Continuous Glucose Monitoring and Tight Glycaemic Control in Critically Ill Patients

Matthew Kent Signal

A thesis submitted for the degree of
Doctor of Philosophy
in
Bioengineering
at the
University of Canterbury,
Christchurch, New Zealand
21 July 2013

Acknowledgements

I would like to thank a number of people who have made my time working on this research during the last few years a rewarding experience.

To my supervisors, Prof. Geoff Chase and Dr. Geoff Shaw, thank you for all of your support, guidance and wisdom over the past 3 years. I have learnt a lot from you both and I appreciate the time you have given me.

To Dr. Chris Pretty, Dr. Aaron Le Compte and all of my colleagues in the Centre for Bioengineering, thank you for your guidance, patience and friendship.

Finally, to my parents and Sarah, thank you for your unconditional love, support and encouragement throughout the duration of my studies. I could not have completed this work without the support you have provided me outside of university.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms and abbreviations

ADA – American Diabetes Association

APE – Absolute Percent Error

BG – Blood Glucose

BGA – Blood Gas Analyser

CDF - Cumulative Distribution Function

CG-EGA – Continuous Glucose Error Grid Analysis

CGM – Continuous Glucose Monitoring

CGMS – Continuous Glucose Monitoring System

CHYLD – Children with HYpoglycaemia and their Later Development

DFA – Detrended Fluctuation Analysis

DM – Diabetes Mellitus

FDA – Food and Drug Administration

FIR – Finite Impulse Response

ICU – Intensive Care Unit

IDM – Infant of a Diabetic Mother

IIR – Infinite Impulse Response

IIT – Intensive Insulin Therapy

IQR – Interquartile Range

LGA – Large for Gestational Age

LMS – Least Mean Square

MAD – Mean Absolute Difference

MAPE – Mean Absolute Percent Error

MARD – Mean Absolute Relative Difference

MC – Monte Carlo

NICU – Neonatal Intensive Care Unit

PC – Personal Computer

POC – Point of Care

PDF – Probability Density Function

RMS – Root Mean Square

SG – Sensor Glucose

SGA – Small for Gestational Age

SPRINT – SPecialised Relative Insulin and Nutrition Tables

STAR – Stochastic TARgeting

TGC – Tight Glycaemic Control

VLBW – Very Low Birth Weight

YSI - Yellow Springs Instruments

Abstract

Critically ill patients often exhibit abnormal glycaemia that can lead to severe complications and potentially death. In critically ill adults, hyperglycaemia is a common problem that has been associated with increased morbidity and mortality. In contrast, critically ill infants often suffer from hypoglycaemia, which may cause seizures and permanent brain injury. Further complicating the matter, both of these conditions are diagnosed by blood glucose (BG) measurements, often taken several hours apart, and, as a result, these conditions can remain poorly managed or go completely undetected. Emerging 'continuous' glucose monitoring (CGM) devices with 1-5 minute measurement intervals have the potential to resolve many issues associated with conventional intermittent BG monitoring. The objective of this research was to investigate and develop methods and models to optimise the clinical use of CGM devices in critically ill patients.

For critically ill adults, an in-silico study was conducted to quantify the potential benefits of introducing CGM devices into the intensive care unit (ICU). Mathematical models of CGM error characteristics were implemented with existing, clinically validated, models of the insulin-glucose regulatory system, to simulate the behaviour of CGM devices in critically ill patients. An alarm algorithm was also incorporated to provide a warning at the onset of predicted hypoglycaemia, allowing a virtual dextrose intervention to be administered as a preventative measure. The results of the in-silico study showed a potential reduction in nurse workload of approximately 75% and a significant reduction in hypoglycaemia, while also providing insight into the optimal rescue dose size and resulting dynamics of glucose recovery.

During 2012, ten patients were recruited into a pilot clinical trial of CGM devices in critical care with a primary goal of assessing the reliability of CGM devices in this environment, with a specific interest in the effects of CGM device type and sensor site on sensor glucose (SG) data. Results showed the mean absolute relative difference of SG data across the cohort was between 12-24% and CGM

devices were capable of monitoring some patients with a high degree of accuracy. However, certain illnesses, drugs and therapies can potentially affect sensor performance, and one particular set of results suggested severe oedema may have affected sensor performance. A novel and first of its kind metric, the Trend Compass was developed and used to assesses trend accuracy of SG in a mathematically precise fashion without approximation, and, importantly, does so independent of glucose level or sensor bias, unlike any other such metrics. In this analysis, the trend accuracy between CGM devices was typically good.

