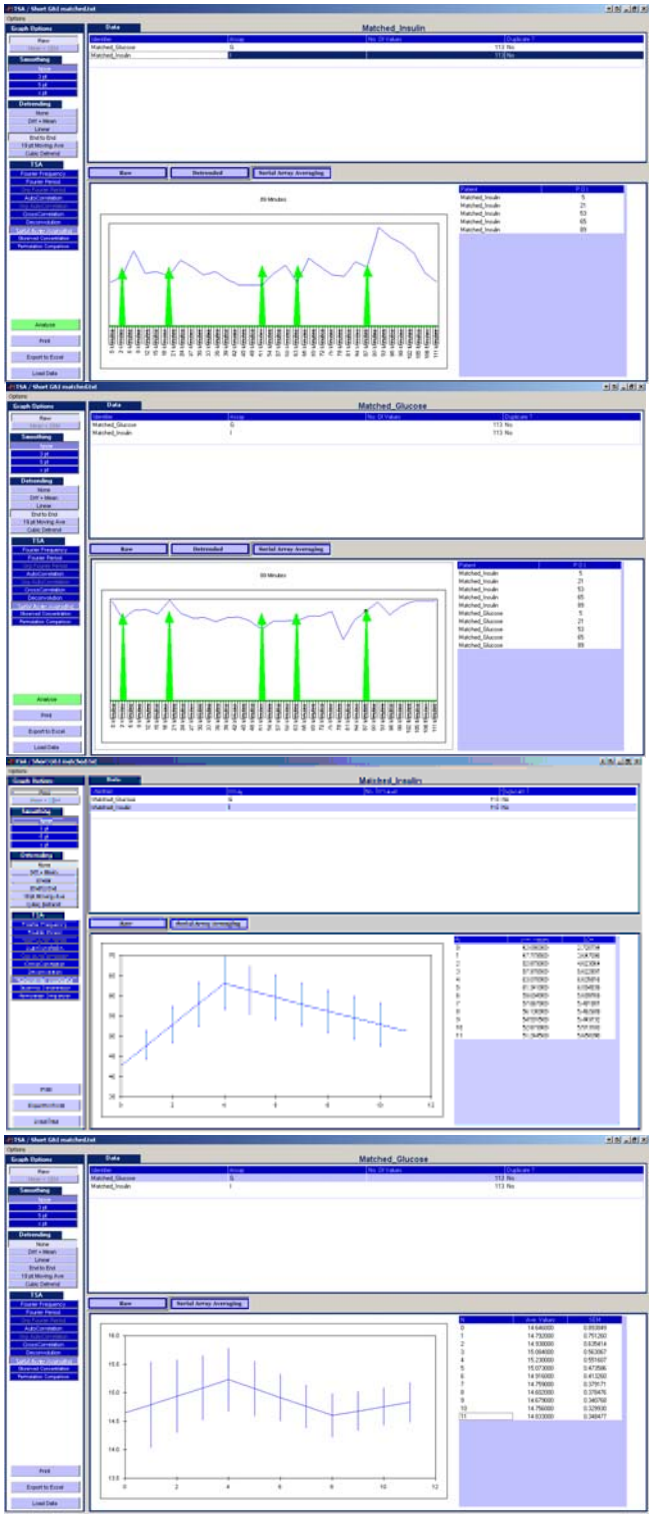


Figure 2.5 – Glucose and insulin matched sampling profile a) Insulin profile with the nadirs before the peaks of interest selected (arrows indicate selected peaks of interest in insulin profile) b) glucose profile with the same peaks of interest superimposed c) Mean + SEM of insulin peaks selected after End to End stationarisation d) Mean + SEM of glucose peaks selected after End to End stationarisation

This form of stationarisation avoids the occasional anomalies of linear stationarisation when hormone concentrations are measured as negative.

2.2.2.2.3 Difference + Mean stationarisation

Differencing + Mean detrending involves the analysis of the sequential difference between data points in an array, the current insulin value – the previous insulin value (x_n - x_{n-1}). This is the differential (or first derivative) of the data [57]. Such stationarisation removes, for instance, complex underlying trends. Since the data are now in the form (change per unit time) the $\frac{\Delta Y}{\Delta X}$ derivative now has no component of the original mean amplitude of the signal, so is no longer possible to determine the concentration of the insulin. The difference of each point is then be added back to the mean for the total dataset to avoid a negative domain and to maintain an absolute power of the signal which reflects the concentration. For example, Two signals with identical periodicity, but different mean values would give identical signal attributes when examined for their first derivative only, but when analysed as Difference + Mean would yield results reflecting the oscillatory signal power adjusted by their mean. As with all methods used in EASY TSA the difference + mean stationarisation assumes that the sampling interval between observations is equal.

This technique allows analysis of data in terms of the periodicity. The effect of a Difference + Mean detrending on the sample Insulin data is to remove the trend from the data. The resulting data set (Figure 2.6) oscillates around the mean concentration of insulin and is indicative of the change over time of the insulin concentrations.

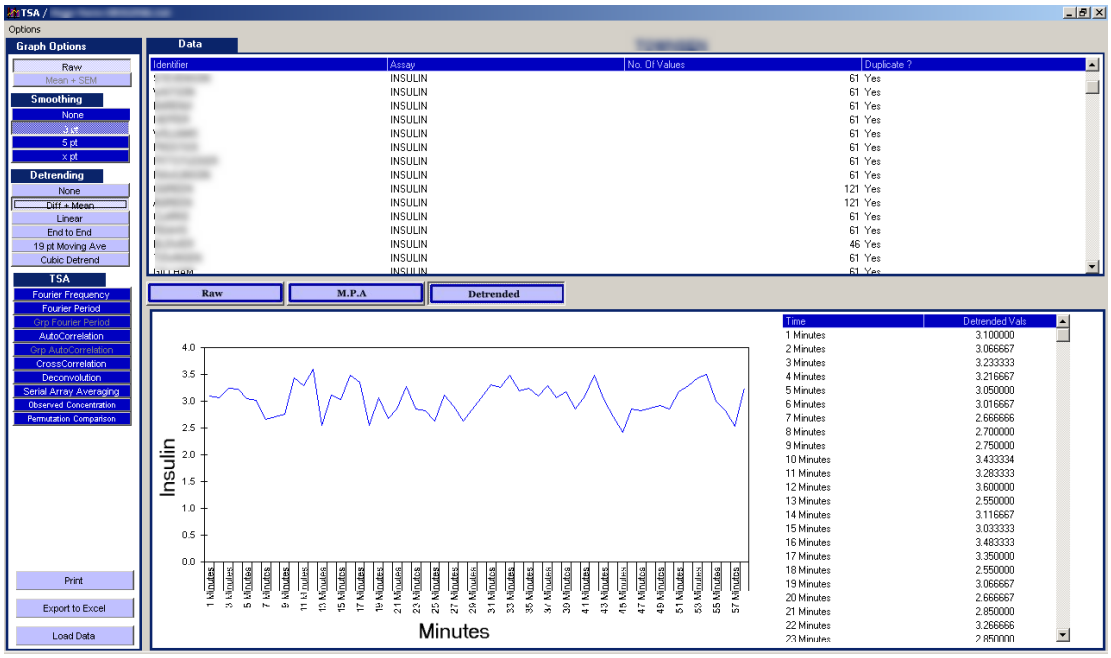


Figure 2.6 - Graph after a 3-point ARMA and Difference+ Mean detrending has been applied to the insulin dataset (previously used in Figure 2.2)

2.2.2.2.4 19-Point ARMA stationarisation

The 19-point moving average works on the same concept as the smaller moving averages that are used for smoothing (Section 2.4.1). As 19-points of information are contributing to each point (38 points if the data are in duplicate) a trend may be removed either by the averaging effect or future points predicted[56]. Application of a 19-point moving average (Figure 2.7) removes the detail from the dataset and the signal of interest may be inadvertently lost, but the data is stationarised. It is preferable to apply a 19-point moving average to long datasets as this technique will removed nine data items from the beginning and end of the profile thus truncating the data by 18. 19-point moving average is an arbitrary selection. But is the one most commonly selected for use in TSA. It is possible to select a different number of data points to stationarise using Easy TSA.

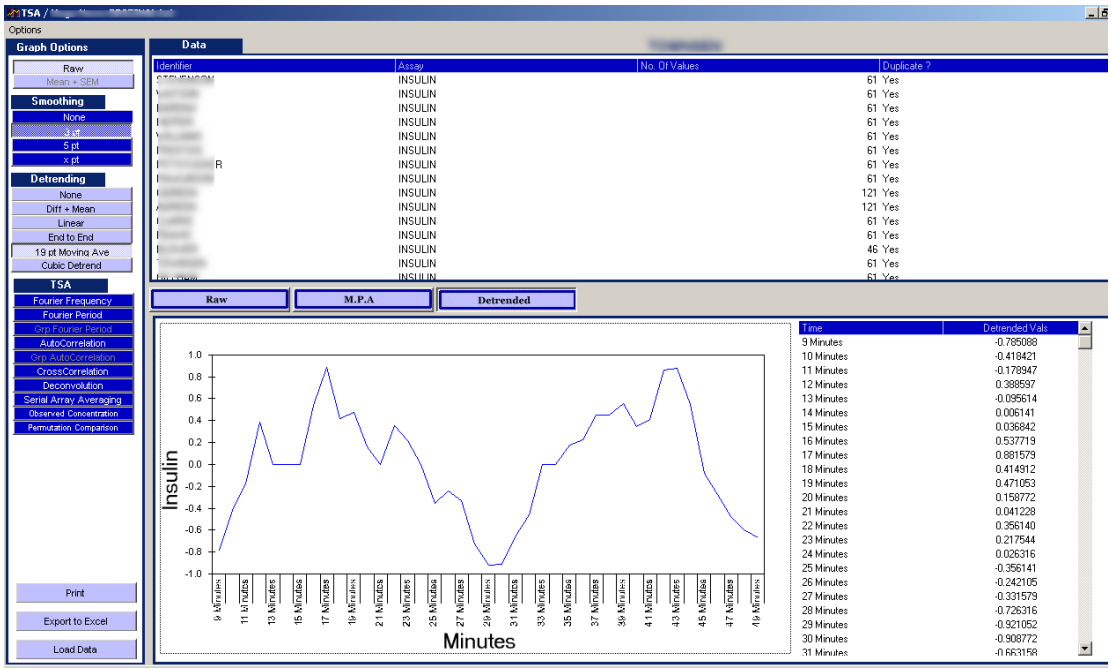


Figure 2.7 Graph after a nineteen point moving average detrending has been applied to an insulin dataset (previously used in Figure 2.2)

2.2.2.2.5 Cubic Polynomial stationarisation

If there is a curvilinear trend in an outcome measure when it is plotted against an independent variable then Cubic Polynomial detrending may be applied, for example insulin concentrations over time. Third order polynomials (EQ 2.5) are applied using the same method as linear stationarisation. The line of best fit is calculated using the least squares[60] method as in linear regression. The stationarisation is achieved by calculating a 3rd order polynomial line of best fit through the data and then subtracting the slope from the dataset[61]

EQ2.5 $$y = b_0 + b_1 x + b_2 x^2 + b_3 x^3$$

The Insulin profile that has been used as an example for the pre-processing techniques is ideally suited to the cubic detrending. It can be determined visually that the trend is not a straight line from high values to low value but is curvilinear in nature. The cubic

detrended data (Figure 2.8) has removed the trend and TSA techniques can now be applied. Cubic detrending is an arbitrary selection of detrending it is feasible to use quartic, quintic, etc. However care must be taken as polynomial detrending may remove the oscillation of interest if the cycle of the signal is equal to the polynomial used.

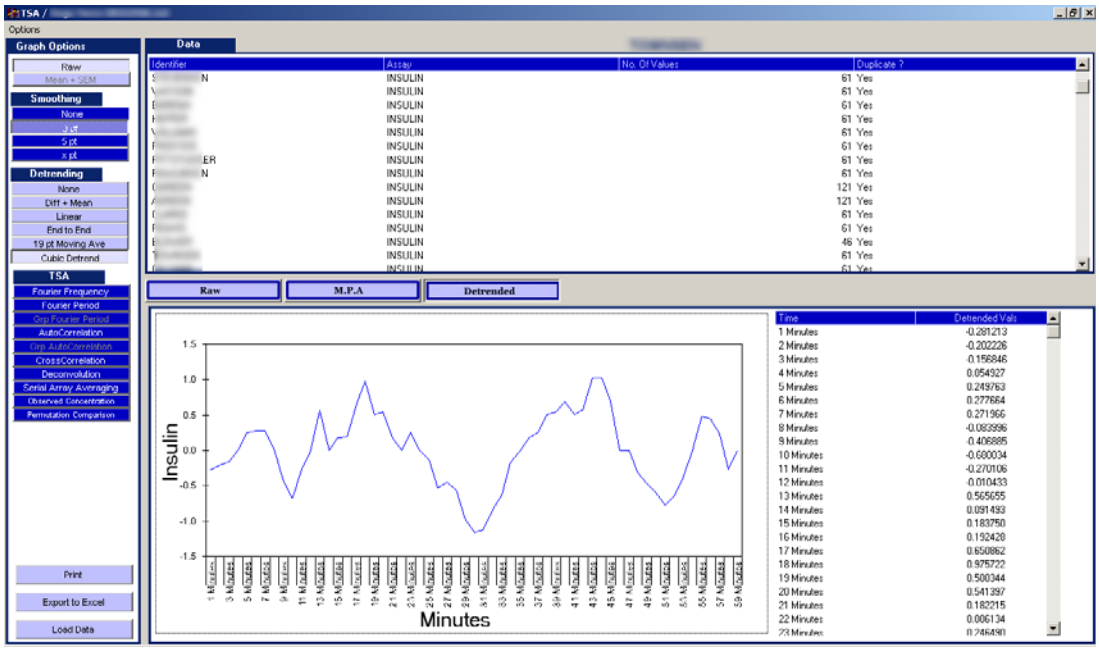


Figure 2.8 Graph of insulin concentrations after 3-point ARMA and Cubic Polynomial detrending has been applied to the dataset (previously used in Figure 2.4b)

2.2.3 Time Series Analysis Techniques

Eight TSA functions have been included in the Easy TSA program. They are:

- Fourier Transform
- Auto-correlation
- Cross-correlation
- Deconvolution
- Probit Analysis
- Serial Array Averaging
- Runs
- Permutation Analysis

The techniques included in Easy TSA were selected for their applicability to clinical research. The TSA methods included enables an examination of multiple profiles to establish common features, an assessment of change when therapies or interventions have been applied; and an examination of dominant features. For example, what is the major oscillation? Are the oscillations regular? Do the oscillations have phasic relationships?

2.2.3.1 Fourier Transforms

The Fourier Transform method of examining the frequency domain has been briefly explored in the introduction. It is now examined in more detail. Fourier Transform provides a breakdown of the data into time-series descriptive components which will describe all the oscillatory activity in terms of its frequency and amplitude. It represents oscillations not as a waveform but as a spectrum of power(i.e. how much of each part of the signal is present)

In 1798 a French Mathematician, Joseph Fourier, was working in Cairo as part of the French civil service. Whilst stationed there he developed the concept of spectral analysis [62]. Joseph Fourier was interested mainly in heat oscillations, and wanted to detect 'hidden' oscillations in the fluctuations that he observed. His mathematical solution represented oscillations not as a waveform but as a spectrum of power(i.e. how much of each part of the signal is present). Joseph Fourier demonstrated that a continuous signal, however complicated, could be represented by a power spectrum and that an apparent random signal may be of several overlapping waveforms[15]. These are known as Fourier Transforms (FT). The default output of a FT is the absolute power present at each frequency. To compare FT's a relative (normalised) FT can be used. A relative FT

displays the power at each frequency as a function of the overall power in the sample. This enables a percentage out of 100% to be generated and the FT's to be compared.

FT enables an analysis of all frequencies within an array. FTs resolve data into a series of sin and cosin components which represent frequencies in the data. The technique is often referred to as spectral analysis. Each frequency has an amplitude attribute which shows the power of that frequency within the data array. Displayed as a histogram such a construct is termed a power spectrum. The spectral peak at a particular frequency or period indicated a greater than random tendency for peaks and troughs in the time series to recur at that frequency or period and may indicate a rhythm.

When a complicated signal is submitted to FT, it yields data that can be displayed as a power spectrum for each component or as a range of frequencies. FTs deconvolute the data to power/amplitude (irrespective of phase) that can be displayed as the amount of how much of the signal is present at each discrete frequency. For example, the FT of a sine wave with eight distinct peaks (Figure 2.9a) establishes the presence of a signal with a frequency of 8 (Figure 2.9b) and an oscillatory period of 6.25 that contains 78% of the total signal power (Figure 2.9c). 78% of the signal represented by one frequency is 'very high' but there is no statistical test available to state this. However, the FT waveform can be contrasted to another FT using a statistical result obtained by comparison of the SEM overlaps. That is the probability of the standard error of the mean (SEM), at any discrete time, encompassing the SEM of the comparative group. If the SEM's do not overlap the probability of these being the same is $p<0.0256$.

a)

b)

c)

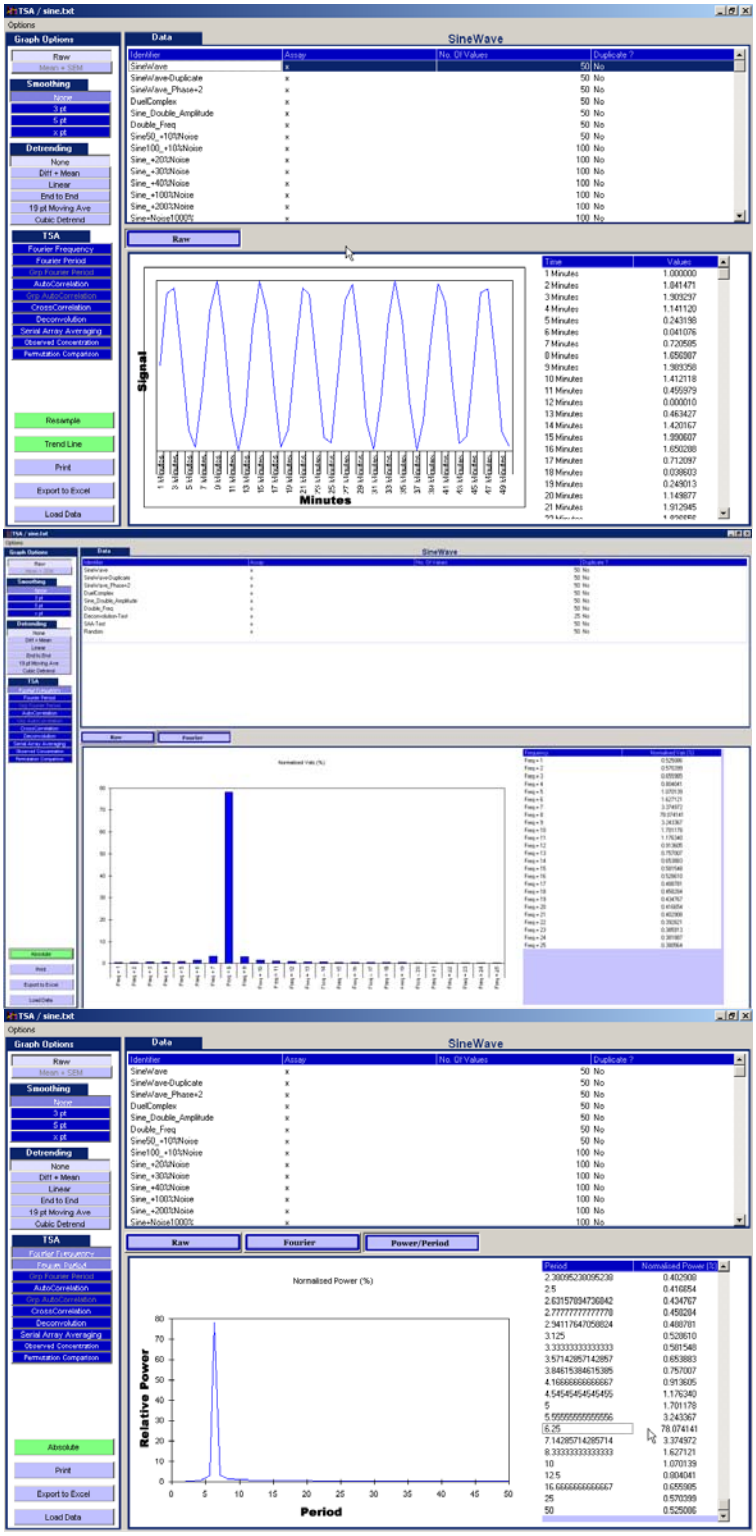


Figure 2.9 – a) 50-point sine wave , b) FT determines maximal power at a frequency of 8, c) A Relative FT power spectrum demonstrating the peaks have a periodicity of 6.25 minutes

FT's can be applied to biological analysis where datasets are 'noisy'. For example, a typical Growth Hormone (GH) profile from a rat model (Figure 2.10a) may contain noise (because it is a biological assay). The profile may contain outliers (as the hormonal kinetics may vary) and the frequency of measurement may not always be constant. This profile once pre-processed using a 3pt moving average (Figure 2.10b) and Cubic detrending (Figure 2.10c) can be submitted to an FT. The FT calculates the power and frequency spectrum (Figure 2.10d) resulting in the identification of a signal in the data with a periodicity of 21 minutes.

a)

b)

c)

d)

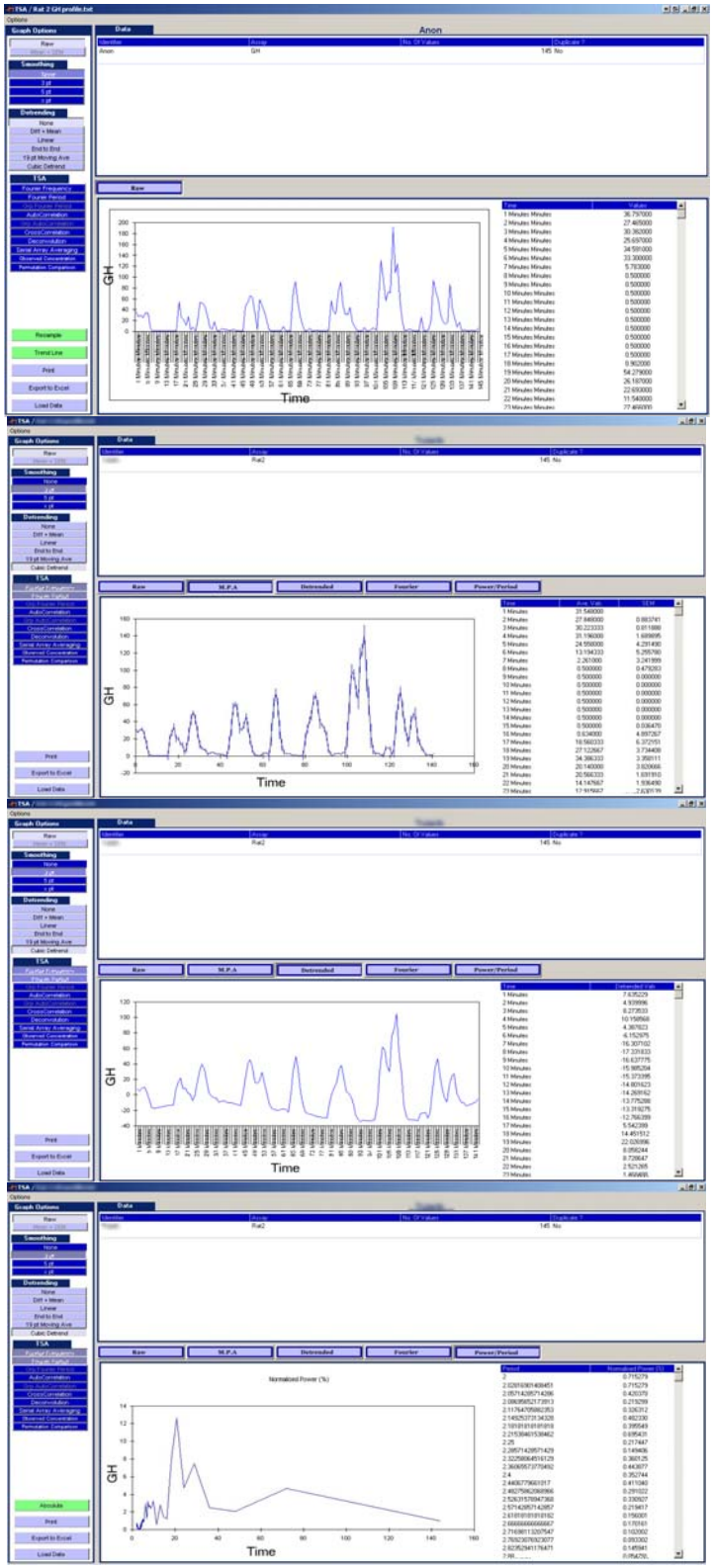


Figure 2.10 – a) Growth Hormone (GH) data from a rat model b) 3pt moving average of the GH data c) 3pt moving average and cubic detrending of the GH data d) Fourier Transform of the pre-processed GH data

TSA and FTs in particular take ‘noisy’ data (noisy for many reasons) and attempt to find relevant characteristics of the data.

FTs can be pooled from the data by calculating the normalised FT’s for each subject and taking the mean. It is important to note that the mean FT may lose precision and harmonics that occur in a single profile may be averaged out.

Another specific example is insulin oscillations. The data series may be noisy, but taking many profiles and using FT allows one to generalise by pooling oscillatory data (Figure 2.11a, b, c) applying a FT (Figure 2.11d, e, f) and determine the mean periodicity In the result (Figure 2.11g) oscillations generally occur at about 12-14 minutes periodicity in these non-diabetic people.

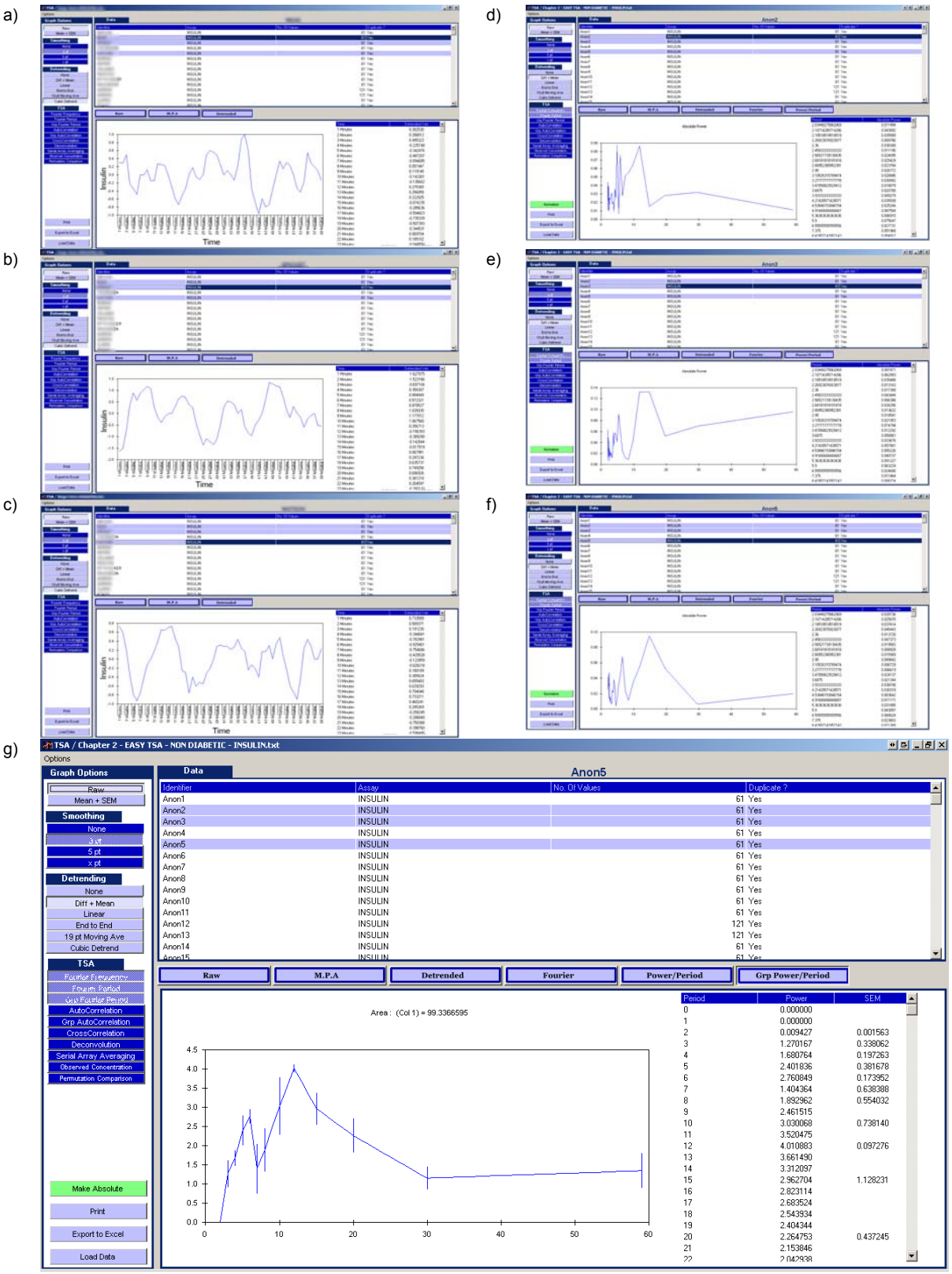


Figure 2.11 a, b, c) Pre-processed (3pt MPA, Diff + Mean Detrended) insulin profiles from people without diabetes. d, e, f) Individual Fourier Transforms for each insulin profile. g) Grouped Fourier transforms for the insulin profiles

As previously stated Easy TSA can produce two types of FTs: absolute and relative. Absolute units are informative of total power (e.g. for monitoring GH concentrations through puberty it is necessary to know if the total concentrations are increasing) while relative (or normalised) FTs are used for comparing oscillatory activity independent of mean concentration, and can be used for comparison between FTs.

The FT of a time series can be calculated using EQ2.6:

EQ2.6

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, \Delta t$$

Where $X(f)$ is the FT of $x(t)$. Frequency (f) is measured in Hertz and t is time. J = 0, …, count of time series-1.The code used to calculate this is in Chapter 1 of the code thesis that accompanies this thesis

2.2.3.2 Autocorrelation

Autocorrelation is a mathematical tool that is used in signal processing for analysing functions or series of values, such as hormonal signals over time. It is the correlation of a signal auto-regressed with itself. Autocorrelation may be implemented as an iterative technique for establishing whether there are regularly recurring waveforms within a data array[15, 63], such as determining the presence of a periodic signal which has been obscured by noise. The result is independent of the mean amplitude of such waves or pulses, but it is affected by pulse-to-pulse variations in amplitude. It is independent of phase (i.e. whether the data begin with a peak or trough) and yields an unbiased estimate of the period and an assessment of its significance. Stationarisation is

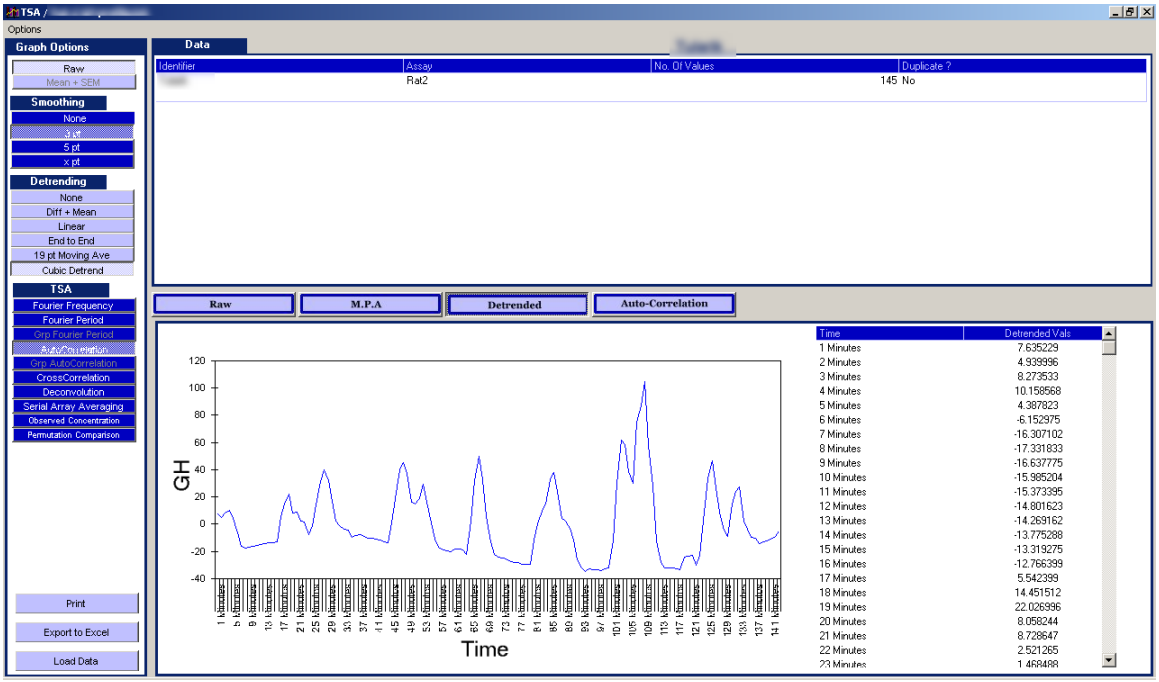
necessary before autocorrelation in order to remove long-term trends that may obscure the oscillatory data and the use of a smoothing algorithm usually produces less noisy results.

For example, analysing a GH profile (Figure 2.12a) when perfectly aligned (Lag = 0) the correlation coefficient (r) will be 1. The correlation coefficient is recalculated as the lag is increased and the new r valued plotted. As the oscillatory datasets move out of phase (where peaks coincide with troughs), r tends to -1. As the peaks in the profile approximately coincide with the peaks in the second profile, the r-value becomes positive. The first significant negative r-value is the ½ phase periodicity (i.e. the peaks coincide with troughs) and the first significant positive r-value represents the periodicity of the dataset. For a regularly occurring waveform, this peak in the r-value will occur again at twice the lag, three times the lag and so on. Noise may obscure the periodicity as the lag increases. When the example is autocorrelated (Figure 2.12b), a peak periodicity of 20 is detected that exceeds the region of significance (delineated by the dotted line).

Biological data will not attain a perfect correlation r=1 (except at Lag 0), but r-values less than 1 may be significant. The r-value is determined to be significant if r exceeds $1.645/\sqrt{N}$ where N='number of values correlated'. The equation $1.645/\sqrt{N}$ is equal to 90% Confidence Interval of the data[60]. Thus data less than $-1.645/\sqrt{N}$ and data greater than $1.645/\sqrt{N}$ have a probability of <=5% or P<= 0.05 of occurring. In the GH profile example the region of significance is delineated by a dotted line on the correlogram (Figure 2.12b) hence the periodicity of 20 is judged significant for this profile. The dotted line has a positive slope as N is decreasing as the Lag increases, the

number of data points being compared is decreasing as the profiles are compared increasingly out of phase.

a)



b)

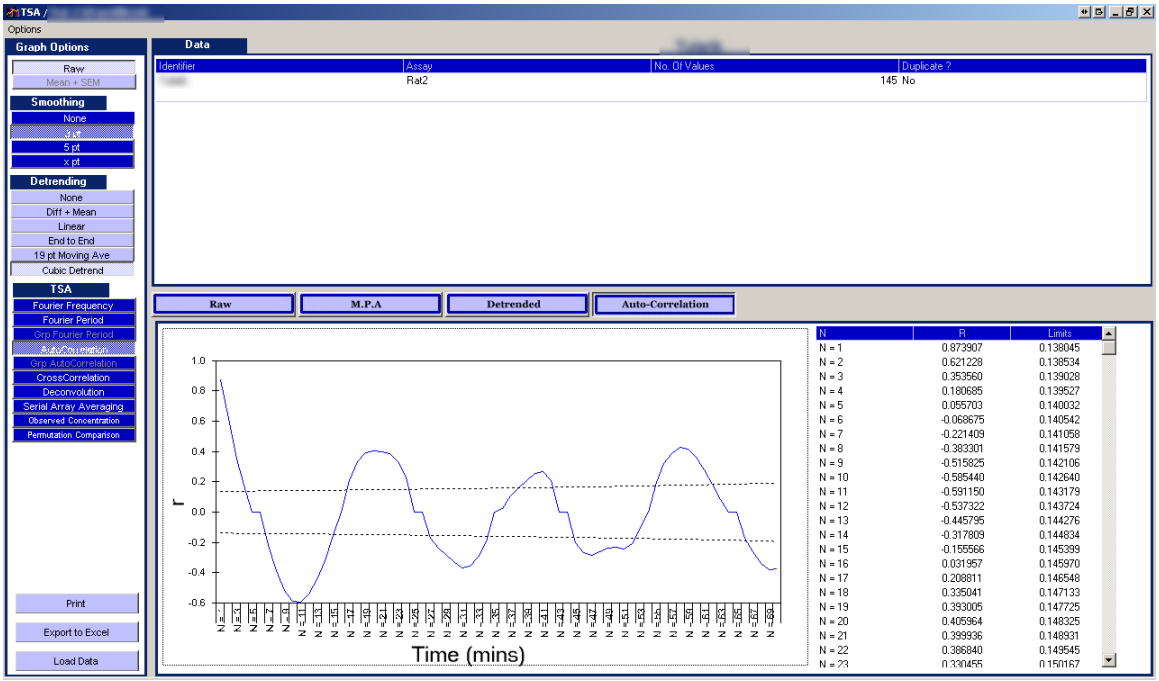


Figure 2.12 a) GH concentration measured every 30 minutes b) Autocorrelation of a) showing a significant r correlation coefficient at a lag of 20.

To obtain the mean of a group of autocorrelation correlograms Fishers z transform [15] can be used to aggregate, average and back transform the correlograms of different profiles. The z transform is used as r values are not normally distributed because they range from -1 to 1 and cannot be compared using parametric statistics.

2.2.3.3 Cross-correlation

The comparison of two signals' phase relationship can be analysed to allow an assessment of biological interactions[64]. Lag in phase may demonstrate an endocrinological hierarchy and show its temporal relationship. However, one needs to be careful about the interpretation as is mentioned below. A relationship may be determined by the cross-correlation of a glucose profile and insulin profile. For this example, a data set was used of length of daylight hours (Figure 2.13a) and temperature (Figure 2.13b) over 2 years. As the daylight hours decreased the temperature dropped. A significant correlation coefficient could be seen (Figure 2.13c).

a)

b)

c)

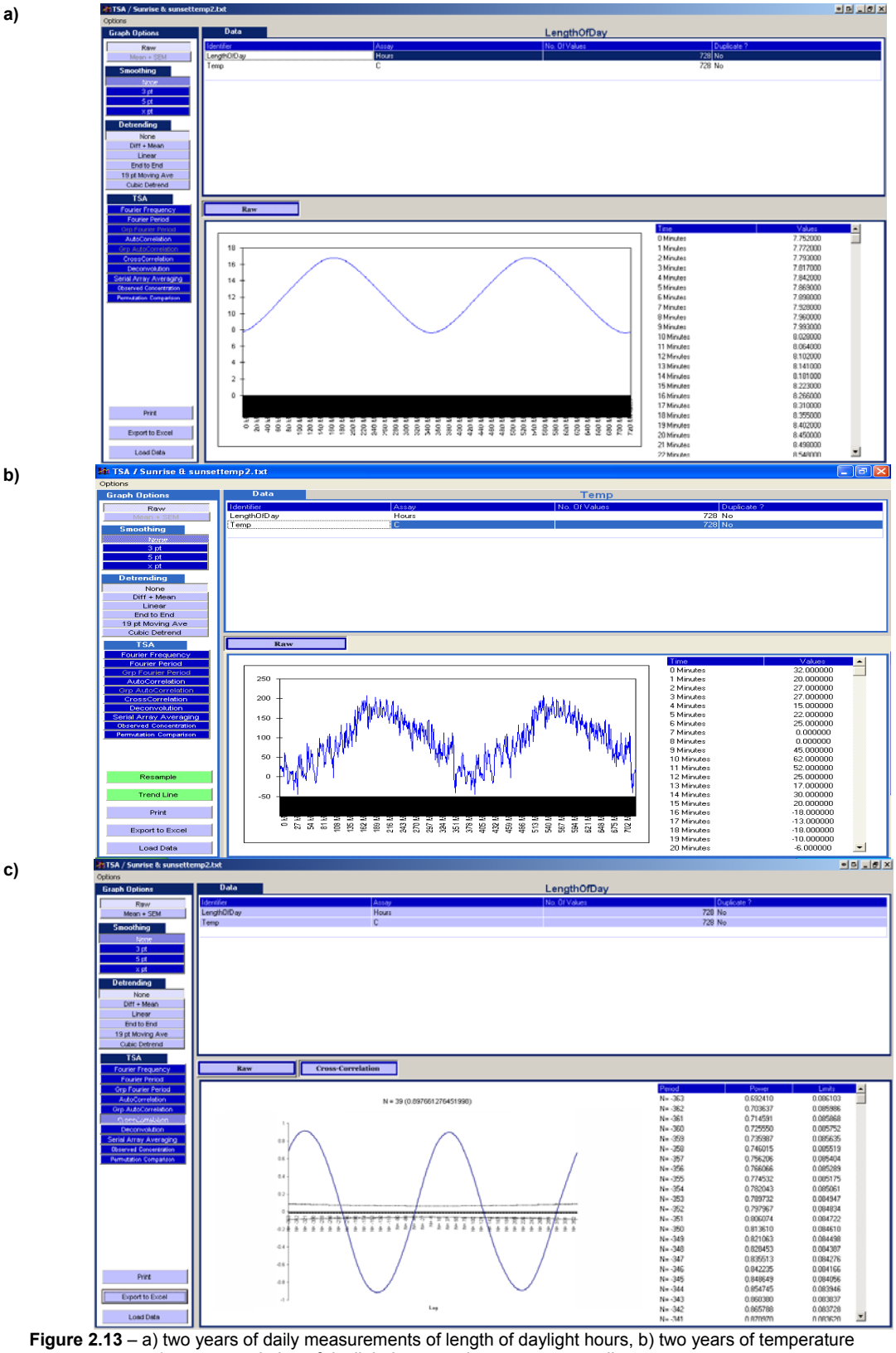


Figure 2.13 – a) two years of daily measurements of length of daylight hours, b) two years of temperature measurements c) cross-correlation of daylight hours and temperature readings

The relationships of interest are ones where the phases of the datasets are the same or they rise and fall at the same rate, or when one is the inverse of the other. In the example results, (Figure 2.13c) r becomes significant at a periodicity of 40. This means that the coldest day occurs 40 days after the longest night. If there were subsequent peaks in the data, they may be multipliers of the first peak. This multiplier can give improved resolution of the actual frequency – for an example see the 'Orbit of the Earth' data example in section 2.3.2

The cross-correlation correlogram (CCC) is plotted in both the positive and negative domain. The technique used for the comparison of two signals is otherwise identical to the autocorrelation that correlates a profile with itself whereas a CC correlates two different profiles. As with the autocorrelation correlograms it is possible to group CCC together using Fishers Z-Transform[15] (aggregate, average and back transform) to generate a mean CCC from a large dataset.

Comprehensive pre-processing of the datasets must be performed. If both datasets contain a similar trend then the CCC may produce a positive result that is in actuality nothing more than the correlation of the trends. If two independent profiles have a similar autoregressive feature then the CCC will produce a series of alternating positive and negative readings. Cross-correlations generate phase information, but do not show causation. Insulin and glucose oscillations may be physically related, but there is no mathematic solution to the problem of whether the insulin reduces the glucose or the glucose stimulates the insulin. Such conclusions need to be made using other physiological criteria. The problem of correlation by chance can be addressed by an increase in the number of profiles assayed and correlated.

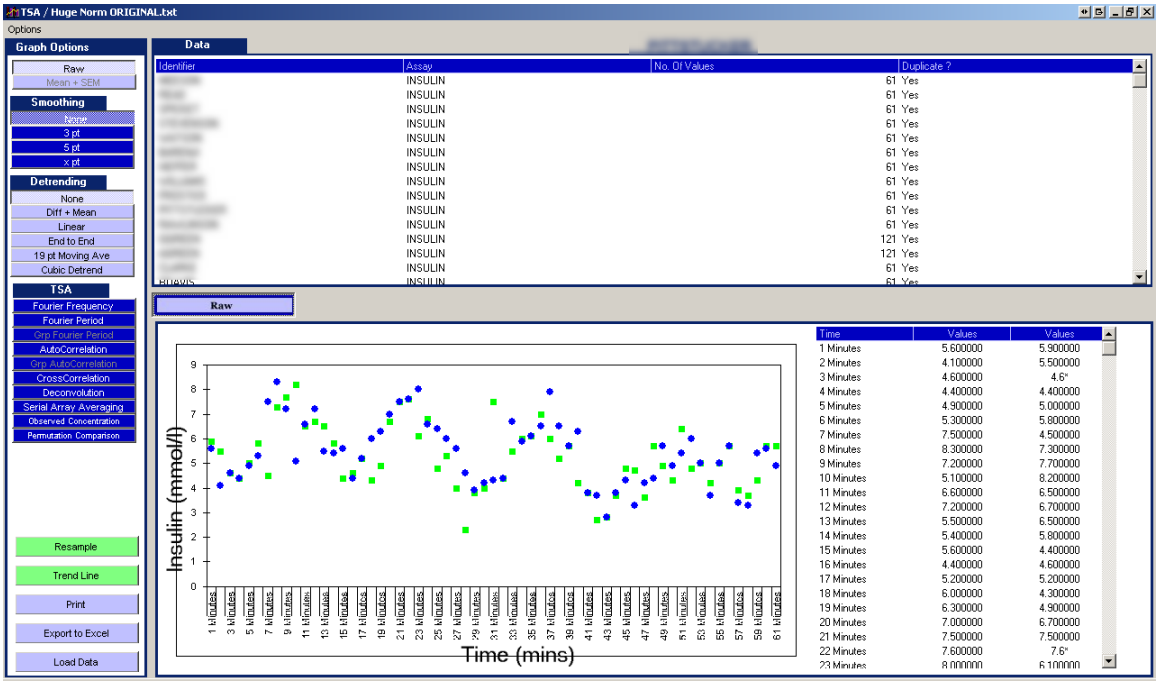
2.2.3.4 Deconvolution

The Deconvolution model used in Easy TSA is a one compartment model. The amount of a hormone measured in the blood represents a balance between secretion, distribution and degradation or clearance. Deconvolution is used as a method of estimating secretion rates from the data. It depends on knowing the half-life of the hormone (or its clearance kinetics) and whether the concentration changes during the observation period.

When the clearance rate of a hormone *in vivo* is slower than the secretion rate the analysis of the datasets must consider this. Deconvolution is the most suitable technique for this.

For example, to determine the secretion rate of an insulin profile taken at 1 minute intervals from a person with type 2 diabetes (Figure 2.14a) it is appropriate to use Deconvolution. The concentration of a hormone being measured will decline dependent on its half-life value. Insulin's half-life is approximately 3.8 minutes. Therefore, if the observed amount of hormone is equal to the decline calculated from such a half-life then no 'new' hormone is being secreted. If it is higher, then there is a secretion still in progress and the difference between the clearance curve and the actual amount of hormone observed is the secretion rate at that time (Figure 2.14b). The process of this 'curve stripping' has previously been depicted by Turner et al in 1979[65]

a)



b)

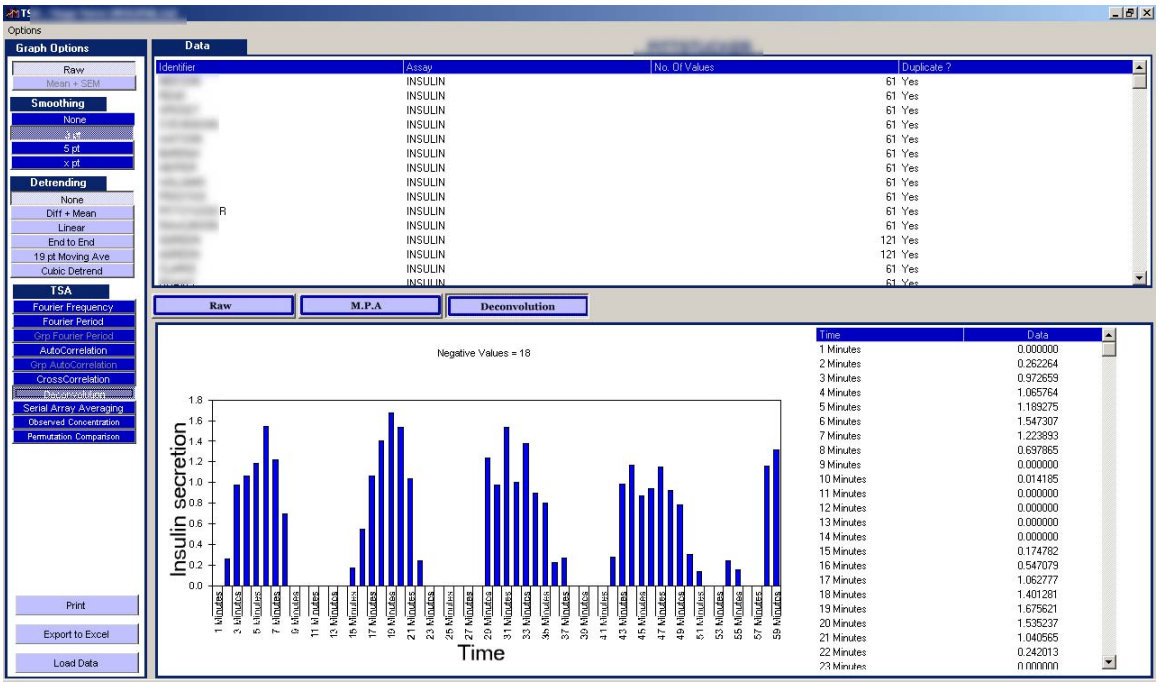


Figure 2.14 a) Insulin concentration measured every minute in duplicate b) Estimated secretion rates of insulin obtained by deconvolution after a 3 point moving average. Eighteen negative values detected and set to 0 for analysis.

Clearance rates *in vivo* are not always static for hormones, so Deconvolution is at best an estimation technique. As the half-life of the hormone is an estimate it is possible for

Easy TSA to report negative secretion rate even though negative secretion do not occur in vivo a count of these is provided on the chart. This can be addressed by modifying the estimate. If it is too conservative this will generate a majority of negative secretions. The half lives of the hormones insulin, c-peptide and GH are pre-configured in the Easy TSA program although these can be changed by the user.

2.2.3.5 Probit analysis function (Observed Concentration)

When analysing a hormonal profile there may be an interest in the peaks and troughs of oscillations occurring in the profile - the dispersion of the data.. Probit analysis uses the totality the measured data by sorting the observed values into a time independent array and ranking the values. Probit analysis also known as Observed Concentration (OC) is a probability calculator[66]. It is a method to quantify the distribution of concentrations present in profiles; this allows an unbiased estimate of attributes present in the profiles.

Traditional methods used to analyse the dispersion of data are mean and standard deviation for normally distributed data and for non-normally distributed data, the data may be log transformed and then analysed, the Median and inter-quartile range may be calculated. The advantage of probit analysis in comparison to the traditional distribution analysis methods mentioned is that the totality of the data are used in probability calculations. The probit function is the inverse of the cumulative distribution function of the normal distribution. The resulting information calculated by the probit function are the parameters of the normal distribution that is assumed to underlie the data. The underlying normal distribution assumption might not be correct. As such the use of Probit calculation would be incorrect.

OC is calculated by sorting the available data into an ascending order and allocating the data to class intervals (bins). The optimum number of bins is approximately 15 as this ensures a wide representation of the dataset [67]. The bins reflect the range of the profile. Using a GH profile as an example (Figure 2.15a) Easy TSA calculates the log distribution of 15 bins for the profile (Table 2.1).

Bin number	Log Serum GH concentration (U/L)
1	<0.4
2	0.41 - 0.63
3	0.64 - 1.0
4	1.01 - 1.58
5	1.59 - 2.51
6	2.52 - 3.98
7	3.99 - 6.3
8	6.31 - 10.0
9	10.01 - 15.8
10	15.9 - 25.1
11	25.2 - 39.8
12	39.9 - 63.1
13	63.11 - 100.0
14	100.01 - 158.5
15	>158.5

Table 2.1 – Log bins selection for the serum concentration of GH

The OC calculates a cumulative frequency distribution for the GH profile. The cumulative frequencies are translated into discrete linear probabilities (linear probits) for each bin, via a look up table extrapolated from Geigy Scientific tables[66]. The probit function is also available in Excel as the NORMINV function. The linear probits are then plotted against the means of the GH bins (Figure 2.15b) but are offset by 5 to allow the function to be positive for all probabilities > 10^{-5} . A linear regression (trend) line can then be applied to this graph allowing the probabilities of concentrations occurring to be calculated. For further information on probability distribution and the probit transformation please see the Geigy Scientific tables[66] reference.

a)

b)

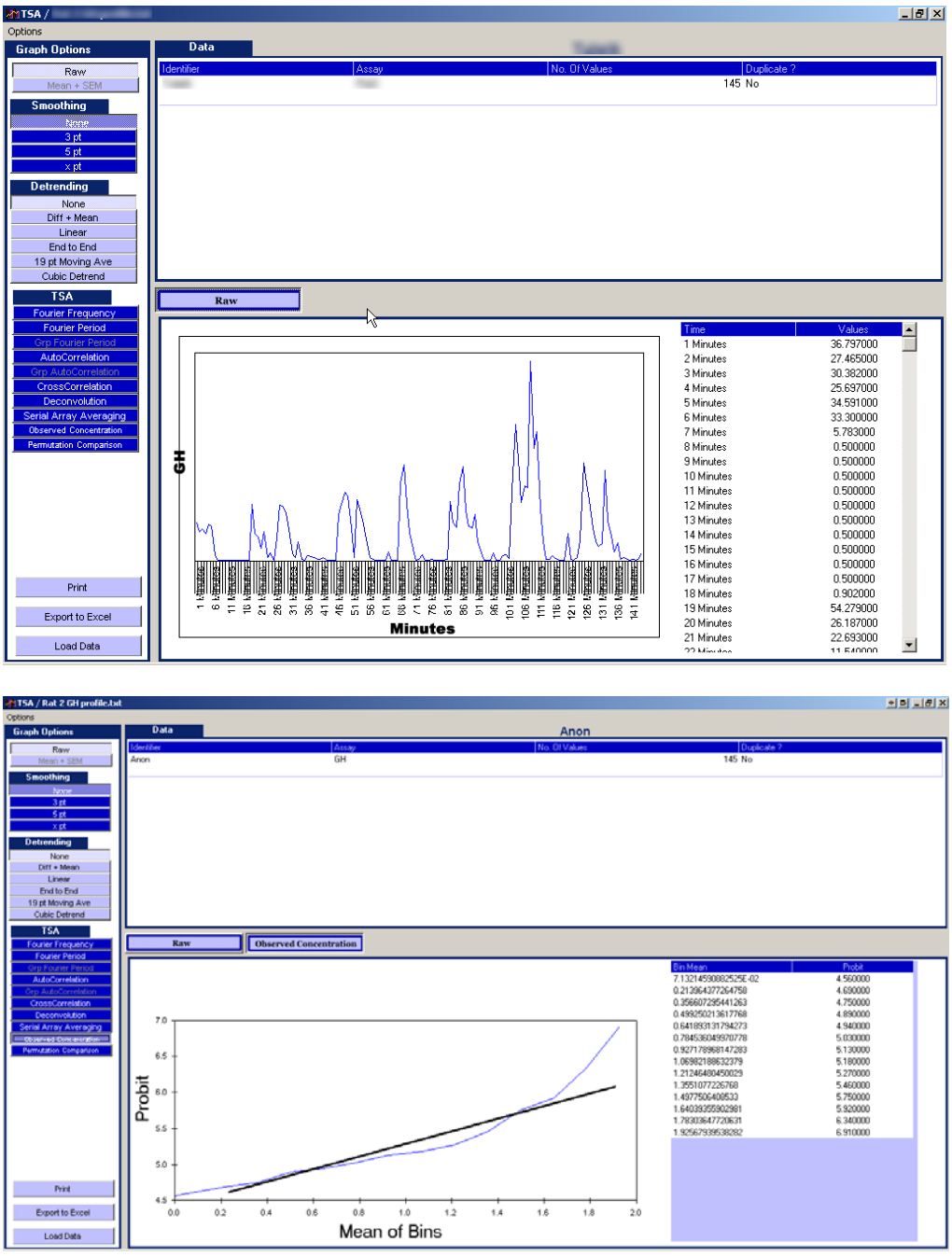


Figure 2.15 - a) Rat Growth hormone profile measurements taken every minute b) Probit regressional analysis graph with trend line.

The equation of this particular regression line (EQ 2.7) is then used to calculate the OC.

EQ 2.7 $$y = 1.0573x + 4.2886$$

The typical reporting methodology for OC is the OC5, OC50, and OC95. The Observed Concentration 50 (OC50) is the value of the 50% probability calculated from the regression line: it is comparable to the median of the data, but is not the same. The OC5 is the value at which 5% probability occurs and can be regarded as the baseline of the data. Similarly, the OC95 is calculated from the 95% probability point on the regression and gives an unbiased estimate of peaks.

The probits are known and can be extrapolated from the Geigy Scientific table [66] of probits. The Probit for 5% of the graph (OC5) is 3.3551.Thus inputting this into the equation EQ2.7:

$$OC5 =$$
$$y = 3.3551 = 1.0573x + 4.2886 \therefore$$
$$3.3551 - 4.2886 = 1.0573x \therefore$$
$$\frac{3.3551 - 4.2886}{1.0573} = x \therefore$$

$$OC5 = 10^{x} \therefore$$
$$OC5 = 10^{-0.88} \therefore$$
$$OC5 = 0.13 \therefore$$

5% of the data is ≤ 0.13

The other OC's were OC50 = 4.7, OC95 = 169.2 respectively.

When analysing a cohort, the datasets can be pooled, the same class intervals must be used for each dataset and the results can then be analysed using standard parametric statistics.

2.2.3.6 Serial Array Averaging

Serial Array Averaging (SAA) is a technique for combining one part of a signal to another part of the signal[68]. It allows the examination of phase relationships (i.e. the comparison of peaks and troughs) between pulsatile regions in the same data array, even when pulses are different durations or amplitude. The result is a normalised pulse indicative both of the mean values (amplitude) and of the mean wave duration. The SAA requires objective starting points the appropriate use of SAA is usually when external stimuli are applied at known times. These times can then be used as the synchronising points between datasets.

Serial/standard array averaging can be a powerful technique for examining phase inter-relationships and for reducing 'noise', because many cycles can be pooled to produce one time series. By using SAA to select appropriate parts of the profile (Figure 2.16a), it is possible to analyse the effect. Eighteen data points have been selected beginning at 10, 27, 49 minutes (Figure 2.16b). The SAA graph (Figure 2.16c) is produced with the standard error of the mean calculated for each point..

a)

b)

c)

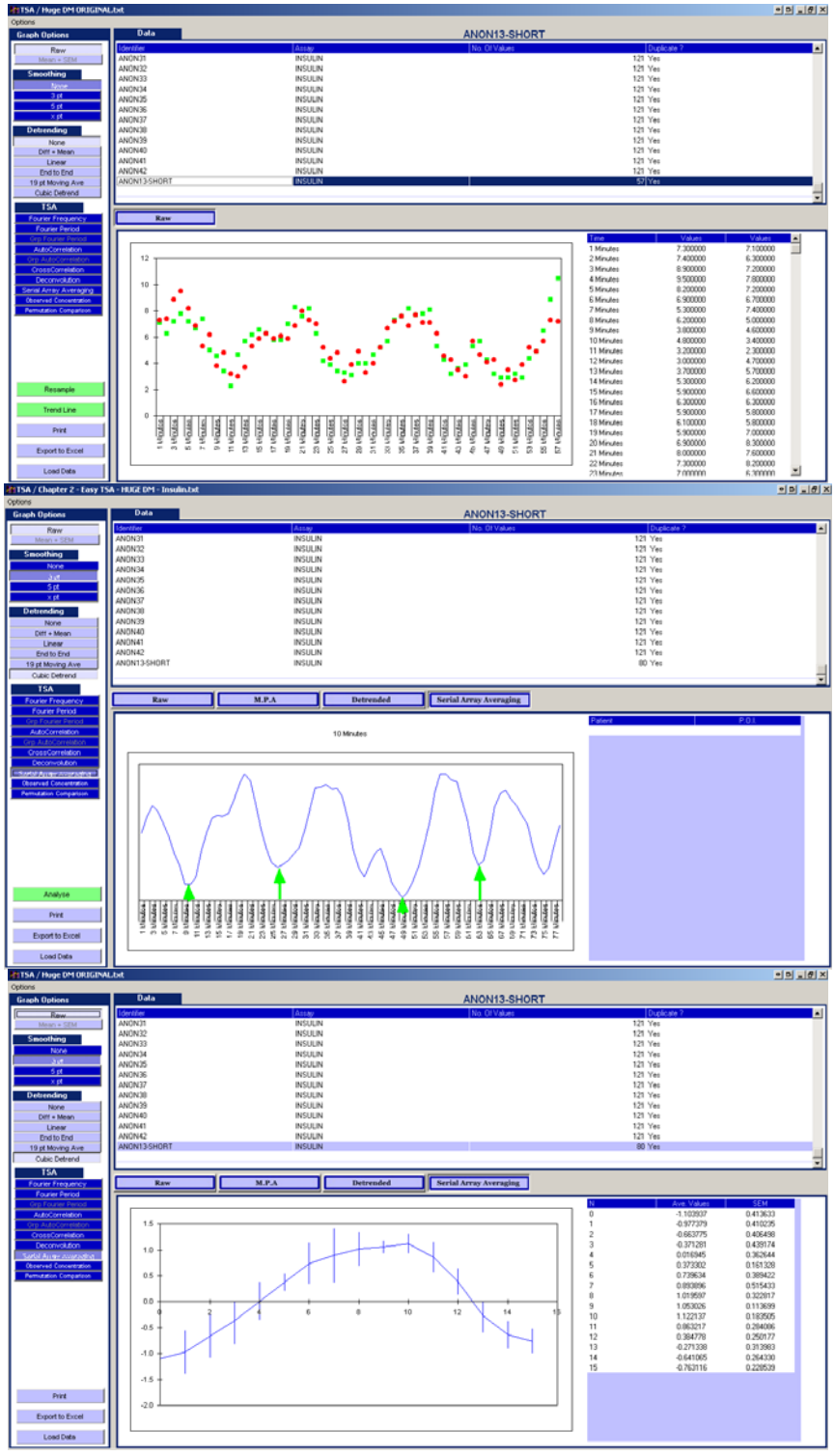


Figure 2.16 a) Insulin profile measured induplicate every minute from a subject with type 2 diabetes b) Insulin profile with 3pt ARMA, cubic detrending with 4 points of interest selected c) SAA of selected 17 minute oscillations peaks.

Pre-processing of data before a SAA is not essential, but if strong trends are, present in the datasets under comparison these can cause problems with interpreting the results. There are no statistical tests available for calculating the significance of a SAA however the waveform of two or more SAA's can be compared in a similar way to FT. If the standard error of the mean bars do not overlap at a discrete time period then the null hypothesis, of there being no difference, is rejected ($p<0.025$).

2.2.3.7 Runs analysis

Analysis of Runs is method of determining whether a dataset contains noise alone or contains a hidden oscillation(s). The Runs method is extrapolated from the paper by F S Swed et al[69] and was incorporated into the Easy TSA program as an easy to use option.

Runs works by grouping a dataset into two categories: those values above the mean and those below it. When two different types of data are arranged in sequence, they will form two or more groups. For example, consider a very short (7 point) GH profile where measurements above the mean are termed a and measurements below the mean are termed b. The sequence aabbbab, has 3 a's, 3 b's and forms 4 runs (aa, bbb, a, b). Therefore if there are m number of a's and n number of b's total number of possible groupings can be expressed by a combinatorial, EQ2.8:

EQ 2.8
$$C_m^{m+n} = C_n^{m+n}$$

The runs equations calculate the randomness of a dataset dependent on the frequency of the oscillations. If u is the number of groups (runs) in the current dataset then the probability of an arrangement of groups yielding u' or fewer groups can be represented by EQ2.9:

EQ2.9

$$P\{u \le u'\} = \frac{1}{C_n^{m+n}} \sum_{u=2}^{u'} \int u$$

Where $\int u = 2C_{k-1}^{m-1} \cdot C_{k-1}^{n-1}$, when $u = 2k$, i.e. u is even

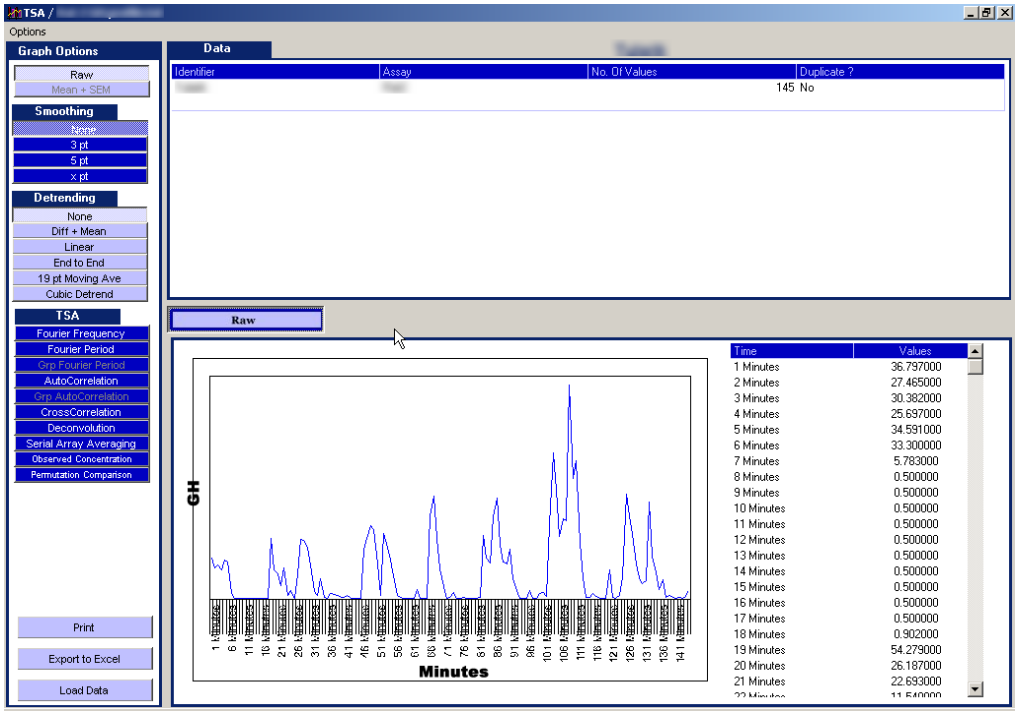
in addition $\int u = C_{k-1}^{m-1} \cdot C_{k-2}^{n-1} + C_{k-2}^{m-1} \cdot C_{k-1}^{n-1}$, when $u = 2k - 1$, i.e. u is odd

for $k = 1,2,...,m+1$

Using the Runs equation it is possible to determine the probability of the dataset not being completely random[70]. If p <= 0.05 then it is concluded that the dataset is non-random and therefore contains a process or signal. This technique is ideally used on duplicate data.

For example, a rat GH profile is pulsatile in nature (Figure 2.17a). The Runs analysis of these data calculates the probability of this profile being non-random (Figure 2.17b) and occurring purely by chance.

a)



b)

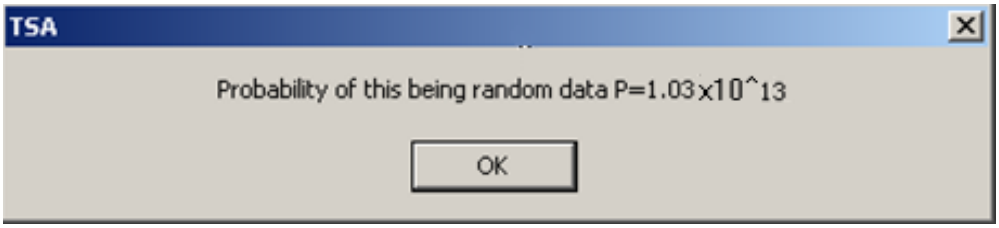


Figure 2.17 a) Rat GH profile is pulsatile b) Runs analysis of the data calculates p =1.03x10^13 of these data occurring by chance.

2.2.3.8 Permutation Analysis

In the 1930's RA Fisher developed the idea of permutation analysis [71]. Permutation analysis sought to determine if a result was random or had an underlying significance. This was achieved by the determination of whether to accept or reject that the profile had a pattern or was purely caused by chance or randomness and, in some cases, to construct confidence intervals for unknown parameters.

With the advent of computers and their ability to perform, repetitive tasks (permutations) at high speed and accuracy, permutation analysis became a viable analytical method for datasets that contain hundreds or thousands of data items. In terms of hormonal analysis with the advent of continuous glucose measuring systems that produce thousands of glucose measurements, permutation analysis is a viable tool.

Permutation analysis tests the null hypothesis (H_0) that there is no difference between two groups. For example, two independent groups of subjects those with type 1 diabetes and those without diabetes would produce glucose samples $z = (z_1, z_2, ..., z_n)$ and $y = (y_1, y_2, ..., y_m)$ from different probability distributions F and G. Therefore

$F \approx z = (z_1, z_2, ..., z_n)$ independently of $G \approx y = (y_1, y_2, ..., y_m)$

Having observed the glucose profiles z and y (Figure 2.18a, Figure 2.19a) the next step is to test the null hypothesis H_0 of no difference between the FT's of the two groups F and G,

EQ 2.10 $H_0 : F = G$

If H_o is true, then there is no difference between the probabilistic behaviour of the first observed glucose profile (z) or the second observed glucose profile (y). If H_0 is False then there is a significant difference between the observations. The permutation analysis that Easy TSA performs is the comparison of the normalised FT's of the two groups F & G. Permutation analysis compares all the possible groupings of the sets of profiles to determine if the initial two groups occurred by chance alone and that there is a significant difference between the two groups FT's.

Group F in the example contains six glucose profiles from people without diabetes (Figure 2.18a), with normalised FT's (Figure 2.18b). Group G contains six glucose profiles from people with type 1 diabetes (Figure 2.19a) with normalised FT's (Figure 2.19b).

The permutation analysis calculates the probability of the FT's generated by the type 1 glucose profiles (in Group G) being the same as the FT's generated by the non-diabetic profile (in Group F) as 0.01 (1.0%) (Figure 2.20). Therefore the null hypothesis is rejected , the two groups are not the same.

a) b)

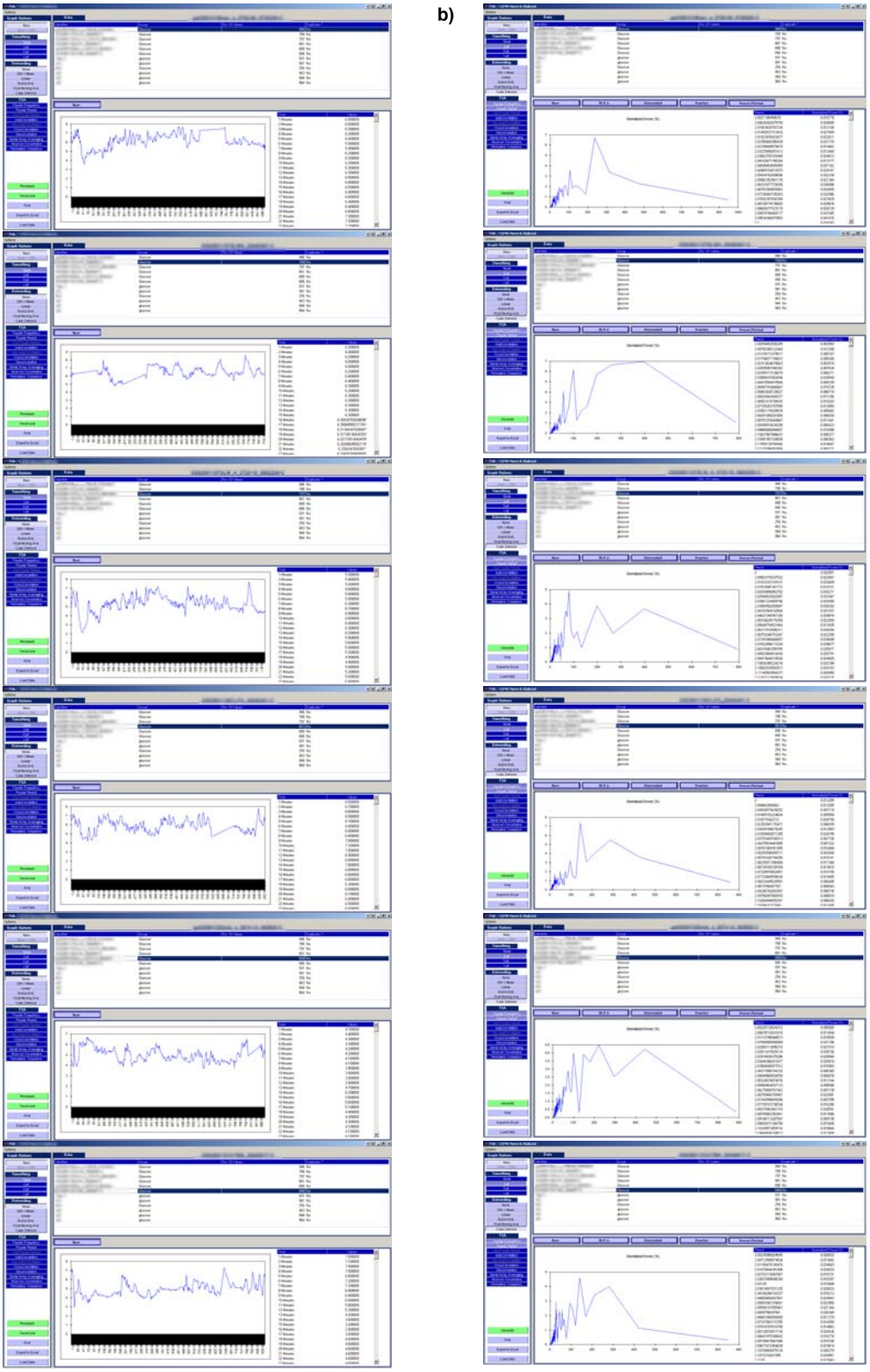


Figure 2.18 – a) Glycaemic profiles of people with type 1 diabetes b) FT's of people with type 1 diabetes that are used in permutation analysis

a) b)

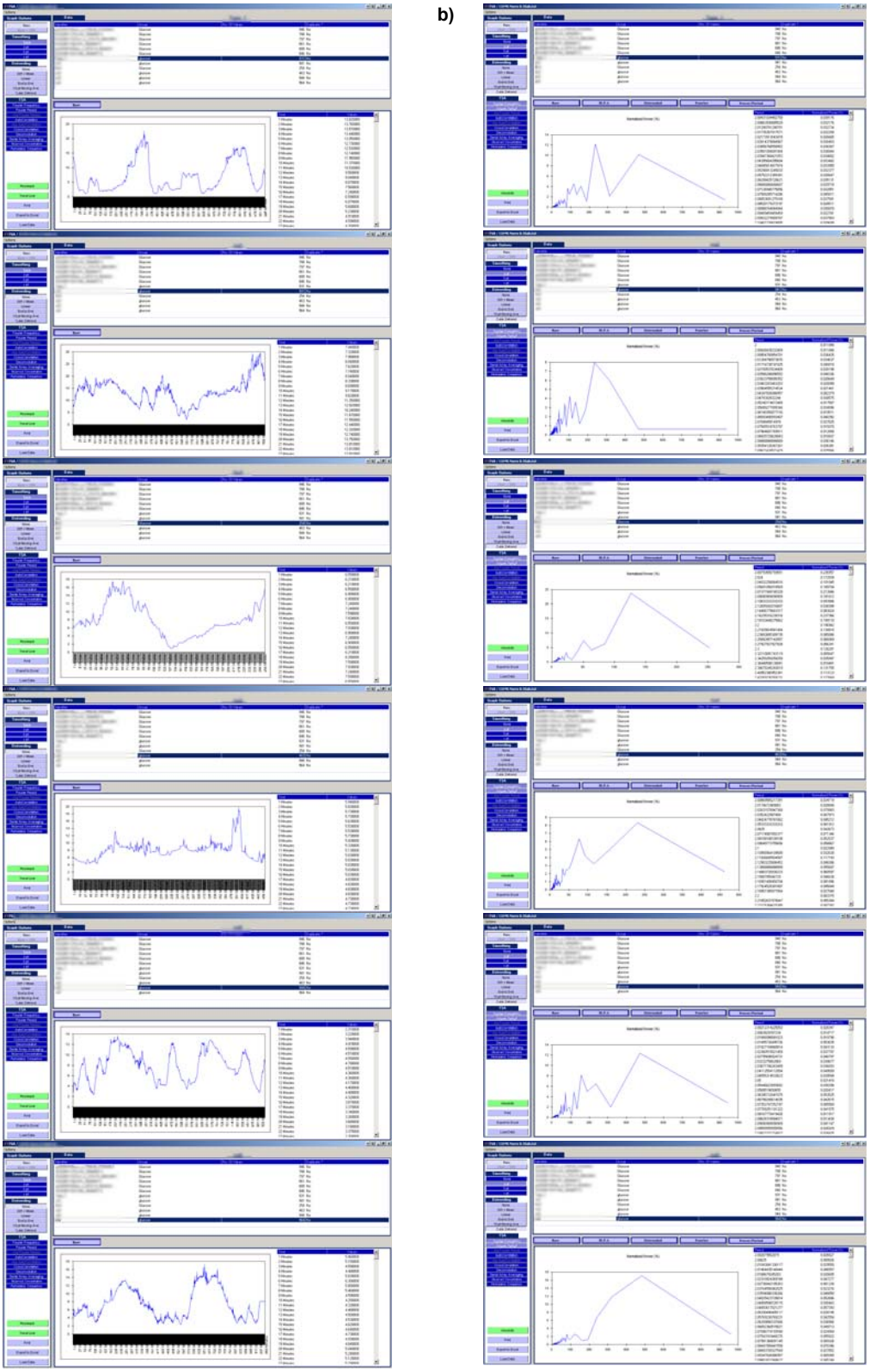


Figure 2.19 – a) Glycaemic profiles of people with no diabetes b) FT's of people with no diabetes that are used in permutation analysis

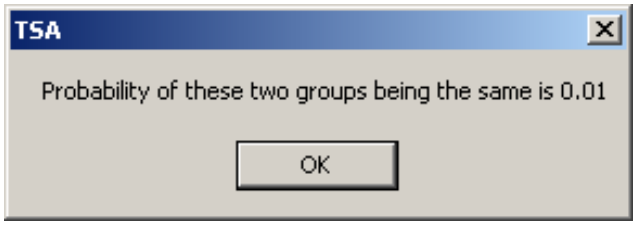


Figure 2.20 – Result from Permutation analysis of whether two groups are the same. 1.0% probability of them being the same is calculated

The Permutation p value provides a result whereby it can be comparatively ascertained if a series of hormonal profiles are significantly different from one another (and had not differed by chance alone).

2.2.4 Conclusions

Time series Analysis techniques are of use when examining megabyte datasets. They enable a reduction in the noise component of the data through a variety of methods and the extraction of user-impendent objective information from the data. It has been demonstrated by a variety of groups how hormonal systems *in vivo* are non steady state[17, 43, 44, 48, 49] and are amenable to analysis using Time Series Analysis (TSA). The Easy TSA software program was developed to examine the oscillations in hormonal concentration. Many hormones are secreted in pulses and the analysis of these can lead to particular insights into physiological systems –e.g. insulin dynamics are less effective when pulsatility is reduced[51].

The methods used within Easy TSA are not new but have been judicially chosen as relevant and useful for investigation into megabyte datasets of biological systems. Permutation analysis of FT and Runs analysis are the first computer based implementations of the mathematical concepts and this is the first TSA tool aimed solely at medical research. The Easy TSA contains no assumptions (embedded constants) on

hormonal profiles, clearance rates, etc and allows the user to select for instance the half life kinetics of insulin, although it does offer a suggestion as to what it should be. Easy TSA interpolates any missing values in the data sets but does require that the sampling is regular and equal in order to perform correctly. Further improvements could be the additional of wizard that guides the user through the correct selection of the appropriate smoothing, detrending and TSA test for the profile under consideration. Easy TSA is due to be released as a freeware program for academic and non-commercial.

The use of TSA for the investigation of diseases and disorders has been undertaken by several groups. Previous studies have shown the relevance of the pulsatile release of hormones. In endocrinology, the control of puberty and reproductive function by GnRH is optimum when GnRH is administered in a pulsatile fashion whereas rapid down regulation of the gonadotroph can be easily achieved by continuous administration of gonadotrophin releasing hormone[48, 49]. This latter effect has been employed in the development of gonadotrophin releasing hormone analogues for treatment of patients with prostatic cancer. Different patterns of GnRH pulsatility regulate individual LH and FSH secretion[50]. A change in the frequency of release and amplitude of release of gonadotropin has been shown to delay growth and puberty with subsequent borderline low sperm count[72]. The development of Easy TSA provides a variety of methods that can aid in the formalisation of the information required from the initial data.

2.3 Validation of the Easy TSA software

Validation of the Easy TSA program was essential before the software was used in a research capacity. This stage was one that was vital to determining whether Easy TSA was functionally correct and robust. Datasets were constructed that consisted of sine waves and supplemental test data that would validate the outcomes from Easy TSA. (The data sets are available on the companion CD). Noise was added to the sine waves to simulate biological data. The noise that was added to the sine waves was calculated using EQ2.11

EQ2.11

If *Rand* generates a number where the last digit is odd then $X_1 = X - n$

If it is even then $X_1 = X + n$

Where, $n = Rand * \left(\left[l / 100 \right] * X_{\max} \right)$

Where X is the data before the noise and X_1 is the data after the noise has been applied, l is equivalent to the percentage power of noise (i.e. 1%), n is equal to the absolute value of noise. The noise could either be added to or subtracted from the data profile depending on whether the random number was odd or even.

2.3.1 Fourier Transform validation

The FT function was tested in a variety of ways. The first test was to assess the ability to extract correctly the frequency from a 50-point sine wave (Figure 2.21a). The amplitude was then doubled (Figure 2.21b) and the data re-tested. The frequency of the original FT was then doubled (Figure 2.21c). A complex sine wave was analysed (Figure 2.21d), consisting of one 50-point wave with another 50-point wave that had double the

amplitude and double the frequency. The complex wave was created to examine the capacity of Easy TSA to resolve back into two components.

	Freq	Amp	Phase
a)	8	1	1
b)	8	2	1
c)	16	1	1
d)	i) 8	1	1
	ii) 16	2	1

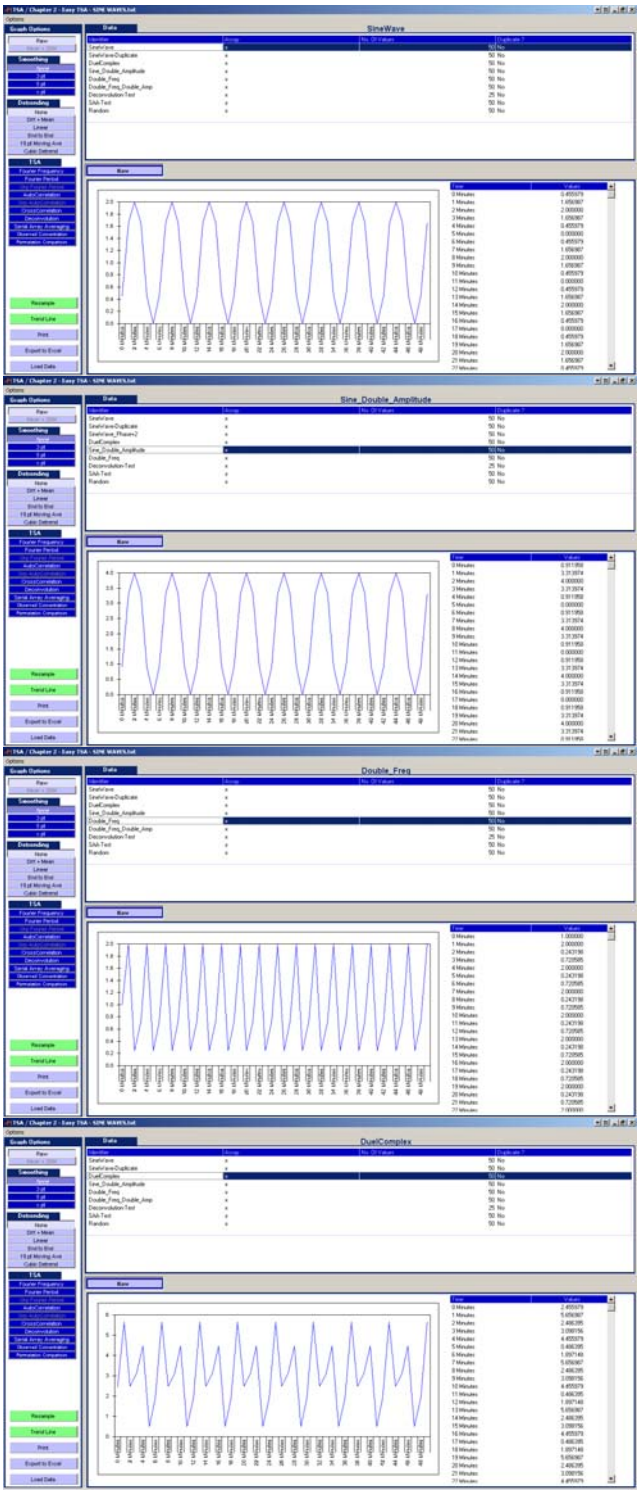


Figure 2.21 –sine waves used in the FT Validation with the breakdown of the frequency, phase and amplitude used.

The FT validation was successful. Doubling the amplitude had no effect on the frequency detected in the relative FT graphs (Figure 2.22a & 2.22b) it did double the power of the frequency in the absolute FT. This is because relative FT's are independent of percentage power contribution (i.e. independent of amplitude) and the two sine wave inputs had identical frequency. Doubling the frequency of the sine wave is detected by the function (Figure 2.22c) and the main spectral period has doubled from a frequency of 6.25 to 12.5.

The analysis of the complex wave, consisting of two sine waves, indicated that the program separated the sub harmonics of the data into separate frequencies (Figure 2.22d).

Raw Data **Relative Fourier Transform**

a)

b)

c)

d)

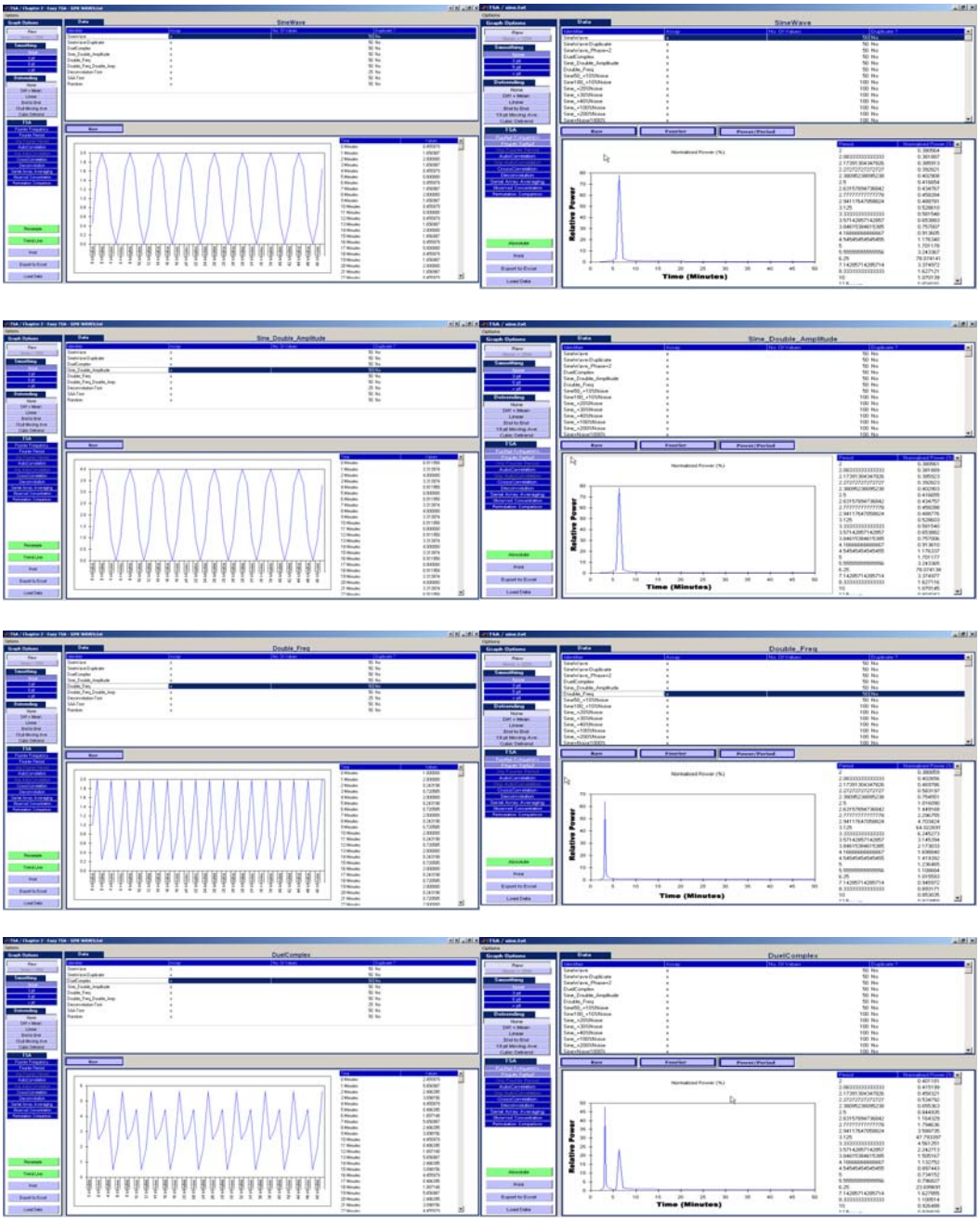


Figure 2.22:- Four sine waves with their relative FT spectral analysis

Further FT tests were to add noise to the 50-point sine wave (Figure 2.23) and to the 50-point complex sine wave (Figure 2.24) to discover the threshold at which no signal extraction was possible. The random noise required before the FT function no longer detected the sine waves for the single 50-point sine was between 250% and 260%. The complex 50-point sine wave had the less powerful frequency obscured by less than 90% of noise.

Figure 2.23 – FT of a Sine wave with randomly generated white noise added with 3pt MPA and difference + mean detrending a) 10% b) 100% c) 200% d) 240% e) 250% f) 260% g) 270%

a)
b)
c)
d)
e)
f)
g)

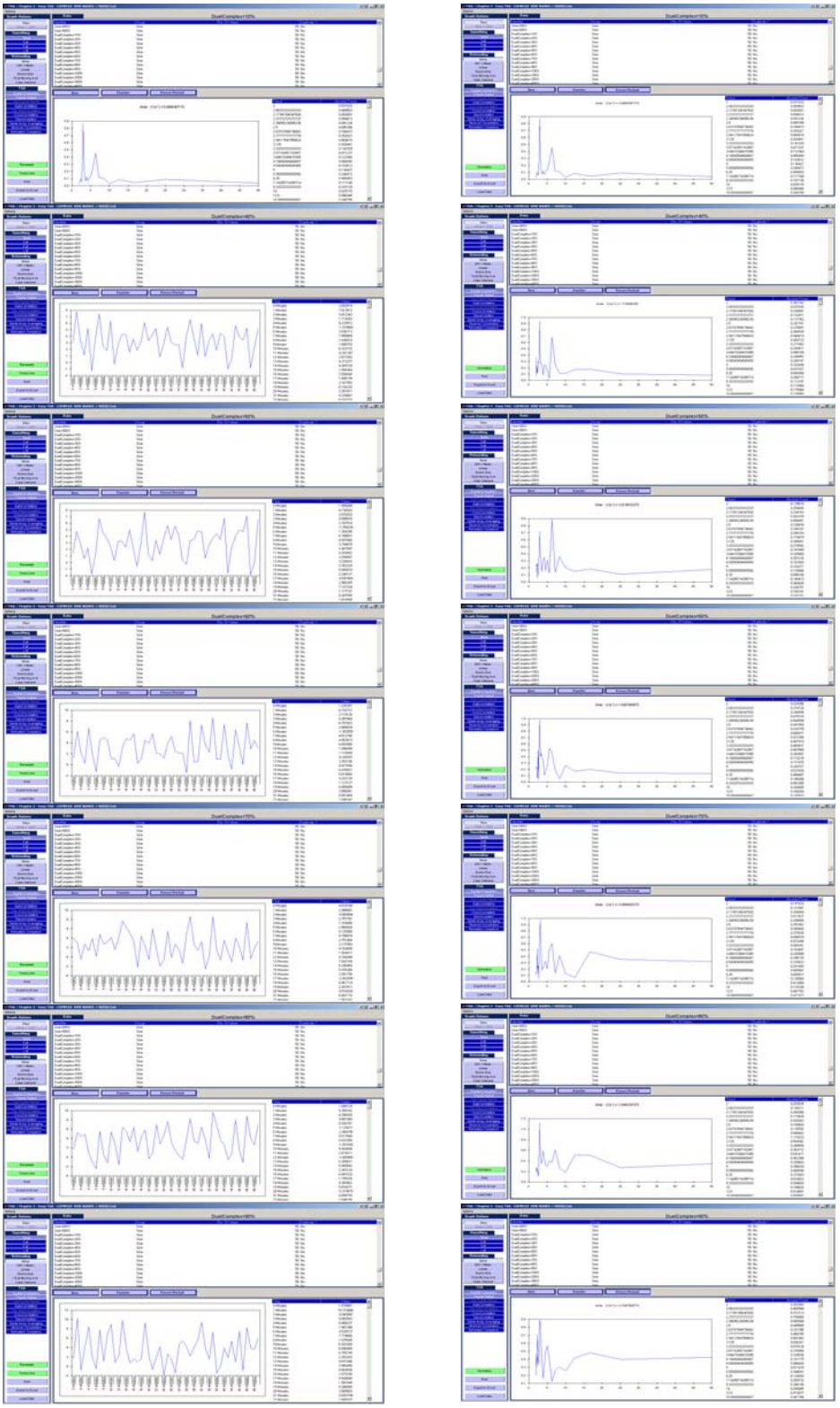


Figure 2.24 - FT of a complex sine wave with randomly generated white noise added a) 10% b) 40% c)50% d) 60% e) 70% f) 80% g) 90%

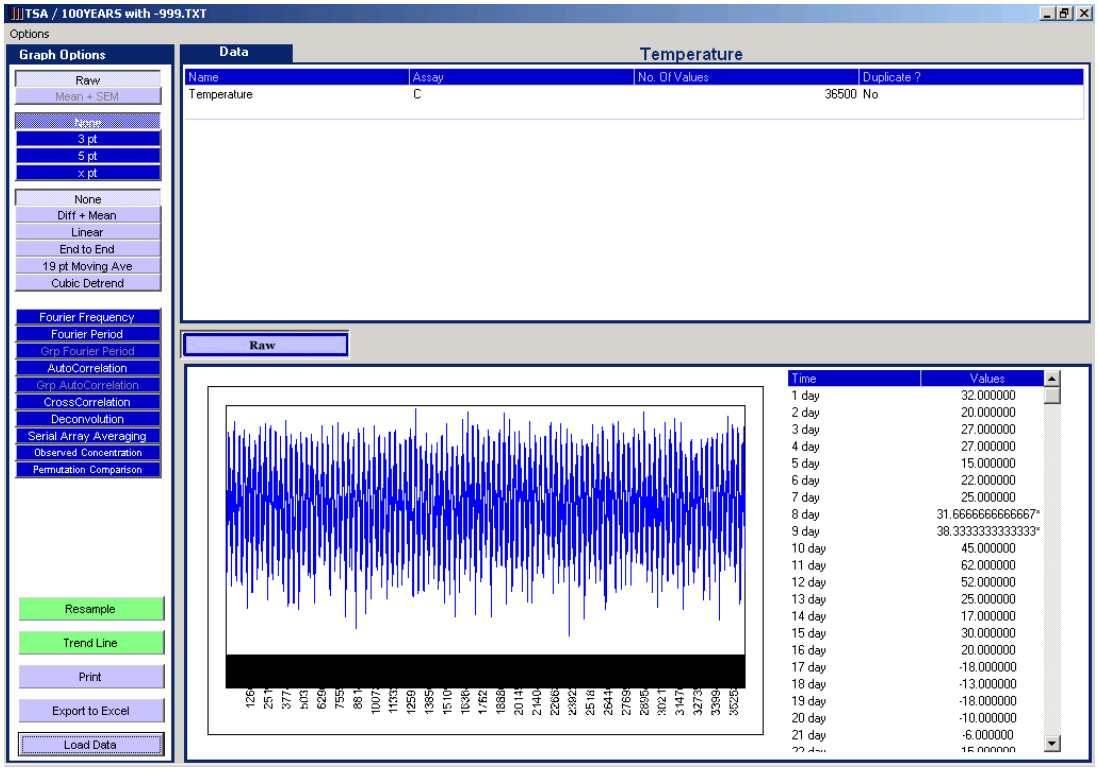
The validation was performed three times to ascertain if the result caused by adding noise was due by chance alone. The additional two outcomes were comparable with the results displayed here. The FT function was seen to be both correct and robust. The results demonstrate that the functionality of the FT processes was as expected. The level of noise needed to obscure the signal was influenced by the clarity of the original oscillations. However, the requirement to add 'white' noise before the signal was obscured was a validation of the robustness of the FT equation and its implementation within Easy TSA.