A recent hypothesis suggesting that glucose complexity is associated with mortality was also investigated using the clinical CGM data. The results showed that complexity results from detrended fluctuation analysis (DFA) were influenced far more by CGM device type than patient outcome. In addition, the location of CGM sensors had no significant effect on complexity results in this data set. Thus, while this emerging analytical method has shown positive results in the literature, this analysis indicates that those results may be misleading given the impact of technology outweighing that of physiology. This particular result helps to further delineate the range of potential applications and insight that CGM devices might offer in this clinical scenario.

In critically ill infants, CGM devices were used to investigate hypoglycaemia during the first 48 hours after birth. More than 50 CGM data sets were obtained from several studies of CGM in infants at risk of hypoglycaemia at the Waikato hospital neonatal ICU (NICU). In light of concerns regarding CGM accuracy, particularly during the first few hours of monitoring and/or at low BG levels, an alternative, novel calibration scheme was developed to increase the reliability of SG data. The recalibration algorithm maximised the value of very accurate calibration BG measurements from a blood gas analyser (BGA), by forcing SG data to pass through these calibration BG measurements.

Recalibration increased all metrics of hypoglycaemia (number, duration, severity and hypoglycaemic index) as the factory CGM calibration was found to be reporting higher values at low BG levels due to its least squares calibration

approach based on the assumption of a less accurate calibration glucose meter. Thus, this research defined new calibration methods to directly optimise the use of CGM devices in this clinical environment, where accurate reference BG measurements are available. Furthermore, this work showed that metrics such as duration or area under curve were far more robust to error than the typically used counted-incidence metrics, indicating how clinical assessment may have to change when using these devices.

The impact of errors in calibration measurements on metrics used to classify hypoglycaemia was also assessed. Across the cohort, measurement error, particularly measurement bias, had a larger effect on hypoglycaemia metrics than delays in entering calibration measurements. However, for patients with highly variable glycaemia, timing error can have a significantly larger impact on output SG data than measurement error. Unusual episodes of hypoglycaemia could be successfully identified using a stochastic model, based on kernel density estimation, providing another level of information to aid decision making when assessing hypoglycaemia.

Using the developed algorithms/tools, with CGM data from 161 infants, the incidence of hypoglycaemia was assessed and compared to results determined using BG measurements alone. Results from BG measurements showed that ~17% of BG measurements identified hypoglycaemia and over 80% of episodes occurred in the first day after birth. However, with concurrent BG and SG data available, the SG data consistently identified hypoglycaemia at a higher rate suggesting the BG measurements were not capturing some episodes. Duration of hypoglycaemia in SG data varied from 0-10+%, but was typically in the range 4-6%. Hypoglycaemia occurred most frequently on the first day after birth and an optimal measurement protocol for at risk infants would likely involve CGM for the first week after birth with frequent intermittent BG measurements for the first day.

Overall, CGM devices have the potential to increase the understanding of certain glycaemic abnormalities and aid in the diagnosis/treatment of other conditions

in critically ill patients. This research has used a range of prospective and retrospective clinical studies to develop methods to further optimise the use of CGM devices within the critically ill clinical environment, as well as delineating where they are less useful or less robust. These latter results clearly define areas where clinical practice needs to adapt when using these devices, as well as areas where device makers could target technological improvements for best effect. Although further investigations are required before these devices are regularly implemented in day-to-day clinical practice, as an observational tool they are capable of providing useful information that is not currently available with conventional intermittent BG monitoring.