2.3.1.1 Fourier Transform calculation of the length of the tropical year

The validation of the FT function was successful and proved the algorithm to be correctly mathematically implemented. It was deemed pertinent to validate the FT function against a 'real world' scenario. By 'real world', is meant that the data used to validate against is noisy, imprecise, with missing values, and may contain sub harmonics of other processes. The validation required that the outcome be known, so it could be determined if the FT was successful. It was decided to calculate the orbit of the earth around the sun using temperature readings alone. The data were obtained from the British Atmospheric Data Centre. The collection of these data began about 400 years ago and were initially taken daily by a brotherhood of monks at Oxford. The analysis using FT used the most recent 100 years readings (Figure 2.25a). Oxford is obviously hotter in summer and cooler in winter dependent upon the angular tilt of the earth on its axis in relation to the sun. The FT described a waveform of maximal power/amplitude (33.0) at a periodicity of 365 days (Figure 2.25b) that indicated a greater than random tendency for peaks and troughs in the time series to recur at a frequency or period of 365 days and indicated a distinct rhythm. The FT provided results in integers, so this

was the most accurate result obtainable using this method. It should be noted that this is not an estimation of the sidereal year, but the tropical year (actually 365.24 days)

a)



b)

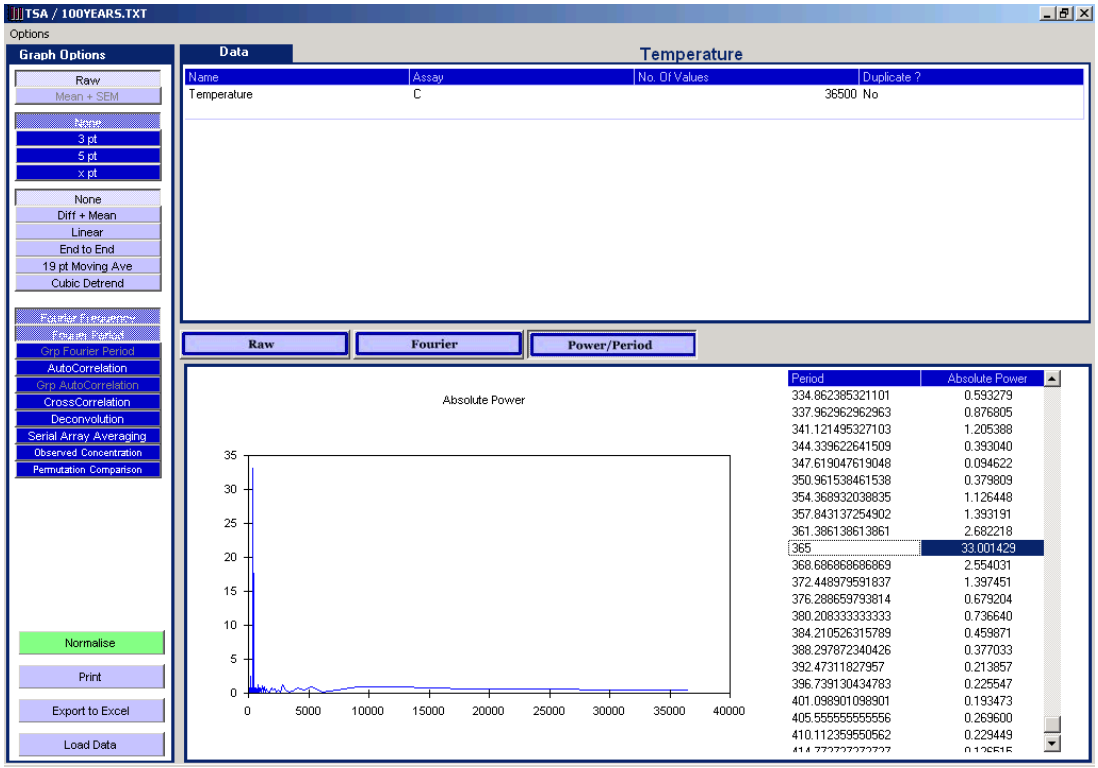


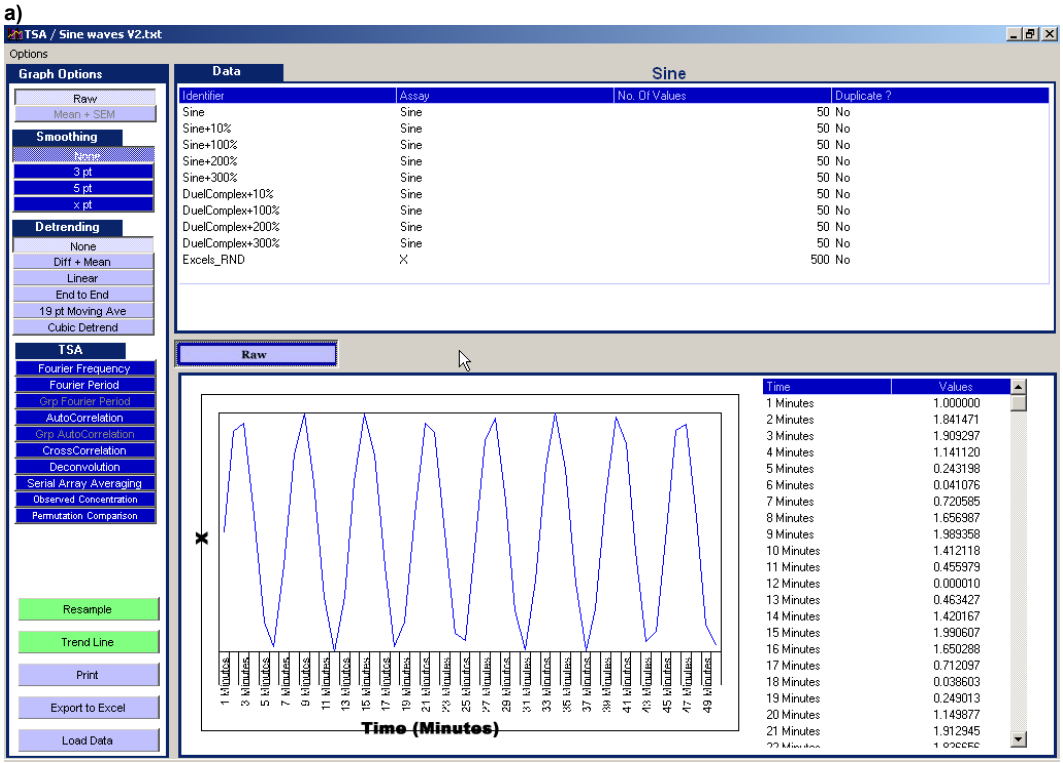
Figure 2.25 a) 100 years of daily temperature readings in Oxford, missing values denoted by * and interpolated using the closest real values b) Fourier Transform of the temperature data denoting a maximal power at a period of 365 days

2.3.2 Autocorrelation & Crosscorrelation validation

The autocorrelation (AC) and the crosscorrelation (CC) functions use the same base method to perform their calculations. It was possible to test AC, and hence validate both AC and CC. The 50-point sine wave used in the FT testing (Figure 2.22a) formed the basis of the validation for the autocorrelation method (Figure 2.26a). The sine wave was run without pre-processing to check the correctness of the function. Increasing amounts of noise were added to the sine wave until the r-value of the correlation lost significance.

There was a complete correlation r=1 when Lag = 6 (Figure 2.26b) which was expected from the original graph (Figure 2.26a) that has an oscillation of 6.25 minutes.

a)



b)

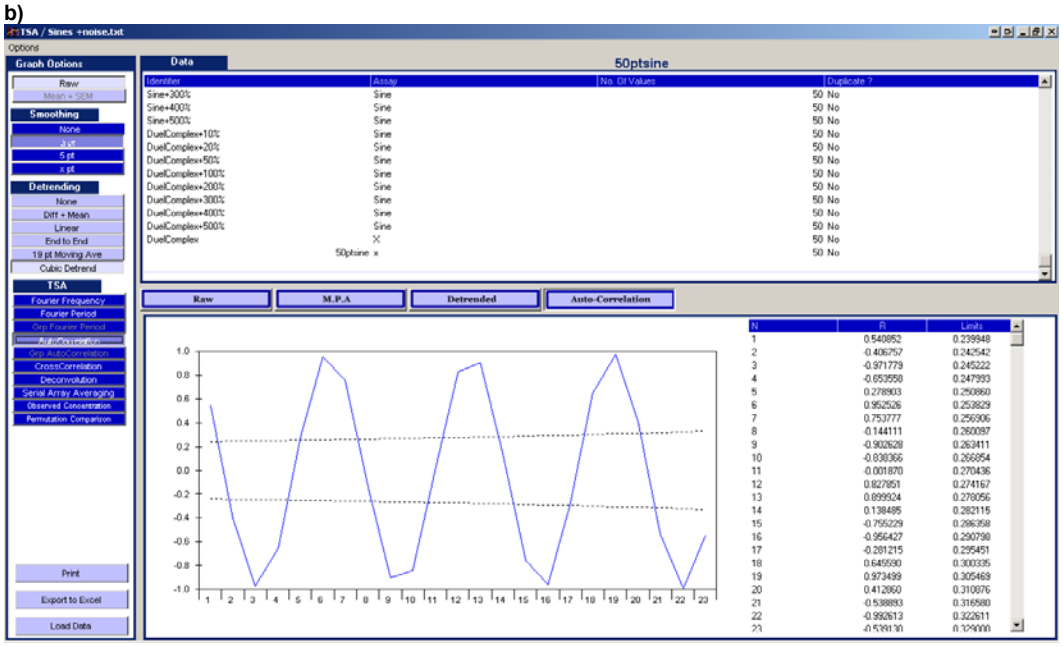


Figure 2.26 - a) 50-point sine wave b) Autocorrelation of the 50 point sine wave

Further validation involved the addition of random noise to the sine wave until the r-value of the correlation became non-significant. Two hundred percent noise was required before the signal was obscured (Figure 2.27).

a)

b)

c)

d)

e)

f)

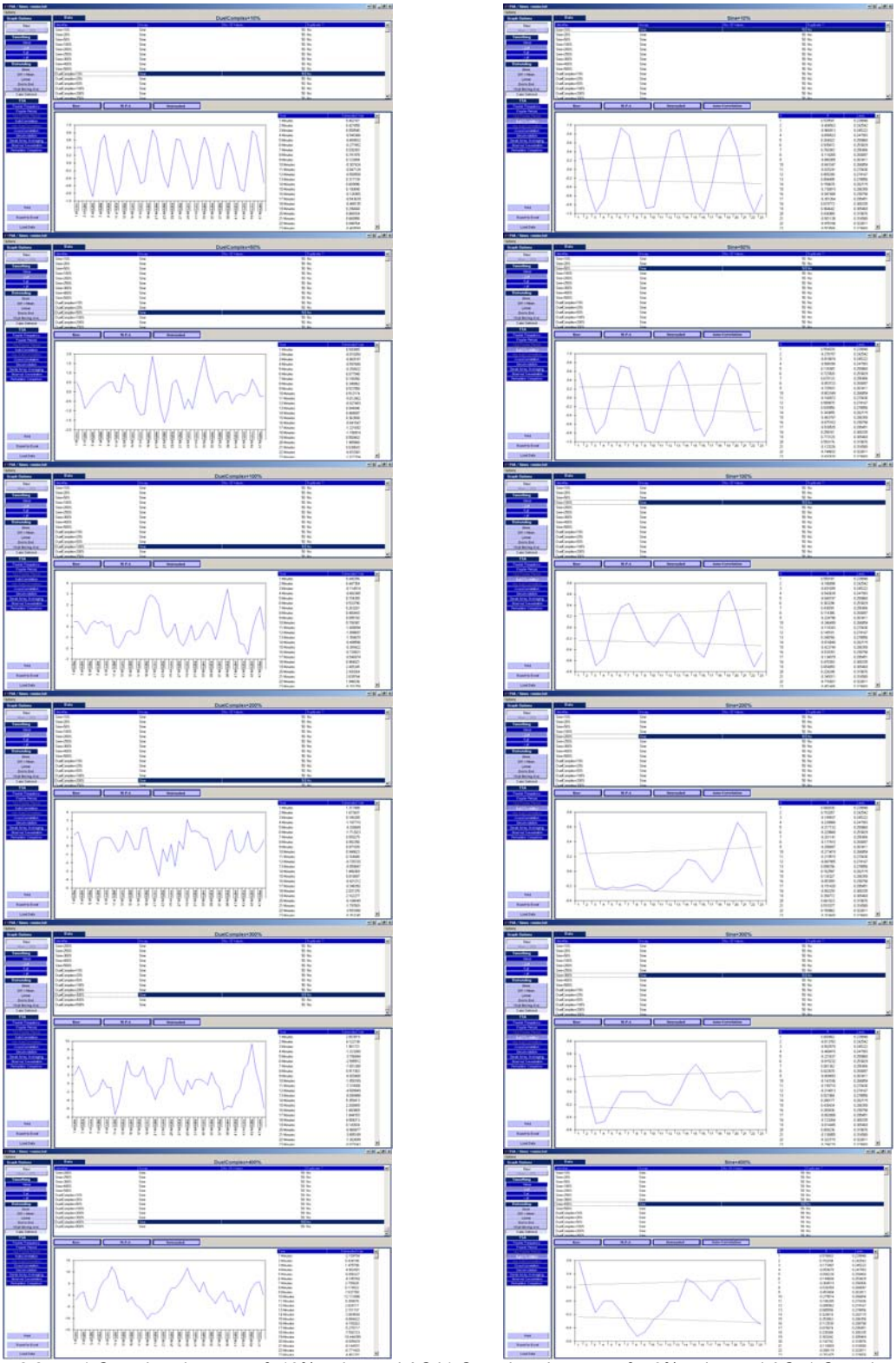


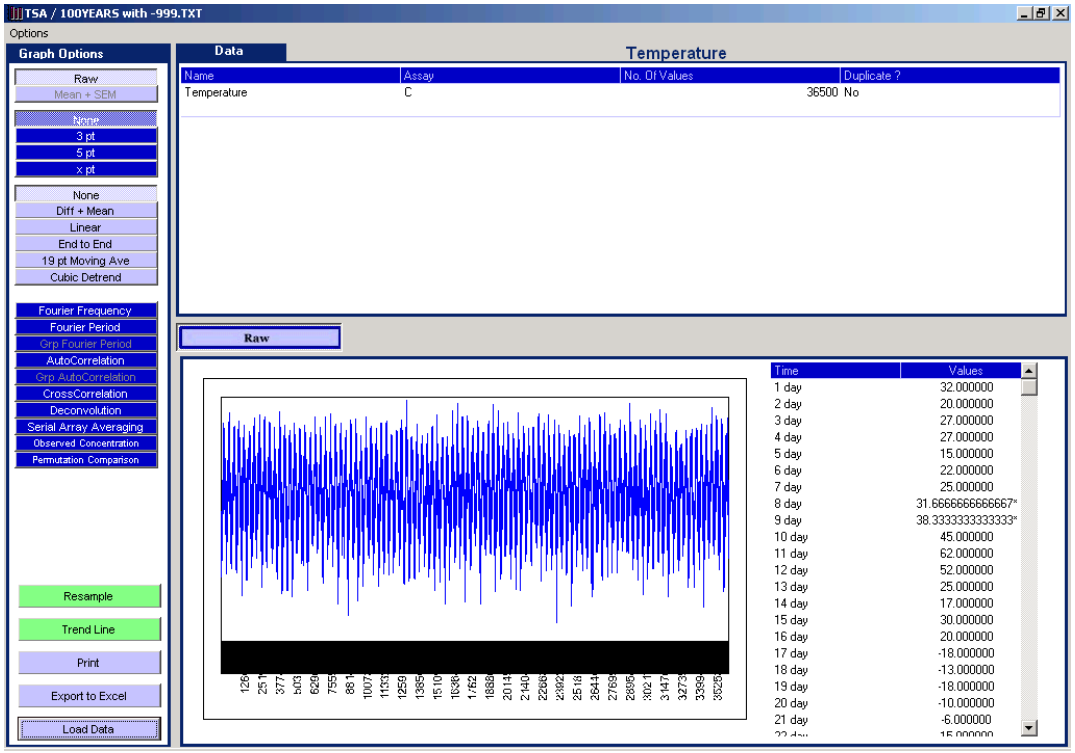
Figure 2.27 – a) Complex sine wave & 10% noise and AC b) Complex sine wave & 50% noise and AC c) Complex sine wave & 100% noise and AC d) Complex sine wave & 200% noise and AC e) Complex sine wave & 300% noise and AC f) Complex sine wave & 400% noise and AC

The validation was performed three times to ascertain if the result caused by adding noise was due to chance alone. The additional two outcomes were comparable with the results displayed here. The level of noise needed to obscure the signal was influenced by the clarity of the original oscillations. Validation of the correlation function was successful. It was interesting to note the amount of 'noise' that was needed to obscure the signal was 200%+ which was comparable to the FT function.

2.3.2.1 Autocorrelation calculation of the length of the tropical year

The calculation of the rotation of the earth using daily temperature readings from the last 100 years (Figure 2.28a) had been successful in the FT Validation. AC is of use as a method for assessing regularly reoccurring waveforms, and so is ideal for applying to daily temperature data. But using AC results it is possible to use the second, third, nth peak to gain precision in the assessment of periodicity. The period calculates to be T_{peak}/n where T_{peak} is the periodicity and n= number of peaks. So for example, if 10 peaks are detected on the AC, another decimal point of accuracy can be added. The ability to provide results that are fractions allows a greater degree of accuracy in the calculation of the rotation of the earth. The autocorrelogram generated from these data (Figure 2.28b) detected a final regular reoccurring waveform at the 49^{th} peak after 17900 days. Thus providing a result of (T_{peak} = 17900 and n=49) 17900/49 = 365.3 days. The ability to calculate an astronomical phenomenon using daily observations of the rise or fall of the level of mercury in a glass tube without the use of astronomical observations is truly remarkable, and illustrates how clear information can be determined from large and noisy data sets.

a)



b)

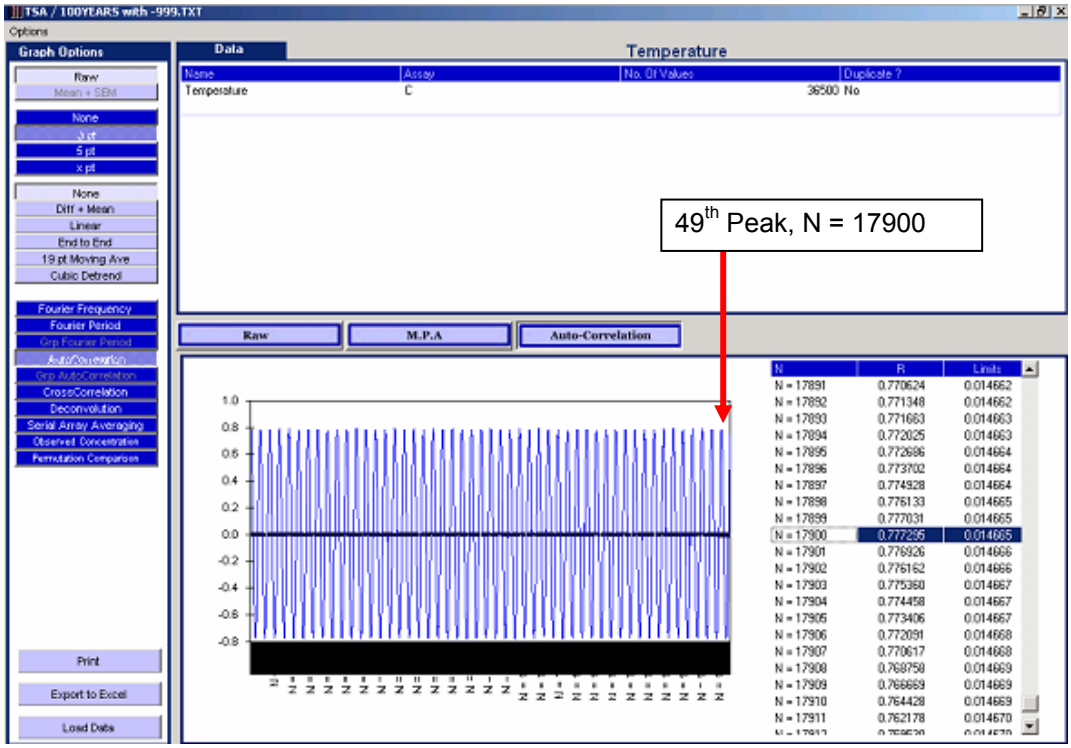
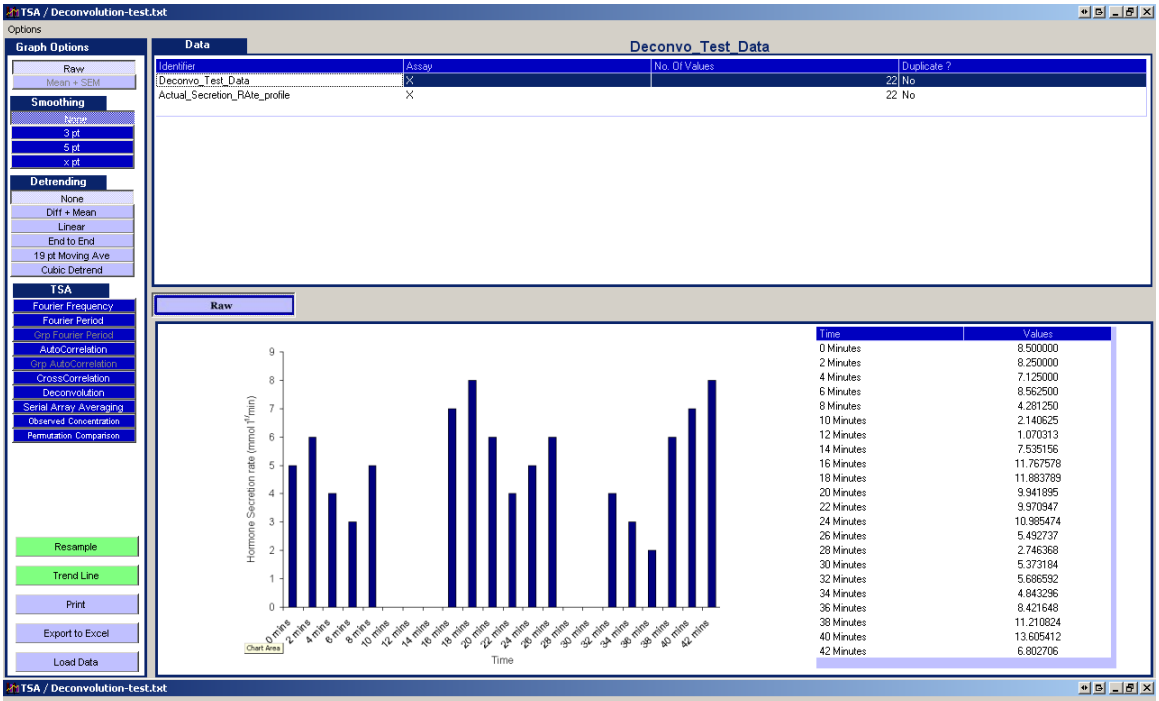


Figure 2.28 a) 100 years of daily temperature readings in Oxford, b) AC of a) maximal power at a period of 17900 days

2.3.3 Deconvolution validation

The validation of the Deconvolution method occurred with an artificial dataset of secretion rates for an imaginary hormone termed ‘Hormone X’. Hormone X was stated to have a half-life of 2 minutes. The secretion rate of Hormone X (Figure 2.29a) generated a hormone concentration profile (Figure 2.29b). To validate the Deconvolution method it required proving that it is possible to obtain the same secretion rate (Figure 2.29a) by measuring only the concentration of Hormone X (Figure 2.29b) and using the Deconvolution method.

a)



b)

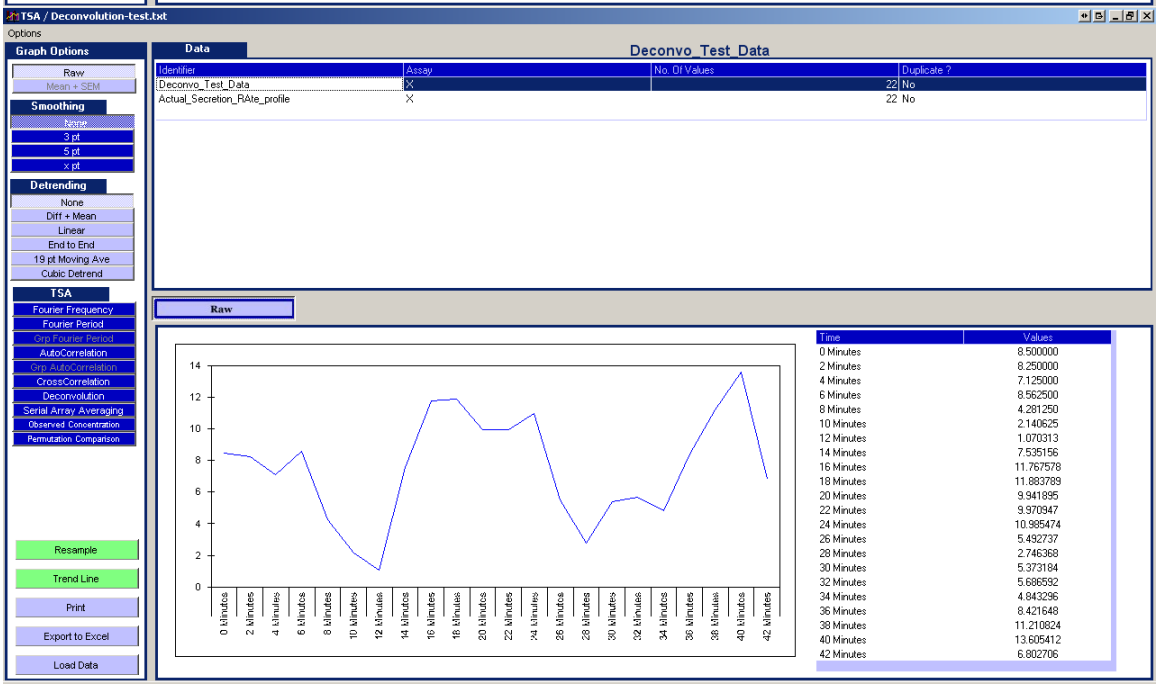


Figure 2.29- a) Secretion rate of hormone X b) Concentration profile generated for Hormone X using secretion rate and a half-life of 2 minutes

The Deconvolution method was applied to the concentration profile (Figure 2.29b). The estimated secretion rate was generated by Easy TSA (Figure 2.30b).The original secretion data (Figure 2.30a) and the estimated data (Figure 2.30b) obtained from the deconvolution method were exactly the same.

a)

b)

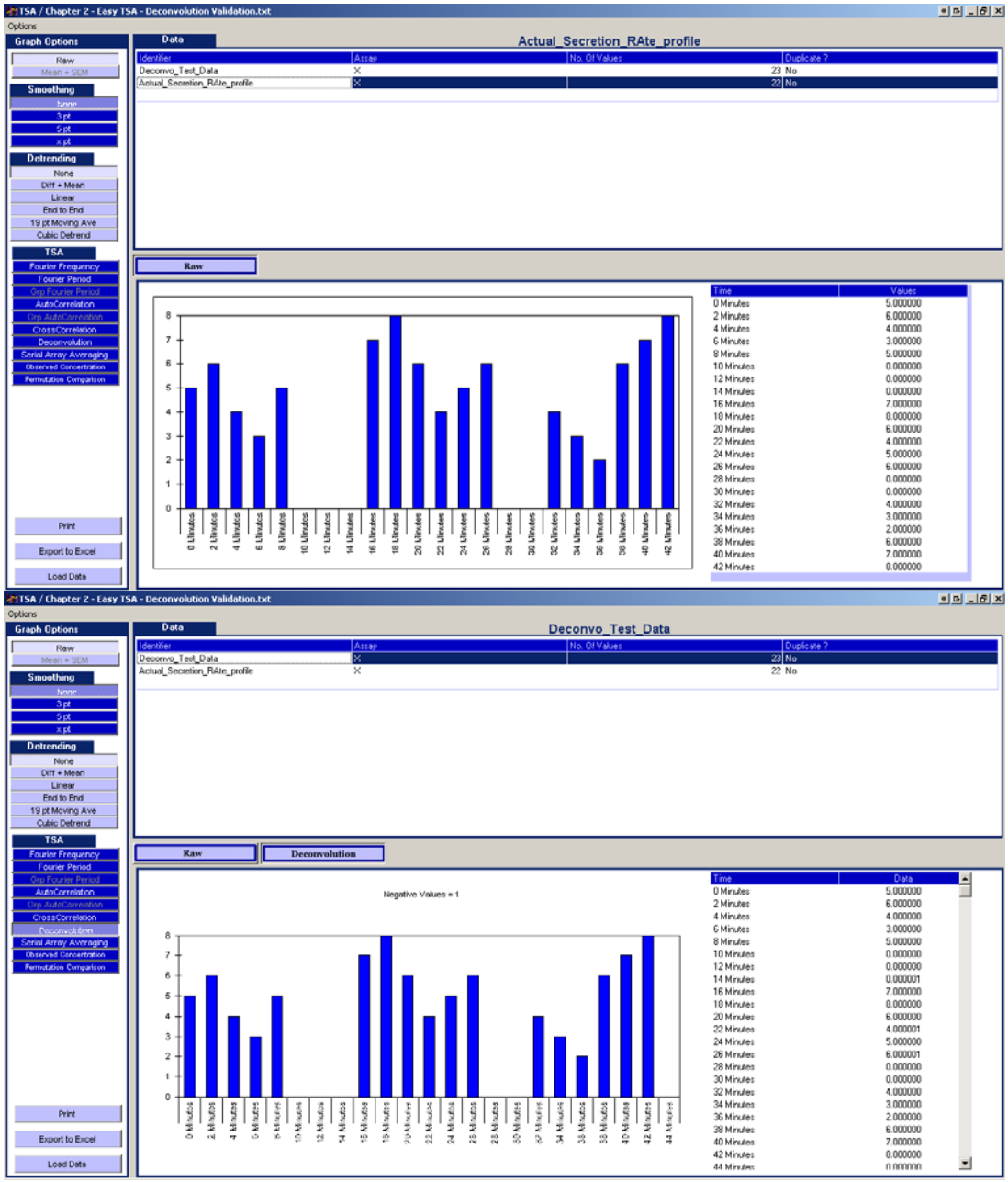


Figure 2.30 - a) Original Secretion Profile of test data b) Deconvoluted Secretion Profile generated from concentration information

2.3.4 Serial Array Averaging validation

Serial Array Averaging (SAA) validation required the construction of an additional artificial dataset. The data set contained information with 3 main pulses of hormone release (Figure 2.31). The three pulses are ratios of one another: The second peak is twice as high as the first peak and the third peak is half as high as the first peak. Each peak occurs over the same period.

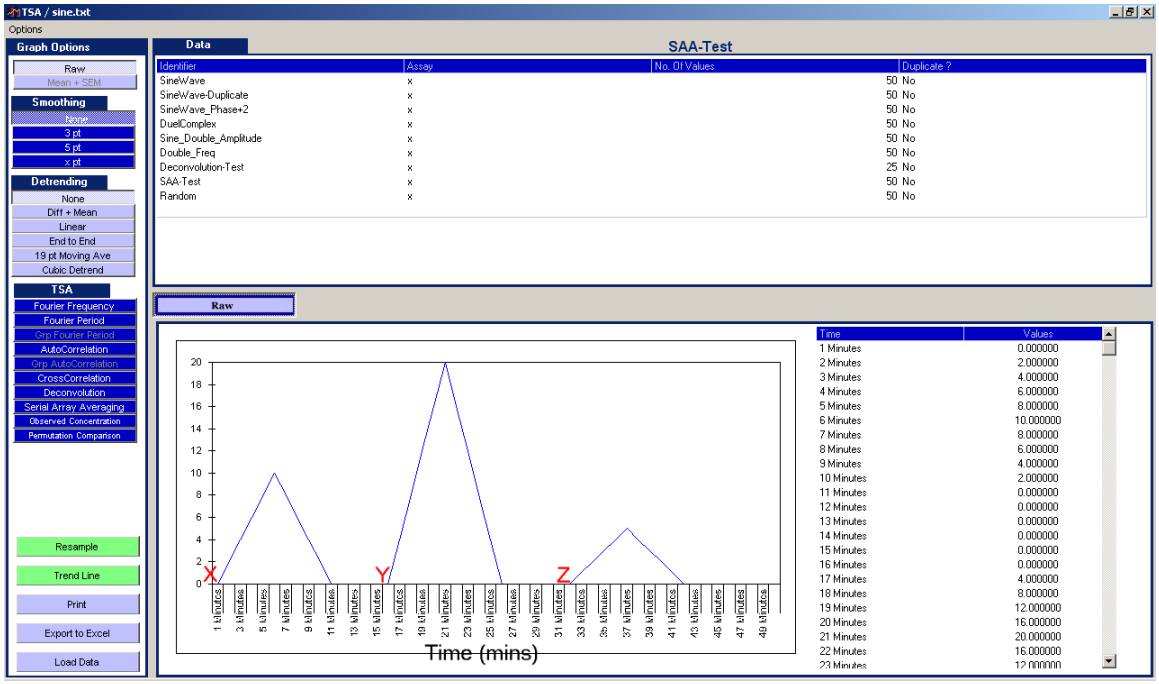


Figure 2.31 - Constructed dataset for Serial Array Averaging validation, using points X, Y, Z as index references.

The SAA was used with X, Y, Z being used as index points, the three peaks were selected for analysis. It performed as expected producing a graph with the correct mean and SEM (Figure 2.32). The forth data item, shown by the red arrow (period 3), was checked to confirm that the mean of the concentrations was calculated correctly and the Standard Error of the Mean bars generated were of the correct height (Table 2.2).

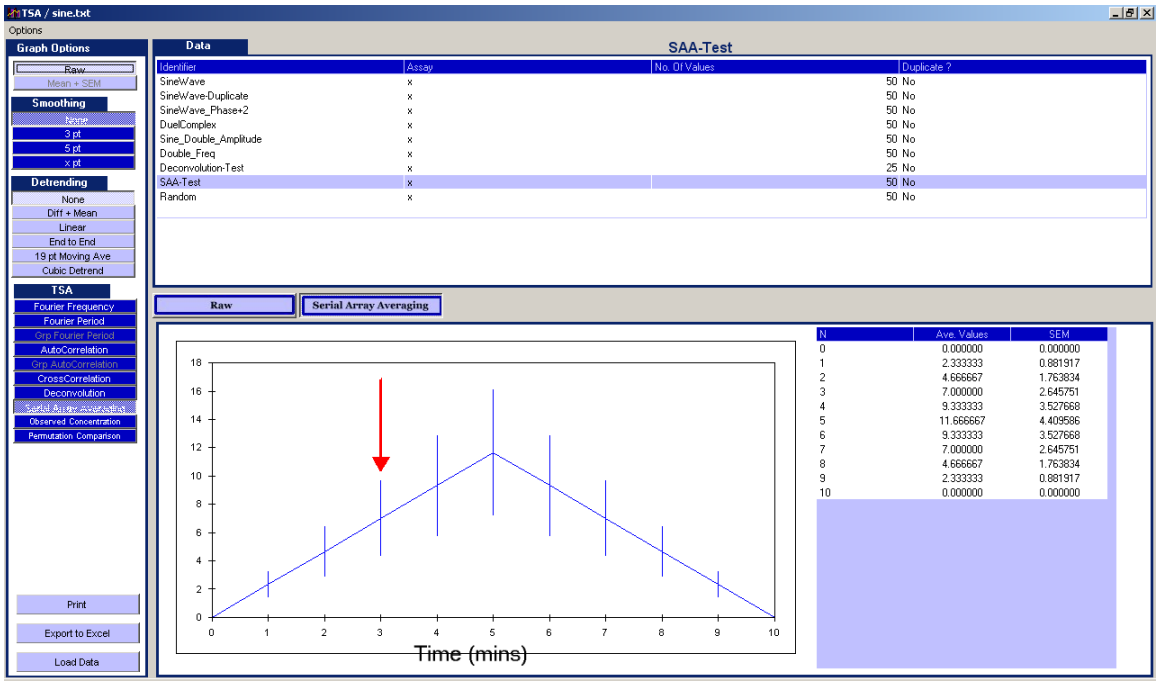


Figure 2.32 - Serial Array Average output of artificial data set

Period 3	
X_4	6
Y_4	12
Z_4	3
Mean	7
SEM	2.6

Table 2.2 - Mean and SEM for period three of the SAA output

2.3.5 Observed Concentration validation

The mathematical validation of Observed concentration (Probit analysis) was performed using excel. An example dataset of 100 data items of GH concentration was used and each step of the algorithm was checked by manually calculating the result using excel.

The dataset was first logged and then the frequency of the concentrations was assigned to 16 bins (Table 2.3a, b, c). The cumulative frequency was calculated for each bin (Table 2.3d). Limits were imposed on the GH concentration data any values in the bottom 2% or top 2% of the cumulative frequency distribution were not included for

analysis as there was insufficient data available to insure an objective analysis. The cumulative frequency was then converted to probits (Table 2.3d) by means of a look up table[66].

		Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9	Bin 10	Bin 11	Bin 12	Bin 13	Bin 14	Bin 15	Bin 16
a	**Upper limit of Bin**	0.12	0.25	0.37	0.49	0.61	0.74	0.86	0.98	1.11	1.23	1.35	1.47	1.60	1.72	1.84	1.97
b	**Mean of bin**	0.06	0.18	0.31	0.43	0.55	0.68	0.80	0.92	1.04	1.17	1.29	1.41	1.54	1.66	1.78	1.90
c	**Frequency**	40	3	1	2	4	2	3	4	2	0	4	6	11	6	8	4
d	**Cumulative Freq.**	40	43	44	46	50	52	55	59	61	61	65	71	82	88	96	100
e	**Probit**	4.75	4.82	4.85	4.9	5	5.05	5.13	5.23	5.28	5.28	5.39	5.55	5.92	6.17	6.75	N/A

Table 2.3 – Observed Concentration calculation table based on example data (available on the CD)

Using this information it was possible to plot in excel a chart of the mean of each bin vs. the probit (Figure 2.33).

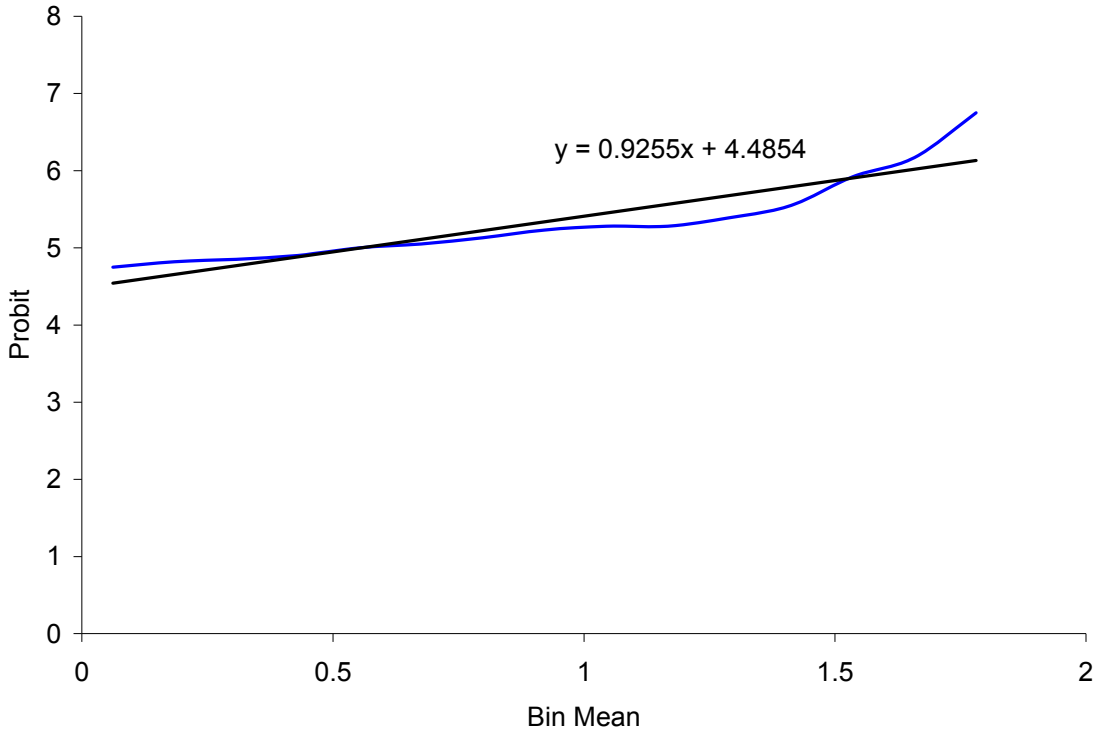


Figure 2.33 – Probit vs. Bin mean for growth hormone profile used in the Observed Concentration validation

A line of best fit was calculated and the equation found. The equation from the line of best fit was then used to calculate the OC5, OC50, OC95 of the dataset (EQ 2.11a, b, c).

OC5	= 10^((3.3551 – Intercept)/Gradient)	EQ 2.11a
OC50	= 10^((5 – Intercept)/Gradient)	EQ 2.11b
OC95	= 10^((6.6449 – Intercept)/Gradient)	EQ 2.11c

Easy TSA was used automatically to calculate the OC and the results were compared (Table 2.4). The validation was successful.

	OC5	OC50	OC95
Excel	0.06	3.60	215.44
Easy TSA	0.06	3.60	215.44

Table 2.4 – Comparison of OC validation results from Easy TSA and excel

2.3.6 Runs Analysis validation

Runs analysis was validated using the 50-point sine wave dataset used in FT and AC validation (Figure 2.33a). The Null hypothesis was that the dataset was random. The first step was to use a non-random dataset and hopefully reject the null hypothesis. The second step was to ascertain at what threshold the signal could not be determined from the noise and was judged to be a random process.

a)

b)

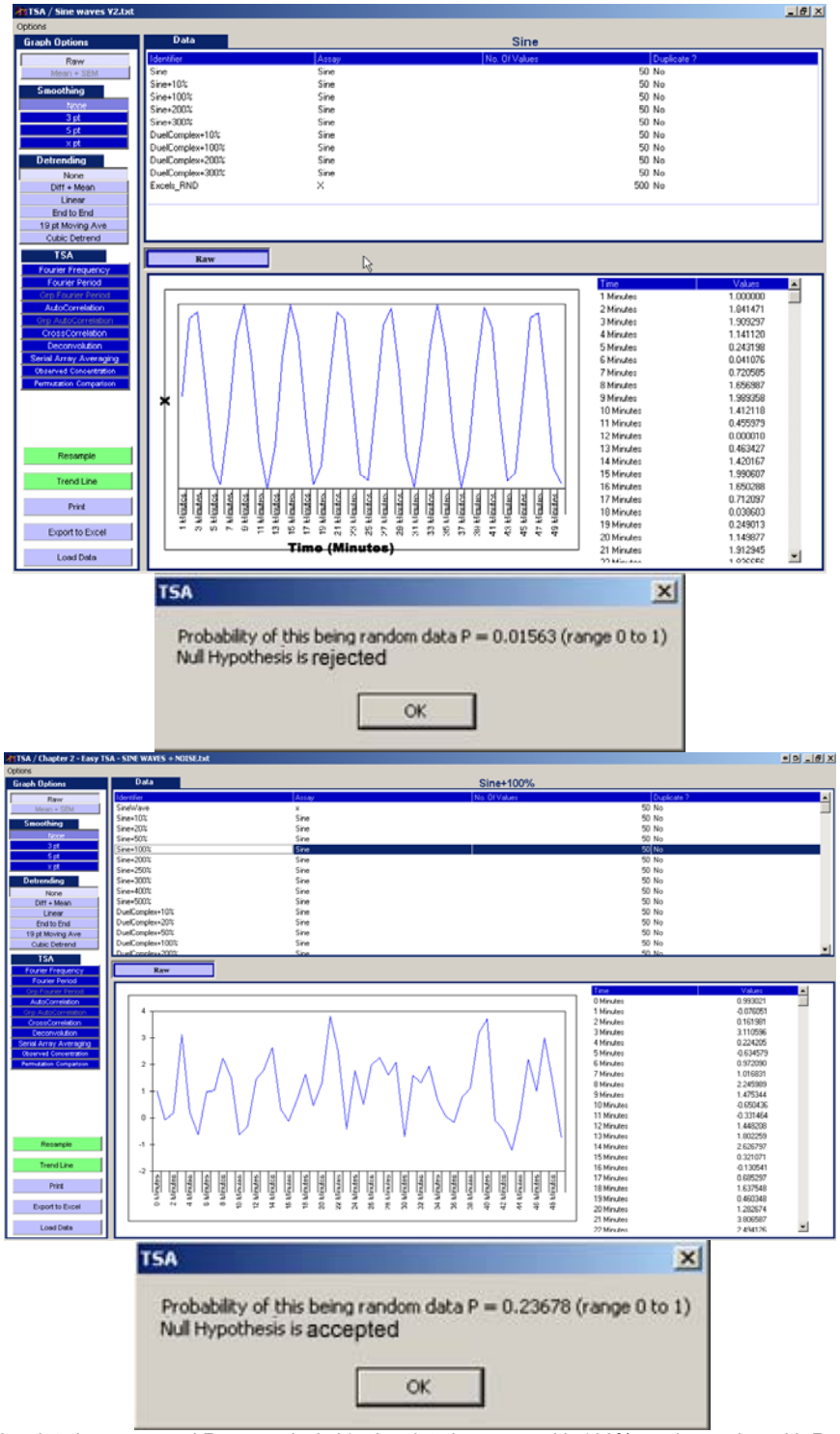


Figure 2.33 – a) 50-point sine wave and Runs analysis b) 50-point sine wave with 100% random noise with Runs analysis

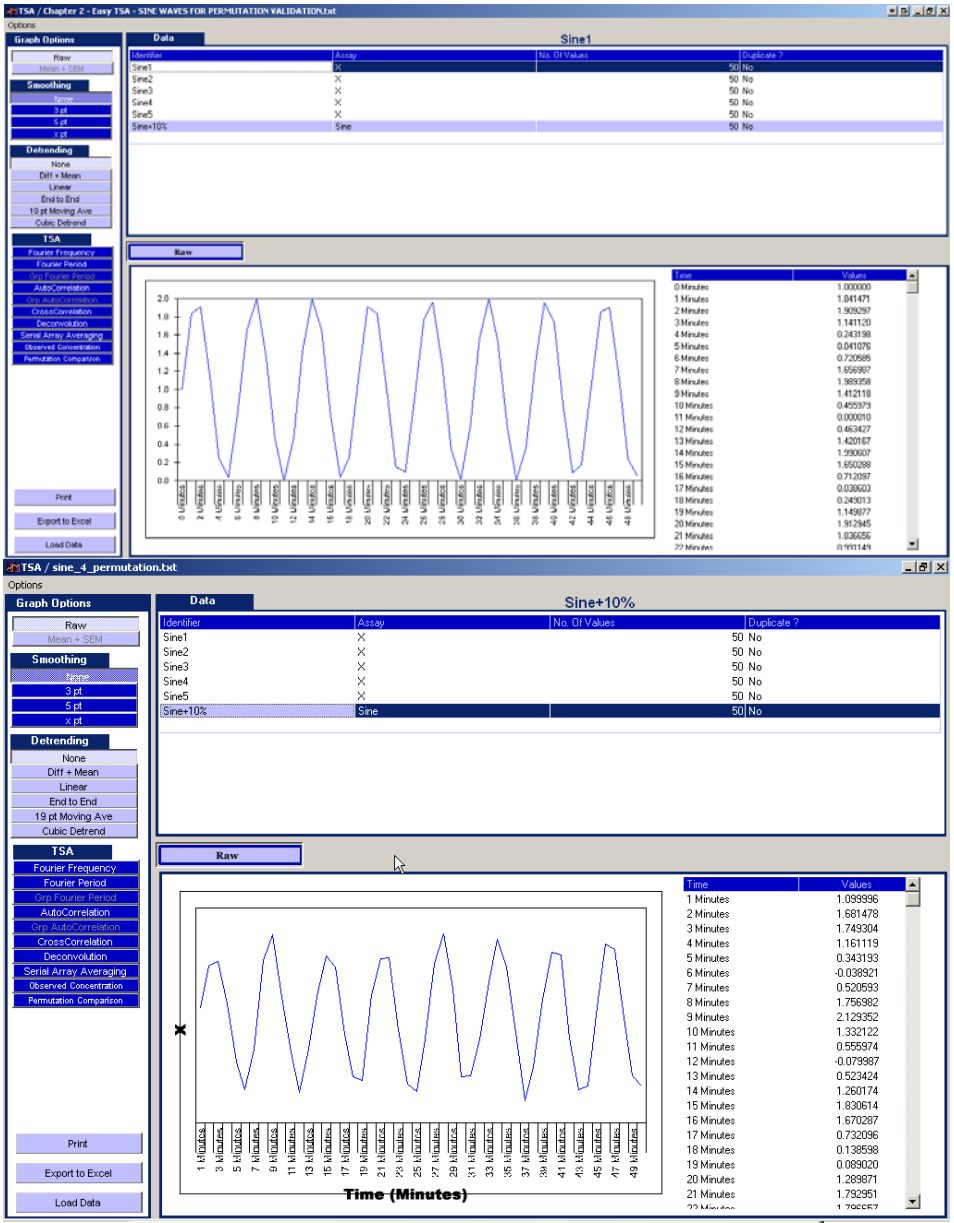
The Runs analysis detected the non-random nature of the sine wave (Figure 2.33a) $p<0.01$. Between 90 and 100% noise (Figure 2.33b) had to be added to the original data for the Null Hypothesis to be accepted and the p value become non significant (0.23) as the information could no longer be determined from the noise. Runs analysis can be problematical as it is dependent upon noise adjacent to individual observations. TSA uses autoregressive processes on the entire dataset.

2.3.7 Permutation analysis validation

Permutation analysis validation needed to reject that the null hypothesis H_0 was true. This equates to the sentence: 'Any patterns appearing in the data are due to chance alone'. If the null hypothesis was rejected then the result was determined to be below the threshold of significance (typically $p<0.05$) and the data was probably random. If the null hypothesis was true then the data was determined to contain non-random information. Permutation analysis could be performed to determine if information is present in a dataset prior to TSA. For the comparison the 50-point sine waves were used. Group 1 consisted of five 50-point sine waves (Figure 2.34a) – identical to those used for validation in other methods and Group 2 was comprised of 1 50-point sine wave with 10% noise added (Figure 2.34b).

a)

b)



c)

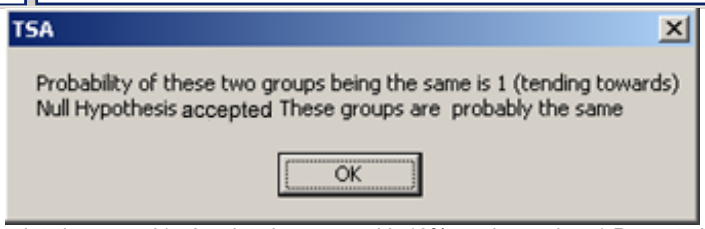


Figure 2.34 – a) 50-point sine wave b) 50-point sine wave with 10% random noise c) Permutation analysis results

The permutation algorithm sought to determine a p value for the non-randomness of the datasets. The Null hypothesis was accepted therefore the datasets where determined to

be probably the same (i.e. there was no significant difference between these groups). By using the Easy TSA program on the accompanying CD, it is possible to recreate these tests and results to confirm this.

2.3.8 Selection of appropriate methods

The selection of the correct detrending, smoothing and TSA routine can influence the outcome of an investigation. There is no one correct choice for any given situation and TSA should only be used with an understanding of the techniques, what they show and when they should be applied. Easy TSA like other commercially available software (SPSS, SAS, EXCEL, etc) allows a user to select and perform inappropriate analysis as it assumes that a basic understanding is present. Easy TSA will ship with a full explanation of the smoothing, detrending and TSA methods that have been set out in detail in the thesis to reduce the subjective element of method selection.

2.4 Clinical application of Easy TSA

Once the validation of Easy TSA was completed, the software could be used in clinical research. The TSA's that were included enabled a research project that sought to recognise and diagnose physiological problems and conditions using TSA of hormonal concentrations. The software was used in a clinical study to examine the relationship between insulin and obesity in people without diabetes.

The following data were originally collected by Helene Flax and David Matthews. The analysis was performed by the author using the Easy TSA program. The results have

been published as a short report in the journal Diabetes, Obesity & Metabolism after peer review[73].

The relationship between insulin secretion, insulin resistance and obesity is well recognised[74-76]. Rabinowitz and Zierler[77], in 1962,were among the first groups to demonstrate insulin resistance in obese subjects, by demonstrating that equivalent insulin concentrations produced less glucose uptake by adipose tissue in obese subjects compared with controls[78]. A variety of groups[79-83] have demonstrated that hyperinsulinaemia and an exaggerated insulin response are characteristics of the obese state. It has been shown that after significant weight loss insulin resistance returns towards normal[84-87]. The improved insulin sensitivity observed has been established by hyperglycaemic and euglycaemic clamp techniques[88, 89].

Regular oscillations of plasma insulin secretion have been shown to confer various physiological advantages. These include heightened receptor responsiveness[46], increased receptor number[90], enhanced efficiency of glucose disposal[90] and improved amplitude modulated signalling to the liver from the portal system[91, 92]. Pulsatile insulin secretion exerts greater metabolic effects than purely continuous hormone delivery[93]. In obese primates Hansen et al[94], demonstrated no change in a Rhesus monkey model. When overfed for a period of 39 to 100 days an increase in their body weight by 20 to 67% above baseline occurred, and plasma sampling demonstrated the presence of regular oscillations in insulin secretion with no change in frequency or amplitude although the insulin concentrations were higher. The same group studied plasma insulin oscillations in eleven obese human subjects[95]. The results supported the findings in the primates, concluding that there was no difference in the frequency and absolute amplitude of their plasma glucose and insulin oscillations compared with a

group of 'normal' weight controls. Zarkovic et al[96] studied the effects of weight loss on pulsatile insulin secretion in obese subjects. The aim was to establish a link between the variability of insulin secretion and changes in insulin sensitivity. The group demonstrated a correlation between the secretion rate of insulin and insulin sensitivity both before and after weight loss.

This study examined oscillatory insulin profiles in non-diabetic subjects and sought to determine the effect of BMI on the secretory pattern of insulin. The TSA methods used, to detect the pulsatile insulin release, were the Fourier Transform (FT) and Auto-correlation (AC).

The FT analysis examines the insulin oscillations from the datasets by detecting both the frequency of the insulin signalling and the strength of the signal[68]. Inter FT comparisons are accomplished by examining the relative power of the signals as a percentage of the overall within-FT power.

To establish the presence of regularly recurring waveforms, within the data array, AC techniques were used.

2.4.1 Methods & Subjects

2.4.1.1 Subjects

Fifty-six non-diabetic subjects (28M, 28F) were studied using data from a number of protocols analysing insulin & Glucose interactions. Diabetes was excluded and defined by fasting plasma glucose greater than 6.0 mmol/l. No subject had a history of diabetes or first-degree relatives with diabetes. Body mass was calculated for each participant

([weight in kg]/ [height in m] ²). The subjects were split by the median BMI of the cohort into two groups. Subjects were in good health and were receiving no medication. There was a significant difference between the subjects demographics (Table 2.5).Group 2's subjects had a higher median age.

Table 2.5 - Subject demographics in the clinical application of Easy TSA

ID	BMI Range	N	Median BMI (Range)	Median Plasma Insulin Concentration (range)	Median Age (Range)
Group 1	<22.8	28	21.4 (19.9 - 22.0)	30.7 (17.7 - 47.6)	22.5 (21.0 - 26.0)
Group 2	>=22.8	28	27.5 (24.7 - 31.8)	42.6 (30.7 - 67.0)	35.5 (23.0 - 47.0)

The difference of the median age of Group 1 in comparison to Group 2 was 13 years. Both insulin secretion amplitude and frequency have been reported to be different in older populations (age >80year)s[97]. Because of the differences between the groups, an additional analysis was required on an age matched subset to determine if the result were due to BMI or if Age was a contributing factor.

2.4.1.2 Methods

Subjects were admitted to a metabolic ward in the morning after an overnight fast from 2200h the night before. On the morning of the study, assessments of BMI's were made. Body weight was measured after voiding, on a standard Avery body-weight scale in indoor clothing and without shoes. Body height was measured to the nearest 0.5 cm with a stadiometer. The study conformed to the declaration of Helsinki and was granted ethical approval by the Central Oxford Research Ethics Committee (COREC no 2227). The clinical investigation was performed by Helen Shipiro and David Matthews at Oxford Centre for Diabetes, Endocrinology and Metabolism in 1984.

On the morning of the studies, basal blood sampling was undertaken after a double-lumen 21G Teflon catheter (Venflon 2, Viggo, Helsinborg, Sweden) was placed under local anaesthetic into a distal forearm vein for continuous automated blood sampling[98]. The hand was heated to 'arterialise' the sampled blood[99, 100]. Blood was sampled through the inner lumen at a constant rate of 2ml.min-1 by a rotary pump (Watson Marlow H.R. flow inducer, Cornwall, UK) into a fraction collector (Redirac, LKB, Bromma, Sweden) which moved forward at minute intervals for a period of up to 120 min to collect samples for insulin assays. The samples were prevented from clotting by the use of calcium-heparin in saline (1250 I.U.ml-1) pumped through the outer lumen in a retrograde manner at a constant rate to give a final sample concentration of 60 U heparin ml-1 blood. Heparin was infused through the outer lumen at 1% of the blood extraction rate. Samples were kept on ice prior to centrifugation and separation immediately. Plasma was then stored at -20ºC prior to batch analysis.

2.4.1.3 Assays

Plasma insulin was measured by radioimmunoassay and using Novo Human Mono-component insulin as standard and mono I-25 (Tyr A14)-human insulin as tracer (Novo Biolabs, Cambridge, UK). Within-batch coefficient of variation by the method of duplicates was 12.5%. Samples from participants were analysed in the same batch.

2.4.1.4 Statistics

Normality of the data was assessed by the Kolmogorov-Smirnov test. Data that were normally distributed were analysed using mean values, and an independent samples t-test was used for parametric comparisons. Data not normally distributed have been presented as median values, and Mann-Whitney 2 independent sample t-test was used to study the differences.

2.4.1.5 Analysis

The signal-noise ratio was improved by using Easy TSA to perform pre-processing of the data before analysis. The assays were measured in duplicate (Figure 2.35a). Easy TSA's 3-point moving average[101] (Figure 2.35b) was applied before analysis. This allowed six data items to contribute to each point used in the analysis, thereby reducing the effect of outliers in the data. Using the built-in cubic polynomial detrending[15] in Easy TSA (Figure 2.35c) confounding factors such as a long-term data trends were removed ensuring short-term oscillations in the data could be analysed.

a)

b)

c)

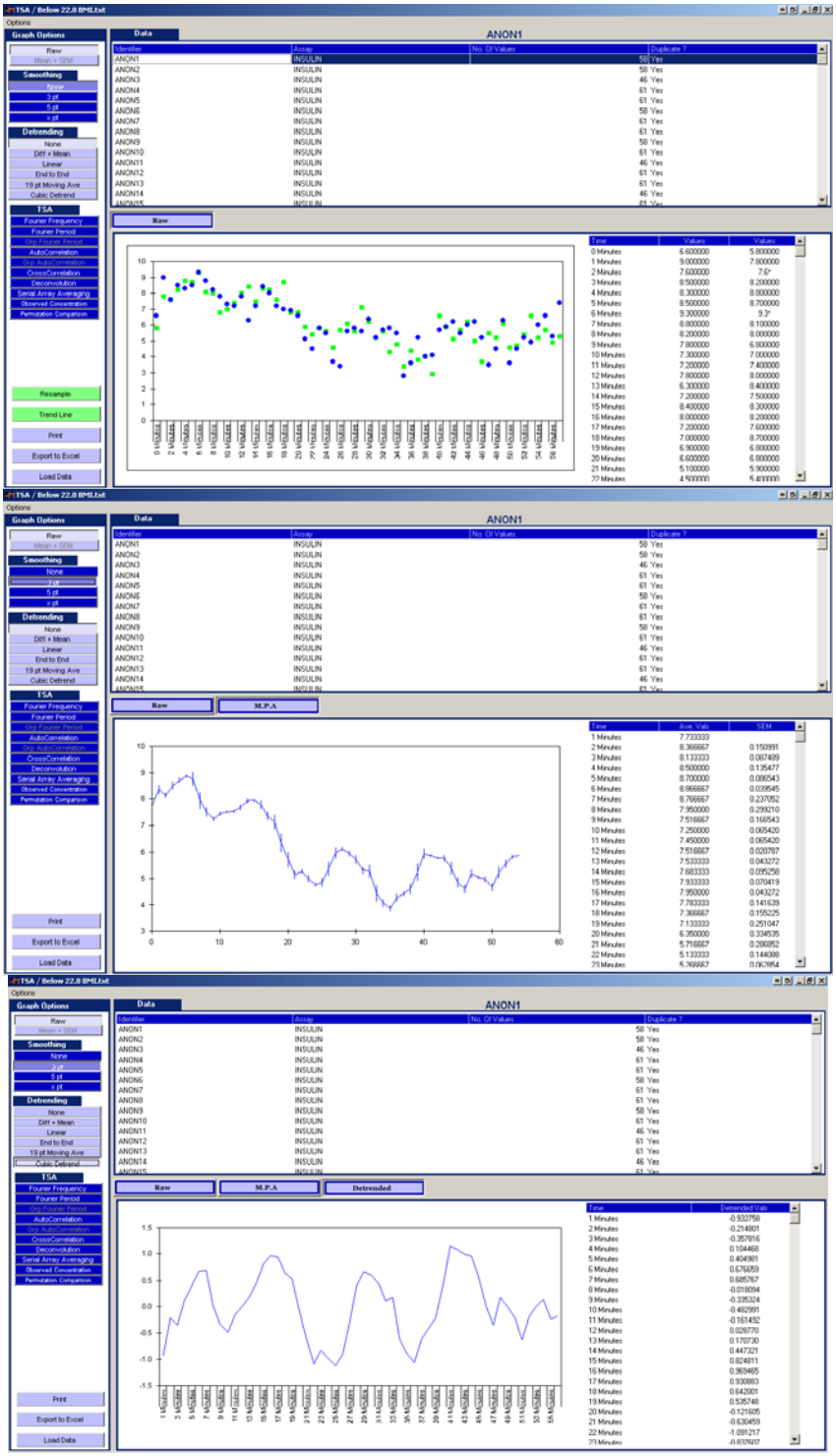


Figure 2.35 – a) An example insulin profile from a subject with a BMI <22.8 b) 3-point ARMA of insulin profile c) 3-point ARMA and cubic detrended insulin profile

Following the pre-processing of the data by Easy TSA an examination of oscillatory data in terms of specific attributes such as amplitude and frequency could be undertaken using two contrasting analytical approaches from the Easy TSA program.

The Fourier Transform (FT) method was used to measure the power of the insulin oscillations of varying periodicities. A relative (normalised) FT was calculated for each subject and each groups FT's were collated and a mean normalised FT produced. The Auto-correlation (AC) technique was used to quantify reoccurrences of insulin waveforms within the data. An AC was performed for each subject and a mean AC was produced for each group. Area under the curve of the absolute FT was calculated automatically by Easy TSA (by the trapezoid rule). The area under the curve allowed a comparison of the total energy/amplitude of the component signals at each frequency to be compared.

Comparisons between the waveforms shape of the time series analysis results was performed by examination of the probability of the standard error of the mean (SEM), at any discrete time, encompassing the SEM of the comparative group.

2.4.1.6 Subset Analysis

To determine if age or BMI alone contributed to the results analysis was performed on a subset of the data to compensate for the disparity in age between the two groups. The age range 20 to 40 years was selected generating a subset of 43 patients. The data was contrasted first by the median age of 23 years and then by the median BMI of 22.

2.4.2 Results

There was a positive correlation between BMI and fasting insulin concentration (Figure 2.36). As the BMI of the subjects increased the fasting level of insulin observed increased, but the highest BMI's generated more outliers. The positive correlation was evident in both groups.

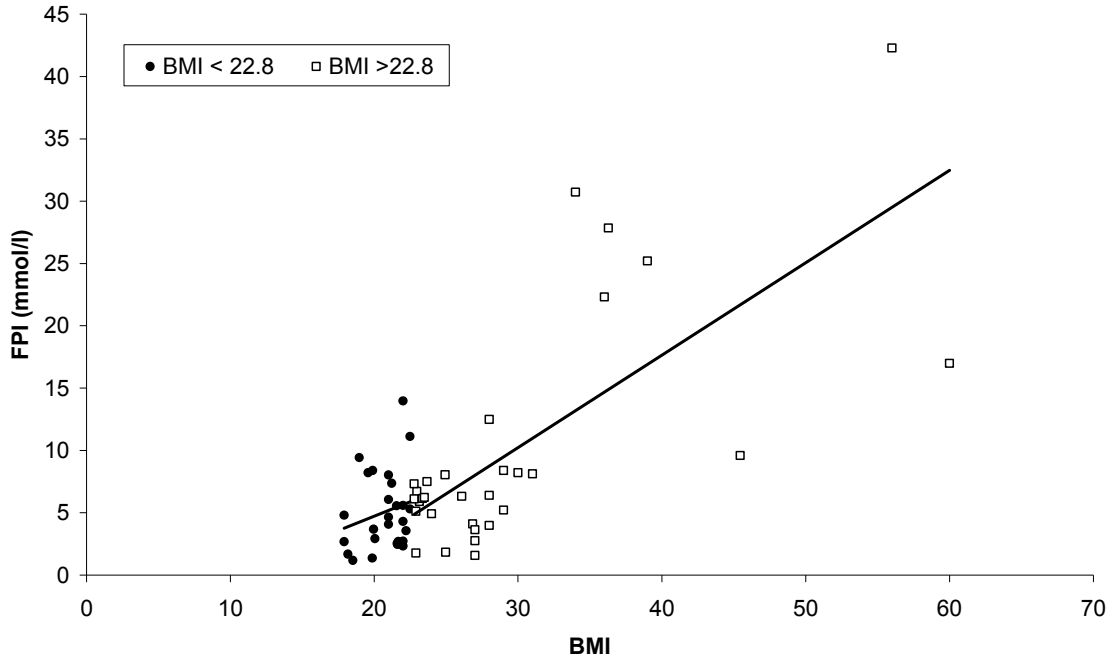


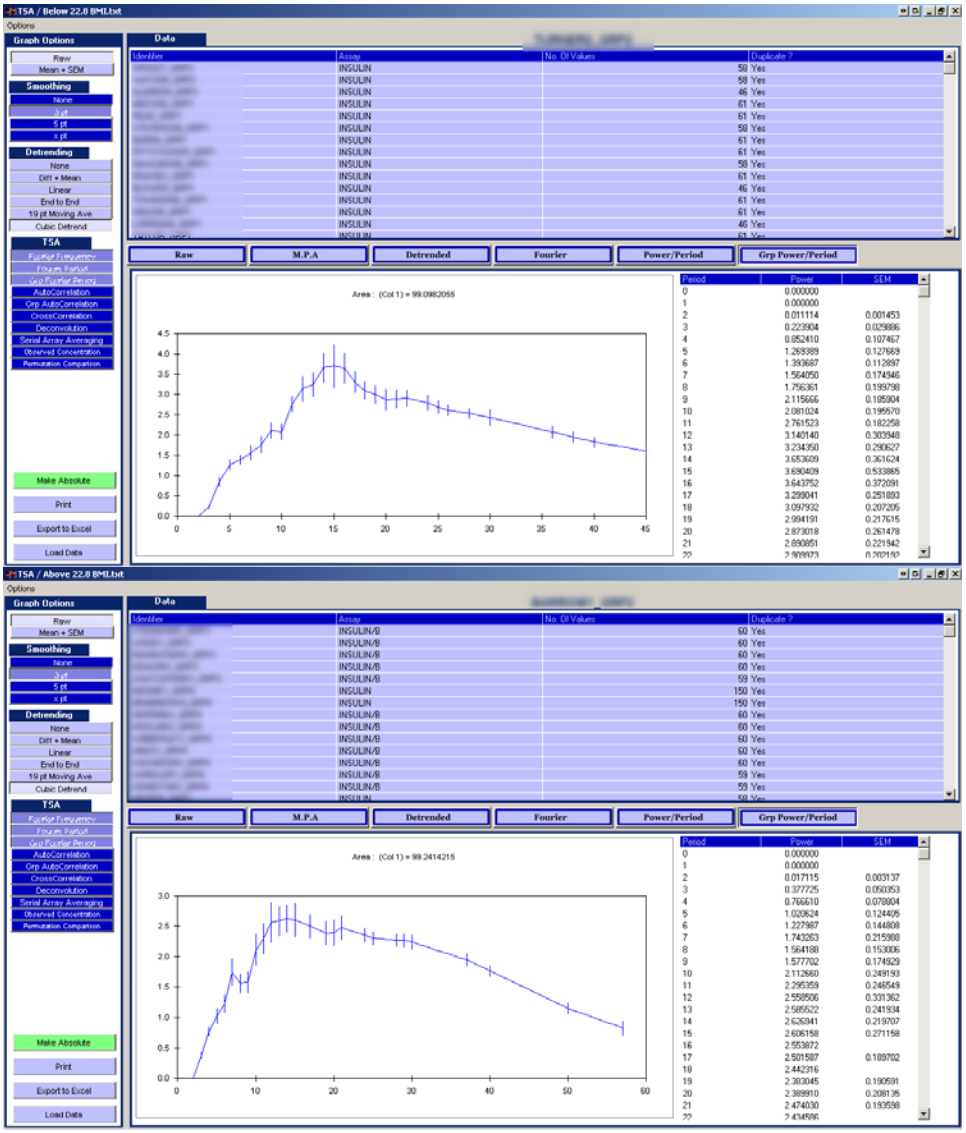
Figure 2.36 - Correlation between BMI and FPI

2.4.2.1 Fourier Transform analysis

The mean normalised FT (Figure 2.37a) of Group 1 produced a dominant spectral peak at 15 minutes (mean percentage power=4.03; SEM=0.60). The total absolute spectral power calculated by area under the curve was 2.80. The mean normalised FT of Group 2 (Figure 2.37b), did not demonstrate dominant spectral peaks over the whole range, total absolute spectral power was 3.98. The shape comparison of the FT's SEM between Group 1 and Group 2 demonstrated a significant difference at the 15-minute periodicity ($p<0.0256$) as they did not overlap (Figure 2.37c).

a)

b)



c)

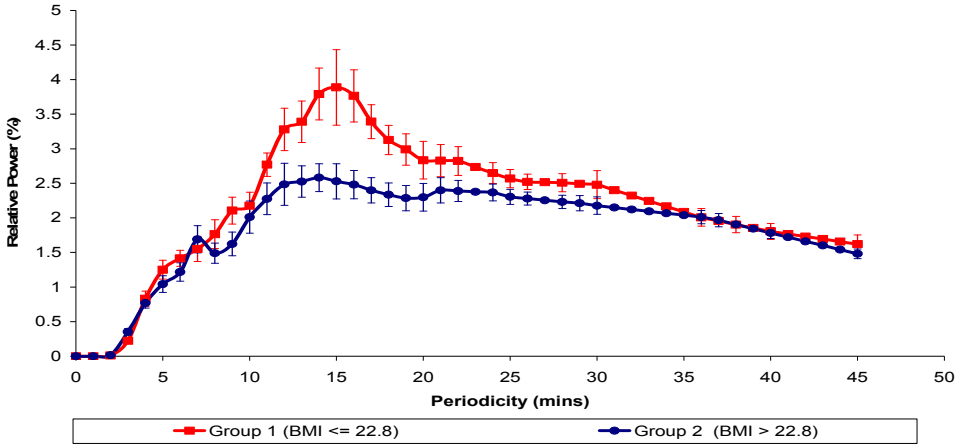
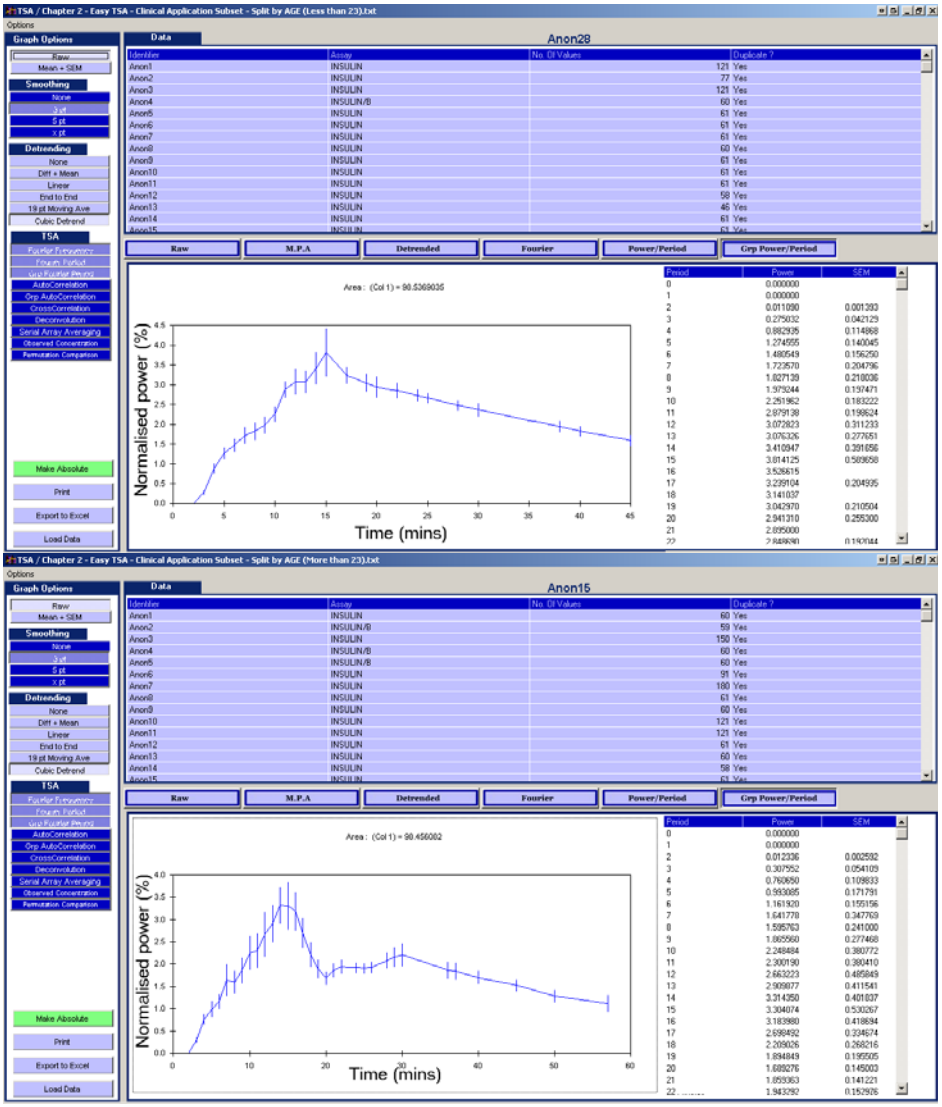


Figure 2.37 – a) Normalised FT of subjects with a BMI <22.8 after a 3-point moving average and a cubic detrending b) Normalised FT of subjects with a BMI <22.8 after a 3-point moving average and a cubic detrending c) Normalised Fourier Transforms of Group 1 & Group 2 overlaid for comparison of waveform.

FT analysis of the age matched cohort was performed to determine if the significance observed at 15 minute periodicity was attributable to age. The cohort was divided by the median age (23 years) and FT analysis was performed after pre-processing. In the mean normalised FT a periodicity of 15 minutes was detected in those < 23 years (Figure 2.38a). In mean normalised FT of those >=23 years a periodicity of 14 to 15 minutes was observed (Figure 2.38b). The SEM's of the two groups encompassed one another at the maximal periodicity indicating no significant difference between the two groups when analysed by FT (Figure 2.38c). The age matched cohort analysis demonstrated that the periodicity was not affected by the age of the subjects.

a)

b)



c)

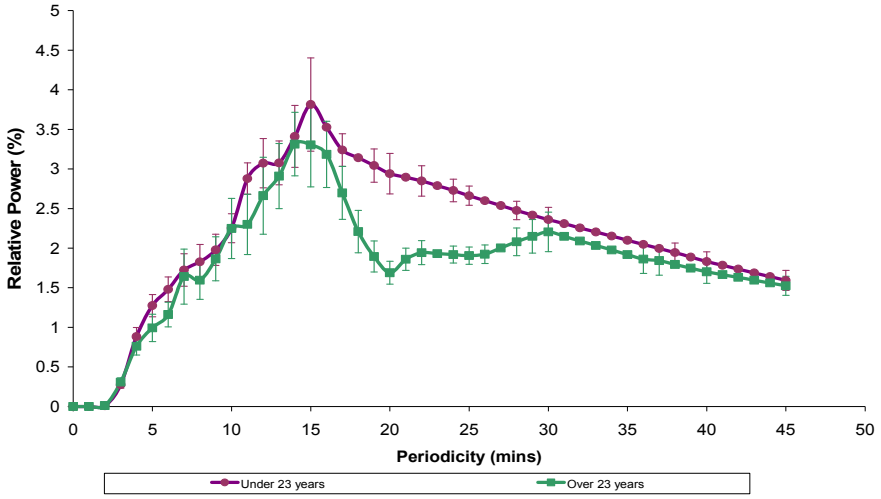
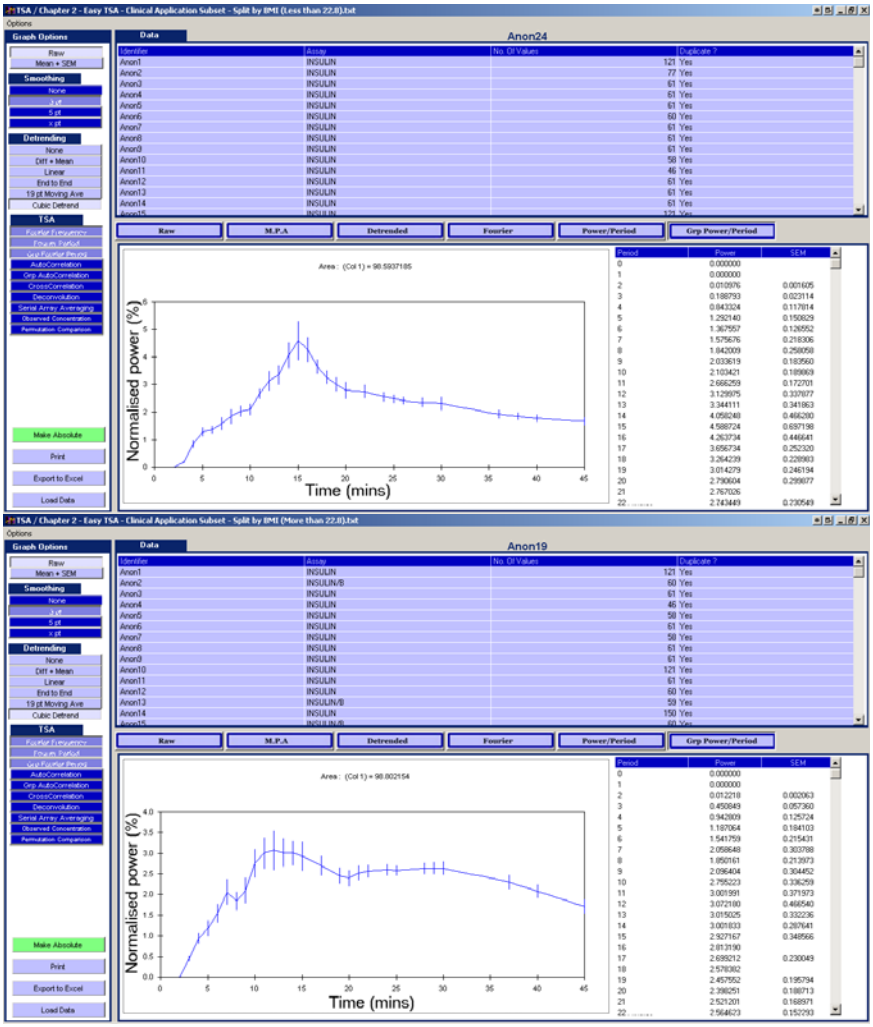


Figure 2.38 – a) Mean normalised FT of age matched subset of those less than 23 years old b) Mean normalised FT of age matched subset of those over 23 years old c) Overlaid mean normalised FT's for SEM waveform comparison.

The age matched cohort was again separated into two groups by the median BMI (22 kg/m^2) for FT analysis. The mean normalised(relative) FT of those with a BMI <22 kg/m^2 (Figure 2.39a) demonstrated a maximal periodicity of 15 minutes. The mean normalised FT of those with a BMI >= 22 kg/m^2 demonstrated no maximal periodicity throughout the profile (Figure 2.39b). The two FT's were overlaid for the SEM comparison at 15 minutes. The two FT's did not intersect (Figure 2.39c) at the 15 minute periodicity ($p < 0.0256$). The age matched cohort analysis demonstrated that the periodicity was affected by the BMI of the subjects.

a)

b)



c)

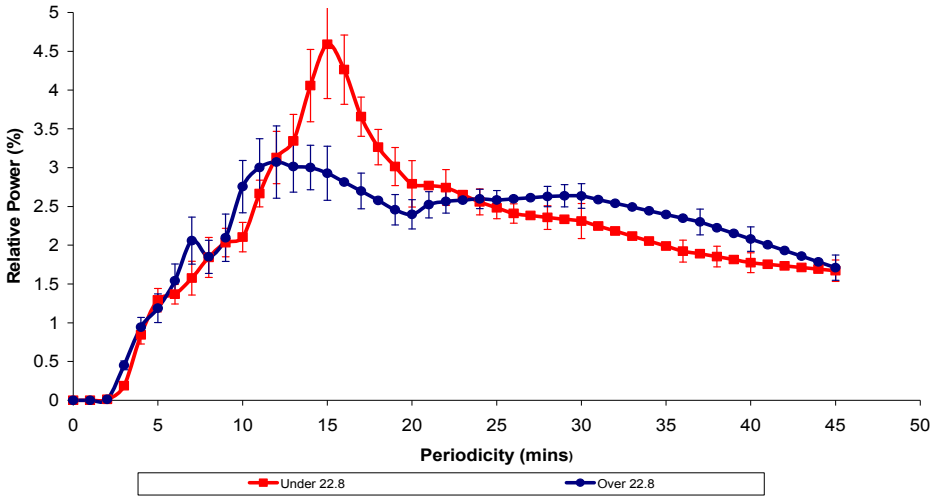
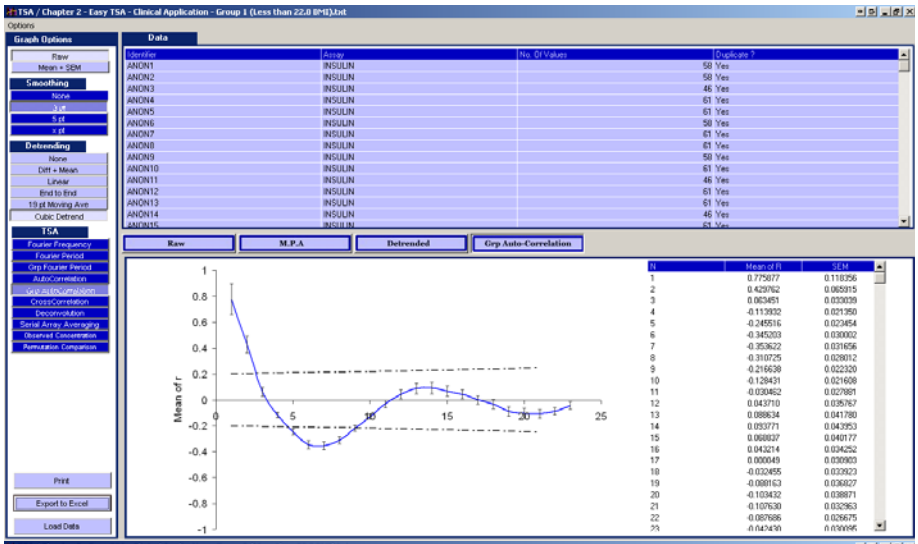


Figure 2.39 – a) Mean normalised FT of age matched subset of those with a BMI <22 b) Mean normalised FT of age matched subset of those with a BMI >= 22 c) Overlaid mean normalised FT's for SEM waveform comparison.

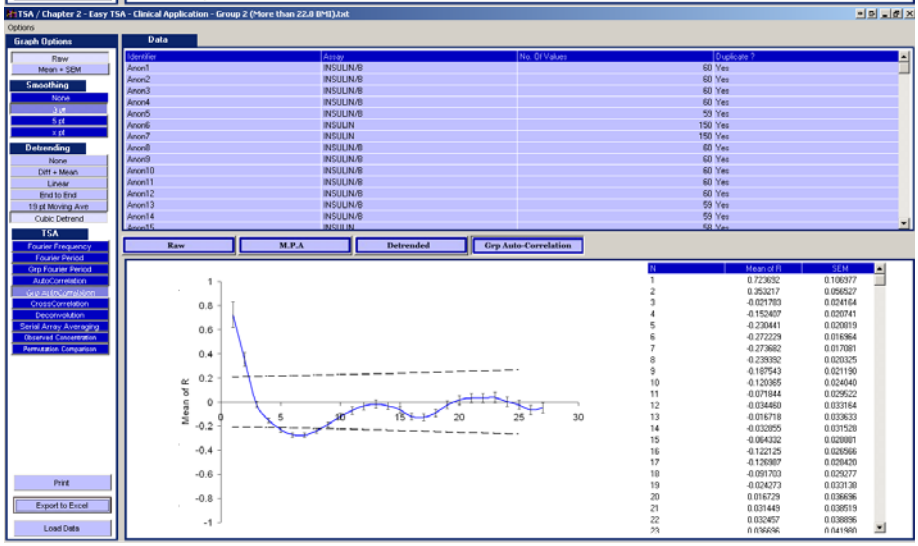
2.4.2.2 Autocorrelation analysis

AC analysis of Group 1 (Figure 2.40a) demonstrated at the 6 minute lag point a significant ($p<0.05$) negative correlation indicating the data were oscillatory. The negative correlation was a significant first nadir at ½ a periodicity. A regular insulin pulsation (1^{st} periodicity) at the 14 minute lag point (R=0.11; SEM=0.05) this did not reach positive significance (as delineated by the dotted line). The mean AC of Group 2 (Figure 2.40b) reached significance ($p<0.05$) with a negative correlation at the 6 minute lag point. Further correlations demonstrated a weak pattern of insulin secretion and a small negative correlation at the 14 minute mark (R= -0.06; SEM=0.03). There was, however, a significant difference ($p<0.0256$) between the autocorrelograms when they were compared by SEM waveform(further information in section 2.4.1.5 Analysis) (Figure 2.40c)

a)



b)



c)

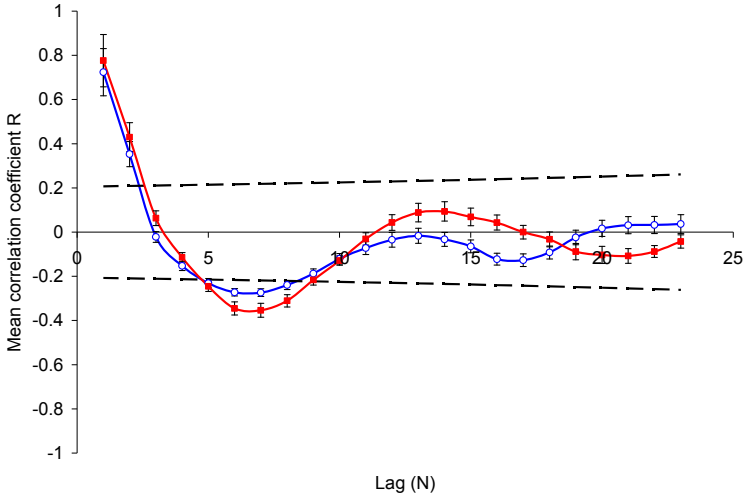
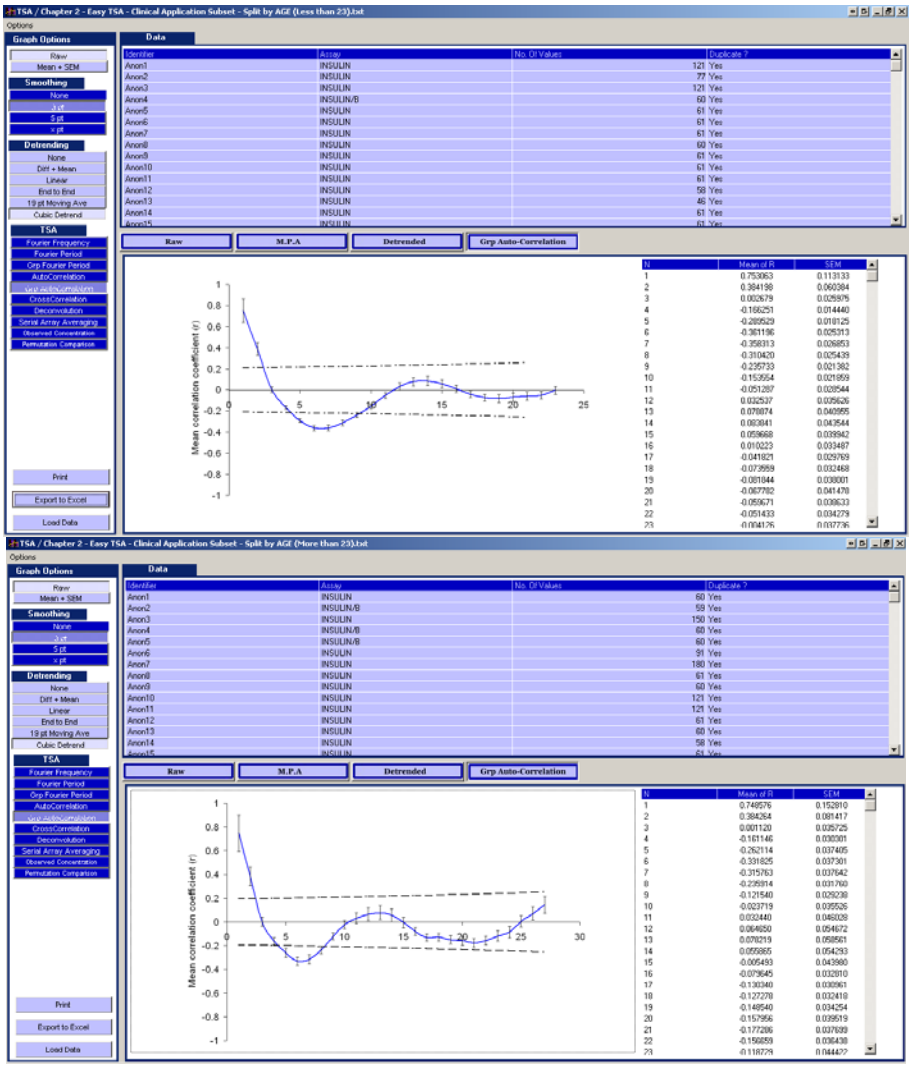


Figure 2.40 – a) Mean autocorrelation for Group 1 (BMI <22.8) with dotted significance bars b) Mean autocorrelation for Group 2 (BMI >=22.8) c) Autocorrelograms overlaid for SEM waveform comparison. Less than 22.8 BMI (■), Greater or equal to 22.8 BMI (○)

Analysis was performed on the age matched cohort (age range 20-40 years). The cohort was split by median age (23 years) into two sets. The first set (<23 years) showed a reoccurring waveform at 14 minute lag (Figure 2.41a). The second set (>=23 years) showed a significant ½ period negative correlation at the 6 minute lag point and a non-significant 1 period waveform at 13 to 14 minute lag (Figure 2.41b).The mean ACs did not reach significance as they did not cross the dotted lines that delineate the $p<0.05$ area. The mean ACs when plotted together demonstrated no significant difference in the waveform analysis (Figure 2.41c). The age matched cohort analysis demonstrated that the reoccurrence of oscillations was not affected by the age of the subjects.

a)

b)



c)

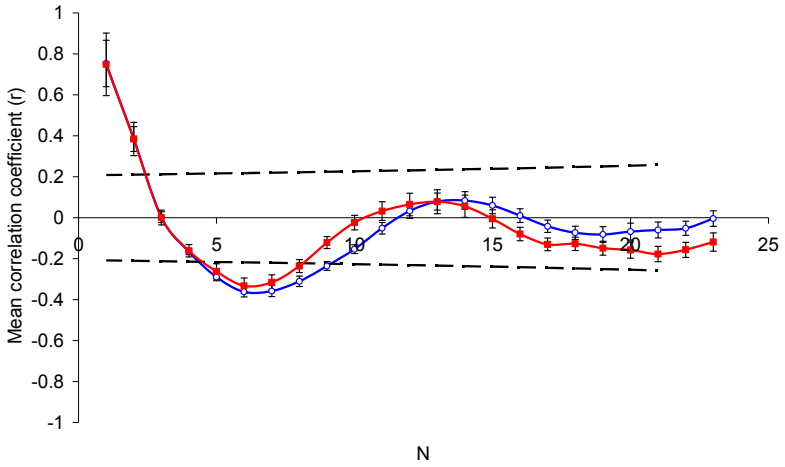
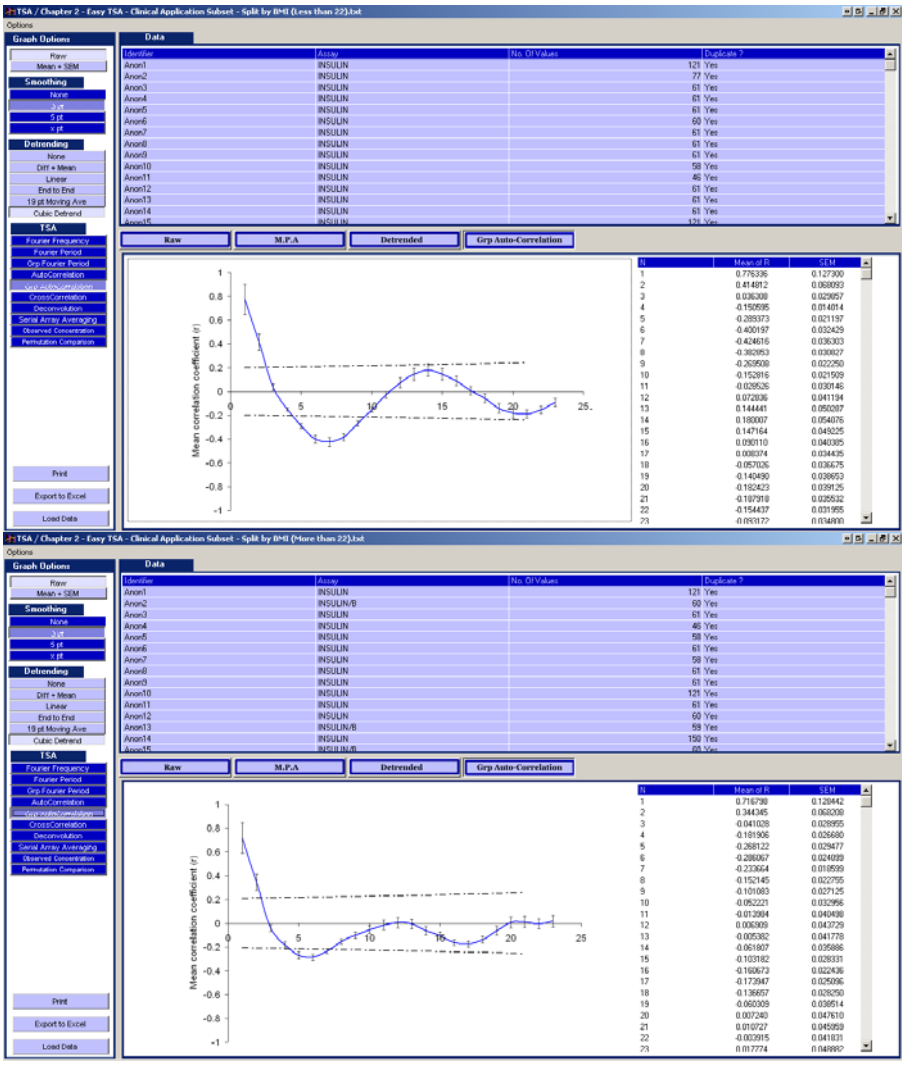


Figure 2.41 – a) Mean autocorrelation for age matched cohort .Group < 23 years - with dotted significance bars b) Mean autocorrelation for age matched cohort. Group >=23 years - with dotted significance bars c) Autocorrelograms overlaid for SEM waveform comparison. Less than 23 years (■), Greater or equal to 23 years (○)

The age matched cohort subjects were split by median BMI (22 kg/m^2) into two sets. The mean AC of the first set (BMI <22) (Figure 2.42a) demonstrated a reoccurring waveform at the 14 minute lag point. The mean AC of the second set (BMI >=22) (Figure 2.42b) demonstrated no reoccurring waveforms. The mean AC's were plotted together for waveform comparison. The SEM's (Figure 2.42c) were significant at the 14-minute lag point ($p<0.0256$). The age matched cohort analysis demonstrated that the reoccurring waveform was affected by the BMI of the subjects.

a)

b)



c)

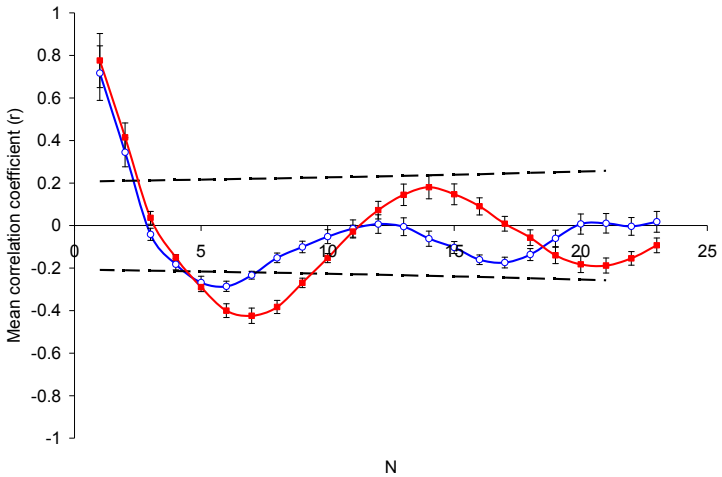


Figure 2.42 – a) Mean autocorrelation for age matched cohort .Group less than 22 BMI - with dotted significance bars b) Mean autocorrelation for age matched cohort. Group greater than 22 BMI - with dotted significance bars c) Autocorrelograms overlaid for SEM waveform comparison. Less than 22 BMI (■), Greater or equal to 22 BMI (○)

2.4.3 Discussion

The study described cross-sectional changes in the oscillatory insulin concentrations of non-diabetic subjects with a range of BMIs. The pulsatile pattern of insulin secretion in the 56 subjects has shown that there was a critical effect of the level of BMI on the frequency, amplitude and regularity of insulin oscillations. The insulin signalling in Group 1 (BMI <22.8 kg.m-2) had a dominant 14 - 15 min period which was missing in Group 2 (BMI >22.8 $kg.m^{-2}$). The fasting plasma insulin concentration was raised in Group 2 and coincident with the appearance of longer-term trends. The analysis of the subset group (age range 20-40 years) was performed to determine if the difference between the groups was a function of age or BMI alone. It was determined that this difference was not caused by the age disparity. This loss of function has been demonstrated to be not of increasing age in this study but a difference in BMI.