Chapter 1. Introduction

Critically ill patients often exhibit abnormal glycaemia due to the severity of their illness. In critically ill adults, stress induced hyperglycaemia, or high blood glucose (BG), has been associated with increased morbidity and mortality (Capes et al. 2000; Krinsley 2003; McCowen et al. 2001; Mizock 1995; Umpierrez et al. 2012). Hyperglycaemia has also been linked to other negative clinical outcomes, including infection (Bistrian 2001), sepsis and septic shock (Branco et al. 2005; Das 2003; Marik & Raghavan 2004), myocardial infarction (Capes et al. 2000), and, polyneuropathy and multi-organ failure (Langouche et al. 2005; van den Berghe et al. 2001). In contrast, in critically ill infants, a common metabolic problem is neonatal hypoglycaemia, or low BG, which may cause seizures and permanent brain injury in affected babies (Stanley & Baker 1999).

Both of these common conditions are diagnosed by BG measurements, typically done by a nurse at the bedside and only taken every few hours at the most frequent. In the critical care environment, frequent BG measurements are often not feasible due to the increase in nurse workload (Aragon 2006; Carayon & Gurses 2005; Holzinger et al. 2005). Furthermore, for very small infants in the neonatal intensive care unit (NICU), blood draw constraints can limit the number of BG measurements permitted each day. Thus, with conventional, intermittent BG monitoring using handheld glucose meters or a local blood gas analyser (BGA), these conditions can remain poorly managed or go completely undetected, ultimately increasing the risk of negative outcomes for patients.

Interestingly, the technology that could potentially resolve the issue of infrequent BG measurements has been around for over 10 years, in the form of continuous glucose monitoring (CGM) devices. CGM devices were originally developed to help people with diabetes manage their condition. They typically offer a 1-5 minute measurement interval, giving information about BG level and BG trends. Several types of CGM devices have been proposed over the years, including subcutaneous glucose sensors (Girardin et al. 2009), micro-dialysis systems (Valgimigli et al. 2010) and a number of methods that offer non-invasive

BG measurements (Vashist 2012). Overall, the most common and widely used type is the subcutaneous sensing CGM devices, which measures glucose concentration just beneath the skin in order to estimate BG concentration. There are several manufacturers of the subcutaneous type CGM device who have approved devices in the market, in contrast to the other approaches.

Although the technology has been developed and is well tested in people with diabetes, there are only a limited number of studies that have investigated CGM devices in critically ill patients, despite their obvious potential benefits (Weiss & Lazar 2007). In the adult intensive care unit (ICU), CGM devices have the potential to increase patient safety through better monitoring, while reducing nurse workload through fewer manual BG measurements required per day. In the NICU, these devices have the potential to help diagnose episodes of hypoglycaemia that would have otherwise gone undetected, while increasing the limited knowledge on the duration and severity of such episodes. However, there are still concerns regarding the accuracy of CGM devices (Rabiee et al. 2009; Vlkova et al. 2009), especially in critically ill patients where illnesses, drugs/therapies and other clinical factors may impede performance (Moser et al. 2010). Hence, a thorough investigation of CGM devices in the intensive care environment, both ICU and NICU, could potentially improve the standard of care for these critically ill patient cohorts.

The research presented in this thesis is primarily aimed at investigating and optimising the use of commonly available off-the-shelf CGM devices in critically ill adults and infants. More specifically, several of the limitations identified by previous studies of these devices are presented and solutions are developed, proposed and validated. The results and findings presented here are largely derived from data collected during previous or ongoing clinical trials of CGM devices. Finally, the findings from this research can be used to improve the standard of clinical care in the Christchurch ICU and the Waikato Hospital NICU, the centres where this research was primarily located, as well as other ICUs in New Zealand by its impact and potential to change clinical practice.

1.1 Preface

This thesis is presented in two parts: *Part I* investigates the use of CGM in critically ill adults, in the ICU and, *Part II* investigates the use of CGM in critically ill infants in the NICU. The chapters in this document are arranged as follows:

Part I - Adult ICU

Chapter 2 provides a thorough background on hyperglycaemia in critically ill patients and an overview of tight glycaemic control, before the potential benefits of CGM in the ICU are discussed.

Chapter 3 contains a review of the literature of CGM error models, robust filtering methods and hypoglycaemia alarm algorithms, all of which can be implemented in-silico.