In vivo regular pulsatile secretion acts within the metabolic system to increase the sensitivity to gluco-kinetics[91], enhance glucose clearance and up-regulate insulin receptor numbers[90],[102].With an increasing level of BMI there was a loss of the fine-tune regulatory system of glucose clearance. The fasting plasma insulin concentration rose in line with BMI. This higher level of background insulin may have increased insulin resistance, but this was not measured. Inter-pulse intervals have been reported to be a primary determinant of insulin resistance[103], but the exact time domain remains unclear. The inter-pulse interval and frequency were both suppressed in Group 2. This may have been due to an , increase in the insulin resistance associated with the higher BMI's.

In general, there was a reduction in maximum amplitude of insulin oscillations for Group 2. This would be in keeping with the postulate that there was economical amplitude-modulated signalling by pulsatile insulin secretion[68]. The significance of these findings was underscored by the change in regularity assessed by the r-values generated from the AC's. The autocorellogram of group 1 noted a significant negative r value of -0.4 at a periodicity of 7 minutes denoting the maximal ½ phase of lag (i.e. peak and trough) a subsequent reoccurring waveform every 14 minutes provides collateral evidence for pulsatility detected by the FT. The 14 minute periodicity does not reach significance but the regularity of the waves was missing in Group 2 due to the insulin pulsatility being suppressed or no longer present.

In conclusion, Easy TSA enabled the detection of abnormalities of insulin amplitude accompanied by a change of oscillatory pattern of insulin secretion with a higher BMI. Insulin release, both in terms of frequency and amplitude, was significantly altered and suppressed as BMI increased.

3.0 Glycaemic assessment in diabetes

Be wary of the man who urges an action in which he himself incurs no risk. - Joaquin Setanti (British poet)

3.1 - Introduction

The availability of megabyte data sets of glycaemic concentrations is increasing. Glucose values in people with diabetes have traditionally been measured a few times a day using home blood glucose monitors. Since the advent of Continuous Glucose Measuring Systems (CGMS) it has been possible to collect thousands of blood glucose samples over a period of days. These megabyte datasets have two dimensions (variables), glucose & time, that require an objective independent clinically relevant assessment.

Blood glucose datasets are currently assessed in a number of ways:

- Manual assessment
- Software assessment
- Time Series Analysis
- Clinical assessment

3.1.1 Manual assessment

The manual method of assessment is by observation. The outliers in the dataset are noted, high and low points, where glucose values above hyperglycaemic or below hypoglycaemic thresholds occur. The properties and attributes of the majority of the data are not assessed. The technique is subjective as the result will vary depending on the assessor.

3.1.2 Software assessment

Software programs can rapidly assess megabyte datasets and determine objective results. Home blood glucose measuring systems and CGMS systems are provided with software to calculate the dispersion of the data (standard deviation) the average values achieved (mean), periods spent in (or out of) the euglycaemic range. The software performs standard analysis of one variable and takes no account of the relationship with the second variable (time). The domain for euglycaemia may vary by manufacturer or subjective input.

3.1.3 Time Series Analysis

Time Series Analysis (TSA) can be used to assess the variables, glucose and time, and their relationship objectively and independently (Chapter 2). The techniques implemented into the Easy TSA program allow an objective, independent assessment of specific aspects of glycaemia but it may not give clinical insight into the appropriate treatment needed to address any problems observed.

3.1.4 Clinical glycaemic assessment

The objective, independent assessment of glucose concentration instability (variability) observed in those with diabetes is important as tight glycaemic control can reduce long term complications and morbidity from micro-vascular and macro-vascular complications [104, 105]. The use of home blood glucose monitoring has been shown to decrease adverse events and improve glycaemia[106-108]. With the advent of continuous glucose

measuring systems[109] patients and clinicians can track the glucose levels at systematic short intervals, but they are faced with an ever-increasing data load.

The clinical approach to the problem of maintaining glycaemic control has tended to be pragmatic. Hypoglycaemia avoidance has been by dose adjustment, by modification of diet, matching diet and exercise and acting cautiously when modifying pharmacological treatment and lifestyle. Physicians and patients may together decide to maintain higher blood glucose levels to avoid severe hypoglycaemia. This pragmatic approach has served well to avoid the worst of the clinical problems. The disadvantage of the approach is that most of the data are not being utilised in a systematic way.

By contrast, for several decades, it has been recognised that blood glucose profiles could be utilised in a systematic way to yield integrated measures of glycaemic control. Such metrics may capture the range and frequency of glucose instability in profiles. The range of glycaemia is illustrated in Figure 3.1. This demonstrates how each subject's glucose values cluster around a glucose concentration with variable measurements and a range of outliers. The clustering, the range, and the percentage of points at unacceptable levels are all important and require independent and objective assessment.

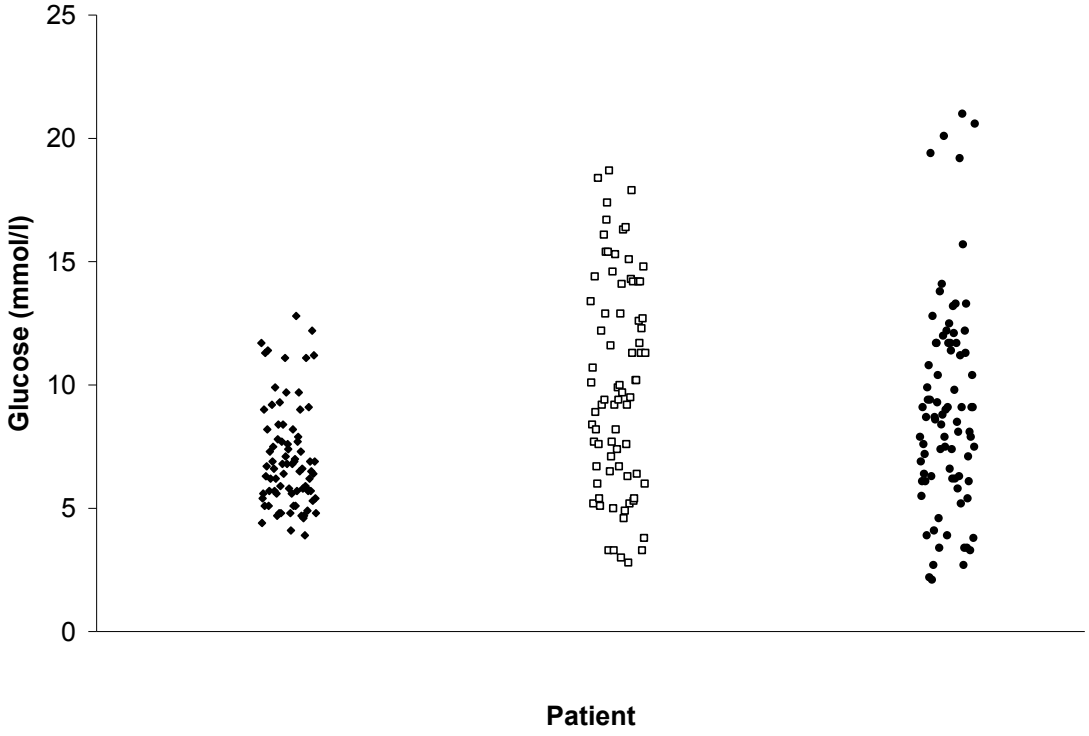


Figure 3.1 – Glucose dispersion diagram from home blood glucose monitoring of three subjects with type 2 diabetes

3.2 Existing methods for glycaemic assessment

A comprehensive literature search was undertaken to find English language methodologies for glycaemic assessment. The search was performed on peer reviewed published original papers obtained from electronic database sources Pubmed (National Centre for Biotechnology Information) and Scopus (Elsevier Group) up to November 2006. The primary search terms used were 'diabetes' and any one of the following: 'glycaemic assessment method', glucose assessment', 'metabolic control assessment', 'hypoglycaemia assessment', 'hyperglycaemia assessment', 'glycaemic risk assessment'.

Original articles that included an explicit methodological process were selected for review. The methods were analysed and compared using an objective process.

Two hundred and ninety nine papers were generated by the search terms of which sixteen entries were relevant. Six papers which were identified as having an explicit methodology were summarised and objectively compared. A critique and summary of each of the papers was undertaken for each method.

3.2.1 M-value 1964

Method

In 1964 Schlichtkrull et al[110] developed the M-value to be a 'reliable, repeatable and sensitive index of the quality of control'. It sought to make an objective appraisal of the quality of a variety of diabetes therapies. The methodology utilised the experience of nine doctors 'familiar with insulin therapy'. Each of the nine were asked to score twenty sets of glucose profiles according to how well controlled they judged the subject to be. A formula was created that matched their mean ranking. The M-value equation was then applied retrospectively to a set of serial glucose readings and a mean for the M-values was calculated.

Critique

The M-Value has not been widely used but it was the first attempt at a mathematical analysis of glycaemic data. The formula was not robust enough in practice[111] and has since been modified many times to attempt to correct these problems[112, 113]. Schlichtkrull et al claimed that the M-value was 'an objective substitute for clinical evaluation'. The uptake of the M-value as an assessment method has been low. This may be partly because it was relatively complex to calculate, and was ambiguously defined. The ambiguity was that a 'fluctuation' term was introduced that was used to determine the range of observed values in a profile without defining the duration of the profile. Moreover, the M-value used the 1960s clinical views, of a small sample (n=9),

about the relevance of hypo- and hyperglycaemia. Views have been considerably modified on hypo- and hyperglycaemia, following the publication of the DCCT in 1991[114] and the UKPDS in 1998[115].

3.2.2 Mean Amplitude of Glycaemic Excursions (MAGE) 1970

Method

The MAGE value is the assessment of the rises and falls in a glucose profile that occur outside one standard deviation of the mean value. It was proposed by Service et al[113] in 1970 in an attempt to encapsulate glycaemic variability by identifying peaks and troughs and calculating their amplitude. MAGE was used to evaluate eight people with diabetes who were analysed continuously for forty-eight hours during 'metabolic balance' studies of six days duration under 'near normal' conditions (fed and ambulatory). Blood was collected continuously at the rate of 6 ml/hour. The results from the analysis of the blood glucose, using MAGE, demonstrated that there were higher values in the more unstable glucose profiles and the authors concluded that MAGE was a 'good quantification method of blood glucose behavioural characteristics'.

Critique

It is not always straightforward to ascertain the definition of a peak or a trough[116], it is difficult to define what constituted a rise or a fall in these analyses . Service et al used the phrase 'appears to be', highlighting the difficulties faced. Different users of the MAGE method may determine different peaks and troughs, in the same profiles, making the MAGE scores subjective. This ambiguity means the method cannot be computerised and can be performed by inspection alone. Examination of the excursions by MAGE attributed the same score to hypo- and hyperglycaemic variations, weighted equally.

3.2.3 Lability Index (LI) 2004

Method

The LI is a comparative measure that calculates, from a glucose profile a measure of lability. The LI score is calculated using the frequency and duration of glucose swings. In 2004 Ryan et al[117] wished to define a measure by which patients for islet transplantation may be considered. A lability formula was used as an index in conjunction with HYPO scores (Section 3.2.3.4). The LI calculation was applied to one hundred subjects. Four weeks of blood glucose results with readings taken at least two times a day (no times specified) were recorded. The LI results were compared to two non-independent (The first and last authors of the paper) blinded clinical assessments. The authors produced a strong (r = 0.87) correlation between the lability calculation and the clinical assessment. The LI was further correlated with the MAGE scores where r = 0.291. The conclusion reached was that 'the lability score gave an objective assessment of the severity of the problems being encountered by the patients.'

Critique

The LI is an absolute mathematical score and is therefore independent of clinical evaluation. The score obtained does not take into account physiological risks from hypo- & hyperglycaemia attributable to the glucose labilities. A clinical evaluation was undertaken by using LI in conjunction with a HYPO score that sought to compensate for the purely mathematical results from LI. The LI was correlated with the opinion of two diabetologists (who where the first and last authors), and this may have affected the published R^2 value of 0.75.

3.2.4 HYPO Score 2004

Method

The HYPO score is a clinically led approach that uses a combination of sources of information. It was developed by Ryan et al[117] in conjunction with their lability index and had the same aim as the lability score to assess patients' suitability for islet transplantation. The score was a combination of properties pertaining to hypoglycaemia (Table 3.1). A scoring system was developed that awarded points when certain criteria were met.

HYPO score parameters	Description
Low glucose values	<2.5 mmol/l or Between 2.5 – 2.9 mmol/l were attributed HYPO scores.
Autonomic symptoms	If these symptoms occurred, no HYPO points were awarded.
Outside help required	If outside help was required then the episode was counted as hypoglycaemia. This was then attributed a HYPO score.
Prior episodes of hypoglycaemia	These were estimated for one year from thirteen weeks of results. Increased episodes were attributed higher HYPO scores.
Ambulance called	If an ambulance was called during a hypoglycaemic episode a higher HYPO score was attributed
Glucagon administered	If glucagon was administered HYPO points were awarded

Table 3.1 Breakdown of the HYPO score criteria

Critique

The development of a scoring system that uses clinical judgement and physiological measures to assess glycaemic variability and control is a novel approach in comparison to the other mathematically based methods. The system used in the HYPO Score was not suitable, as a method to judge glucose control and hypoglycaemia, for a number of reasons:

Low glucose values: - Counting low glucose values is non-comparable with other studies as the threshold for hypoglycaemia is set at different levels for individual studies[118-122].

Autonomic symptoms: - Autonomic symptoms are suppressed with increasing duration of diabetes[123, 124] but this was not specifically discussed in the HYPO score.

Outside help required: If a subject did not, or was unable to, call anyone and rectified a severe hypoglycaemic situation on their own they did not achieve a HYPO score.

Prior episodes of hypoglycaemia: Prior occurrences of hypoglycaemia over the last twelve months in this HYPO score were estimated from four weeks of information. The estimation of the number of episodes may lead to clinically relevant under and over estimations as it depends on patient recall.

Ambulance called: An ambulance may be called when not necessary or may not be called by a subject who corrects the glycaemia by himself or herself. This may cause problems with the sensitivity and specificity of using this as a measurement.

Glucagon administered: Patients score if glucagon is administered. This may be problematic, as a measure, because it is dependent on the availability of glucagon and the experience of the subject or carer.

The measurements would be affected by whether the patient lives alone or with a carer. Measurement methods for hypoglycaemia are always complex. Hypoglycaemia severity continues to be a problematic issue because of its disabling nature and this prevents accurate observations at the time they are required.

3.2.5 Lag plot 2005

Method

A lag plot analysis is the equivalent of the coefficient of variation and measures the variability of the data around the mean. The analysis of glycaemic profiles by use of a lag plot is one of the most recent developments. It was published in 2005 by Albisser et al[125] and was developed to answer the question 'How good is your glucose control?'. The method involved producing the lag plots of seven patients 'spanning a range of diabetes'. The range of diabetic patients analysed was determined to be those that

would have a series of different glucose profiles. The results demonstrated that people with dispersed glucose profiles produce dispersed lag plot and those with tight glucose profiles produced tight lag plots. The conclusion reached by the authors was that the lag plot 'empowers providers to identify easily in a visually informative method problems in glycaemic control, take proactive treatment, evaluate the outcomes, adopt beneficial strategies, and rule out interventions with no benefits'.

Critique

The lag plot is a representation of dispersion of data. Mathematically it is the visual representation of N=1 of an autocorrelogram, or the coefficient of variation. The Lag plot offers no additional information when used as a tool for glycaemic assessment. There is no differentiation between the hyperglycaemic and hypoglycaemic states by the lag plot and hence it is not possible to ascribe a clinical insight into glycaemic variations.

3.3 Development of a new metric for glycaemic assessment

The methods reviewed from the literature search exposed a variety of differing views on what was important when examining glycaemia. The assessment of the level of glycaemic control of a person with diabetes is a critical aspect in optimising diabetes care. Selection of the correct method for the analysis enables a clinically relevant diagnosis and treatment to be made. Selection of a method with 1960's views on glycaemia, the M-Value, or using interpreted results, the HYPO score, or other previously discussed shortcomings may lead to incorrect diagnoses and treatments.

There was a need for a new methodology to report an integrated glycaemic risk score from glucose profiles that would complement summary measures of glycaemia, such as the HbA1c. A novel method was developed that encapsulated the totality of hypo-hyperglycaemic risk to a patient relevant on their glycaemic profile. It enabled rapid objective, independent assessment of the megabyte datasets.

The method was developed by seeking an objective assessment of the relevant importance and clinical risk of blood glucose variations and dwell times. To determine this metric fifty diabetes health care professionals (Table 3.2) assigned risk values to a range of forty blood glucose (BG) concentrations. The health care professionals each had varied levels of seniority and experience, including paediatric practitioners whose overriding concerns included the effects of hypoglycaemic episodes on growth and development. No clinical context was specified for the assessment of glucose values via questionnaire. The glucose values in the questionnaire encompassed a wide interval of glycaemic values. The forty BG values were presented in a random order, including repeated values, using a structured questionnaire. It is important to note that the GRADE score was developed using clinicians views of diabetes in 2007 these views may change in the future.

Professional Category	N	Years Of Experience Median (Range)
Professors of diabetes	3	29 (26 - 32)
Consultants	5	18(16-20)
Specialist Registrars specialising in diabetes	18	11 (8-14)
Senior House officers	5	5 (4-7)
Nurses specialising in diabetes	7	7 (3-15)
Others specialising in diabetes (Dieticians, Statisticians, Research Fellows)	12	6 (1-17)

Table 3.2 - Summary of the fifty health care professionals and researchers in diabetes care indicating years of experience.

Their subjective assessment (range 'dangerously low' to 'dangerously high') was made using a visual analogue scale (Figure 3.2).

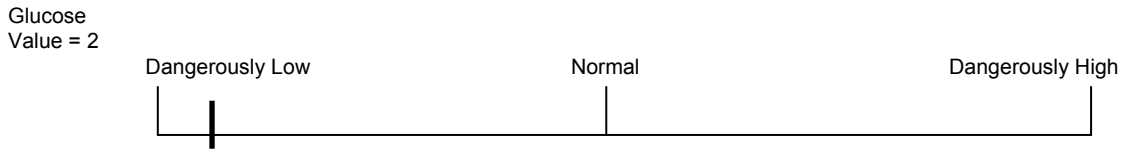


Figure 3.2 - An example of the visual analogue scale used by the diabetes health care professionals to grade a blood glucose value of two.

A score was calculated according to where the line was drawn ranging from 0 (no risk) to 50 (maximal risk). For example, blood glucose (BG) concentrations of greater than 30 mmol/l (540 mg/dl) drawn near 'Dangerously high' were scored tending to 50 and likewise BG concentrations near 1 mmol/l (18 mg/dl) drawn near 'Dangerously Low' scored tending to 50.

The median responses for each BG value were used to develop an equation relating risk to glycaemia. The function was then applied to calculate the risk values for individual glucose concentrations within a glucose profile.

3.3.1 - Glycaemic Risk Analysis in Diabetes Equation (GRADE)

This median approximation of risk values was termed the Glycaemic Risk Assessment Diabetes Equation (GRADE) score. The GRADE score generated by this process allowed the assessment of each discrete point in a glycaemic profile. The contribution of risk attributable to hypoglycaemia, euglycaemia and hyperglycaemia to the GRADE score can be expressed as a percentage contribution: GRADE (hypoglycaemia%, euglycaemia%, hyperglycaemia %). The GRADE, the closest approximation to the median estimation, evolved to be a second order polynomial with a log-log transformation (Figure 3.3).

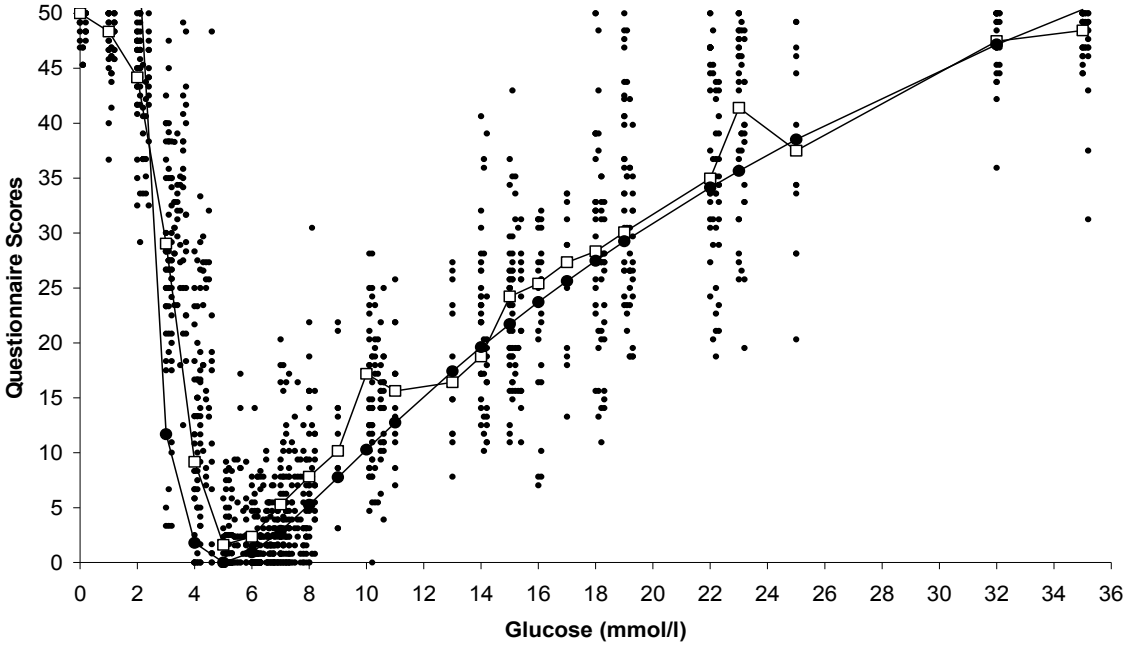


Figure 3.3 - Questionnaire scores of the assessed glucose values by 50 health care professionals and researchers. Median Glucose Assessment (□); GRADE approximation (●)

The structure of the formula was designed to give a continuous curvilinear approximation with a nadir at 4.96 mmol/l (90 mg/dl) and high adverse weighting for both hyper- and hypoglycaemia. The GRADE value can be calculated from either a mg/dl or a mmol/l profile of BG.

mmol/l	$GRADE\ value = 425 * (\log(\log(x)) + 0.16)^2$	EQ 3.1
mg/dl	$GRADE\ value = 425 * (\log(\log(x*18)) + 0.16)^2$	EQ 3.2

Where x is the observed glucose value and the log is base 10.

GRADE operates for BG ranges between 2.06 (37 mg/dl) and 33.42 mmol/l (630 mg/dl). Values outside this range were ascribed a GRADE value of 50. The GRADE score is the median of the GRADE values in a profile. The relative contribution to the GRADE risk score of hypo- and hyperglycaemia can be calculated (EQ 3.3).

EQ 3.3 - Equation for calculation of contribution to GRADE score as a percentage. BG measured in mmol/l

Hypoglycaemia (BG<3.9)	Euglycaemia (3.9≥BG≤7.8)	Hyperglycaemia (BG≥7.8)
$\frac{\sum GRADE\ BG < 3.9}{\sum All\ GRADE\ Values} \times 100$	$\frac{\sum GRADE\ 3.9 \geq BG \leq 7.8}{\sum All\ GRADE\ Values} \times 100$	$\frac{\sum GRADE\ BG > 7.8}{\sum All\ GRADE\ Values} \times 100$

A GRADE score < 5 reflected the euglycaemic range (Figure 3.3). This euglycaemic range was used to define the ranges used in the equation (EQ 3.3). The glycaemic range 3.9 mmol/l (70 mg/dl) to 7.8 mmol/l (140 mg/dl) was termed euglycaemic, the range < 3.9 mmol/l (70 mg/dl) was termed hypoglycaemia and the range >7.8 mmol/l (140 mg/dl) termed hyperglycaemia.

In reporting a GRADE score, percentages are used to report weighted risk calculated from hypoglycaemia, euglycaemia, and hyperglycaemia. For example, a GRADE score

of 10 indicates an increased risk of microvascular complications, macrovascular events and cardiovascular morbidity. The percentage risk contribution deconvolutes (EQ 3.3) to (18%, 80%, 2%) indicating elevated risk of hypoglycaemic episodes. Conversely, a GRADE score of 10 (2%, 48%, 50%) would be typical of someone with poorly controlled hyperglycaemia. Obviously, these percentages sum to 100%

3.4 Comparison of GRADE to existing metrics of glycaemic assessment

The properties of the glycaemic assessment methods are compared (Table 3.3, 3.4). To allow a comparison of the methodologies and GRADE a set of five self-reported blood glucose results (Figure 3.4), taken four times daily, for twenty one days, were analysed using each of the methods. In addition a blinded assessment was made by five clinicians in diabetes on how well controlled they judged the profiles to be (1=Well controlled, 10=poorly controlled). Further, a Coefficient of Variation (CV), a measure of the dispersion, was calculated for each of the glycaemic profiles.

Method	Independent of observer bias	Mathematically robust	Includes clinical judgments	Weighted for hypoglycaemia	Weighted for hyperglycaemia	Data frequency independent	Normative ranges defined
M-Value[110]	YES	NO	YES	YES	YES	NO	YES
MAGE[113]	NO	NO	NO	YES	YES	NO	NO
Lability Index[117]	YES	YES	NO	NO	NO	YES	NO
HYPO Score[117]	NO	NO	YES	YES	NO	YES	YES
Lag Plot[125]	YES	YES	NO	NO	NO	YES	NO
GRADE[126]	YES	YES	YES	YES	YES	YES	YES

Table 3.3 – Glucose assessment methods: comparison of properties

HbA1c

7.7%

6.8%

7.0%

9.1%

7.0%

Figure 3.4 – Home blood glucose measuring profiles for 21 days, 4 readings daily**.**

The results from the application of these techniques to the five data sets demonstrated that the outputs from each method share a common theme ‘the score given for glycaemia increases as the profile worsens’. The M-Value and GRADE define the threshold between a ‘good’ and a ‘poor’ glycaemic profile, the other methods do not.

† Subject ID	Age	Diabetes Type	BMI	HbA1c	Clinicians score	Cv	M Value	MAGE	Lability	HYPO Score	Lag Plot	GRADE Score	Breakdown Hypo/Eugly/Hyper
A	52	1	24.0	7.7	7.0	0.42	34.4	37.5	7.8			11.5	5.9 / 4.1 / 90.0
B	80	1	25.9	6.8	6.8	0.47	35.5	29.7	6.2	Clinical based score not assessable from profiles	Graphical Output Only	10.4	17.5 / 5.6 / 76.9
C	59	1	27.7	7.0	7.0	0.62	56.4	0.0[a]	7.7			21.0	64.5 / 2.6 / 32.9
D	60	2	32.8	9.1	6.0	0.33	47.4	35.1	8.4			16.4	1.9 / 1.0 / 97.1
E	70	1	28.0	7.0	3.25	0.30	10.7[b]	18.1	2.1			3.7[b]	___ / ___ / ___

Table 3.4 – Glucose assessment method results from five sample datasets

Subject A and Subject B achieved M-Values that were evaluated as being 'poor', high glycaemic excursion scores from MAGE, above average lability score and high GRADE scores with the majority of the risk attributed to Hyperglycaemia. The experts in diabetes scored the glucose profiles 7.0/10 and 6.8/10 respectively and the CV was 0.42 and 0.47 respectively. The methods operating in different ways reached a consensus that these are both poorly controlled glycaemic profiles.

Subject C was given a 'very poor' M-Value. The glycaemic excursion could not be calculated by MAGE due to a large standard deviation. The lability index was above average and the GRADE score was very high with the majority of microvascular & macrovascular risk attributable to hypoglycaemia. The experts in diabetes scored the glucose profile a 7.0/10. The CV was 0.62.

Subject D had a large M-Value, and an increased number of glycaemic excursions measured by MAGE. Subject D had the most labile glycaemic profile according to the LI score and the GRADE score was high with the risk attributable to hyperglycaemia. The experts in diabetes scored the glucose profile a 6.0/10. The CV was 0.33. Although the experts in diabetes rated this, profile the second best of the five. The other methods all disagreed. The experts in diabetes may have considered that a higher level of overall hyperglycaemia was not as important as hypoglycaemia.

Subject E was well controlled according to the M-Value, the glycaemic excursions (MAGE) were estimated to be high, but were the lowest of the group. Subject E had a profile that was not labile and achieved a good GRADE score. A GRADE score <5 did not require analysis of where glycaemic risk was attributable. The experts in diabetes

scored the glucose profile a 3.25/10. The CV was 0.3. MAGE was the only process that scored this profile as unsatisfactory this might have been due to the problem of deciding where a trough begins and a peak ends.

It has been shown in the data sets, used for comparison, that it was possible for two subjects (C, E) to have a HbA1c of 7.0 yet have contrasting glucose control. HbA1c alone was not sufficient to assess glycaemia. A glycaemic assessment method would be required to measure control and the clinical risk it represents to the subject. Now that the technology exists to measure glucose transcutaneously every few minutes by biosensor technologies a credible risk could be apportioned to every glucose measurement taken. The assessment of glucose profiles could be extended by allowing a quantitative clinical measurement of risk, from both hypoglycaemia and hyperglycaemia, to be obtained.

3.4.1 Conclusions of the Glycaemic Assessment Methods

Clinicians and scientists have sought to understand more completely glucose control and to quantify it for several decades. Each method has a relatively low citation count (max 65 –for MAGE[113] accessed October 2006) indicating that the uptake of glucose assessment methods has been low. This may be attributed to the problems associated with assessing megabyte glycaemia datasets using these methods. Alternatively, it may be that the assessment of glucose profiles has previously been underrated as a central part of diabetes clinical care. The correct assessment of glycaemic control may allow associated risk from microvascular, macrovascular complications and morbidity to be reduced.

3.5 Retrospective application of GRADE assessment

The creation of the new metric, GRADE, allowed an independent and objective assessment to be made on the megabyte datasets obtained from those with diabetes. To determine if GRADE was clinically relevant, in addition to being objective and independent, it was applied to BG data sets obtained from Menarini Ltd (Florence, Italy), Medtronic Inc (Northridge, USA), London Centre for Paediatric Endocrinology, and Oxford Centre for Diabetes Endocrinology and Metabolism.

The BG data sets analysed included continuous glucose monitoring data and four-times daily home BG monitoring for one week. It was necessary that the period be specified for purposes of comparison. Measurements taken at the time of hypoglycaemic episodes were included since they defined excursions from euglycaemia. A period of 72 hours was used for the continuous sampling from transcutaneous glucose sensors when comparisons were being made.

The analysis occurred on profiles from non-diabetic subjects from four main ethnic groups (Caucasian, Hispanic, Asian and African) and profiles from those with type 1 and type 2 diabetes.

The median GRADE scores, and the contribution from hypo- hyperglycaemia, derived from continuous blood glucose monitoring in the non-diabetic population was calculated (Table 3.5). In the populations available for study medians of the GRADE scores ranged between 0.8 and 1.19. There were no significant differences between the non-diabetic groups. In the non-diabetic Asian population the technique identified a subject with a high GRADE score whose glucose profile contained episodes of hyperglycaemia and

which could be considered as impaired glucose tolerance or impaired fasting glucose. The non-diabetic population had a GRADE score of <= 5 that represented a blood glucose range between 3.8 mmol/l (70 mg/dl) and 7.8 mmol/l (140 Mg/dl).

					GRADE scores (50 Max)							
										Contribution (%)		
Subject group	Assessed by:	N (subjects)	Total No of observations	Mean Glucose (SD)	Min	Lower Quartile	Median	Upper Quartile	Max	Hypo.	Eugl.	Hyper.
Non-diabetic Caucasian	Continuous Monitoring	31	770	5.0 (0.7)	0.37	0.73	0.92	1.39	2.92	0,	100,	0
Non-diabetic Hispanic	Continuous Monitoring	11	798	5.1 (0.6)	0.27	0.46	0.78	1.03	1.26	0,	100,	0
Non-diabetic Asian	Continuous Monitoring	23	721	5.5 (0.7)	0.22	0.55	0.80	1.21	5.32	0,	100,	0
Non-diabetic African	Continuous Monitoring	3	740	5.0 (0.8)	0.25	0.94	1.19	2.40	3.62	0	100,	0
Non-diabetic Other	Continuous Monitoring	4	677	5.1 (0.9)	0.46	0.62	0.80	1.43	2.35	0	100,	0
Type 1 Diabetes	Continuous Monitoring	36	885	8.6 (2.0)	2.24	5.91	8.09	12.68	37.66	20,	9,	71
Type 2 Diabetes	Home Glucose Monitoring	26	82	9.9 (2.1)	4.46	6.83	9.97	12.15	21.57	1,	7,	92

Table 3.5 - Median GRADE scores for non-diabetics subjects and people with both type 1 and type 2 diabetes.

As expected patients with type 1 and type 2 diabetes clearly had higher GRADE scores than the non-diabetic population. The contribution of hypo- and hyperglycaemic episodes, and how they relate to the GRADE scales, were obtained from the sample data sets (Table 3.5)

3.6 Prospective application of GRADE assessment

The correct independent and objective assessment of megabyte glycaemic datasets may allow associated risk from microvascular, macrovascular complications and morbidity to be reduced by assessment of the glycaemic control a subject has. The new metric, GRADE, could be used as a tool to supply an integrated view of clinical risk relevant to a patient.

A clinical study was designed to apply GRADE in a clinical context, to measure the statistical reproducibility of using the GRADE value in clinical practice. In addition the study sought to determine a statistical measurement of the usefulness of the GRADE value.

3.6.1 Study Design

The GRADE study was conducted at the Oxford Centre for Diabetes, Endocrinology & Metabolism (OCDEM) and was an investigation of existing outpatients currently treated with a range of differing insulin regimes. Subjects were identified as meeting the inclusion criteria by use of our diabetes register. Individuals were approached by their clinician or a research nurse to seek informed consent for participation in this study. An information sheet was provided for the patient and informed consent was taken at the

clinic visit. Patients' were given as much time as they need to make a decision whether to participate in the study and the individual had the opportunity to ask questions.

The subjects were asked to measure their blood glucose at four specified times daily for twenty-one days. Subjects were to continue their 'normal' daily habits throughout the duration of the trial. The participants were contacted by telephone during the second week of the trial to discuss any questions or queries that may have arisen. Data received and gathered for each individual was identified by a unique number. The study operated under ethical approval COREC 06/Q1605/38.

3.6.1.1 Inclusion Criteria

- Diagnosis of diabetes (type 1 or type 2)
- Age >18 years.
- Duration of diabetes > 2 years.
- Informed consent given.

3.6.1.2 Exclusion Criteria

- Major psychiatric disease.
- History of irregular home blood glucose monitoring; failure consistently to check glucose levels.
- History of non prescribed drug or alcohol abuse
- Renal failure (creatinine above 150 μmol/l)
- Abnormal liver function tests (> 1.5 x upper limit of reference interval)
- Pregnancy

- Any other condition or circumstance, which, in the opinion of the investigator, would affect the patient's ability to participate in the protocol.

3.6.2 Subjects

Of the sixty patients recruited fifty-three (Male 26, Female 27) patients completed the full twenty-one days of sampling. The data from the seven patients who withdrew from the study were not used in analysis. The fifty-three recruited patients had a mean BMI of 28.3(5.3) with a mean HbA1c of 7.7(1.2) at baseline. Subjects were receiving a variety of insulin regimes (mixtures, basal + bolus, analogues, isophane insulin's). Thirty-one of the patients had type 1 diabetes the remainder had type 2 diabetes.

3.6.2.1 Statistical Considerations

The power calculation for statistical significance in the trial required Fifty-three participants.. Study Power was calculated to have a 90% power based on detecting a 10% difference in means between HbA1c and GRADE values from a home blood glucose monitoring pilot study of sixteen subjects to a significance of $p<0.05$. Sixty subjects were recruited.

Statistical reproducibility of the GRADE value was be measured by inter-subject Coefficient of Variation and the subjects Discriminent Ratio[127] in comparison to the group.

The usefulness of GRADE and the correlation coefficient between GRADE, HbA1c, and hypoglycaemia was analysed using standard exploratory processes recommended by

Chatfield[15]. It was noted that the analyses were likely to be parametric for the representation of the relationship with hypoglycaemia and HbA1c and non-parametric for the GRADE score.

3.6.3 Results

The glycaemic profiles for each of the patients from the four daily measurements over the twenty-one days were analysed (Table 3.6). The blood glucose of measurements recorded by the subjects varied between a mean of 6.3 mmol/l and 15.6 mmol/l. The median GRADE results calculated for each of the patients varied between 1.3 and 22.7. The breakdown, of the GRADE score, in terms of contribution to risk by hypoglycaemia, euglycaemia and hyperglycaemia was performed in GRADE scores >5.

		Glucose			GRADE						
ID	HbA1c	Mean	SD	CV	Median	Range			Hypo	Eugly	Hyper
CRU034	5.6	7.0	2.7	38.5%	2.0	0.5	-	6.3			
CRU027	5.9	6.3	3.2	50.4%	5.3	1.4	-	12.1	45%	20%	34%
CRU026	6.1	6.3	3.2	50.6%	5.0	1.1	-	12.6			
CRU071	6.2	6.7	2.3	35.0%	3.5	1.1	-	8.5			
CRU061	6.3	7.0	2.9	40.9%	4.5	1.6	-	11.0			
CRU019	6.5	6.6	3.5	53.7%	6.3	1.4	-	13.9	47%	4%	48%
CRU070	6.5	6.0	1.3	21.0%	1.3	0.5	-	2.7			
CRU039	6.6	9.0	4.1	45.0%	10.8	3.8	-	17.1	11%	4%	85%
CRU062	6.7	7.8	2.9	36.5%	5.8	1.4	-	10.0	7%	15%	79%
CRU065	6.7	8.0	2.5	31.1%	4.3	0.9	-	10.0			
CRU067	6.7	7.2	2.5	34.7%	2.8	1.3	-	5.2			
CRU011	6.8	9.0	4.2	46.6%	8.0	3.8	-	15.4	25%	7%	68%
CRU066	6.8	7.8	1.8	23.5%	4.3	2.0	-	8.1			
CRU069	6.9	7.8	2.5	31.8%	4.8	2.1	-	8.3			
CRU06	7.0	7.0	2.1	29.6%	2.2	0.5	-	5.8			
CRU012	7.0	7.9	4.9	61.8%	10.3	3.1	-	22.3	37%	3%	60%
CRU063	7.0	9.5	2.6	28.0%	8.3	5.1	-	11.2	0%	8%	92%
CRU074	7.0	6.9	4.3	61.4%	10.5	4.7	-	24.9	52%	2%	46%
CRU014	7.1	10.2	4.9	48.5%	14.5	6.0	-	22.3	18%	4%	78%
CRU055	7.1	7.5	3.7	49.7%	6.0	1.7	-	11.3	25%	15%	60%
CRU057	7.2	8.3	4.3	52.3%	8.2	3.3	-	15.8	27%	14%	59%
CRU024	7.3	7.1	4.0	56.2%	9.4	2.1	-	18.2	61%	4%	36%
CRU029	7.3	9.2	3.2	35.1%	7.3	2.2	-	13.6	0%	11%	89%
CRU017	7.4	9.4	4.4	46.9%	11.7	4.9	-	18.3	21%	3%	76%
CRU028	7.4	11.4	4.9	42.7%	12.7	6.9	-	18.6	3%	3%	94%
CRU054	7.4	7.9	3.7	47.0%	6.1	1.4	-	13.5	17%	16%	67%
CRU068	7.4	9.9	11.6	117.3%	7.5	3.5	-	10.4	0%	13%	87%
CRU059	7.6	7.7	4.6	59.6%	8.0	4.0	-	17.9	36%	5%	59%
CRU010	7.7	10.1	4.3	42.4%	10.5	4.3	-	20.0	7%	6%	87%
CRU072	7.7	8.2	2.6	31.3%	6.1	0.8	-	10.3	0%	8%	92%
CRU036	8.0	9.6	4.8	49.9%	8.4	2.6	-	17.9	11%	6%	83%
CRU040	8.0	8.9	3.3	37.3%	7.8	2.6	-	15.0	7%	11%	82%
CRU044	8.0	8.6	2.0	22.6%	6.3	4.0	-	10.0	0%	14%	86%
CRU049	8.0	10.4	3.3	31.6%	11.3	4.8	-	17.6	2%	5%	93%
CRU060	8.0	10.3	3.9	38.3%	11.8	2.9	-	18.3	4%	4%	93%
CRU022	8.1	9.2	2.2	24.2%	9.0	4.9	-	12.4	0%	8%	91%
CRU047	8.1	8.3	3.5	42.5%	6.3	2.2	-	13.5	15%	10%	74%
CRU021	8.3	8.9	4.9	54.4%	8.9	1.8	-	18.6	20%	3%	76%
CRU050	8.4	10.1	6.0	59.8%	10.5	4.8	-	21.1	10%	2%	88%
CRU001	8.6	7.0	2.0	28.6%	2.0	0.5	-	4.9			
CRU015	8.7	6.8	3.4	51.0%	6.7	3.0	-	13.0	42%	5%	53%
CRU053	8.7	13.2	4.8	36.4%	19.8	11.3	-	26.6	11%	1%	88%
CRU064	8.7	7.8	1.8	23.7%	5.5	3.8	-	8.5	6%	35%	59%
CRU018	8.8	9.6	4.6	48.2%	7.7	3.3	-	15.2	3%	6%	91%
CRU043	8.8	8.9	3.8	42.2%	7.5	3.0	-	12.9	4%	11%	85%

CRU037	8.9	10.4	2.4	22.9%	10.3	7.5	-	14.5	0%	2%	98%
CRU073	8.9	9.3	3.8	40.3%	10.3	2.7	-	13.7	3%	9%	89%
CRU031	9.0	12.6	4.7	37.1%	15.1	7.8	-	24.7	0%	3%	97%
CRU013	9.1	12.5	4.1	33.1%	19.9	10.5	-	22.7	2%	1%	97%
CRU020	9.1	11.1	5.1	46.4%	11.9	6.6	-	21.7	6%	3%	91%
CRU041	9.7	8.7	3.7	42.2%	5.5	2.0	-	11.2	2%	14%	84%
CRU025	10.5	15.6	2.6	16.3%	22.7	19.6	-	25.6	0%	0%	100%
CRU016	11.9	14.1	4.2	30.2%	21.1	11.0	-	25.5	0%	1%	99%

Table 3.6 - GRADE scores and risk contribution comparison table *LQ = Lower Quartile, UQ = Upper Quartile

The Discriminant Ratio was calculated to be 3.76. GRADE's with scores <= 5 were detected in patients with HbA1c near to clinical target levels, and in those with raised HbA1c's

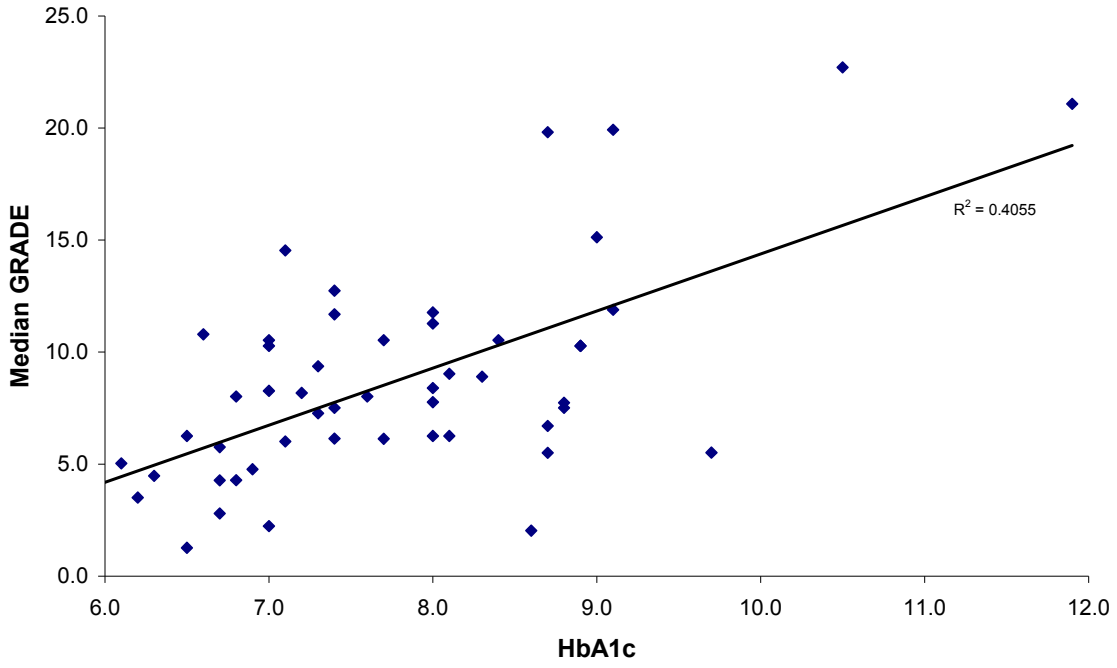


Figure 3.5 –Median GRADE scores and HbA1c at beginning of study with regression line.

Median GRADE and HbA1c were compared for each of the subjects (Figure 3.4). There was subjects who had HbA1c of 7.0 at baseline had median GRADE values ranging

from 2.2 to 10.5, but the correlation (R^2) between HbA1c and Median GRADE values was 0.41.

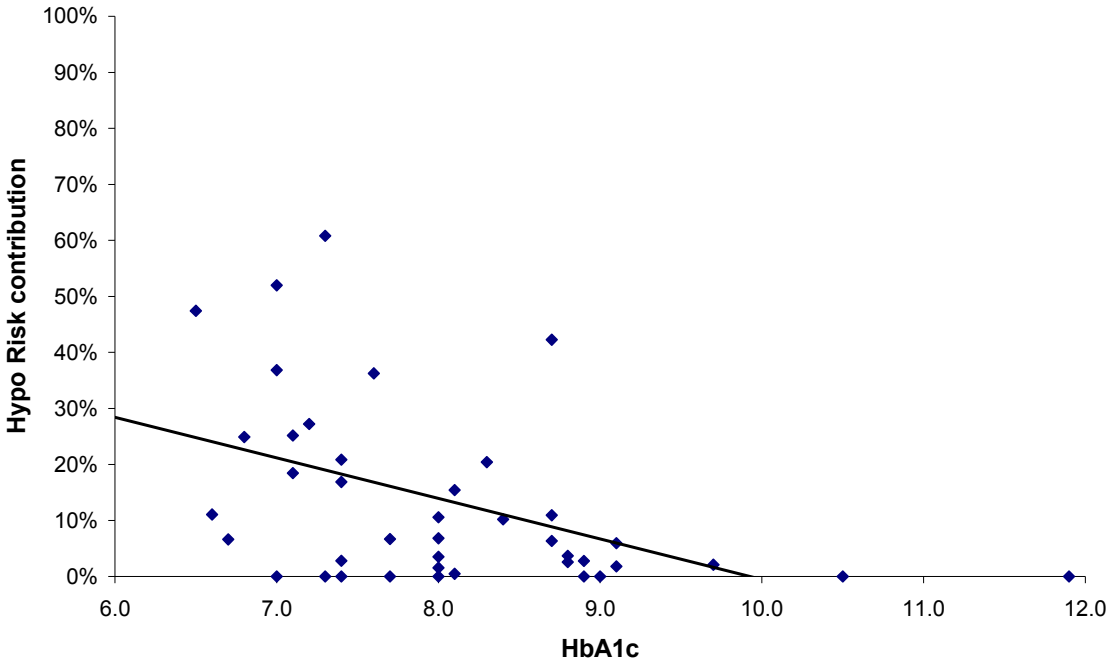


Figure 3.6 – HbA1c and Risk attributable to Hypoglycaemia with regression line.

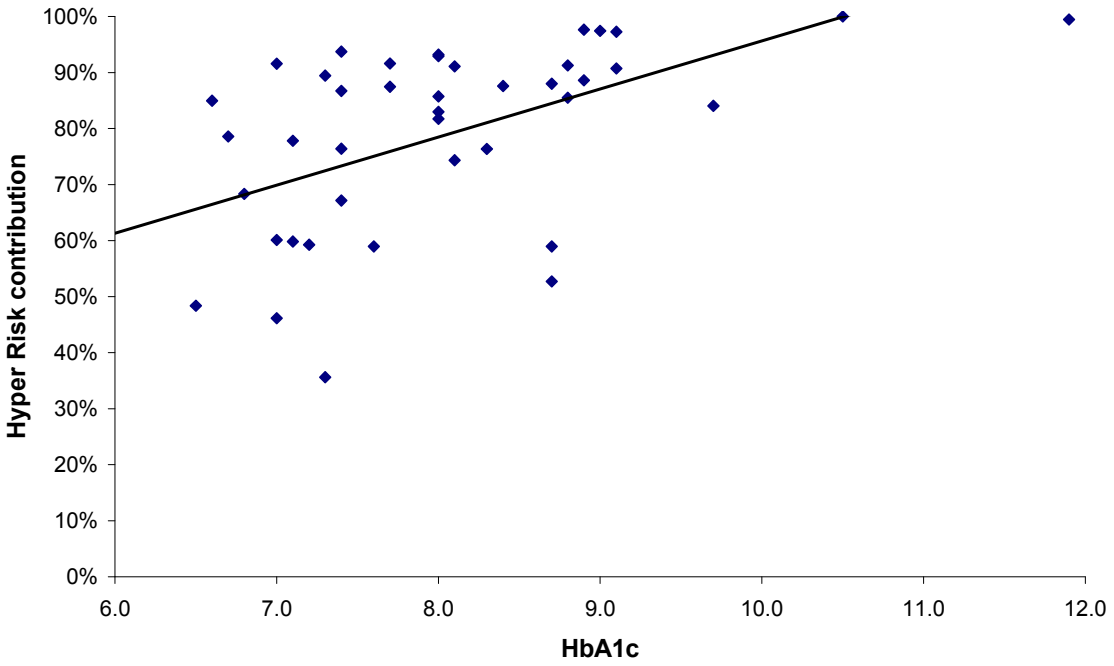


Figure 3.7 - HbA1c and Risk attributable to Hyperglycaemia with regression line.

The correlation between HbA1c and the 'Clinically Assessed Risk' (extracted from the GRADE equation) was determined (Figure 3.6 and Figure 3.7). Lower HbA1c's correlated with an increased risk of hypoglycaemia and higher HbA1c's correlated with an increased risk of hyperglycaemia.

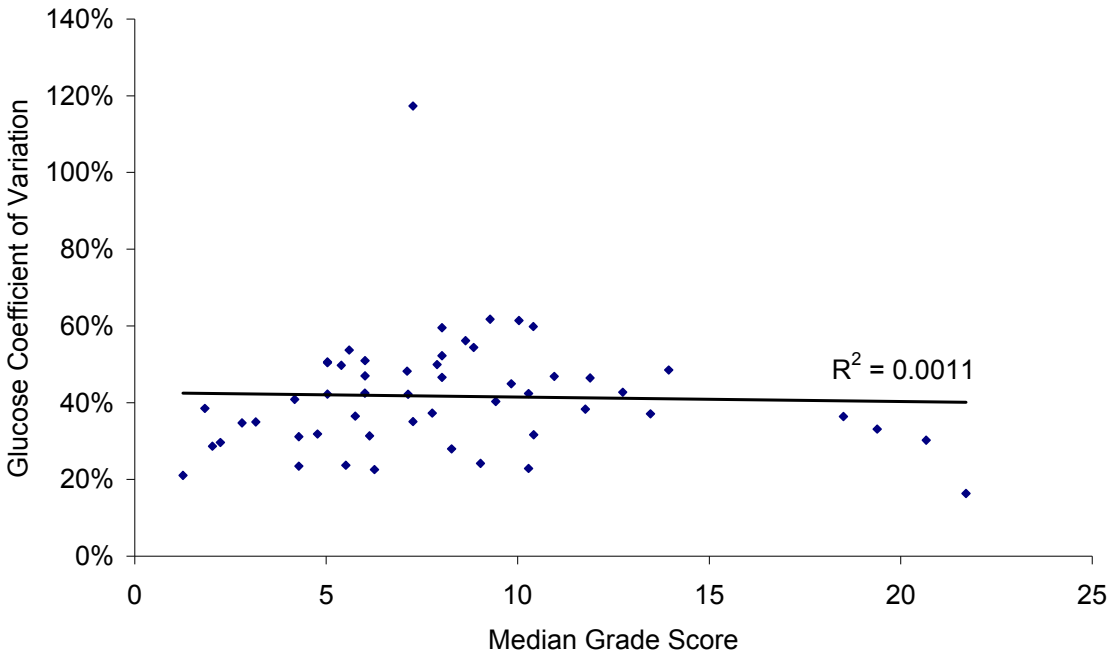


Figure 3.8 – Median GRADE score vs. Glucose Coefficient of Variation

As expected there was no correlation between CV and the median GRADE score. The GRADE assessed the CV in relation to the overall glycaemia and the assessed risk ascribed by the experts in diabetes.

3.6.4 Discussion

Glucose Risk Assessment Diabetes Equation (GRADE) is a simple robust method of quantifying glycaemic profiles into a clinical risk score. GRADE uses all the blood

glucose information available to provide a single value defining the assessed clinical risk to which a subject was exposed. It was noted in non-diabetic and well-controlled people with diabetes that glucose profiles yield GRADE scores <= 5. GRADE scores >5 indicate periods of clinically relevant hypoglycaemia or hyperglycaemia. This has been refined further (in parenthesis) as weighted risk contributions from hypoglycaemia, euglycaemia, and hyperglycaemia. The contribution to the GRADE score of hypoglycaemia, euglycaemia, and hyperglycaemia gives a clinically meaningful measure of glycaemic risk attributable to a specific cause

Non-diabetic groups with GRADE scores <= 5 have hypoglycaemic (BG <3.9 mmol/l) readings. Such readings do occur in those without diabetes without necessarily incurring recognisable clinical signs or symptoms. Indeed such readings may be outliers. It should be emphasised that the weighted risk contribution percentages for hypoglycaemia, euglycaemia and hyperglycaemia are not dwell times in glucose domains. The rationale for reporting the risk divisions of the GRADE score was to guide clinical decisions. If the GRADE score was low (<= 5) then the risk was judged as low and the parenthetic percentages are of no relevance clinically: e.g., a score using relative percentages of GRADE 1 (55%, 35%, 10%) equates to absolute weighted risks of 1 = 0.55 + 0.35 + 0.1.

The calculation of a GRADE score depends on systematic recording of data. Using the GRADE score on data collected at a time when a subject was feeling unwell, due to inter-current illness would generate an inappropriately high GRADE score. This may result in inappropriate clinical action if the content of the recording period is not clearly defined. Similarly, a regular series of post-prandial blood glucose values would indicate a higher risk than justified by skewing the data to hyperglycaemic situations. The ideal glycaemic profiles for a GRADE scoring are near continuous glucose measurement

taken at defined intervals over a period of 24h or more. Continuous glucose measurement systems and multiple self-administered tests generate a large amount of data, and these large data sets are difficult and time consuming to assess clinically[128]. Such data are ideal for GRADE scoring. Continuous glucose monitoring covers both fasting and post-prandial periods and the GRADE score can be calculated without the need to ascertain when these occurred.

The Discriminant Ratio (DR), of a data set, is defined as the ratio of the between-subject variance and the within-subject variance[127]. A DR of 3.76 was obtained on analysis of the prospective data from the fifty -three participants. This ratio indicated that GRADE has the ability to discriminate between different subjects and allow comparison between GRADE scores. The CV is a measure of the dispersion of a probability function. The subjects measured their glucose over twenty-one days and their profiles CV's ranged between 31.6% and 49.7%. The contributing factors to these CV's may be due to the biological variations in people with diabetes and the loss of beta cell function and increase of insulin resistance.

Home blood glucose monitoring (HBGM) is the most common form of blood glucose measuring, but there are many associated problems[129-131]. It has been shown that GRADE scores can be applied when there are sufficient data to be representative of glycaemic control. In the prospective study, it was possible to calculate a GRADE score using twenty-eight blood glucose measurements within a one-week period. The GRADE calculation should be used on data with regular patterns of measurement to get comparable results. For example measuring blood glucose at the same time consistently or always pre/post-prandially Caution should be used when comparing GRADE scores

from continuous glucose measuring with HBGM. This is because pattern of measurements between the two methods is not comparable.

There was a distinct separation of the groups, when quantified by a GRADE score, between those with and without diabetes. In the non-diabetic euglycaemic subjects the spread of GRADE scores was, as expected, tight and well defined. It was noted that well controlled people with diabetes have GRADE scores <= 5, but the majority of those examined had scores ranging from 5 to 13 reflecting appropriately their weighted risk from hypo and hyperglycaemic states. As GRADE scores increase the clinically assessed risk of microvascular complications, macrovascular events and cardiovascular or morbidity are raised.

The contribution to the GRADE scores for people with type 2 diabetes came almost entirely from hyperglycaemia, 92%, with a 1% contribution from hypoglycaemic excursions. Those with type 1 diabetes had a significantly higher contribution from hypoglycaemia (20%) these results correlate with the results from the DCCT [4]. HbA1c is considered to be a indicator of future long term complications[115]. The GRADE score can be used as an adjunct to HbA1c to examine the level of risk associated with different glycaemic control. Over-obsessional attempts to lower HbA1c values carry with them the detriment of increased risk of hypoglycaemia.

The GRADE score was developed as a measure of clinically assessed risk to a subject from hypoglycaemia and hyperglycaemia. While it was possible to have a low HbA1c, corresponding tight glycaemic control does not necessarily follow. Subject numbers CR006 & CR074 (from the Prospective Study) have an identical HbA1c. Subject CR006 achieved a ‘low risk’ GRADE score of 2.2 and subject CR074 had an ‘at risk’ GRADE

score of 10.5. A decomposition of subject CR074's GRADE score attributes risk from both hypo- & hyperglycaemia at approximately the same percentages. A GRADE score > 5 defines poor glycaemic controlled, despite a HbA1c of 7.0, and may be benefited by a more in depth clinical assessment or a change in treatment regimen. Subject CR006 was at low risk and therefore no decomposition of the GRADE score was necessary.

The correlation (R^2) between HbA1c and GRADE and was calculated as 0.41. The correlation indicated a general trend that a higher HbA1c indicated increased glycaemic risk to the subject. Patients with identical HbA1c's and yet different GRADE scores highlight the problem of using HbA1c alone as a measure of individual glycaemic control.

The Prospective Study demonstrated that lower HbA1c's increased the glycaemic risk attributable to hypoglycaemic episodes and conversely as HbA1c's increased the risk from hyperglycaemic episodes increased. This was an expected result as patients with a lower HbA1c are closer to the hypoglycaemic range. If they were closer to this range, it would be easier for the oscillatory nature of a glucose profile to dip into this range and similarly with a higher HbA1c into the hyperglycaemic range.

In conclusion, the GRADE score can be used as an adjunct to other glycaemic assessment methods, such as HbA1c. GRADE can offer clinically applicable insight into the level of control. The GRADE score summarizes diverse glycaemic profiles into a single assessment of risk. Well-controlled glucose profiles yield GRADE scores <= 5 and higher GRADE scores represent increased clinical risk from hypo or hyperglycaemia.

GRADE's use as a method to both assess risk and examine the causes of that risk from hypo- and hyperglycaemia ideally suits it for use as an adjunct to HbA1c.GRADE was been published in Diabetes Medicine in 2007 [132].

4.0 Retinopathy Assessment

Hindsight is always twenty-twenty - Billy Wilder (1906 - 2002)

4.1 Computer vision and retinal images

The analysis of megabyte datasets in this thesis has so far concentrated on the investigation of 2 or less variables of data (i.e. glucose concentration vs. time). Information systems of biological data can have a multiplicity of variables and megabytes of data. The methods needed to extract the information from these datasets may not always have been developed. With the increase in computer processing power it has become feasible to automate the analysis of multi-variable datasets. Retinal images have at least the 3 variables (dimensions) of colour (red, green, blue). If the retinal images are compared longitudinally there is an additional dimension, time.

The information in retinal images may be obscured by noise in the form of blurring of the image, resolution limits of the camera taking the image, artefacts in the form of dust and hairs, astigmatic deformations caused be the lens of the eye and many more. Similar methods to those used in pre-processing 1 and 2 dimensional datasets are needed to reduce the amount of noise in the retinal images in order to evaluate objectively any retinopathy in the images.

Retinopathy is a serious potentially sight threatening complication of diabetes and because of this the National Service Framework (NSF) working party recommends the use of camera technology, using either film or digitisation, to assess retinal images [133]. The English Diabetic Screening Project Advisory Group (EDSPAG) has been established to steer the development of a National Screening Programme for Sight-Threatening Diabetic Retinopathy. The standardisation being implemented, by EDSPAG, of retinal image storage and retrieval is becoming more relevant as the traditional use of

high quality film has gradually given way to digital camera technology. The standardisation is needed for quality assurance of the images being taken as well as the ability to cross-reference with other systems. A set of images may be stored for each patient, and stored images allow for audit and documentation of progression.

The scale of the management task for handling photographs is huge: current estimates are that there are between 2% and 4% of the adult UK population affected by diabetes. Using the conservative estimate of 2% affected by diabetes, this implies over 1,000,000 ophthalmic screening visits every year. If this were to be undertaken photographically, the minimum number of images generated in the UK alone would be 2,000,000 photographs (since each eye needs independent photography). If photographs are collected more frequently, in greater numbers than one per patient, or screening were to be extended to impaired fasting glucose patients, (up to 4% of the population in total) then the estimate may be in excess of 4,000,000 photographs annually in the UK alone.

The problem with accumulating large numbers of photographic images - either on film or in digital form - is that they need review and/or grading to allow interpretation that may influence management and reduce the risk of intraocular deterioration. It is not sufficient to grade a subset of these images, as they all require interpretation. One of the problems of reviewing retinal photographs is that the task is both highly skilled and very repetitive. Skilled ophthalmologists do not have the time for laborious photograph examination and the training of graders has generally only been undertaken for use in trials - for example the UKPDS grading centre set up by Professor Eva Kohner and Steve Aldington at the Hammersmith Hospital. The work is intense and arduous. Trained graders are difficult to recruit, train and retain because of the nature of the work. Grading occurs in dimly lit rooms, and rest periods need to occur frequently owing to the

problems associated with retinal grading, computer systems have been proposed and developed by a number of groups to assist or replace the grader.

4.1.1 Computer analysis of multi-dimension datasets

The use of computers to help automate the task of screening for retinopathy is a relatively new development. This is in part due to the technologic requirements of scanning retinal images. One of the earliest uses of computers to aid the clinician in retinal screening was proposed in 1964 by Monnier et al[134] who analysed the response of the retina to difference frequencies of light . The techniques needed for automatic retinal screening have now progressed substantially. The latest methodologies can to count the number of microaneurysms with a high degree of sensitivity[135]. The sensitivity and specificity of each of these technologies incrementally advance and may soon meet the sensitivity of 80% and specificity of 95% that is required by the National Institute for Clinical Excellence (NICE) guidelines for clinicians[136] .An automated system would need to achieve this in a correctly powered fully randomised prospective study.

A variety of technologies have been proposed to examine these complicated images[135, 137 , 138 , 139]. These computer recognition packages have generally performed poorly in clinical trials by being both complex and non-intuitive[140 , 141]. The aim of the majority of these software solutions has been to identify retinopathy and attempt to quantify it automatically for the user. The use of neural networks[142 , 143] which try and ‘teach’ computing systems to recognise patterns have not achieved a high enough sensitivity and specificity with the result that a considerable amount of energy

has been expended on systems which are not sufficiently reliable to be used in clinical practice.

The use of computers to aid the detection of retinopathy has been accomplished in as numerous different ways as there are groups performing the work. The detection of micro-aneurisms (MA) is being performed by a number of groups[135, 144, 145] as the appearance of MA's is a definitive indicator of background retinopathy and is one of the earliest signs that is detected through retinal screening. Early attempts at automated MA detection have relied on morphological and threshold techniques[144, 145]. The threshold needed to achieve high sensitivity comes at the expense of specificity.

The monitoring of diabetic retinopathy requires the longitudinal analysis of patient images, to differentiate progression from regression or no change. This has been recently addressed by one group[135] . The automated measurements recorded by the software proved more consistent than the manual measurements that showed large variations, but regression of MA's (those that disappeared) were more technically difficult for the system to identify.

The segmentation (by colour differences) of these early warning lesions and quantification of their development is being investigated by other groups[137, 146]. Detection and segmentation of fundus lesions are being investigated, as development of lesions or vessels within the fundus may result in a loss in visual acuity[3]. Confirming the absence of retinopathy, or detecting the location of a single or a small number of artefacts is time consuming when investigating early stages of retinopathy, due to time invested in consideration of borderline objects of doubtful identity. Nicolai Larsen et al[147] achieved an automated lesion detection rate of 90.1% for patients with

retinopathy and correctly identified 81.3% without leisons, progressing automated computer analysis closer to achieving the requirements of the NICE guidelines

Feature extraction techniques are being developed[148] that allow the tracing of retinal vasculature and cross-over (where vessels cross each other on the retina) using region growing and edge detection algorithms. There is a requirement for reliable processing techniques to use with non-ideal images. Feature extraction is a fundamental step in a fully automated analysing system, as the computer needs a mathematically exact definition of the data to be analysed. The feature extraction techniques allow the registration and fusing of retinal images, but retinal image registration is complicated. The images are of a curved surface taken from a wide range of viewpoints sometimes using an un-calibrated camera with artefacts arising from reflection, refraction and dispersion. Involuntary body and eye movements compound these effects.

Usher et al[146] have developed a neural net-based system that attempts to segment and categorise retinal lesions, but a maximum sensitivity of 95.1% could be achieved at the expense of a reduction in specificity to 46.3%, due to the complex nature of the images.

In a related area of image analysis, compression of retinal images is being investigated by other groups[149, 150] and the effect that compression techniques have on the quality of retinal images. By reducing file size, it is possible to send retinal images quickly across computer networks as long as no clinical detail is lost through compression.

Whichever approach used, the common aim is to detect progression of retinopathy early enough to avoid visual loss in the future, by seeking to abstract, extract and reduce the information in a non-steady state pathological process and investigate suitable timed treatment to prevent the progression of retinopathy.

4.2 Development of an Automated Retinal Image Differencing (ARID) tool

4.2.1 The aim in developing ARID

The research plan was to analyse retinal images using the concept of assessment of change rather than specific lesion recognition. This was fundamentally different from the current technologies outlined in the Section 4.1, which are aiming to identify the specific lesions. This may be a more achievable solution due to the complex nature of retinopathy development and the plethora of information available in an image. The aim was to develop a fully Automated Retinal Image Differencing (ARID) computer system to highlight change between retinal images over time for use in primary and secondary care and by retinal screening centres.

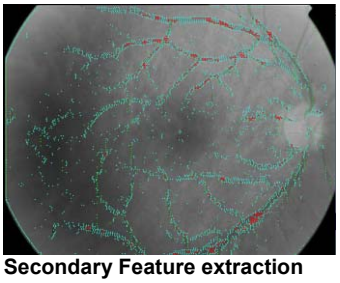
4.2.2 Ethos in developing ARID

ARID was designed with two distinct problems to overcome the registration (matching up of the images) and the differencing (blending the images together and cancelling out items that occur in both images). The first iteration of the ARID program required user input in order to function, this has since been automated. The difference image is now available immediately on accessing a subject's file.

4.3 Methodology in developing ARID

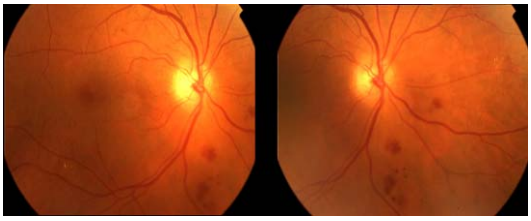
The development of ARID was an incremental process that involved the use of formalised methods and object orientated designs and methodologies. The final version, available on the accompanying CD, is the sum of these processes and may be used as an effective tool in the identification of the progression, regression, or stabilisation/no change in diabetic retinopathy. Each stage of the automation is outlined in Figure 4.1 and discussed in detail in this section. The full code of ARID is available in Thesis Volume II the companion thesis.

Initial Feature extraction



ARID extracted consistent aspects of the retinal image to use as the basis for comparison. This was achieved using a novel edge detection algorithm.

Secondary Feature extraction



ARID Compared pixel intensities around points identified to give a matching score. These scores were optimized to find a maximum likelihood model of the best fit of the images.

Registration & Differencing

Registration (rotation and translation) was performed using matrices of the images using the matching score

Enhancement

In ARID there were three available methods of enhancement. Each one enhances a different aspect of retinopathy.

Figure 4.1 – Processes involved in the ARID software assessment of change.

The development of ARID occurred in a Visual Studio environment and was designed to be run on a Microsoft Windows™ PC.

4.3.1 Selection of images for development

The initial stage in the development of a program to register and difference images was to select a suitable pair of test images. The images needed to be of the same size and

magnification and the colour and contrast had to match as closely as possible. This was to make the initial test as simple as possible. The similar starting conditions contributed to smallest amount of image manipulation that needed to occur when differencing. Future iterations of the ARID software addressed these artificial assumptions and compensated for them. It was decided to use one simple image of a small size (213 pixels x 152 pixels) of the retina located above the fovea. This image was then rotated by 13 degrees and the result was used as the second image (Figure 4.2).

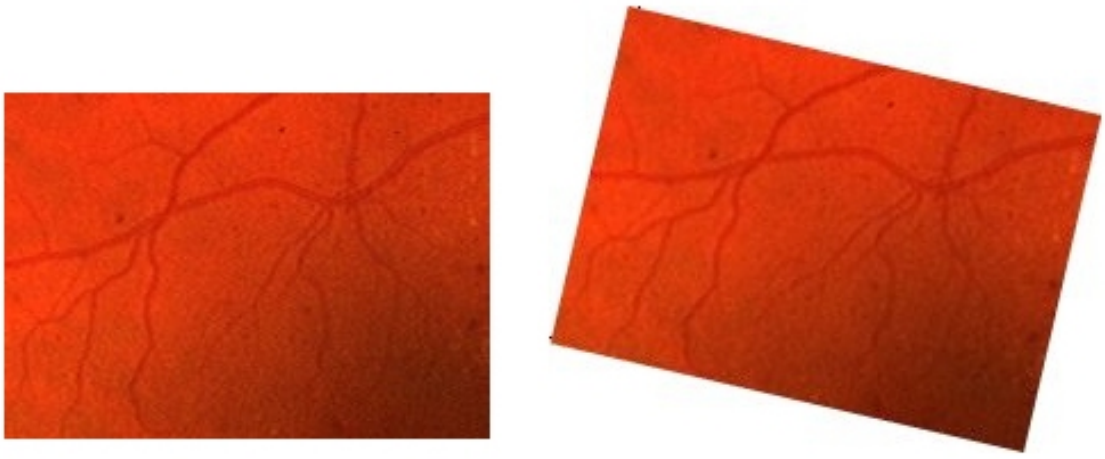


Figure 4.2 – Simple retinal image sample

4.3.2 Development of manual assisted registration of images

In the initial version of ARID, a user chose corresponding point pairs at regions of interest (ROI) (Figure 4.2). The point pairs chosen were at the bifurcation of the vessels, as these were easy to identify and could be matched up with their twin in the second image by the human eye. A minimum of two pairs needed to be selected to enable triangulation.

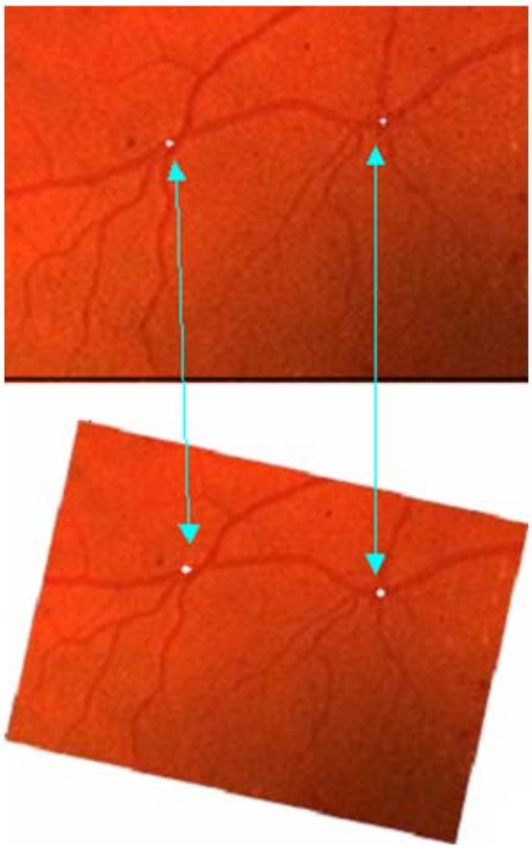


Figure 4.2 – Selection of point pairs

The manual selection of corresponding point pairs introduced the problem of human error (inaccuracy) into the program. The closer the points matched their partners in the corresponding image the more absolute the registration, but it was difficult to select correctly the exact same point in both images.

The solution was to get the program to compensate and recalculate the most likely correct choice for the points identified on the user's initial selections, which was a first step towards removing a need for human interaction. Figure 4.2 shows the initial human selection of the points. From the initial selection, the program takes an area of 20^2 pixels around the first points selected in both images (Figure 4.3); this area is shown as a blue box.

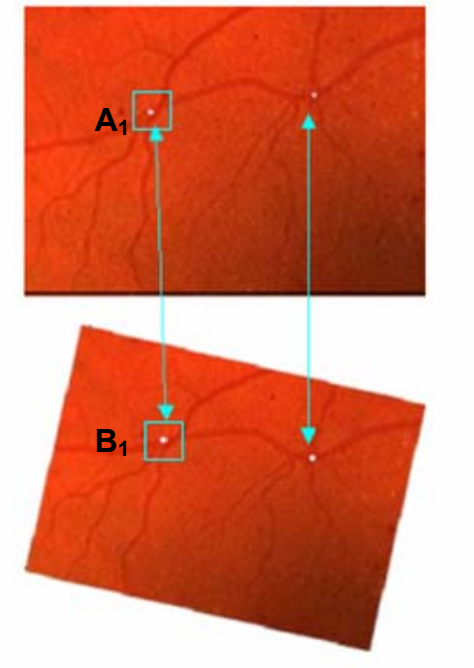


Figure 4.3 - Selection of point areas

Differencing of the colour channels then occurs for these areas. The more exact the match between the areas the smaller the difference between them was.

Consider the point A_1 in the first image to have co-ordinates (X, Y) and Point B_1 in the second image to have co-ordinates (X+i, Y+j). Both i and j are altered incremented from -10 to + 10 and the best match was where $A_1 – B_1 = 0$. The co-ordinates that generate a value closest to zero are then stored as the optimum selected variable. If a more accurate match was identified for the A_1 then a new A_1 was set.

4.3.3 Development of Matrix Rigid Registration

After selection, the registration of the image was performed using the pre-selected points as landmarks for the registration. A translation was performed on the first selected point

pair, Point A1 onto Point B1, using a translation matrix. Mathematically this may be defined as adding 1x2 matrix to another 1x2 matrix, see example. The resulting matrix was the new co-ordinates for the point.

For example if Point A1 had co-ordinates of [20, 20] within Image 1 and Point B1 had co-ordinates of [30, 30] within Image 2. The maths would look as follows:

Point + Translation Matrix = New Point

Point A1 = [20 20]
Translation = [i j]
Point B1 = [20 +i 20+j]
AND
Point B1 = [30 30]
∴
[20 +i 20+j] = [30 30]
i = 10; j = 10
Translation Matrix = [10 10]

The translation matrix, as calculated by the initial point pair, was then used to translate every pixel in Image 1 to its new coordinates including the points selected.

Following translation, image rotation occurred. Point A_1 was used as the origin. The point pairs in Image 1 were rotated by 1° clockwise, EQ 4.1.

EQ4.1

$$x' = \cos(\theta) * x - \sin(\theta) * y$$

$$y' = \sin(\theta) * x + \cos(\theta) * y$$

Theta is the angle of rotation. Co-ordinates (x, y) are rotated to (x', y').The results of the equations are the new position of the points in Image 1.

The reciprocal distance between the new rotated co-ordinates of the selected points in Image 1, and the unchanged co-ordinates of their twins in Image 2 was then calculated using a least distance algorithm. The results were stored for later comparison. The points in Image 1 were then rotated by another 1°, and the new reciprocal distance

recorded. This occurred for 20° both in the clockwise and anticlockwise directions. The smallest reciprocal distance between the point pairs was identified and the corresponding rotation was stored as the optimum one. Image 1 was then rotated by the optimum angle of theta.

The manual registration worked but when presented to screening staff it was felt that it would benefit greatly by not having the human input of selecting corresponding point pairs. The development of an automated version was then designed and produced.

4.3.4 Development of automated registration for ARID

The development of a novel method to identify the vasculature in a retinal image was an integral part of the process for fully automating the ARID program. The vasculature of the retina is a relatively fixed object and was considered a candidate starting location for comparison between images. There is movement of vessels on the retina over time but crossover points of vessels remain fixed. Using the crossover points, the structure of the vasculature tree was comparable over time.

4.3.4.1 Development of pre-processing techniques used in ARID

The colours of vessels on the retina tend to be slightly darker than the surrounding background (Figure 4.4). This difference in intensity of colour was a starting point to allow a difference to be distinguished and to specify that a vessel may be present. The difference in intensity was not an axiom and retinal images where the pupil was not fully dilated may have excessive shadowing of the image.

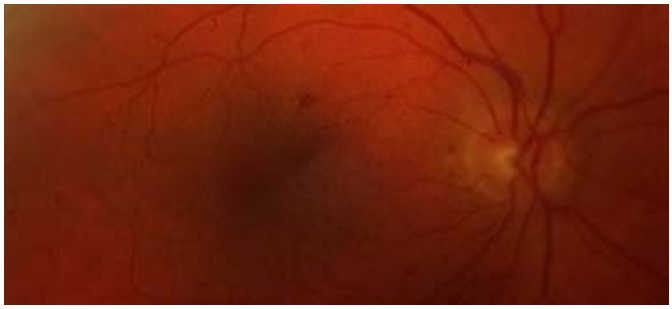


Figure 4.4 – Vessel Arcade

The retinal images under comparison are rarely of the exact same field of the eye. The most common field to be examined is the fundus in 45 degree field[151]. The 45 degrees under observation are rarely the same 45 degrees observed on the previous image. The vasculature extracted from the two images may be from different parts of the retina. Alignment of the images (by registration) needed to occur before differencing could be undertaken.

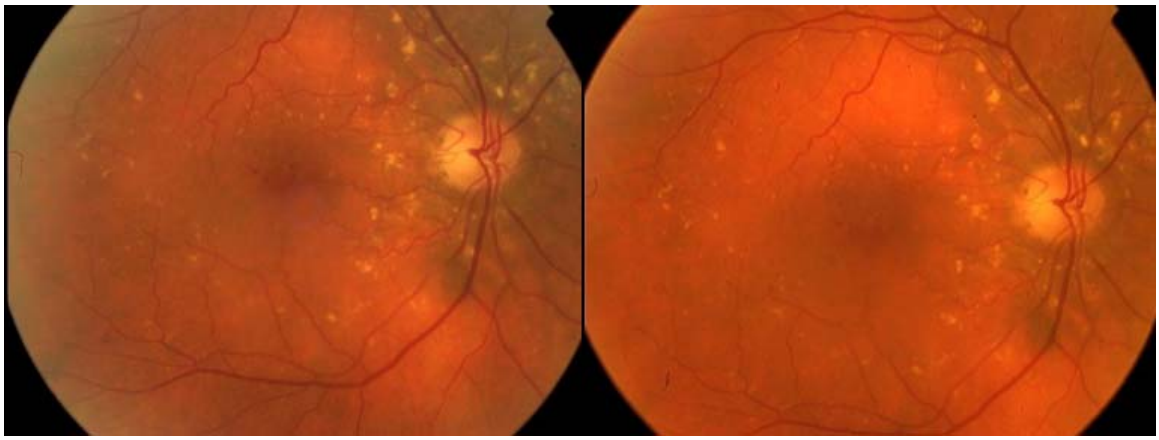


Figure 4.5 – 45° fields from the same person 1 year apart

Artefacts in the image need to be taken into account either artificial such as a hair or a reflection, which may either occlude part of the vasculature or be mistaken for part of it. These are compensated for by ARID's pre-processesing stage as they fail to qualify for the rules governing the identification of a vessel and are not included.

Figure 4.6 – Artefacts in a retinal image

Pre-processing of the images occurs to obtain the optimum amount of information[152], and to standardise the images. This was a crucial phase in the automated registration process. To overcome the multiplicity of problems an enhancement of the obscured information had to occur. ARID pre-processing follows three stages:

- The de-shadowing of images allowed 'hidden' information to be used in the calculations for correctness of fit.
- Equalisation of resolution and dimensions allow like to be compared with like.
- Compensation for artefacts in the images

4.3.4.1.1 Development of a De-shadowing technique for images

When pupil dilation fails to occur to a suitable level the quantity of light reaching the back of the retina is reduced. This lack of light is one cause of the darker areas (or shadows) that occur in retinal photography (Figure 4.7). There was less useable information

available in these areas, due to the very low intensity levels, but it may be pertinent to use all the information available in the image. This 'hidden' information may sometimes be critical to the successful mapping of one image onto another was little other data to use.

Figure 4.7 – Shadowed retinal image

The method of de-shadowing was a two-stage process. The non-shadowed was examined and then an algorithm adapted the shadowed areas to bring them into line with the lighter quartiles of the image. The de-shadowing process was achieved by selecting random pixels from the image (up to 10%) whose intensity levels were in the top three quartiles of the image intensity. These pixels formed the basis of the estimated appearance of the remaining the image. The frequency of occurrence of intensity ranges was calculated by assigning the pixels collected into 'bins'. The bins each covered an equal length range of intensity. The frequency of the pixels assigned to each bin represented the spread of pixel intensities throughout the non-shadowed areas of the image.

The shadowed areas were then selected, possible shadow pixels (that occurred in the bottom quartile of colour intensities) fitting this description were identified and assigned

to an array. Once these ‘seed’ points were established each was assessed according to the pixels surrounding it. If an adjacent pixel qualified in the bottom quartile of intensities the probability of the first pixel being part of a shadow was increased. Each seed pixel was assessed to a depth of five pixels in every direction. The pixels that qualified at 60% certainty of being a shadow pixel were then assigned to bins exactly like the first phase.

A lookup table was created where the distribution of intensity of the shadowed bins were remapped to match the distribution in the first phase. The resulting image had considerably less shadowing present (Figure 4.8) and revealed part of the retina that could be used for the auto-registration of the images. The extrapolation of these pixels meant they were less informative than the un-extrapolated areas and less importance was placed on them during registration.

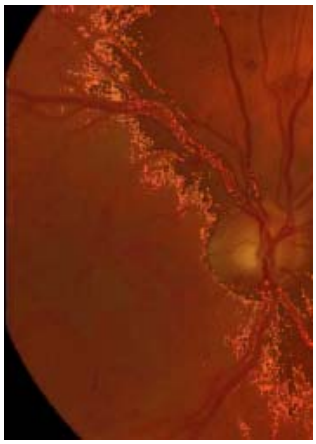


Figure 4.8 – De-shadowed area of the retina

Retinal images are rarely directly comparable as they may have been taken at different resolutions on different cameras and hence are different sizes. The images are reduced using the visual studio resample function. The interpolation and compositing quality are set to maximum to avoid the loss of information. The original images are kept for the

differencing method and the size and resolution matched images are used for comparison of the vessel arcades.

4.3.4.2 Vessel detection algorithm

The next step was vessel edge identification. The identification of the vessels was a multilayer process, the first step being the identification of the edges of the vessels. Vessels are normally darker than the background of the retina. This was not always true but it gave a starting point in the identification. The edge of a vessel was identified as a higher intensity moving to a lower intensity and staying low for a predefined width or starting low and moving to a high intensity (Figure 4.9). This method incurs false positives but it does capture the vessel edge.

Figure 4.9 – Intensity Map of a retinal Vessel

4.3.4.2.1 Vessel detection algorithm Stage 1 – width detection of vessels

The widths of the lower intensities were detected and the vessels width was defined by the image size and resolution. Vessels in the retina may vary in their width and length and a lower limit was set on the specific width of a detection must be to qualify as a vessel, expressed as a percentage of the image size. Conversely, an upper limit was set for vessel width, and this further reduced the number of false positives as the differentiation between vessels and areas of haemorrhage was now more clearly defined.

4.3.4.2.2 Vessel detection algorithm Stage 2 – Edge detection of vessels

After obtaining a candidate point that was estimated to define the edge of a vessel, the surrounding environment of that pixel was examined, in both the horizontal and vertical domains. The pixels directly surrounding the candidate were examined in exactly the same way as the proposed pixel. If they qualified for potential vessel edge status then the candidate pixels probability of identifying a vessel edge was increased and this information/pixel was stored for later analysis. If the surrounding pixels failed to conform to the intensity and width checks then the candidate pixel was discarded.

These processes meant that microaneurysms, flame haemorrhages and other types of retinopathy, including localised bleeding, were not mistaken for the edge of vessels by failing one of the initial tests for edge detection. False positive identification did occur as some features in the images satisfied the rules governing the identification of a vessel wall (approximately 5%).

An array of candidate pixels for the edge of the vessels had been identified. There were false positives present due to artefacts, shadows, or alternative retinal features that qualified for the tests defined for vessel wall detection. The next stage refined the process further. The theory behind the refinement was that the process expected a few false positives but by fulfilling those criteria it made them less likely to pass the last filter unless they were a definite vessel wall pixel.

The final stage was the proximity of candidate pixels to each other. The detection of vessel walls may generate a number of pixels on both sides of a vessel (Figure 4.10). By examining the number of candidates surrounding the current selection it was possible to

increase the likelihood of the candidate being a definite vessel wall. The higher the number of other candidate pixels within a certain area, the higher the probability of being a correct identification.

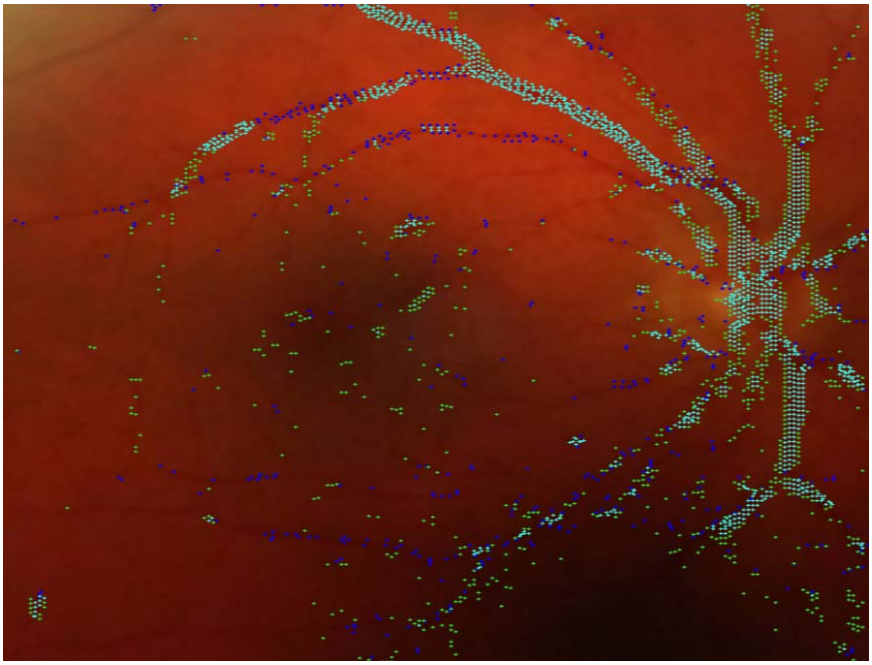


Figure 4.10 – Candidate points for possible vessel edges

These final candidate points were then re-identified as ‘definite’ points. The mid point between the definite vessel wall pixels was then taken as the centre of the vessel. The points left after final filtering had the highest probability of being the vessels and formed the basis of the automated registration of the vascular tree. The identification of the vasculature was then considered ‘good enough’[153] as the majority of pixels identified were on the vasculature and the remaining false positive were few and sparse. This total processing time of the image to determine the ‘definite’ points took approximately 13 seconds.

4.3.4.3 Fast normalised cross-correlation algorithm

Cross-correlation is a standard method of estimating the degree to which two series are similar[64, 154]. An image may be considered as a matrix of different intensities and saturations. The degree to which each member of the matrix correlates with another matrix was calculated. The normalised form of correlation (normalised correlation coefficient) used with images, for template matching, does not have a simple and efficient frequency domain interpolation. The calculation of the correlation was performed in the spatial domain initially, but this process increases computation time. A normalised cross-correlation from a transform domain convolution provides an order of magnitude increase in computation time over spatial domain computation. The normalised cross correlation equation implemented was the same equation used in making the film 'Forest Gump'[155], to register templates to frames in the film .The Fast Normalised cross-correlation (FNCC) seen in EQ4.2.

EQ4.2

$$C(u,v)=\frac{\sum_{xy}\left(I[u+x,v+y]-\bar{I}uv\right)\hat{T}(x,y)}{\sum_{xy}\left(I[u+x,v+y]-\bar{I}uv\right)^2\sum_{xy}\hat{T}(x,y)^2}$$

Where I is image one, and $\hat{T}$ is image two. An increased correlation (C) the more exact the match.

Using the FCNN equation enabled the two pre-processed images to be compared rapidly at different angles and translations. The 'definite' vessel pixels (identified in the

previous filtering) were correlated and the optimum match was identified in a mean time of 13.4 seconds.

4.4 Differencing of images using ARID

The two images were now registered as closely as possible using the normalised cross-correlation. The next objective was to subtract (difference) mathematically the two images. This was accomplished by inverting the colours of one of retinal pictures (Figure 4.11). The two images would then cancel out and display grey where the inverse colour in one image overlaid the unchanged colour in the other image.

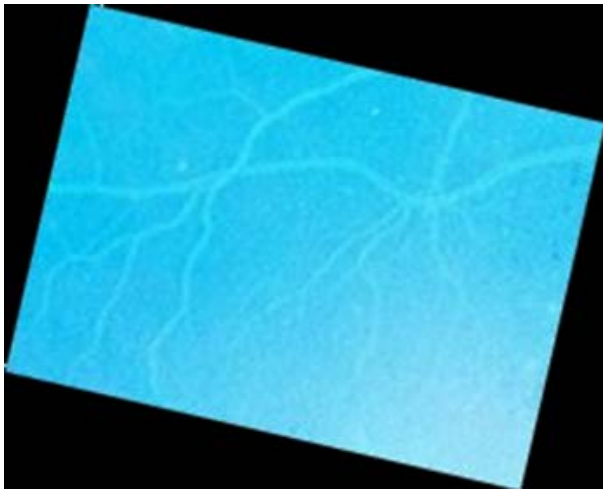


Figure 4.11 – Colour Inversion of retinal image

The images were converted using the RGB colour model where Red, Green or Blue have a minimum value of 0 and a maximum value of 255 and the mixture of these 3 numbers identifies a colour. There is a fourth member of the RGB model the Alpha channel, the number for this represents how transparent the colour is.

An example of the colour ‘Bright Red’ would be:

RGB = [255, 00, 00]

R = 255

G = 0

B = 0

And with an Alpha number of 255 for completely solid

The inversion of the pixels' colour (Figure 4.11) was accomplished by taking the three RGB numbers for Red, Green and Blue and subtracting them from 255. This created a colour that was exactly the inverse of the prior colour. The inverted image was then placed over the other image, after registration, with the Alpha levels of the inverted image being set to exactly 50% or 128. This meant that 50% of the colour in each pixel in the inverted image would be visible and 50% of the unchanged colours in the other image would be visible.

When the pixel colours are exactly the inverse of the pixel colour in the other image they cancel each other out and grey would be displayed (Figure 4.12). Differences in the images do not have a corresponding match in the other image so they are easy to distinguish, as they do not disappear.

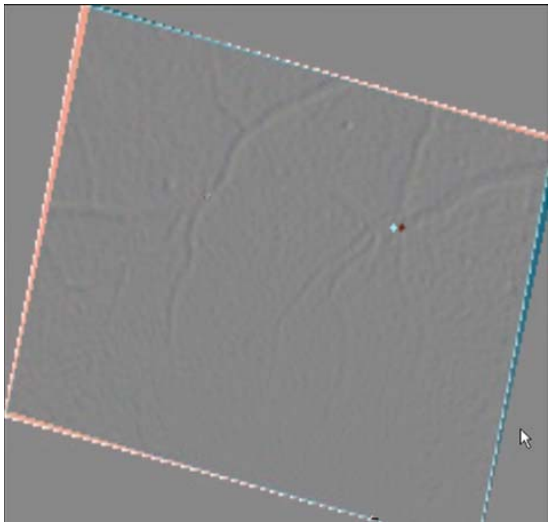


Figure 4.12 – Differenced or alpha blended retinal Image

4.5 Validation of the ARID software

The testing of the ARID program was an integral part of the development. Before use, the software needed to be both sufficiently robust and accurate for use in a diagnostic and therapeutic setting. Therefore, the software needed to match up and difference a range of images. The differenced images, in this validation, were not graded, as the purpose was to assess the robustness, reproducibility, and accuracy of the ARID program.

Ten image pairs were selected in a blinded random manner from an existing anonymous dataset held at Oxford Centre for Diabetes, Endocrinology and Metabolism (OCDEM), Oxford, UK. The main features of a retinal image (Fig 4.13).The image pairs were centred on the fovea and covered a 45-degree field. Each image pair was loaded into the ARID program and the automated registration and differencing algorithms were selected. The results were then exported for manual visual comparison.

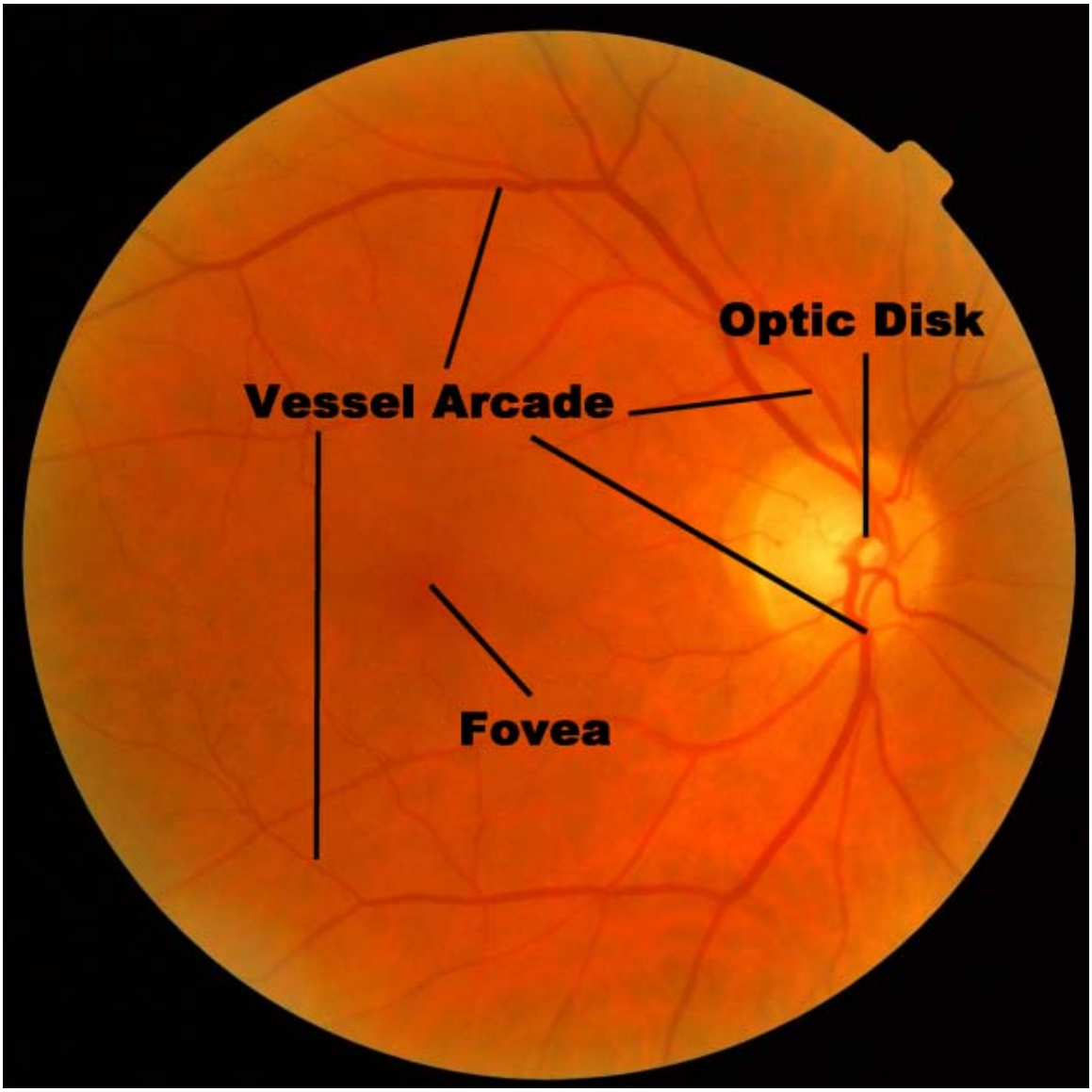


Fig 4.13 – Retinal image with major landmarks indicated

4.5.1 Results of validation

The results from the testing are graphical in nature and are displayed below. It was possible to ascertain at a glance whether the software has been successful or not.