Chapter 4 presents an in-silico feasibility study to determine whether CGM devices can be couple with an existing glycaemic control protocol to improve patient safety and reduce nurse workload, while maintaining control of glycaemia. *Material in Chapter 4 has been published in: Signal, M, Pretty, CG, Chase, JG, Le Compte, AJ and Shaw, GM (2010). "Continuous Glucose Monitors and the Burden of Tight Glycemic Control in Critical Care: Can they cure the time cost?" Journal of Diabetes Science and Technology (JoDST), Vol 4(3), pp. 625-635, ISSN: 1932-2968.*

Chapter 5 discusses CGM data analysis and highlights the distinction between point-to-point accuracy and trend accuracy. A novel trend accuracy metric, the *Trend Compass*, is developed and validated.

Chapter 6 presents the results of an observational 10 patient pilot clinical trial at Christchurch Hospital's ICU assessing the reliability of off-the-shelf CGM devices in critically ill patients. *Material in Chapter 6 has been published in: Signal, M Fisk, L, Shaw, GM and CG, Chase (2013). "Concurrent continuous glucose monitoring in critically ill patients: Interim results and observations," Journal of Diabetes Science and Technology (JoDST), Vol 6(7), (invited).*

Chapter 7 tests a new hypothesis on glucose complexity using CGM data collected during the pilot clinical trial at Christchurch Hospital's ICU. *Material in Chapter 7 has been published in: Signal, M, Thomas, F, Shaw, GM and Chase, JG (2013). "Complexity of continuous glucose monitoring data in critically ill patients: CGM devices, sensor locations and DFA methods," Journal of Diabetes Science and Technology, in press (Oct. 2013).*

Chapter 8 summarises the research on CGM in critically ill adults, presented in Chapters 2-7.

Part II - Neonatal ICU

Chapter 9 provides an in-depth background on neonatal hypoglycaemia and its resulting complications, before the potential benefits of CGM in the NICU are discussed.

Chapter 10 investigates one of the key technological aspects of CGM devices, the calibration algorithms that convert raw sensor signals into useful glucose data. An alternative calibration scheme is developed and the impacts it has on metrics used to classify hypoglycaemia are assessed. *Material in Chapter 10 has been published in: Signal, MK, Le Compte, AJ, Harris, DL, Weston, PJ, Harding, JE and Chase, JG for the CHYLD Study Group (2012). "Impact of Calibration Algorithms on Hypoglycaemia Detection in Newborn Infants Using Continuous Glucose Monitoring," Diabetes Technology & Therapeutics (DTT), 14:10, ISSN: 1520-9156.*

Chapter 11 assesses the impact of timing delays and measurement errors in calibration BG measurements on output SG data. Specifically, how these sources of error affect metrics used to classify hypoglycaemia.

Chapter 12 develops a method of detecting and classifying unusual episodes of hypoglycaemia in SG data. The classifications are presented graphically by colour coding a time series plot of SG data. *Material in Chapter 12 has been published in: Signal, MK, Le Compte, AJ, Harris, DL, Weston, PJ, Harding, JE and Chase, JG for the CHYLD Study Group (2012). "Using Stochastic Modelling to Identify Unusual*

Continuous Glucose Monitor Measurements in Newborn Infants," BioMedical Engineering OnLine, 11:45, (open access), ISSN: 1475-925X.

Chapter 13 investigates hypoglycaemia in neonates that are at risk of developing the condition. Results determined using conventional intermittent BG monitoring methods are compared to results from analyses of CGM device data.

Chapter 14 summarises the research on CGM in critically ill infants, presented in Chapters 9-13.

Chapter 15 contains a discussion of future work, both in the adult ICU and the neonatal ICU.

PART I - CGM IN ADULT ICU

Chapter 2. Background - Adult ICU

This chapter provides background on hyperglycaemia in critically ill patients and its resulting complications. In particular, one proposed method of reducing hyperglycaemia is tight glycaemic control (TGC). An overview of the literature on this subject, with specific reference to several key studies, is presented and some of the risks associated with TGC are highlighted. Finally, the application of CGM devices to TGC is presented as a means of enhancing glycaemic control.