Subject A

Fig A_1 - Right eye – 11/99

Fig A_2 - Right eye – 04/99

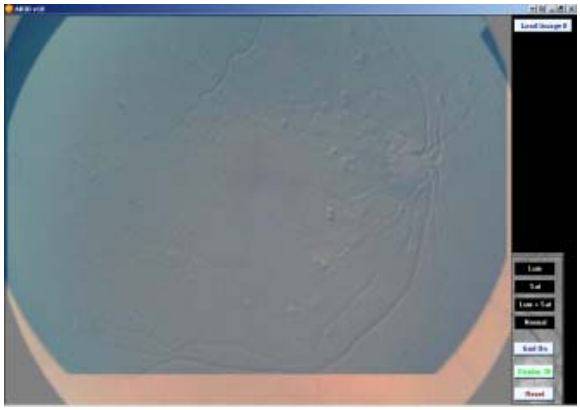


Fig A_3 - Final registered & differenced image

Subject B

Fig B_1 - Right eye – 09/01

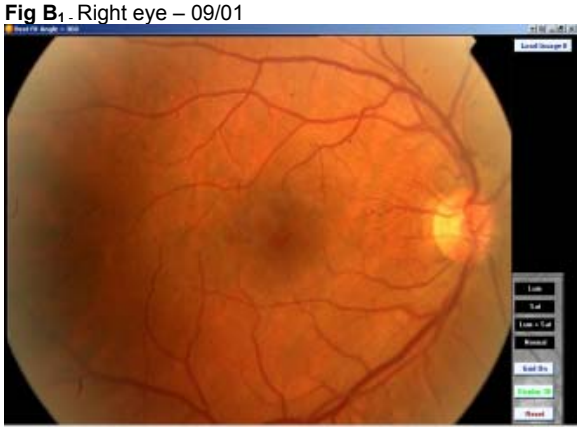


Fig B_2 - Right eye – 03/00

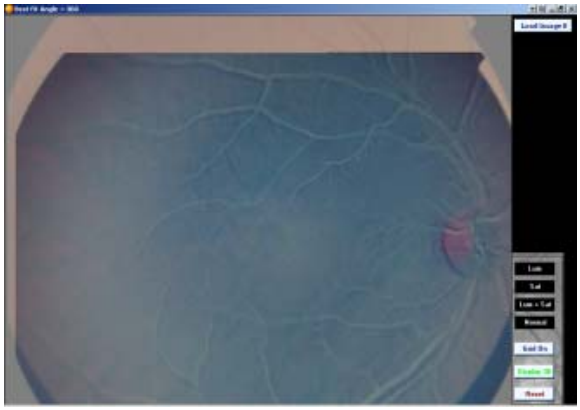


Fig B_3 - Final registered & differenced image

Subject C

Fig C_1 - Right eye – 07/00

Fig C_2 - Right eye – 07/01

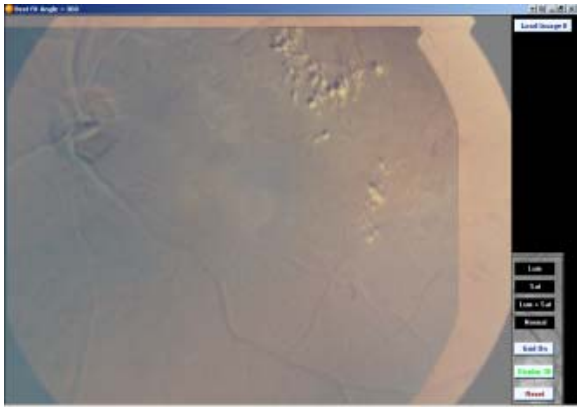
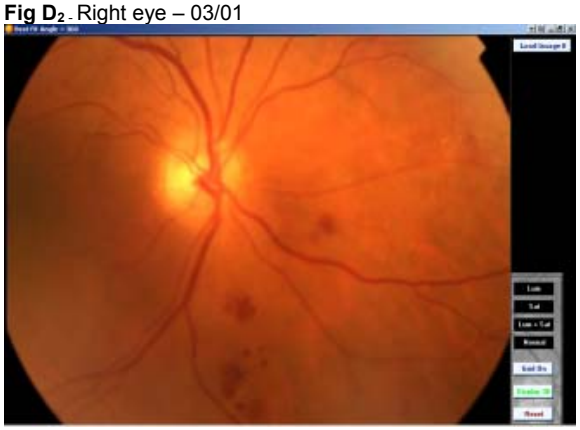


Fig C_3 - Final registered & differenced image

Subject D

Fig D_1 - Right eye – 06/00

Fig D_2 - Right eye – 03/01



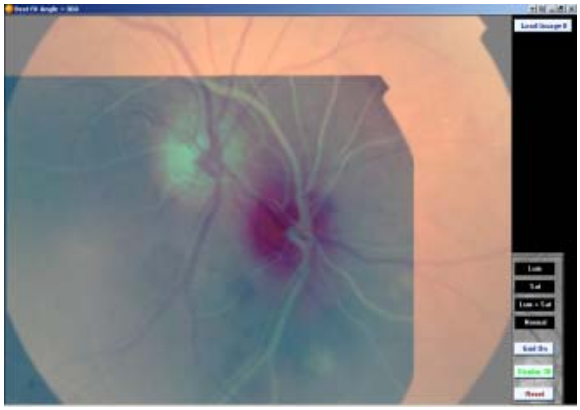


Fig D_3 - Final registered & differenced image

Subject E

Fig E_1 - Right eye – 07/00

Fig E_2 - Right eye – 07/01

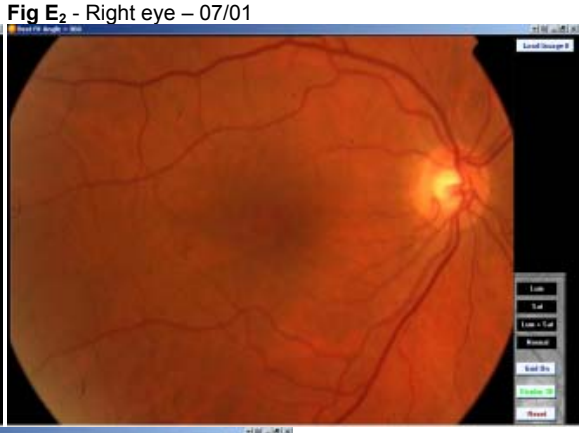


Fig E_3 - Final registered & differenced image

Subject F

Fig F_1 . Right eye – 02/00

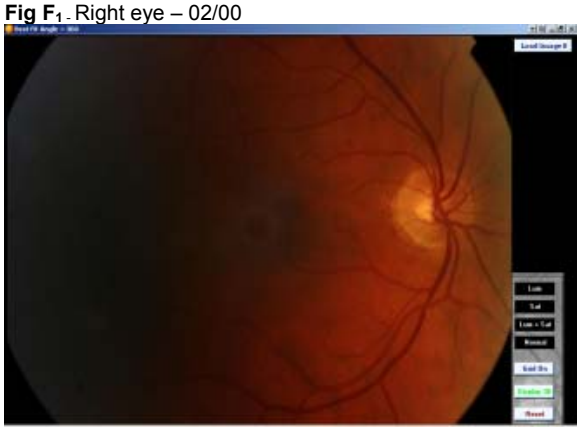
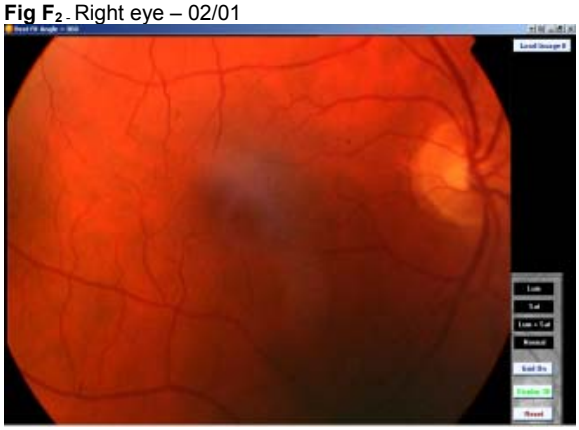


Fig F_2 . Right eye – 02/01



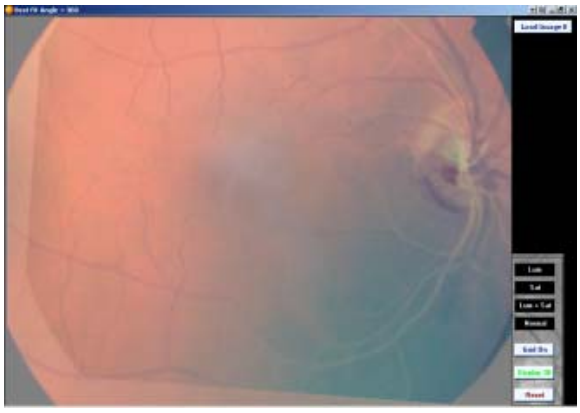


Fig F_3 . Right eye – Final registered & differenced image

Subject G

Fig G_1 . Right eye – 03/01

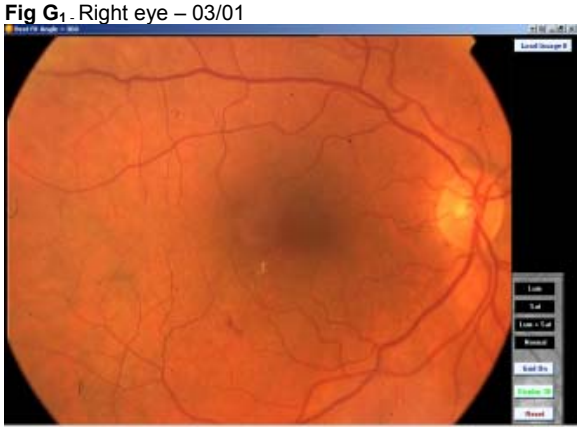


Fig G_2 . Right eye – 03/03

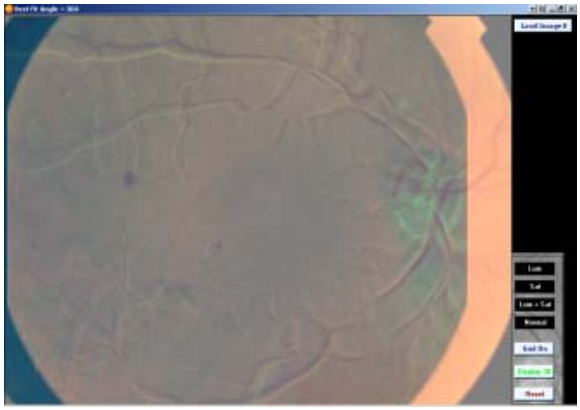
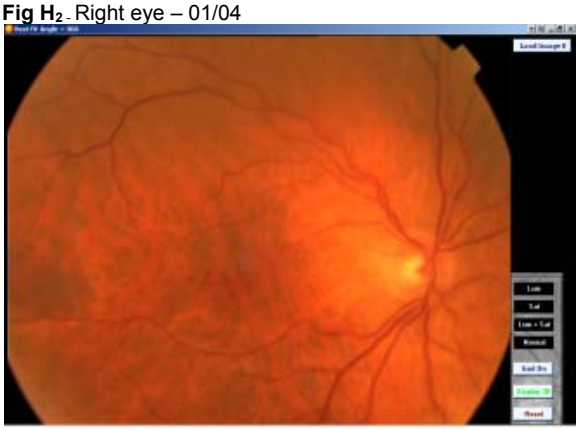


Fig G_3 - Right eye – Final registered & differenced image

Subject H

Fig H_1 - Right eye – 01/99

Fig H_2 - Right eye – 01/04



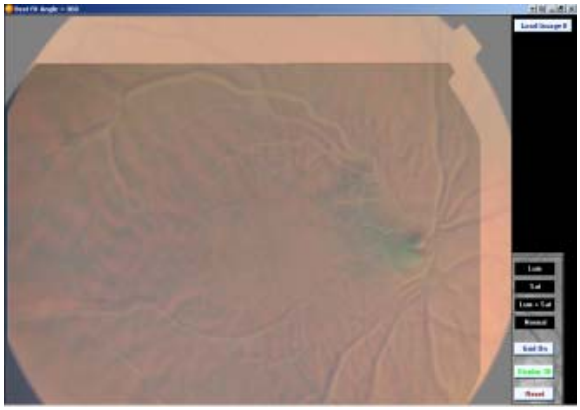


Fig H_3 - Right eye – Final registered & differenced image

Subject I

Fig I_1 . Left eye – 01/99

Fig I_2 . Left eye – 01/00

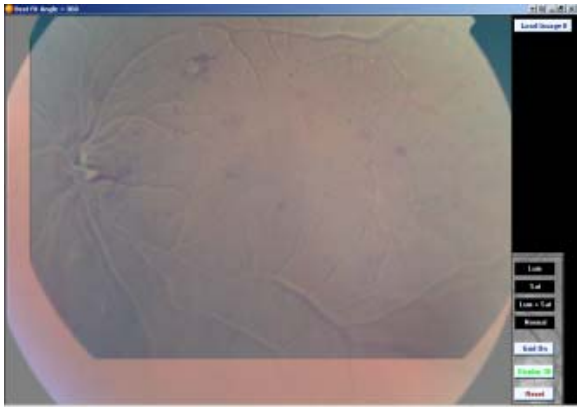


Fig I_3 - Left eye – Final registered & differenced image

Subject J

Fig J_1 - Left eye – 03/01

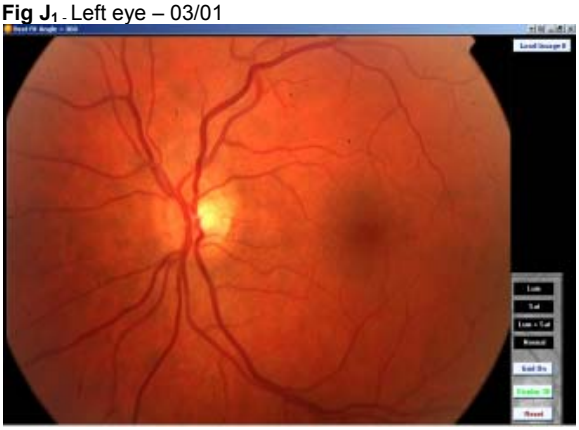


Fig J_2 - Left eye – 03/02

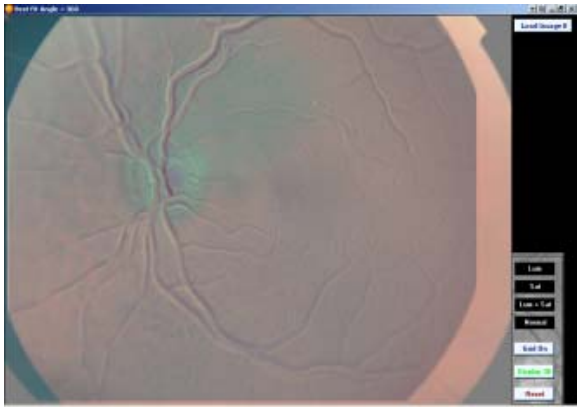


Fig J_3 - Left eye – Final registered & differenced image

4.5.2 Discussion of validation

The ten patients testing achieved successful automated registration in 90% of image pairs. Out of the 10 image pairs selected Subject D failed to register and difference correctly. This may have been due to one of the images of the pair not being correctly centred on the fovea. It was possible to register the images using the manual method and selecting corresponding point pairs, but approximately 33% of the image for Subject D was available for comparison due to the lack of overlapping fields. This illustrates the need for high quality photographs with a similar field of view.

The ARID program provided identical results when run a second time supporting system reproducibility. This demonstrated an advantage over human assisted registration and differencing, as the latter results depend on human operations that are unlikely to achieve 100% reproducibility.

4.6 Clinical application of ARID

The ARID system analysed similarities in images, matched them up, and identified the changes between them. This demonstrated what had changed to the user. It was not intended to replace the retinal image screener, but to reduce the workload of images that require grading. ARID was designed to be a proof of concept of the idea that it is possible to take two images of the same eye, at different times, and to difference them leaving the residual of what has changed. In essence a very simple idea, yet in practice a highly complex problem. The resulting image would show the differences between the images and the matching architecture of the vessel arcade would cancel out. This would make it possible for the screener, grader, and ophthalmologists to see instantly if there had been any change in the state of the retina since the last image was taken.

A retrospective study sought to determine if the use of the ARID software was an aid to the retinal screener. It was thought that ARID might act as a filter to reduce the number of images that require grading and classifying, reduce in the subjectivity of grading by producing an image of change that is operator independent. The reduction in the number of images requiring grading would save time and money.

4.6.1 Methodology of clinical study

4.6.1.1 Subjects

In 1995/6, 3611 patients were recruited following a clinical audit in Gloucestershire primary care. The details of the audit have been discussed elsewhere[156]. The Gloucestershire Diabetic Eye Screening Programme developed a research dataset from this audit. From these data 160 images (80 image pairs) were obtained. Forty image pairs were selected that had no apparent change between the two images and forty

pairs were selected that displayed a change in the state of the retinopathy. These two groups were selected in a random blinded manner. The 80 subjects comprised those with type 1 and type 2 diabetes who were over the age of 16 years. Exclusion criteria included previous record of laser therapy, low quality images and those with significant cataracts. The existing research dataset was approved by the local medical ethics committee. Participants gave written informed consent according to the Helsinki Declaration after receiving full information of how their images were to be used.

4.6.1.2 Photography

A Topcon TRC-NW55 (TOPCON, Tokyo, Japan) interfaced with a Sony 3-chip video attachment with a pixel resolution of 768x576 and a file size of 1.26 MB was used with the IMAGEnet 2000 computer system (TOPCON, Tokyo, Japan). A non-mydriatic-type digital fundus camera was used as poor iris dilation even following pharmacological mydriasis is a feature of diabetes[157] but all patients were imaged following instillation of the mydriatic Tropicamide 1.0%.

Digital images were captured in true colour (16,777,216 colours), labelled and stored in JPG format on the IMAGEnet computer system. One 45-degree field retinal image, centred on the macular, was selected from each patient visit. Pairs of visits for each patient were then masked and randomly coded as A or B for the first or second visit. The coding information was not available to assessors.

4.6.1.3 Manual grading

The dataset was graded by Retinopathy Grading Centre staff Helen Lipinski, Dawn Groves, Dianne Hodds at the Retinopathy Grading Centre (RGC), Imperial College,

London. An assessment of the quality of the photograph was made and photographs were marked as N/A (Non Assessable) if the image quality made it difficult to determine whether a characteristic was present in a field. Assessable images were then scored according to the Early Treatment of Diabetic Retinopathy study (ETDRS) scale [158]. The image pairs were finally categorised according to how each type of lesion had progressed, regressed, or not changed between image A and image B. Lesions assessed were: microaneurysms (MA), haemorrhages (haem), hard exudates (HE), cotton wool spots (CWS), intraretinal microvascular abnormalities (IRMA), venous beading (VB), new vessels at disc (NVD) or elsewhere (NVE and fibrous proliferations at disc (FPD) or elsewhere (FPE). Three independent graders from the RGC graded the digital images in a masked fashion. Retinal images were viewed on 19-inch CRT monitors with settings of 1024x768 and 32-bit colour (16,777,216 colours).

4.6.1.4 Automated Differencing

The ARID software was used to perform the automated differencing of the sequential image pairs. The same three graders who were involved in the manual grading of the image pairs performed the grading of the ARID images in a masked fashion. The ARID grading was performed after a 30-day gap from the initial grading to minimise crossover. The differenced images provided by the ARID software were graded without image enhancement. The image pairs when differenced using ARID produced an image map (Figure 4.14) of the change between the longitudinal retinal photographs. These ARID images were used to assess whether the retinopathy had changed or remained unchanged between the image pair. Following all grading, the image pair orders were unmasked for analysis of results.

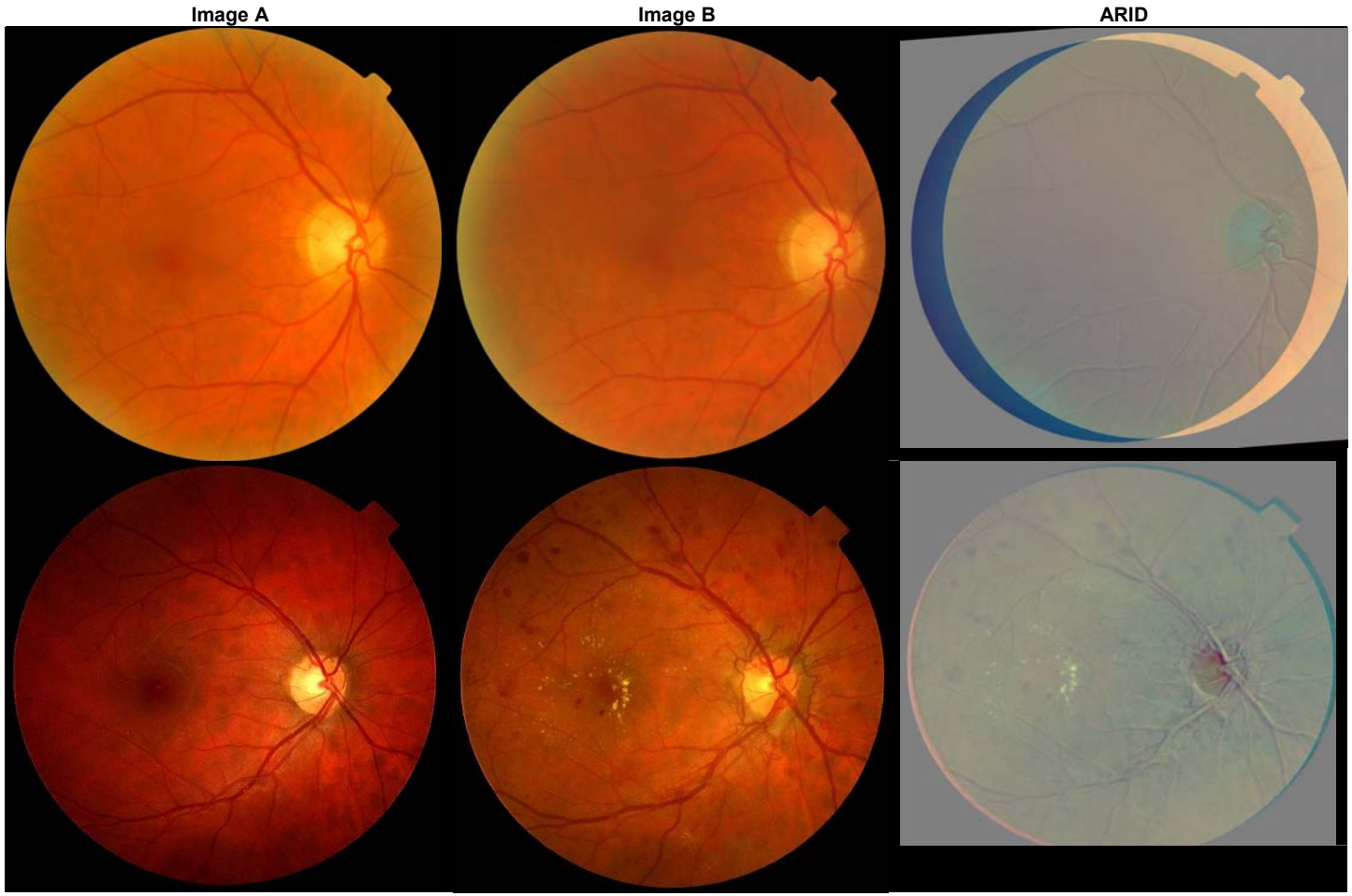


Figure 4.14 – Example of the images produced by ARID for use in assessment of change

4.6.1.5 Statistical analysis

Subjects were categorised as 'Progressed', 'Regressed' or 'No change' depending on the results of the comparison between the image pairs. Kappa statistics (level of agreement), sensitivity and specificity (EQ 4.1) were performed to analyse the within-grader variability and the between-grader variability whilst using the ARID software and compared to the conventional methods.

Sensitivity

$$sensitivity = \frac{TP}{TP + FN}$$

Specificity

$$Specificity = \frac{TN}{TN + FP}$$

EQ 4.1 – Sensitivity and specificity equations where TP = True positive, FP = False positive, TN = True negative , FN = False negative

A power calculation could not be processed as the assessment of change in retinal images had not been attempted before. Eighty image pairs were selected arbitrarily to cover a range of retinal complications. McNemar Chi square was used to check equality of sensitivities and specificities. A p value of <0.05 was the criterion for significance.
The sensitivity, specificity and kappa assessment were performed on the results using manual grading contrasted to the results obtained using the ARID software.

As there was no comparison to determine if the manual grading was performed correctly the 'Gold' standard for comparison was the agreement of all 3 graders on the direction of change (split by lesion type) on the manually assessed images. The sensitivity, specificity, kappa could then be calculated between this 'Gold' standard and the ARID assessed image results.

4.6.2 Results

A total of 80 patients (80 eyes, 160 images) were photographed and their retinopathy status assessed by visual grading by three retinal graders. Four patients were excluded from the analysis due to the images being classified as ungradeable by the visual grading system, leaving 76 image pairs assessed using the manual method and the ARID method.

4.6.2.1 Manual Grading

Consensus between the three graders occurred on 55 (72.3%) of the images. The between grader sensitivity was 84.1% and the specificity was 56.3%. The specificity statistic was influenced by a false-positive classification (i.e. patients with no retinopathy but identified by the grader as retinopathy being present) of 18.4% (14/76) (Table 4.1).

Lesion Type	Count
Microaneurysms Only	4
Microaneurysms + Hemorrhages	1
Microaneurysms + Hard Exudates	2
Microaneurysms + Cotton Wool Spots	2
Hemorrhages Only	2
Hard Exudates Only	1
Hard Exudates + Cotton Wool Spots	1
Cotton Wool Spots Only	1
TOTAL	**14**

Table 4.1 – False-positive detection of retinopathy from the manual classification of three graders

The sensitivity (84.1%) was influenced by the false classification of an image pair as having ‘No change’. The false-negative classification accounted for 9.21% (7/76) of the dataset (Table 4.2).

Lesion Type	Count
Microaneurysms Only	3
Microaneurysms + Hemorrhages	0
Microaneurysms + Hard Exudates	0

Microaneurysms + Cotton Wool Spots	0
Hemorrhages Only	2
Hard Exudates Only	1
Hard Exudates + Cotton Wool Spots	0
Cotton Wool Spots Only	1
TOTAL	**7**

Table 4.2 – False-negative detection of retinopathy from the manual classification of three graders

4.6.2.2 Automated analysis

The comparison of manual grading to grading using ARID (Table 4.3) demonstrated an increased sensitivity and specificity. The mean sensitivity of ARID (87.9%) was increased significantly in comparison to manually grading sensitivity (84.1%) ($p<0.05$). The specificity of the automated analysis (87.5%) increased significantly from the specificity (56.3%) achieved by manually grading ($p<0.05$).

	Grader A	Grader B	Grader C	Mean (SD)
Sensitivity	85.0%	84.1%	94.5%	87.9%
Specificity	86.1%	90.6%	85.7%	87.5%
Kappa (95%CI)	0.71 (0.55-0.87)	0.73 (0.58-0.89)	0.80 (0.65 - 0.95)	0.75

Table 4.3 - Sensitivity, Specificity and Kappa of three graders. Comparison between manual grading and Automated assistance grading (N=76)

The Kappa statistics assessed the grader agreement. In general, it is agreed that a kappa of 0.6-0.8 may be regarded as 'substantial' and a kappa 0.8-1.0 may be regarded as 'almost perfect' [159]. The agreement was between the manual assessment and the assessment using ARID. A mean kappa of 0.75 was achieved indicating 'substantial' agreement.

In the manual grading (without using ARID) consensus was reached between the three graders on 55 image pairs. To remove the between-grader variability an analysis was performed on this subset. The manual grading results were compared to the ARID analysis (Table 4.4).

	Grader A	Grader B	Grader C	Mean (SD)
Sensitivity	86.5%	89.2%	100.0%	91.9%
Specificity	88.9%	94.4%	88.9%	90.7%
Kappa (95%CI)	0.72 (0.53-0.91)	0.80 (0.64-0.97)	0.92 (0.80-1.0)	0.81

Table 4.4 - Sensitivity, Specificity and Kappa of three graders. Comparison between manual grading and Automated assistance grading (N=55)

The subset's mean sensitivity of manual grading compared to ARID was increased from 87.9% to 91.9%. The mean specificity was increased from 87.5% to 90.7%. The increase in sensitivity and specificity was significant ($p<0.05$). The mean Kappa increased from 0.75 to 0.81 progressing from a 'substantial' to an 'almost perfect' agreement.

4.6.2.2.1 Results by grader in the subset

The sensitivity of Grader A (86.5%) was affected by a 9.1% (5/55) false negative count (i.e. the images assessed using ARID were ascertained to be lesion-free when they had been marked as having retinopathy present using manual grading). The five false negative assessments were images containing only microaneurysms. The specificity (88.9%) was achieved with a false positive count was 3.6% (2/55). One false positive assessment was due to incorrect haemorrhage identification and the other a false assessment of a microaneurysm.

The sensitivity (89.2%) and specificity (94.4%) for Grader B (Table 4) was generated by a false negative count of 7.3% (4/55) and a false positive count of 1.8% (1/55) respectively. The false negative classification was due to two microaneurysm/ haemorrhage images; one microaneurysm image and one haemorrhage image being

missed by the assessment using ARID. The false positive was due to the single misclassification of a microaneurysm image.

Grader C achieved 100% sensitivity as no false negatives were scored. The specificity was 88.9% as the classification of false positives was 3.6% (2/55). The false positives were generated by one image falsely classified as having microaneurysms and cotton wool spots when it was previously classified in the manual grading as having no retinopathy and the other image misclassified as having microaneurysms.

4.6.3 Discussion

A variety of groups have sought to automate the grading of retinopathy with varying levels of success on retrospective data. Of these groups none have used algorithms to assist the retinal graders in analysing longitudinal change in retinal images. The aim of these groups has been to assess fully the retinopathy in order to achieve the National Institute for Clinical Excellence guidelines[136] and to replace the human element of grading. In this study, computer algorithms were used to process sequential retinal digital images. ARID located the major retinal landmarks and the vascular tree from the image pairs. Using the information obtained on the vascular tree the images were cross-correlated and an optimum match obtained. The images were then differenced without human interaction. The images generated by the ARID software did not rely on the hardware used to capture the images and the output was the same if different resolution images were used. The differenced images demonstrated the change in the intervening time between the capturing of the two images. The time taken to produce the ARID change image was not related to the resolution of the image. All image resolutions took

the same time to process as the images were scaled down to determine the vasculature and optimum correlation and then the originals were used for differencing.

Unlike the other studies of grading the aim of ARID was the absolute assessment of the presence/absence of retinopathy by its clinical grading in ETDRS. National guidelines of sensitivity and specificity[136] are not relevant here because those guidelines apply to the selection of lesions, not to the detection of change.

The sensitivity and specificity of all the graders using the manual grading of the image pairs was 84.1% and 56.3% respectively compared to the 'Gold' standard of complete agreement between all 3 graders. The specificity in particular was unexpectedly low but demonstrate the subjective assessment of manually grading retinal images. The lower rates may have been due to the quality of the images that were used for the grading. The images were 768x524 pixels in resolution. Modern retinal cameras can take pictures that are 3000x1960 or greater. The relatively lower resolution of the images would make it more difficult to identify the point haemorrhage/microaneurysm or other minor retinopathy where the image resolution was low.

The within-grader assessment, whereby grader A using the manual method was compared to grader A using the ARID method, produced a mean sensitivity of 87.9% and the mean specificity of 87.5%. The comparison of these results to the between-grader manual assessment (84.1% sensitivity, 56.3% specificity) demonstrated an increase in sensitivity. This significant ($p<0.05$) increase meant that more retinopathy had been detected using ARID than by using manual inspection. The significant specificity increase ($p<0.05$) reflecting a decrease in false positives. If this increase in

sensitivity was replicated in clinical practice there would be a resulting reduction in the number of images needing detailed grading, as fewer images would be incorrectly assessed to have progressing retinopathy.

The subset of images (N=55) was comprised of those image pairs where the three graders agreed on the direction of change (split by lesion type) using the manual grading. The within- grader sensitivity was significantly increased ($p<0.05$) from 87.9% to 91.9% and the specificity significantly increased ($p<0.05$) from 87.5% to 90.7%. The use of ARID as an aid to graders would reduce the need to grade images where no retinopathy had occurred and where existing retinopathy had not progressed. The aim was to decrease the workload of the graders by allowing those images demonstrating no change to be triaged to allow graders to concentrate on images where retinopathy might be progressive. ARID achieved a high specificity (90.7%), and as the majority of patients with diabetes in most screening populations are likely to have no retinopathy (and therefore no change), the application of ARID to this population would substantially reduce the burden of manual grading. The risk of overlooking a single or few microaneurysms may be inconsequential in clinical practice where substantial change is the trigger for the need to refer onwards to expert ophthalmologic review.

The within-grader Kappa statistics indicated that there was a ‘substantial’ (0.75) agreement between using ARID and using the manual method, and this increased to an ‘almost perfect’ (0.81) agreement when analysing the subset of images in which there was no disagreement between the expert assessors.

False positives and false negatives may reduce with the availability of a differenced image, and personal settings of enhancement capabilities may help graders. The missing of lesions (false-negatives) may occur due to a random function or the graders systematically repeating a mistake. If the latter is true, it may be possible to identify markers of when detection is poor and consider this as a topic for individual re-training schedules.

In conclusion ARID demonstrated an increased sensitivity (91.9%) and specificity (90.7%) of retinal grading compared to analysis by purely manual inspection. The automatic display of an ARID differenced image where sequential photographs are available would allow rapid assessment and appropriate triage.

5.0 Non-Linear dynamical systems (chaos)

It will take time to restore Chaos [in Iraq] – George W Bush, April 13, 2003

5.1 Introduction to non-linear dynamical systems

The techniques and methods used in the analysis of megabyte datasets within this thesis have examined data in terms of the number of variables (dimensions) and the quantity of data. For example, Fourier Transform sought to extract information obscured in the data by analysing the oscillations in terms of spectral power. The methods used throughout make a basic assumption that the data are both predictable and reproducible. This assumption is not always correct. Biological systems exist that are deterministic yet are neither predictable nor repeatable. Instead they exhibit chaos, where a small change in the initial conditions produces a wholly different outcome.

According to the Oxford English dictionary "Chaos" is defined as:
'Behaviour of a system which is governed by deterministic laws but is so unpredictable as to appear random, owing to its extreme sensitivity to changes in parameters or its dependence on a large number of independent variables; a state characterized by such behaviour.' [14]

Chaos analysis is part of the field of non-linear dynamical systems analysis, the study of systems that evolves in time. Non-linear dynamical (Chaos) systems can be observed using the simplest experiments. A study by Shaw [160] in 1984 examined a dripping tap. Shaw recorded the amount of drips using a microphone. As the drip rate was increased the linear periodic predictable drip became irregular with periods between the drips doubling or halving in length - in essence chaotic in behaviour. A minute adjustment in the initial conditions changed the outcome.

It seems inherently unlikely that biology is anything other than deterministic (cause and effect). A number of comparative studies have been performed [161-168] examining cause and effect of influencing insulin secretion or sensitivity. It is possible to analyse the insulin secretion or sensitivity in the Cartesian way (i.e. line of best fit) but the system is influenced by several hormones which are in turn influenced by glucose and other hormones. The insulin secretion has unpredictable elements. To analyse such feedback systems in biology a different set of tools may be needed to the ones used in the analysis of megabyte datasets so far. Chaos theory can be used in the assessment of biological phenomena or systems that are governed by deterministic laws yet appear to be unpredictable due to their sensitivity to initial starting conditions.

A further example is heart rate it is controlled by a natural pacemaker that produces steady regular heartbeats. On extended observation occasionally beats are dropped. If this rhythm becomes increasingly irregular, the change in timing of one beat may be magnified in the change of the next beat. If arrhythmia continues it may lead to morbidity. Are regular predictable oscillations therefore desired? The answer is: 'not quite'. Several studies have looked at heart rates in a variety of pathophysiological situations[169-172] and noted varying oscillations and perhaps chaos. Pincas et al in particular examined children who have had an aborted sudden infant death episode. They published that there was an association with more regular heart rates than the 'completely normal' controls[173]. Regular predictable rhythms, with low chaos, were not the desired. The analyse of approximate entropy of the heart rates provided a insight into a potential risk factor for sudden infant death syndrome.

5.2 Chaos analysis and glucose homeostasis

The glucose regulatory system is a dynamic system that maintains glucose homeostasis through the feedback mechanism of glucose, insulin, and contributory hormones [90]. Diabetes, a multi-factorial disease, affects glucose homeostasis [174] whereby subjects are unable to produce enough insulin or unable to use the insulin they produce effectively. Use of chaos analysis techniques to explore the information produced from the observation of this system may aid in the understanding of glucose regulation. In order to implement chaos methods in the study of glucose regulation, the glucose regulatory system can be thought of as three separate terms:

- The signal: Glucose is the signal that elicits the triggering of feedback mechanisms, like insulin secretion, which traditionally are thought of as a homeostatic relationship[175]

- The chaos: The glucose regulatory system contains elements of chaos, deterministically fractal (further information in chapter 1.2.6), where the relationship between glucose and the feedback mechanisms may repeat over time but the detail tends to infinity.

- The noise: The relationship invariably contains stochastic processes and erratic 'noise' from other biological processes, or assay measurement (further information in Chapter 2 on noise).

Insulin secretion and action has been examined here employing chaos modelling. It is known that insulin and glucose, analysed independently, can be characterised by

chaotic features[176-178]. The inherent chaos observed in the glucose concentrations observed by Katayama[177] enabled the creation of a prediction algorithm that in turn enabled a tighter control of glucose levels. The amount of Chaos in the system may be either beneficial or detrimental and is a property that is often overlooked because it is difficult to understand.

Since insulin is the main anti-hyperglycaemic factor and its main role is to ascertain proper energy homeostasis and euglycaemia, its interaction with other factors can be simplistically examined at an effects level, namely the insulin – glucose phasic relationship. The glucose regulatory system attempts to maintain a homeostatic 'norm', to settle at a point, which in chaos analysis is termed an 'attractor'. Feedback processes correct towards an attractor by making high excursions lower and low excursions higher. The domain within which the oscillations of feedback occur is termed a 'limit-cycle'. For example, when postprandial glucose levels are raised, the pancreatic β-cells secrete insulin to reduce the glucose level[3]. If glucose levels drop below a threshold then additional glucose is released by the system. The process of actual release of energy is multi-factorial. In simplified terms, the glucose, amino acids and fatty acids are converted into ATP (adenosine triphosphate) by cells that contain mitochondria through an intricate cycle of processes. The potential energy of the metabolically active substances is released when the ATP is broken down into smaller units. This feedback system has a limit cycle with biological threshold levels (at which glucose and insulin are released) that oscillates around the attractor of the optimal blood glucose level.

Individually isolated endocrine cells have their own inherent secretion pattern and if an external stimulus (i.e. glucose) with a different periodicity is imposed then the ensuing

hormone secretion will depend on the relationship between these two periods. Secretion may resume with a partial harmonic of the stimulus, but in other cases apparently random secretion will take place giving irregular or chaotic patterns. There is latency or lag within the glucose regulatory system. In people without diabetes the first phase response of insulin secretion to glucose stimulation occurs within 30 seconds with a periodicity of 12-13 minutes [36]. With diabetes the first phase response to stimulus may be delayed or smaller and the periodicity altered[36]. In theory, if the latency is altered (e.g. gain in the feedback loop increased) then the system may become unstable and oscillate periodically.

Systems may oscillate outside a limit cycle, and have the possibility of changing to a new attractor. An example for such a change in glucose regulation is the onset of gestational diabetes. The insulin, glucose feedback system is disrupted during pregnancy and may not return to the same homeostatic point or may set a different attractor and develop into diabetes.

The glucose regularity system can be visualised at rest at a single moment in time using the HOMA model [179] that calculates the ß-cell function and Insulin resistance of a subject. To visualise the complex homeostatic relationship over a period of time when not at rest, glucose and insulin may be described as a point on a graph with co-ordinates (x, y). At the next time period the measurements of the system may have changed and a new point can be plotted with coordinates (x_1,y_1). If a series of these points are plotted over a period, a graph of the history of the glucose insulin phasic relationship can be constructed (Figure 5.1).

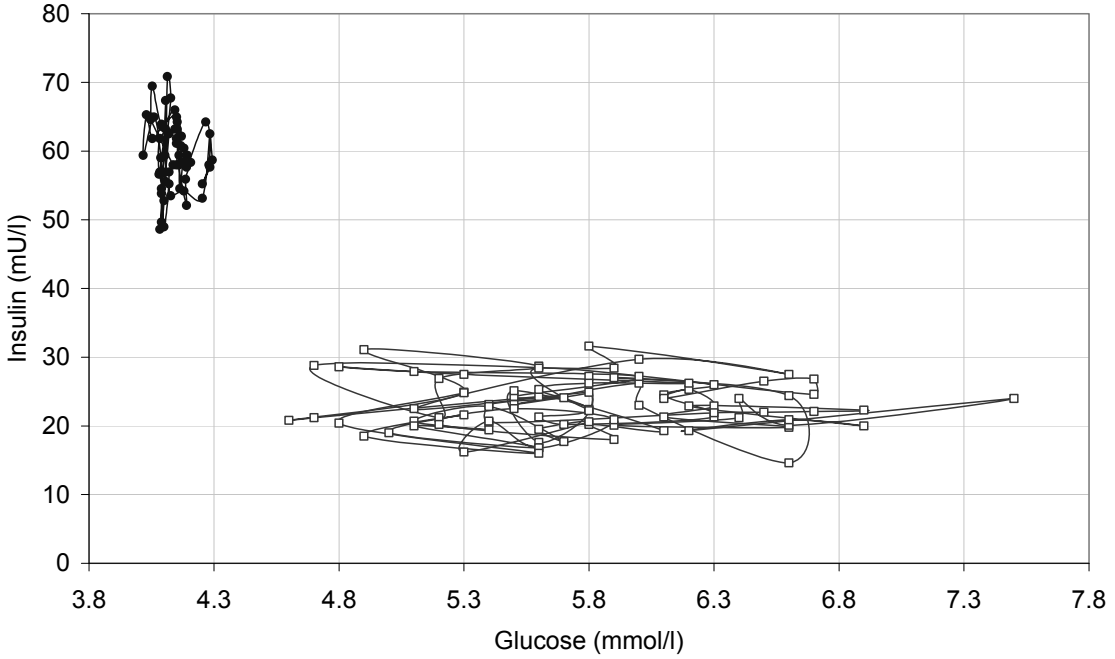


Figure 5.1 – Diagram of the glucose insulin feedback system measurements taken every 3 minutes for 180 minutes in a person without diabetes (●), and a person with type 2 diabetes (□).

5.3 Short term chaos analysis of homeostasis

Of the feedback systems in vivo it was the aim to investigate the applicability of the non-linear methods available in chaos analysis to the glucose regulatory system. Chaos analysis of the phasic relationship between glucose and insulin has not been previously studied. To investigate this relationship a method to assess the extent of chaos characteristics was created.

5.3.1 Subjects and Methods

An audit of a megabyte dataset from an existing study was performed. The data were generated by Prof David Matthews during an investigation into the effects of

somatostatin on pulsatility *in vivo*. Thirty normal subjects and four people with type two diabetes were studied. Insulin and glucose values were measured for 120 minutes with sampling every minute. The basal period was sixty minutes before administration of somatostatin and it was this period that was used.

5.3.1.1 Chaos analysis techniques

The subject's glucose regulation relationships were plotted in time to allow the assessment of homeostasis to be made.

5.3.1.2.3 Assessment of attractor point

The attraction point was the co-ordinates around which the insulin-glucose relationship plot oscillated. For each subject the mean value for glucose and the mean value for insulin were calculated. These values were then plotted as the attraction point of the homeostatic systems.

5.3.1.2.4 Assessment of Limit Cycle

The limit cycle of each subject's homeostatic system was the phasic plot of insulin and glucose. Each phasic plot described a circumference around the attractor as the concentrations of the substrates were adjusted by the metabolic pathways. To define the limit cycle of a subject a 95% confidence interval (CI) for glucose and insulin was used.

5.3.1.2.5 Normalised Area of Attraction

The normalised area of the Attraction was calculated by defining an oval using the 95% CI of glucose & Insulin (the limit cycle). The 95% CI measures the probability that 95% of the data are within the range it defines from the mean. Using the 95% CI for glucose and insulin it was possible to calculate an oval centred on the attractor and that used the glucose 95% CI as width and insulin 95% CI as height.

The insulin data were non-normally distributed and were therefore logged to normalise the distribution. This log transform converted the distribution in the insulin domain to one comparable to the distribution observed in the glucose domain. Subsequently the glucose was multiplied by a factor of 0.5 to normalise a 1 unit change of this substrate to a 1 unit change Insulin. The factor of 0.5 was derived from the ratio between the 95%CI of glucose and 95%CI of normalised insulin calculated from the subjects without diabetes. The conversion factor was needed as without this a 1 unit change in insulin concentration would increase the area disproportionately compared to a 1 unit change in glucose. The areas of the limit cycles could then be robustly compared. The equation used to determine the area of the oval was in EQ5.1.

EQ 5.1

$$Area = Height * Width * \frac{\pi}{4} \quad where,$$
$$Height = \log(\bar{y} + Y_{CI}) - \log(\bar{y} - Y_{CI}) \quad and$$
$$Radius = (X_{CI} * 0.5) \therefore$$
$$Width = 2 * (X_{CI} * 0.5) \therefore$$
$$Width = X_{CI}$$

Where $\bar{y}$ = mean insulin, y_{CI} =95%CI of insulin, x_{CI} = 95%CI of glucose

5.3.1.2.6 Lyapunov exponent

A key element of deterministic chaos is the sensitive dependence on the initial conditions. The Lyapunov exponent (λ)[180] is a quantitative measure of this dependence. It is a measure of the mean rate at which points in a dynamical system move with respect to each other (i.e. their trajectories). The absolute value of the exponent indicates the degree of stability.

- $\lambda < 0$ The system is stable and periodic..
- $\lambda = 0$ Lyapunov exponent of zero indicates that the system is in some sort of steady state.
- $\lambda > 0$ The system is unstable and chaotic.

The number of independent variables affecting the subjects' glucose regulatory system, and therefore the outcome for each subjects system cannot be predicted in terms of resultant glycaemia. However if a series of analyses were preformed to compare changes in the relative trajectories of the points in the system over time -in other words one calculated accumulated errors in predicting one outcome from the other – it would then be possible to discern if there arose:

- a steady state single point attractor.
- a periodic attractor.
- a aperiodic strange attractor.

The Lyapunov exponent was calculated using the normalised ratio between glucose and insulin.

5.3.2 Results

The analysis of the phasic relationship between insulin and glucose in the glucose regulatory system (Figure 5.2) indicated a smaller range for glucose and insulin excursions with the non-diabetics subjects.

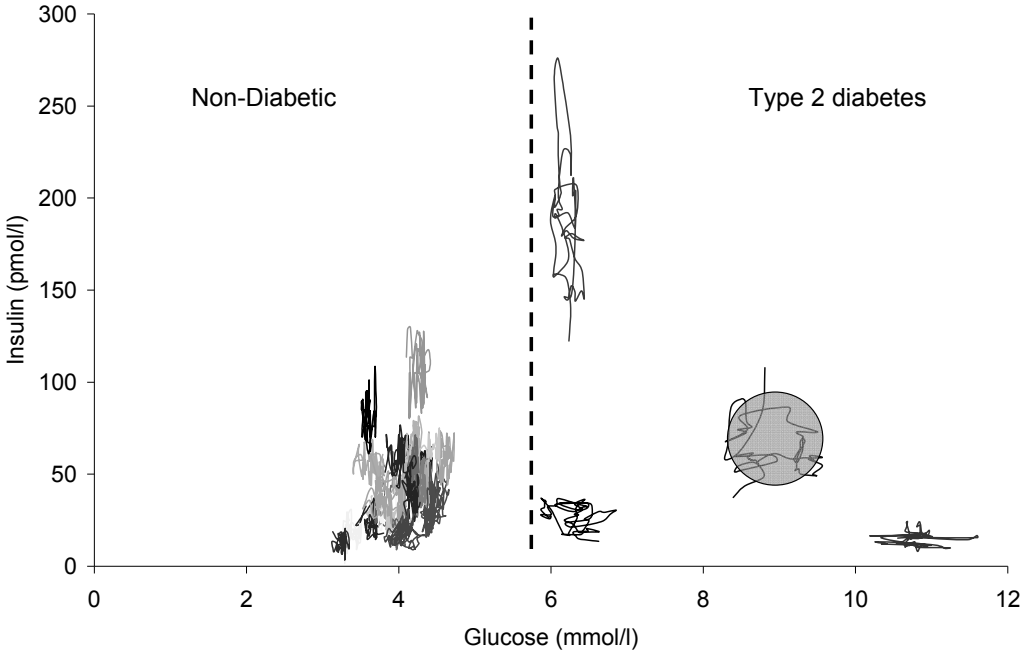


Figure 5.2 – Relationship of glucose and insulin feedback mechanism 30 non-diabetic subjects with measurements taken every 1 minutes for 60 minutes and 4 people with type 2 diabetes with measurements taken every 1 minutes for 60 minutes. Area calculated by oval encompassing 95%CI of insulin and glucose

The attraction points of the subjects were calculated as the mean glucose and mean insulin. The ranges of the glucose and insulin excursions were then assessed as the 95% CI, Figure 5.3. The insulin data was displayed on a logged scale to determine the Area of the limit cycle domain.

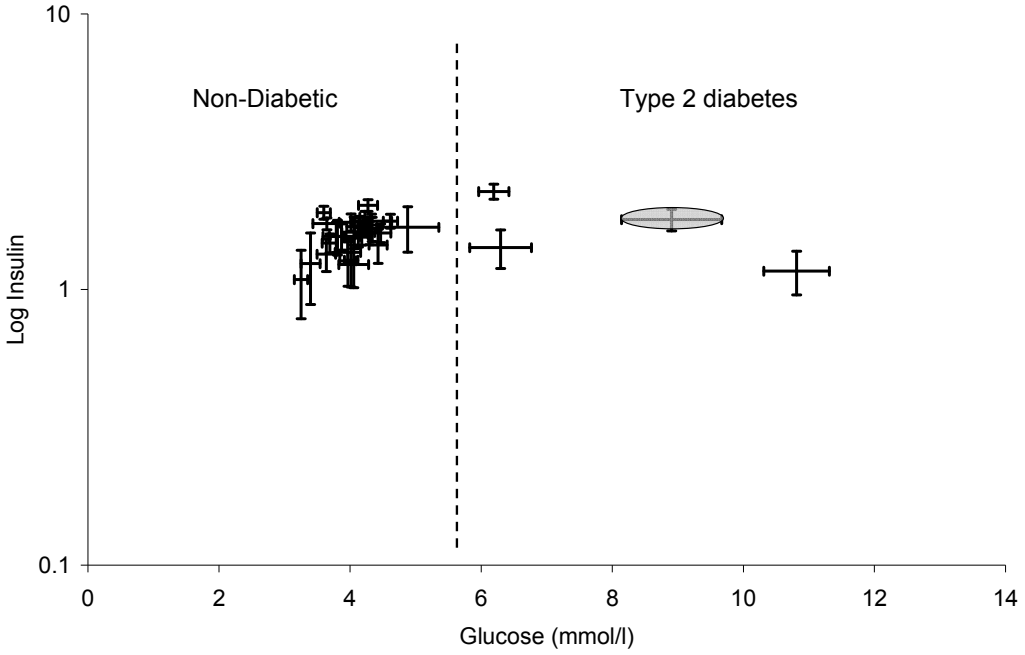


Figure 5.3 – Attraction points of the subjects defined by mean glucose and mean insulin with 95% CI for both substrates. Insulin is log transformed.