2.1 Hyperglycaemia in critically ill patients

Critically ill patients often experience stress induced hyperglycaemia, which has been associated with increased morbidity and mortality (Capes et al. 2000; Krinsley 2003; McCowen et al. 2001; Mizock 1995; Umpierrez et al. 2012). Hyperglycaemia has also been linked to other negative clinical outcomes, including infection (Bistrian 2001), sepsis and septic shock (Branco et al. 2005; Das 2003; Marik & Raghavan 2004), myocardial infarction (Capes et al. 2000), and, polyneuropathy and multi-organ failure (Langouche et al. 2005; van den Berghe et al. 2001). BG measurements are used to diagnose hyperglycaemia, but diagnostic criteria can vary between critical care centres. Typically, 1-2 BG measurements above a threshold of ~8-10mmol/L constitutes hyperglycaemia (Singer et al. 2009; Umpierrez et al. 2012).

Stress induced hyperglycaemia is prevalent in both diabetic and non-diabetic critically ill patients. The primary cause of hyperglycaemia is increased levels of counter-regulatory hormones, released in response to critical illness (McCowen et al. 2001). During the acute phase of critical illness, several hormones including glucagon, growth hormone, catecholamines and glucocorticoids are significantly elevated in response to physiological stress (Turina et al. 2005). These hormones increase glucose production by stimulating the metabolic pathways for gluconeogenesis, glycogenolysis and lipolysis. Gluconeogenesis is the formation of glucose from lactic acid and certain amino acids, glycogenolysis is the breakdown of glycogen into glucose, and, lipolysis is the breakdown of lipids into

its constituents, including glycerol, which is subsequently converted into glucose by the liver. Hence, these avenues represent forms of unregulated endogenous glycaemic production, raising BG levels. In addition, these hormones increase insulin resistance, which reduces insulin mediated glucose uptake and further increases BG concentration (Chase et al. 2011; Gearhart & Parbhoo 2006).

Exogenous factors, such as the administration of some medications and/or the use of some therapies, can interact with physiological processes, increasing the severity of hyperglycaemia. In particular, glucocorticoid steroids, the catecholamines epinephrine and norepinephrine, and β-blockers are reported to reduce insulin sensitivity and thus increase BG levels in critically ill patients (Pretty et al. 2011). Other sources of increased BG levels include excess calories from parenteral and enteral nutrition, as well as dextrose infusions used for fluid resuscitation and drug delivery (Mizock 1995; Nylen & Muller 2004).

As critical illness evolves, hyperglycaemia can remain persistent for days or even weeks, increasing the risk of negative outcomes. Although the exact mechanisms are not completely understood, it is thought that long term exposure to high BG levels induces inflammatory changes and oxidative stress, which harm cardiovascular and endothelial function (Ellahham 2010). Furthermore, elevated BG levels have a detrimental effect on immune function and may significantly increase a patient's susceptibility to infection (Van den Berghe et al. 2003). Thus, it was hypothesised that controlling BG levels to a euglycaemic range might reduce some of the negative outcomes associated with hyperglycaemia.

2.2 Tight glycaemic control in critically ill patients

In 2001, Van den Berghe et al. published the results of a prospective, randomised, controlled trial that used a TGC protocol to reduce hyperglycaemia in critically ill cardiac surgery patients (van den Berghe et al. 2001). Patients were randomly assigned to either the intensive insulin therapy (IIT) treatment group whose BG was maintained between 80-110mg/dL, or, the conventional treatment group whose BG was maintained between 180-200mg/dL. The

primary outcome was ICU mortality, and the secondary outcomes included hospital mortality as well as several measures of patient morbidity.

The trial was a success and the results showed a reduction in ICU mortality of 18-45%, depending on length of admission, for patients assigned to the IIT treatment group. Furthermore, significant reductions in hospital mortality and several measures of morbidity were also observed in the IIT group (van den Berghe et al. 2001; Van den Berghe et al. 2003). These measures of morbidity included a reduction in the duration of ventilator support, a reduction in renal impairment, a reduction in bloodstream infection, a reduction in polyneuropathy, a reduced need for blood transfusions and fewer therapeutic interventions. Overall, the researchers concluded that TGC reduces morbidity and mortality among critically ill patients in the surgical ICU.