The attraction points and limit cycles of the subjects were examined, Table5.1, with the novel metric Normalised Area of Attraction (NAA) establishing a significant ($p<0.01$) difference between the glucose regulatory systems of the two groups. A Lyapunov Exponent evaluation demonstrated that the maximum exponent was positive in all subjects in both groups but was not different between the two groups ($P=0.16$).

	Control subjects	Subjects with type 2 diabetes	*p*
N	30	4	
Mean Glucose (Attractor)	4.05	22.7	0.002
Glucose 95%CI (Limit cycle)	0.16	0.49	9.76×10^{-7}
Log Mean Insulin (Attractor)	1.58	1.66	0.55
Log Insulin 95%CI (Limit cycle)	0.17	0.19	0.45
Normalised Area of Attraction (SD)	2.28(2.16)	7.65 (3.36)	1.17×10^{-4}
Lyapunov exponent (λ) (SD)	+0.42(0.21)	+0.36(0.22)	0.16

Table 5.1 – Chaos metrics.

Point Attractors, Limit Cycles, Normalised Area of Attraction and Lyapunov exponent comparison between subjects with type 2 diabetes and controls using Pearson's independent samples t-test [assumed equivalent variance t-test where Levene's Test for Equality of Variances was significant].

The examination of the NAA determined a method by which it was possible to contrast the reaction time of insulin to the glucose concentration , the secretion rate of insulin, and the release of glucose in response to energy needs. Non-diabetic subjects had varying homeostatic 'norms' that the feedback mechanisms *in vivo* sought to maintain. When this was contrasted to the 4 subjects with type two diabetes the insulin resistance and/or insulin production did not function as well or as quickly and the phase domain determined by the area of the limit cycle was increased. One subject in the normal group had a large NAA with a higher basal glucose concentration this may be indicative of impaired fasting glucose of impaired glucose tolerance.

.

5.4.4 Discussion

The regulation of glycaemic control (the outcome) can be seen to be dependent upon glucose and insulin (the initial starting conditions) in a non-linear and sensitive way. The attraction point, limit cycle and NAA methods used to assess the feedback mechanism enabled an evaluation of the scope of response of the glucose regulatory loop. The interaction of these variables in a feedback system was demonstrated to be chaotic by the observation of the limit cycles. The short term analysis of the glucose regulatory system determined the boundaries within which insulin secretion and glucose regulation fluctuated were quite broad, and this was something observable in subjects across the whole spectrum of glucose metabolism, in the non-diabetic subjects and the subjects with type 2 diabetes. The phasic glucose insulin relationship, in the control subjects maintained a narrower limit cycle around a fixed attractor point, whereas the limit cycle of people with diabetes was increased. This may be due to less insulin either being produced and/or being used less effectively so that the feedback mechanism observed by the phasic relationship increased in area. This would be in keeping with the knowledge that diabetes is characterised by insulin resistance and/or decreasing insulin production.

The Lyapunov exponent established chaos as being present in both groups indicating a sensitivity to initial starting conditions it did not reach significance between the groups. The sensitivity to initial starting conditions was equivalent between groups this may have been due to the small group size or the treatment of the diabetes group. The Normalised Area of Attraction (NAA) sought to encapsulate the chaos component by quantifying the range of excursions, adjusted to an equal distribution and concentration. The area of attraction was normalised so a 1 unit change in glucose was equivalent to a

1 unit change in insulin. In the 'basal' state the glucose concentrations decreased for the duration of the study. This decrease in glucose concentration had an effect on the area of the 95%CI for glucose (as N increased so did the area). It was critical to compare NAA's that are generated using the same number of data points (relative NAA's). The relative NAA's allowed a direct comparison of NAA's between subjects without bias. Using this model of chaos, it was demonstrated that the degradation in chaos boundary characteristics had changed by a factor of three. Examining individuals' profiles and their NAA may be a tool to establish homeostatic integrity and itself may have prognostic implications.

The chaos analysis method of evaluating glucose metabolism in terms of glucose homeostasis and not in term of cut-off values being determined by a national standard, such as NICE guidelines [104, 105] , although theoretical, may enable a more realistic approach to the effective treatment and prevention of diabetes and its complications. The mismatch of treatment and objective conditions, which becomes clinically evident in the form of hypoglycaemia, is well known and represents the major obstacle in every effort of treatment intensification. It is known from physiology that maintaining euglycaemia is a dynamic process which involves interaction of mechanisms with an ever-changing "milieu" (environment), inner and outer. The introduction of chaos boundary analysis may help, allowing individualisation of treatment regimens. In that context, it may be advisable to seek therapeutic interventions that restore mechanisms which can ascertain the maintenance of euglycaemia and not simplistically try to counterbalance hypoglycaemia and hyperglycaemia. Interventions that utilise the information obtainable by chaos analysis may in the long term favourably affect the other parameters of the disease, and hence protect b-cell secretory capacity, protect from formation of glycated end-products, and prevent cardiovascular disease. These

parameters are known to be linked not only to overall glycaemic control, but to glycaemic excursions as well [181] .The creation of the linear method GRADE (further information in Chapter 3) sought to assess this using a single variable glucose. The possible increase in cardiovascular morbidity has been attributed to a variety of reasons including the formation of advanced glycation end products. It has therefore been advocated that long term and transient hypoglycaemia should be treated as aggressively as possible.

The use of chaos methods on non-linear feedback systems in vivo as an information system may be useful. However, the use of chaos analysis in diabetes research is still in its infancy (<50 papers, PUBMED Nov 2007). The method described is not feasible for routine use of assessing of treatment efficacy, because of its relatively intensive investigation and the need for multiple assays. However its use in an experimental setting for comparison of therapeutic interventions is something that may well be justified and may illuminate the pathological processes.

6.0 Discussion and Conclusion

Research is the process of going up alleys to see if they are blind. - Marston Bates

6.1 Discussion

The tools and techniques described and constructed during the analysis of non-steady state physiological and pathological processes have been developed to analyse the data rich multi-variable megabyte datasets used in clinical practice and research. The datasets used were varied but the underlying methodology was to use information systems (systems modelling) to extract objective independent information. The theory behind the approach is comparable to meteorology. Billions of bits of data on weather are assessed by computer algorithms to determine wind speed, direction, atmospheric pressure and so on to produce one simple image that can be understood by everyone: sunshine or rain.

The aim was to develop and implement methods that are able to manipulate megabyte data sets from subjects with metabolic disorders that were beyond the assessment of traditional methods, in order to produce clinically relevant, objective, user-independent information and outcomes. In the production of the thesis two novel software programs (Easy TSA, ARID) and two novel methods GRADE and Chaos NAA have been developed, validated and clinically applied to achieve this aim.

6.1.1 Data vs. Information

The challenge one faces in analysing megabyte datasets is the confusion between the number of variables, the number of data points and the repeated measure of data points. The three components of data (initially introduced in Chapter 1) are discussed separately to delineate the differences between them. The correct understanding of the components has been critical to the correct application of the information systems.

6.1.1.1 Number of variables

The datasets analysed ranged from a single variable (used in GRADE) to three variable (four variables when examined longitudinally) retinal images examined by ARID. Multi-variable datasets are common in clinical practice and research. For example, hormonal substrates may be measured at defined periods and are therefore two variable datasets (substrate and time). The use of multi-variable Information systems to analyse the multi-variable data is not common. The lack of use could be due to a number of factors but may be because information systems can be mathematically complicated and therefore difficult to understand. Time series Analysis has been used for many years in physics and astronomy but the programs developed for use under these conditions have needed datasets that were often continuous and tending to infinity in length. The techniques traditionally used in clinical research and practice examine a single variable and hence may 'miss' information that could be critical in the analysis. One could consider an equivalent example as someone being colour blind and examining a painting. They may comprehend what the painting was about and be able to answer questions on aspects which the colour blindness had no influence (i.e. how large is the picture? Who painted it?). However, they would 'miss' information about the interplay of colour and contrast which may be critical to the understanding of the painting. The information systems used to analyse the datasets in the thesis were developed to examine all the variables in context with the information that was being sought.

6.1.1.2 Number of data points

The availability of cheap and large data storage devices for computers has meant that more data can be recorded and stored than in any other point in human history. The automation in measuring and capture devices (e.g. CGMS, retinal cameras) and advances in technology has meant that there has been an increased uptake of the cheap and large data storage. Data sets from photographic images in particular contain millions of data points that require analysis preferably using an objective information system. The increase in data is often seen as an end in itself. A recent example is that the National Service Framework (NSF) made a recommendation that everyone with diabetes should have an image taken of their eyes. This proposal equates to approximately 4,000,000 images to be recorded annually. An increase in data analysis will be needed to obtain the information from these images. An increase in data does not necessarily equate to an increase in information, a fact often overlooked.

6.1.1.3 Repeated measurements

It is preferable to have repeat measures of data points to reduce the effect of noise (by an averaging effect) and reduce the influence of outliers on the data. Indeed assaying is sometimes undertaken in duplicate to achieve this. As the number of repeated measurements increases so does the cost. A balance must be reached between the effect repeated measures have on noise and accuracy and the financial constraints.

6.1.2 Historical context of data vs. information in vivo

The collection of data and variables is increasing – for reasons previously covered. Historically the data that could be collected was limited by the science of the day but relevant information was still obtained. For example in 1616 William Harvey is credited

with describing how the circulatory system works. The data was collected from one person although various animals' circulatory systems were observed primarily. As time has progressed the data and variables collected have increased and systems specifically designed to extract the information from the data have evolved. In the 20th centaury statistical methods were developed to enable the effective analysis of increasing large data sets. Recent clinical trials have included up to 1 million participants[182] each with a plethora of variables and data points. Statistical methods still have a role in the analysis of megabyte datasets but more complex and more subtle methods are needed as both data, variables and subject numbers increase.

6.2 Information from data

The use of systems modelling to examine data sets and obtain information is not a new phenomenon. Information systems such as the computation of compound interest appeared in ancient Mesopotamia. The data sets commonly used in the 21st century are multi-variable data-rich megabyte datasets that have particular challenges associated with these attributes (further information in Chapter 1).

The advantages of specific systems models (information systems) have been discussed in each chapter of the thesis. The aim throughout has been to obtain information that can translate into a direct clinical or research outcome. The Information systems used enabled relevant objective information to be extracted from data sets. The different information systems approaches shared a focus in addressing this primary aim. The information systems used have sought to overcome the challenges in the analysis of data from biological datasets and in the particular areas of glycaemic variability using GRADE, hormonal oscillations using Easy TSA, *in vivo* feedback mechanisms using Chaos analysis and retinopathy using ARID.

6.2.1 GRADE

The premise was to develop an objective, independent measure that was categorical (risk 0 to 50). The information system the Glycaemic Risk Assessment in Diabetes Equation (GRADE) took a single variable (blood glucose) and sought to abstract a single value of risk (of developing microvascular and macrovascular complications). The modelling aspect was not performed on glycaemic values but was of the fifty experts in diabetes assessments of risk presented by different glycaemic values. The single value output from GRADE could be deconvoluted into risk from hypoglycaemia and hyperglycaemia. The information system was used to standardise the assessment of glycaemia and to provide a method by which it was possible to use the totality of the data available and not just the extreme excursions.

6.2.2 Easy TSA

The Time series Analysis (TSA) program Easy TSA is a collection of a number of information systems that were collated to be used in the 2-variable datasets collected when examining metabolic systems signalling. The information systems sought to extract information from the data that could not be recognised using traditional analytical single variable processes (e.g. mean, SD). The commonality of the modelling was to detect and evaluate the 'hidden' information in the data. The information systems examined both variables (the analyte and time) and their inter-relationship. The technique extracted information that was not detectable when analysing both variables individually.

6.2.3 Chaos Analysis

Non-linear dynamic systems analysis, of which chaos modelling is a component is a paradigm shift in the use of information systems. The paradigm shift is that data that appears to be random may contain information and it is just that traditional techniques are unable to evaluate it. There are numerous methods available to examine data to determine if it is chaotic, fractal or periodic and to quantify to what degree that is true. The information systems used in this thesis attempt to model aspects of the data that were not considered to contain relevant or 'useful' information (further information in Chapter 5). A specific method used was developed to analyse the feedback controlled glucose regulatory system and sought to quantify the system in terms of how 'stable' it was. The concept was to categorise a phenomenon that was previously un-categorical in terms of traditional analysis.

6.2.4 ARID

The assessment of retinopathy in people with diabetes was addressed using the Automated Retinal Image Differing (ARID) program. The information system provided an aid to the user to answer the Boolean question: Is there a change in retinopathy? Yes or No? The modelling was of the longitudinal retinal images that were multi-variable and data-rich. The technique was different from other systems in that it did not provide a definitive objective output but sought to assist the user in achieving that objective. The reasoning behind the decision was that retinal images are too noisy, data-rich, different and complicated to be assessed by computer alone. To make an incorrect assessment may lead to blindness. The information system developed was unique in this aspect. The final resulting outcome was the user's subjective interpretation and not an

information systems objective interpretation. Ideally it would be preferable for the result to be objective, repeatable, and independent but currently the computing processes and machine vision algorithms developed are insufficiently effective to be relied upon entirely.

6.3 Conclusions

The thesis advocates the use of information systems and systems modelling when analysing megabyte datasets. Traditionally the use of information systems has been small. The minimal model was one of the first information systems to be used to aid understanding in metabolic control and has a citation count of 384 (from scopus.com December 2007). HOMA is the only model to be widely used with a citation count of 4686 (from scopus.com December 2007) but is the exception to the trend. There are numerous factors as to why information systems models are not more commonly used. They can be complicated, difficult to understand, and mathematically complex and are therefore not understood and/or not used by clinicians and researchers. The methods and techniques used throughout the thesis have been described in detail and can be used by anyone with a PC. A complete understanding of 'how' the modelling works is not needed in order to 'make' it work. It is hoped that the simplicity of use and the benefits engendered allow the assessment of non-steady state physiological and pathological processes to become a critical aspect of understanding of the metabolic system and its disorders.

The production of models, information systems and software solutions has involved a steep learned curve. Throughout the production of the software, the running of the trials , the writing of the papers and the publishing of the thesis I have learnt an immeasurable

amount about scientific method and the nuances involved in carrying out research within the NHS and Oxford university. I have gained an understanding of diabetes ,endocrinology their physiologies and pathologies.

6.4 Further work

The information systems developed for this thesis are not abstract concepts that are mathematically interesting but of no real use. One hopes that the uptake of systems modelling will increase as the benefits inherent in the techniques become more widely known. The Easy TSA and ARID programs will be released as free software and publications citing their use will be produced. The production of the software and methods within the thesis are a beginning and not the end of the journey. There were many problems that were not overcome in the creation of the information systems as it would be impossible to build the perfect model. However they have been designed and built to solve specific problems/questions and have performed successfully in clinical trials and investigations. It is hoped that the use of information systems to model whole body physiology may be possible in the future.

7.0 References

1. Bacon, S.F., *Religious Meditations, Of Heresies.* 1597.
2. Pickup, J.C.W.G., ed. *Textbook of Diabetes*. Third ed. 2003, Blackwell Publishing: Oxford.
3. Watkins, P., *ABC of Diabetes*. 5th ed. 2002: Blackwell BMJ Books.
4. Home, P., *Contributions of basal and post-prandial hyperglycaemia to micro- and macrovascular complications in people with type 2 diabetes.* Curr Med Res Opin, 2005. **21**(7): p. 989-98.
5. IDF, I.D.F., *Diabetes Atlas*. 2003.
6. Himsworth, H.P., *Diabetes mellitus: its differentiation into insulin-sensitive and insulin-insensitive types.* Lancet, 1936. **i**: p. 127-130.
7. Kurrer, M.O., et al., *Beta cell apoptosis in T cell-mediated autoimmune diabetes.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 213-8.
8. EURODIAB ACE Study Group, *Variation and trends in incidence of childhood diabetes in Europe. .* Lancet, 2000. **355**(9207): p. 873-6.
9. American Diabetes Association, *Type 2 diabetes in children and adolescents. .* Diabetes Care, 2000. **23**(3): p. 381-9.
10. Gruber, A., et al., *Diabetes prevention: is there more to it than lifestyle changes?* Int J Clin Pract, 2006. **60**(5): p. 590-4.
11. Heaton, P.C. and S.M. Frede, *Patients' need for more counseling on diet, exercise, and smoking cessation: results from the National Ambulatory Medical Care Survey.* J Am Pharm Assoc (Wash DC), 2006. **46**(3): p. 364-9.
12. Reaven, P., *Metabolic syndrome.* J Insur Med, 2004. **36**(2): p. 132-42.
13. WHO, *Diagnosis and classification of diabetes mellitus.* World Health Organisation, Geneva, 1999.
14. OED, *Oxford English Dictionary*. 2006, Oxford University press.
15. Chatfield, C., *The analysis of time series (6th edition)*. 6th edition ed. 2004, London: Chapman and Hall.
16. Goldberger , A.L. and B.J. West, *Chaos in physiology: health or disease*. 1987: Plenum Press.
17. Fehm, H.L. and K.H. Voigt, *Pathophysiology of Cushing's disease.* Pathobiol Annu, 1979. **9**: p. 225-55.
18. Schmitz, O., et al., *High-frequency insulin pulsatility and type 2 diabetes: from physiology and pathophysiology to clinical pharmacology.* Diabetes Metab, 2002. **28**(6 Suppl): p. 4S14-20.
19. Recabarren, S.E., et al., *Pulsatile leptin secretion is independent of luteinizing hormone secretion in prepubertal sheep.* Endocrine, 2002. **17**(3): p. 175-84.
20. Juhl, C.B., et al., *High-frequency oscillations in circulating amylin concentrations in healthy humans.* Am J Physiol Endocrinol Metab, 2000. **278**(3): p. E484-90.
21. Gumbiner, B., et al., *Abnormalities of insulin pulsatility and glucose oscillations during meals in obese noninsulin-dependent diabetic patients: effects of weight reduction.* J Clin Endocrinol Metab, 1996. **81**(6): p. 2061-8.
22. Lorenz, E.N., *Deterministic Nonperiodic Flow.* Journal of the Atmospheric Sciences, 1963. **20**(2): p. 130-141.
23. Pincas, S. and A. Goldberger, *Physiological time-series analysis. What does regularity quantify?* Am J Physiol, 1994. **266**: p. H1643-1656.
24. Mandelbrot, B., *The fractal geometry of nature.* 1983, New York: Freeman.

25. Glenny, R. and H. Robertson, *Fractal properties of pulmonary blood flow heterogeneity.* J Appl Physiol, 1991. **70**: p. 1024-1030.
26. Yeragani, V., et al., *Fractal dimension of heart rate time series in an effective measure of autonomic function.* J Appl Physiol, 1993. **75**: p. 2429-2438.
27. Katz, M. and E. George, *Fractals and the analysis of growth paths.* Bull Math Biol, 1985. **47**: p. 273-286.
28. WHO, *Diagnosis and classification of diabetes mellitus updated*, in *World Health Organisation, Geneva*. 2007.
29. Bergman, R.N., L.S. Phillips, and C. Cobelli, *Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose.* J-Clin-Invest, 1981. **68**(6): p. 1456-67.
30. Schelter, B., et al., *Testing for directed influences among neural signals using partial directed coherence.* J Neurosci Methods, 2006. **152**(1-2): p. 210-9.
31. Scargle, J.D., *Studies in Astronomical Time Series Analysis: IV. Modeling Chaotic and Random Processes with Linear Filters.* The Astrophysical Journal, 1990. **359**: p. 469-482.
32. Ariznavarreta, C., et al., *Circadian rhythms in airline pilots submitted to long-haul transmeridian flights.* Aviat Space Environ Med, 2002. **73**(5): p. 445-55.
33. Reebs, S.G. and P. Doucet, *Relationship between circadian period and size of phase shifts in Syrian hamsters.* Physiol Behav, 1997. **61**(5): p. 661-6.
34. Jorgensen, J.O., et al., *Pulsatile versus continuous intravenous administration of growth hormone (GH) in GH-deficient patients: effects on circulating insulin-like growth factor-I and metabolic indices.* J Clin Endocrinol Metab, 1990. **70**(6): p. 1616-23.
35. Li, Y. and A. Goldbeter, *Pulsatile signaling in intercellular communication. Periodic stimuli are more efficient than random or chaotic signals in a model based on receptor desensitization.* Biophys J, 1992. **61**(1): p. 161-71.
36. Matthews, D.R., *Oscillatory insulin secretion: a variable phenotypic marker.* Diabetic Medicine, 1996. **13**((Supplement)): p. S53-S58.
37. Vanhorebeek, I. and G. Van den Berghe, *Hormonal and metabolic strategies to attenuate catabolism in critically ill patients.* Curr Opin Pharmacol, 2004. **4**(6): p. 621-8.
38. Goodner, C.J., et al., *Insulin, glucagon, and glucose exhibit synchronous, sustained oscillations in fasting monkeys.* Science, 1977. **195**(4274): p. 177-9.
39. Lang, D.A., et al., *Cyclic oscillations of basal plasma glucose and insulin concentrations in human beings.* New Eng J Med, 1979. **301**: p. 1023-1027.
40. Matthews, D.R., et al., *Greater in vivo than in vitro pulsatility of insulin secretion with synchronized insulin and somatostatin secretory pulses.* Endocrinology, 1987. **120**(6): p. 2272-8.
41. Lang, D.A., et al., *Pulsatile, synchronous basal insulin and glucagon secretion in man.* Diabetes, 1982. **31**(1): p. 22-6.
42. Stagner, J.I., E. Samols, and G.C. Weir, *Sustained oscillations of insulin, glucagon, and somatostatin from the isolated canine pancreas during exposure to a constant glucose concentration.* J-Clin-Invest, 1980. **65**(4): p. 939-42.
43. Honda , Y., K. Takahashi, and S. Takahashi *GH secretion during sleep in normal subjects.* J Clin Endocrinol Metab, 1969. **29**: p. 20-29.
44. Parker , D., et al., *Rhythmicities in Human GH Concentrations in Plasma*. Endocrine Rhythms, ed. K. DT. 1979, New York: Raven Press.
45. Krieger, D., *Rhythms in CRF, ACTH, and Corticosteroids*. Endocrine Rhythms, ed. D. Krieger. 1979, New York: Raven Press.

46. Clayton, R.N. and K.J. Catt, *Gonadotropin-releasing hormone receptors: characterization, physiological regulation, and relationship to reproductive function.* Endocr-Rev, 1981. **2**(2): p. 186-209.
47. Ead, H., J. Green , and E. Neil, *Comparison of the effects of pulsatile and non pulsatile blood flow through the carotid sinus on the reflexogenic activity of the sinun baroreceptors in the cat.* J Physiol, 1952. **118**: p. 509-519.
48. Belchetz, P.E., et al., *Hypophysial responses to continuous and intermittent delivery of hypopthalamic gonadotropin-releasing hormone.* Science, 1978. **202**(4368): p. 631-3.
49. Crowley, W., et al., *The physiology of gonadotropin releasing hormone (GnRH) secretion in men and women.* Recent Prog Horm Res, 1985. **41**: p. 473-531.
50. Marshall, J. and R. Kelch, *Gonadotrophin-releasing hormone:role of pulsatile secretion in the regulation of reproduction.* N Engl J Med, 1986. **315**: p. 1459-1468.
51. Matthews, D.R., et al., *Physiology of insulin secretion: problems of quantity and timing.* Neth J Med, 1985. **28 Suppl 1**: p. 20-4.
52. Jansson, J., et al., *Circumstantial evidence for the role of the secretory pattern of GH in control of body growth.* Acta Endocrinol, 1982. **99**: p. 24-30.
53. Clark, R. and I. Robinson, *GH responses to multiple injections of a fragment of human GH-releasing factor in conscious male and female rats.* J Endocr, 1985. **106**: p. 281-289.
54. Jeffrey, S., et al., *The episodic secretory pattern of GH regulates liver carbonic anhydrase III: studies in normal and mutant GH-deficient dwarf rats.* Biochem J, 1990. **266**: p. 69-74.
55. Box, G. and F.M. Jenkins, *Time Series Analysis: Forecasting and Control*. Second ed. 1976, Oakland, CA: Holden-Day.
56. Nelson, C.R., *The prediction performanceof the FRB -MIT-Penn model os the US economy.* Journal of the Royal Statistical Society, 1974. **Series A**(137): p. 131-165.
57. Chatfield, C., *The analysis of time series (3rd edition)*. 3rd edition ed. 1984, London: Chapman and Hall.
58. Granger, C., *The typical shape of an econometric variable.* Econometrica, 1966. **34**: p. 150-161.
59. Neter, J., et al., *Applied Linear Regression Models*. 1996: McGraw-Hill.
60. Dowdy, S. and S. Wearden, *Statistics for Research*. 2 ed. Vol. 1. 1985: Wiley-Interscience. 623.
61. Kleinbaum, D.G., et al., *Applied Regression Analysis and Multivariable Methods*. Vol. 3rd Edition. 1997: Duxbury Press.
62. Polyanin, A.D., *Handbook of Integral Equations*. Vol. 1st edition. 1998: CRC. 816.
63. Diggle, P., *A biostatistical introduction*. Time series. 1990, Oxford: Oxford University Press.
64. Duda , R.O. and P.E. Hart, *Pattern Classification and Scene Analysis*. Vol. 2nd Edition. 2001: John Wiley and Sons.
65. Turner, R.C., et al., *Insulin deficiency and insulin resistance interaction in diabetes: estimation of their relative contribution by feedback analysis from basal plasma insulin and glucose concentrations.* Metabolism, 1979. **28**: p. 1086-96.
66. GEIGY, *GEIGY Scientific Tables.* 1982. **Two**.
67. Wand M D, J.M.C., *Kernal Smoothing.* 1995.
68. Matthews, D.R., *Time series analysis in endocrinology.* Acta Paediatr Scand Suppl, 1988. **347**: p. 55-62.

69. Swed, F.S. and C. Eisenhart, *Tables for testing randomness of grouping in a sequence of alternatives.* The Annals of Mathematical Statistics, 1943. **14**(1): p. 66-87.
70. Schefler, W.C., ed. *Statistics for the biological sciences.*, ed. H. JW. 1969, . Addison-Wesley Publishing Co.
71. Efron, B. and R.J. Tibshirani, *Introduction to the Bootstrap*. 1994: Chapman & Hall/CRC. 456.
72. Lin, L., et al., *A homozygous R262Q mutation in the gonadotropin-releasing hormone receptor presenting as constitutional delay of growth and puberty with subsequent borderline oligospermia.* J Clin Endocrinol Metab, 2006. **91**(12): p. 5117-21.
73. Hill, N.R., H. Flax, and D.R. Matthews, *Higher body mass index is associated with irregular and suppressed insulin pulsatility.* Diabetes Obes Metab, 2007. **9**(4): p. 603-4.
74. Faber, O.K., et al., *Decreased insulin removal contributes to hyperinsulinemia in obesity.* J-Clin-Endocrinol-Metab, 1981. **53**(3): p. 618-21.
75. Ferrannini, E., et al., *Insulin resistance and hypersecretion in obesity. European Group for the Study of Insulin Resistance (EGIR).* J Clin Invest, 1997. **100**(5): p. 1166-73.
76. Camastra, S., et al., *beta-cell function in morbidly obese subjects during free living: long-term effects of weight loss.* Diabetes, 2005. **54**(8): p. 2382-9.
77. Rabinowitz, D. and K.L. Zierler, *Forearm metabolism in obesity and its response to intra-arterial insulin. Characterization of insulin resistance and evidence for adaptive hyperinsulinism.* J Clin Invest, 1962. **41**: p. 2173-2181.
78. Karpe, F. and G.D. Tan, *Adipose tissue function in the insulin-resistance syndrome.* Biochemical Society Transactions, 2005. **33**(5): p. 1045-1048.
79. Volk, A., et al., *Insulin action and secretion in healthy, glucose tolerant first degree relatives of patients with type 2 diabetes mellitus. Influence of body weight.* Exp Clin Endocrinol Diabetes, 1999. **107**(2): p. 140-7.
80. Bodkin, N.L., et al., *Central obesity in rhesus monkeys: association with hyperinsulinemia, insulin resistance and hypertriglyceridemia?* Int J Obes Relat Metab Disord, 1993. **17**(1): p. 53-61.
81. Scheen, A.J., *From obesity to diabetes: why, when and who?* Acta Clin Belg, 2000. **55**(1): p. 9-15.
82. Duman, B.S., et al., *The interrelationship between insulin secretion and action in type 2 diabetes mellitus with different degrees of obesity: evidence supporting central obesity.* Diabetes Nutr Metab, 2003. **16**(4): p. 243-50.
83. Preeyasombat, C., et al., *Racial and etiopathologic dichotomies in insulin hypersecretion and resistance in obese children.* J Pediatr, 2005. **146**(4): p. 474-81.
84. Newburgh, L.H., et al., *A new interpretation of diabetes mellitus in obese middle-aged persons: Recovery through reduction in weight.* Transactions of the Association of American Physicians, 1938. **53**: p. 245-257.
85. Long, S.D., et al., *Weight loss in severely obese subjects prevents the progression of impaired glucose tolerance to type II diabetes. A longitudinal interventional study.* Diabetes Care, 1994. **17**(5): p. 372-5.
86. Holte, J., et al., *Restored insulin sensitivity but persistently increased early insulin secretion after weight loss in obese women with polycystic ovary syndrome.* J Clin Endocrinol Metab, 1995. **80**(9): p. 2586-93.

87. Mori, Y., S. Mamori, and N. Tajima, *Weight loss-associated changes in acute effects of nateglinide on insulin secretion after glucose loading: results of glucose loading on 2 consecutive days.* Diabetes Obes Metab, 2005. **7**(2): p. 182-8.
88. Jimenez, J., et al., *Effects of weight loss in massive obesity on insulin and C-peptide dynamics: sequential changes in insulin production, clearance, and sensitivity.* J-Clin-Endocrinol-Metab, 1987. **64**(4): p. 661-8.
89. Colman, E., et al., *Weight loss reduces abdominal fat and improves insulin action in middle-aged and older men with impaired glucose tolerance.* Metabolism, 1995. **44**(11): p. 1502-8.
90. Matthews, D.R., et al., *Control of pulsatile insulin secretion in man.* Diabetologia, 1983. **24**(4): p. 231-7.
91. Bratusch Marrain, P.R., M. Komjati, and W.K. Waldhausl, *Efficacy of pulsatile versus continuous insulin administration on hepatic glucose production and glucose utilization in type I diabetic humans.* Diabetes, 1986. **35**(8): p. 922-6.
92. Sindelar, D.K., et al., *Basal hepatic glucose production is regulated by the portal vein insulin concentration.* Diabetes, 1998. **47**(4): p. 523-9.
93. Paolisso, G., et al., *Insulin effects on glucose kinetics in non-insulin-dependent diabetic patients with secondary failure to hypoglycaemic agents: role of different modes and rates of delivery.* Eur J Med, 1992. **1**(5): p. 261-7.
94. Hansen, B.C., K.L. Jen, and R.A. Wolfe, *The effects of over-feeding and obesity on rapid oscillatory patterns of insulin secretion in rhesus monkeys.* Fed Proc, 1979. **38**: p. 878.
95. Hansen, B.C., et al., *Influence of nutritional state on periodicity in plasma insulin levels in monkeys.* Am-J-Physiol, 1982. **242**(3): p. R255-60.
96. Zarkovic, M., et al., *Effect of weight loss on the pulsatile insulin secretion.* J Clin Endocrinol Metab, 2000. **85**(10): p. 3673-7.
97. Meneilly, G.S., J.D. Veldhuis, and D. Elahi, *Disruption of the pulsatile and entropic modes of insulin release during an unvarying glucose stimulus in elderly individuals.* J Clin Endocrinol Metab, 1999. **84**(6): p. 1938-43.
98. Matthews, J.N., et al., *Analysis of serial measurements in medical research.* BMJ, 1990. **300**(6719): p. 230-5.
99. Abumrad, N.N., et al., *Use of a heated superficial hand vein as an alternative site for the measurement of amino acid concentrations and for the study of glucose and alanine kinetics in man.* Metabolism, 1981. **30**(9): p. 936-40.
100. Nauck, M.A., R.W. Blietz, and C. Qualmann, *Comparison of hyperinsulinaemic clamp experiments using venous, 'arterialized' venous or capillary euglycaemia.* Clin Physiol, 1996. **16**(6): p. 589-602.
101. Chatfield, C., *The analysis of time series: theory and practice*. 1975, London: Chapman and Hall.
102. Paolisso, G., et al., *Greater efficacy of pulsatile insulin in type I diabetics critically depends on plasma glucagon levels.* Diabetes, 1987. **36**(5): p. 566-70.
103. Peiris, A.N., et al., *Body fat distribution and peripheral insulin sensitivity in healthy men: role of insulin pulsatility.* J Clin Endocrinol Metab, 1992. **75**(1): p. 290-4.
104. NICE, *Management of Type 2 Diabetes. Managing blood glucose levels.*, in *NICE Guidelines*. 2002.
105. NICE, *Diagnosis and management of type 1 diabetes in children, young people and adults*, in *NICE Guidelines*. 2006. p. http://www.nice.org.uk/page.aspx?o=CG015NICEguideline, Accessed 26th Oct 2006.

106. Franciosi, M., et al., *The impact of blood glucose self-monitoring on metabolic control and quality of life in type 2 diabetic patients: an urgent need for better educational strategies.* Diabetes Care, 2001. **24**(11): p. 1870-7.
107. Cohen, M. and P.Z. Zimmet, *Home blood-glucose monitoring: a new approach to the management of diabetes mellitus.* Med J Aust, 1980. **2**(13): p. 713-6.
108. Kibriya, M.G., et al., *Home monitoring of blood glucose (HMBG) in Type-2 diabetes mellitus in a developing country.* Diabetes Res Clin Pract, 1999. **46**(3): p. 253-7.
109. Laurell, T., *A continuous glucose monitoring system based on microdialysis.* J Med Eng Technol, 1992. **16**(5): p. 187-93.
110. Schlichtkrull, J., O. Munck, and M. Jersild, *[M-Value, an Index for Blood Sugar Control in Diabetics].* Ugeskr Laeger, 1964. **126**: p. 815-20.
111. Wojcicki, J.M., *Mathematical descriptions of the glucose control in diabetes therapy. Analysis of the Schlichtkrull "M"-value.* Horm Metab Res, 1995. **27**(1): p. 1-5.
112. Mirouze, J., et al., *[Insulin Efficiency Coefficient. M Coefficient of Schlichtkrull Corrected and Simplified by the Continuous Blood Glucose Recording Technic.].* Diabete, 1963. **11**: p. 267-73.
113. Service, F.J., et al., *Mean amplitude of glycemic excursions, a measure of diabetic instability.* Diabetes, 1970. **19**(9): p. 644-55.
114. DCCT Research Group, *Epidemiology of severe hypoglycaemia in the Diabetes Control and Complications Trial.* American Journal of Medicine, 1991. **90**: p. 450-459.
115. UKPDS Group, *Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33).* The Lancet, 1998. **352**: p. 837-854.
116. Hauffa BP, S.H., *Receiver-operated characteristic curve analysis of two algorithms assessing human growth hormone pulsatile secretion (PULSAR, CLUSTER): Comparison of peak detection efficacy.* Horm Res, 1994. **41**: p. 169-176.
117. Ryan, E.A., et al., *Assessment of the severity of hypoglycemia and glycemic lability in type 1 diabetic subjects undergoing islet transplantation.* Diabetes, 2004. **53**(4): p. 955-62.
118. Alkalay, A.L., et al., *Population meta-analysis of low plasma glucose thresholds in full-term normal newborns.* Am J Perinatol, 2006. **23**(2): p. 115-9.
119. NHS Borders, *Diabetes Service NHS Borders - Hypoglycaemia - Definition, Causes and Action*. 2004.
120. Mol, A. and J. Law, *Embodied Action, Enacted Bodies. The Example of Hypoglycaemia*. 2004. p. http://www.lancs.ac.uk/fass/sociology/papers/mol-law-embodied-action.pdf, Accessed 26th Oct 2006.
121. Alberti, K.G. and P.Z. Zimmet, *Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation.* Diabet Med, 1998. **15**(7): p. 539-53.
122. Patel, V., et al., *Oxygen reactivity in diabetes mellitus: effect of hypertension and hyperglycaemia.* Clin Sci (Lond), 1994. **86**(6): p. 689-95.
123. Grimaldi, A., et al., *Unawareness of hypoglycemia by insulin-dependent diabetics.* Horm Metab Res, 1990. **22**(2): p. 90-5.
124. Fisher, S.J., et al., *Insulin signaling in the central nervous system is critical for the normal sympathoadrenal response to hypoglycemia.* Diabetes, 2005. **54**(5): p. 1447-51.

125. Albisser, A.M., et al., *How good is your glucose control?* Diabetes Technol Ther, 2005. **7**(6): p. 863-75.
126. Hill, N.R., et al., *A method for assessing quality of control from glucose profiles.* Diabet Med, 2007. **24**(7): p. 753-8.
127. Levy, J., et al., *Discrimination, adjusted correlation, and equivalence of imprecise tests: application to glucose tolerance.* Am J Physiol (United States), 1999. **2**(1): p. E365-75.
128. Kerssen, A., H.W. de Valk, and G.H. Visser, *Day-to-day glucose variability during pregnancy in women with Type 1 diabetes mellitus: glucose profiles measured with the Continuous Glucose Monitoring System.* Bjog, 2004. **111**(9): p. 919-24.
129. Mlcak, P., et al., *A continuous glucose monitoring system (CGMS) - a promising approach for improving metabolic control in persons with type 1 Diabetes mellitus treated by insulin pumps.* Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 2004. **148**(1): p. 33-8.
130. Jeha, G.S., et al., *Continuous Glucose Monitoring and the Reality of Metabolic Control in Preschool Children With Type 1 Diabetes.* Diabetes Care, 2004. **27**(12): p. 2881-2886.
131. Jovanovic, L., *Continuous glucose monitoring during pregnancy complicated by gestational diabetes mellitus.* Curr Diab Rep, 2001. **1**(1): p. 82-5.
132. Hill, N.R., et al., *A method for assessing quality of control from glucose profiles*. 2007, Diabetic Medicine.
133. DOH, D.o.H., *National service framework for diabetes: standards*. 2001, Department of Health.
134. Monnier, M., O. Vatter, and L. Hoesli, *Action of Colour Flicker Stimuli on the Electrical Responses of Retina and Optic Cortex in Man.* Doc Ophthalmol, 1964. **18**: p. 207-20.
135. Goatman, K.A., et al., *Automated measurement of microaneurysm turnover.* Invest Ophthalmol Vis Sci, 2003. **44**(12): p. 5335-41.
136. NICE, N.I.f.C.E., *Management of Type 2 Diabetes. Retinopathy - Screening and early management*, in *NICE Guidelines*. 2002. p. 11.
137. Smith, R.T., et al., *Automated detection of macular drusen using geometric background leveling and threshold selection.* Arch Ophthalmol, 2005. **123**(2): p. 200-6.
138. Morgan, J.E., et al., *Digital imaging of the optic nerve head: monoscopic and stereoscopic analysis.* Br J Ophthalmol, 2005. **89**(7): p. 879-84.
139. Larsen, M., et al., *Automated detection of fundus photographic red lesions in diabetic retinopathy.* Invest Ophthalmol Vis Sci, 2003. **44**(2): p. 761-6.
140. Ege, B.M., et al., *Screening for diabetic retinopathy using computer based image analysis and statistical classification.* Comput Methods Programs Biomed, 2000. **62**(3): p. 165-75.
141. Teng, T., M. Lefley, and D. Claremont, *Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy.* Med Biol Eng Comput, 2002. **40**(1): p. 2-13.
142. Frame, A.J., et al., *A comparison of computer based classification methods applied to the detection of microaneurysms in ophthalmic fluorescein angiograms.* Comput Biol Med, 1998. **28**(3): p. 225-38.
143. Sivakumar, R., et al., *Diabetic Retinopathy Analysis.* J Biomed Biotechnol, 2005. **2005**(1): p. 20-27.
144. Hipwell, J.H., et al., *Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool.* Diabet Med, 2000. **17**(8): p. 588-94.

145. Cree, M.J., et al., *A fully automated comparative microaneurysm digital detection system.* Eye, 1997. **11**((Pt 5)): p. 622-8.
146. Usher, D., et al., *Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening.* Diabet Med, 2004. **21**(1): p. 84-90.
147. Larsen, N., et al., *Automated detection of diabetic retinopathy in a fundus photographic screening population.* Invest Ophthalmol Vis Sci, 2003. **44**(2): p. 767-71.
148. Staal, J., et al., *Ridge-based vessel segmentation in color images of the retina.* IEEE Trans Med Imaging, 2004. **23**(4): p. 501-9.
149. Conrath, J., et al., *Evaluation of the effect of JPEG and JPEG2000 image compression on the detection of diabetic retinopathy.* Eye, 2006.
150. Newsom, R.S., et al., *Effect of digital image compression on screening for diabetic retinopathy.* Br J Ophthalmol, 2001. **85**(7): p. 799-802.
151. Van Heuven, W.A.J., A. Kassoff, and J.I. Krepostman, *Diabetic retinopathy study. Report Number 6. Design, methods, and baseline results. Report Number 7. A modification of the Airlie House classification of diabetic retinopathy.* Investigative Ophthalmology and Visual Science, 1981. **21**(1): p. 149-226.
152. Glasbey, C.A. and G.W. Horgan, *Image Analysis for the Biological Sciences*. 1995: John Wiley & Sons. 230.
153. Yourdon, E., *When Good Enough Software Is Best.* IEEE Software, 1995. **12**(3): p. 79-81.
154. Gonzalez, R.C. and R.E. Woods, *Digital Image Processing*. 1992: Addison-Wesley.
155. Lewis, J.P., *Fast normalized cross-correlation.* Vision Interface, 1995: p. 120-123.
156. PCCAG, P.a.C.C.A.G., *An Audit on the Care of Adult Diabetic Patients in Gloucestershire*. 1996, Primary Care Clinical Audit Group: Gloucester.
157. Negi, A. and S. Vernon, *An overview of the eye in diabetes.* Journal of the Royal Society of Medicine, 2003. **96**(6): p. 266-272.
158. ETDRSG, *Grading diabetic retinopathy from stereoscopic color fundus photographs--an extension of the modified Airlie House classification. ETDRS report number 10. Early Treatment Diabetic Retinopathy Study Research Group.* Ophthalmology, 1991. **98**(5 Suppl): p. 786-806.
159. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.
160. Shaw, R., *The Dripping Faucet as a Model Chaotic System*. 1984, Santa Cruz: Aerial Press. 113.
161. Nagasaka, S., et al., *Comparison of pioglitazone and metformin efficacy using homeostasis model assessment.* Diabet Med, 2004. **21**(2): p. 136-41.
162. Yamanouchi, T., et al., *Comparison of metabolic effects of pioglitazone, metformin, and glimepiride over 1 year in Japanese patients with newly diagnosed Type 2 diabetes.* Diabet Med, 2005. **22**(8): p. 980-5.
163. Sato, J., et al., *Comparison of the effects of three sulfonylureas on in vivo insulin action.* Arzneimittelforschung, 2001. **51**(6): p. 459-64.
164. Natali, A. and E. Ferrannini, *Effects of metformin and thiazolidinediones on suppression of hepatic glucose production and stimulation of glucose uptake in type 2 diabetes: a systematic review.* Diabetologia, 2006. **49**(3): p. 434-41.
165. Matthews, D.R., et al., *Long-term therapy with addition of pioglitazone to metformin compared with the addition of gliclazide to metformin in patients with*

type 2 diabetes: a randomized, comparative study. Diabetes Metab Res Rev, 2005. **21**(2): p. 167-74.
166. Belcher, G., et al., *Safety and tolerability of pioglitazone, metformin, and gliclazide in the treatment of type 2 diabetes.* Diabetes Res Clin Pract, 2005. **70**(1): p. 53-62.
167. Betteridge, D.J. and B. Verges, *Long-term effects on lipids and lipoproteins of pioglitazone versus gliclazide addition to metformin and pioglitazone versus metformin addition to sulphonylurea in the treatment of type 2 diabetes.* Diabetologia, 2005. **48**(12): p. 2477-81.
168. Patlak, M., *New weapons to combat an ancient disease: treating diabetes.* Faseb J, 2002. **16**(14): p. 1853.
169. Guevara, M.R., L. Glass, and A. Shrier, *Phase locking, period-doubling bifurcations, and irregular dynamics in periodically stimulated cardiac cells.* Science, 1981. **214**(4527): p. 1350-3.
170. Findley, L.J., G.A. Farkas, and D.F. Rochester, *Changes in heart rate during breathing interrupted by recurrent apneas in humans.* J Appl Physiol, 1985. **59**(2): p. 536-42.
171. van Woerden, E.E., et al., *Fetal heart rhythms during behavioural state 1F.* Eur J Obstet Gynecol Reprod Biol, 1988. **28**(1): p. 29-38.
172. Souza Neto, E.P., et al., *Assessment of cardiovascular autonomic control by the empirical mode decomposition.* Methods Inf Med, 2004. **43**(1): p. 60-5.
173. Pincas, S., T. Cummins, and G. Haddad, *Heart rate control in normal and aborted - SIDS infants.* Am J Physiol, 1993. **264**: p. R638-646.
174. Wallace, T.M. and D.R. Matthews, *The assessment of insulin resistance in man.* Diabet Med, 2002. **19**(7): p. 527-34.
175. Bernard, C., *Leçons de physiologie expérimentale appliquée à la médecine faites au Collège de France.* Baillère et Fils, 1855: p. 296-313.
176. Nielsen, K., et al., *Sustained oscillations in glycolysis: an experimental and theoretical study of chaotic and complex periodic behavior and of quenching of simple oscillations.* Biophys Chem, 1998. **72**(1-2): p. 49-62.
177. Katayama, T., T. Sato, and K. Minato, *A blood glucose prediction system by chaos approach.* Conf Proc IEEE Eng Med Biol Soc, 2004. **1**: p. 750-3.
178. Sturis, J., et al., *Phase-locking regions in a forced model of slow insulin and glucose oscillations.* Chaos, 1995. **5**(1): p. 193-199.
179. Matthews, D.R., et al., *Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man.* Diabetologia, 1985. **28**(7): p. 412-9.
180. Sprott, J., *Chaos and Time-Series Analysis*. 1 ed. 2006, Oxford: Oxford University Press. 507.
181. Decode Study Group on behalf of the European Diabetes Epidemiology Group, *Glucose Tolerance and Cardiovascular Mortality: Comparison of Fasting and 2-Hour Diagnostic Criteria.* Arch Intern Med, 2001. **161**(3): p. 397-405.
182. *The Million Women Study: design and characteristics of the study population. The Million Women Study Collaborative Group.* Breast Cancer Res, 1999. **1**(1): p. 73-80.

8.0 Publications

Abstracts and poster presentations concerned with the methods developed in this thesis

Year	First Author	Title	Meeting	Location
2006	Hill, N.R.	Automated identification of change in retinal images	EASD	Copenhagen
2006	Hill, N.R.	Automated Retinal Screening	IDF	South Africa
2005	Hill, N.R.	Glycaemic Risk Assessment	IDF	South Africa
2005	Hill, N.R.	Automated identification of change in retinal images	DUK	B'Ham
2005	Hill, N.R.	Automated Retinal Image Differencing	DTT	San Francisco
2005	Hill, N.R.	An integrated index of Glycaemic Control obtained from glucose readings in Clinical Practice	ESPE LWPES	Lyon
2005	Hill, N.R.	A practical method of assessing glucose profiles	EASD	Athens
2004	Hill, N.R.	Retinopathy assessment	DTT	Philadelphia
2004	Hill, N.R.	Automated Retinal Image Differencing	EASD	Munich

Papers concerned with the methods developed in this thesis

2008 M M El-Kasti, H C Christian, I Huerta-Ocampo, M Stolbrink, S Gill, P A Houston, J S Davies, JChilcott, **N R Hill,** D R Matthews, D A Carter & T Wells , The pregnancy-induced increase in baseline circulating growth hormone in rats is not induced by ghrelin, *Journal of Neuroendocrinology in press*

2007 Hill, N.R., Matthews, D.R. Pioglitazone and metformin *Drugs of Today* 43 (7), pp. 443-454

2007 Hill, N.R., Hindmarsh, P.C., Stevens, R.J., Stratton, I.M., Levy, J.C., Matthews, D.R. A method for assessing quality of control from glucose profiles *Diabetic Medicine* 24 (7), pp. 753-758

2007 Hill, N.R., Flax, H., Matthews, D.R. Higher body mass index is associated with irregular and suppressed insulin pulsatility *Diabetes, Obesity and Metabolism* 9 (4), pp. 603-604

2006 Lin, L., Conway, G.S., **Hill, N.R**., Dattani, M.T., Hindmarsh, P.C., Achermann, J.C. A homozygous R262Q mutation in the gonadotropin-releasing hormone receptor presenting as constitutional delay of growth and puberty with subsequent borderline oligospermia *Journal of Clinical Endocrinology and Metabolism* 91 (12), pp. 5117-5121