During the following years, several studies attempted to reproduce the positive results of Van den Berghe. In particular, two studies achieved similar results with TGC reporting reductions in hospital mortality of 29-35% (Chase et al. 2008c; Krinsley 2004). However, several others failed to achieve this result (Brunkhorst et al. 2008; Finfer et al. 2009; Preiser et al. 2009). Another significant outcome of the study by Chase et al. was the reduction in severe hypoglycaemia with TGC, which has been independently associated with increased mortality in critically ill patients (Egi et al. 2010; Hermanides et al. 2010b). This latter result was in contrast to increased hypoglycaemia with TGC seen in almost all other studies (Griesdale et al. 2009).

Interestingly, successful TGC studies also report significant economic benefit, with savings averaging approximately $1500-3200USD per patient treated, at the time of publication (Krinsley & Jones 2006; Van den Berghe et al. 2006). These savings were primarily attributed to reduced ICU length of stay, fewer ventilator days, less imaging and fewer laboratory requirements (Krinsley & Jones 2006). Hence, there are economic motivations, as well.

As noted, in addition to successful TGC trials, there have been several studies that have failed to match the positive results (Brunkhorst et al. 2008; De La Rosa Gdel et al. 2008; Finfer & Delaney 2008; Griesdale et al. 2009; Meijering et al. 2006; Preiser et al. 2009; Treggiari et al. 2008). Most notably, one study was stopped early due to unintended protocol violations (Preiser et al. 2009), and another study was stopped for safety concerns due to the level of hypoglycaemia in the TGC treatment group (Brunkhorst et al. 2008). Following the publication of the NICE-SUGAR study, which reported higher mortality and hypoglycaemia with TGC (Finfer et al. 2009), a meta-analysis of several TGC studies showed that most TGC studies failed to achieve a result either way (Griesdale et al. 2009). However, due to large variations in the number of centres, cohorts and ICU types, comparisons between studies are very difficult. Thus, TGC remains a highly contentious subject with the benefits/risks still heavily debated.

When drawing conclusions about the benefits and risks of TGC it is important to ensure results are interpreted correctly. There is potential for negative results to be falsely negative due to unintended patient crossover or cohort overlap, which can lead to misinterpretation of results. For example, it should not be assumed that all patients in the TGC treatment group achieved TGC, while patients in the conventional treatment group did not have BG levels in the TGC target glycaemic band.

Thus, there are two separate questions that need to be answered. The first is a physiological question: Is normoglycaemia associated with better outcomes in critically ill patients, irrespective of how it comes about? The second is a clinical question: Is glucose control achievable, consistently and reliably, in a clinical setting?

Evidence in response to the first question suggests that normoglycaemia is associated with lower mortality, irrespective of how it is achieved (Signal et al. 2012b). The second question, regarding the difficulties involved with implementing a successful and consistent TGC protocol, potentially highlights one of the causes of inconsistent findings in TGC studies to date. Achieving TGC

on a per-patient basis is extremely difficult, largely due to the inter-patient variability in illness and response to insulin treatment (Chase et al. 2011; Lin et al. 2008; Suhaimi et al. 2010). Further complicating the matter, each patient's condition evolves over time and the TGC protocol must adjust in response to any intra-patient variability. It is likely that failing to manage variability will result in poor control and an inability to separate the cohorts of a randomised controlled trial, ultimately making it difficult to show the benefits of TGC (Chase et al. 2011).

Overall, the aim of any TGC protocol is to reduce elevated BG levels with minimal or zero hypoglycaemia. However, protocols or clinical practices that utilize large insulin doses can suffer from high glycaemic variability and/or excessive hypoglycaemia (Chase et al. 2011; Meijering et al. 2006), both of which have been independently associated with mortality in critically ill patients (Bagshaw et al. 2009; Egi et al. 2006; Egi et al. 2010; Hermanides et al. 2010a; Hermanides et al. 2010b; Krinsley 2008). Furthermore, the frequency of BG measurements required for TGC needs careful consideration, as a trade-off often exists between TGC performance/safety and nurse burden. Due to this trade-off, BG is typically only measured every 1–4 h and more frequently only if the levels are already at or near hypoglycaemia. More frequent measurements, and even 1–2 hourly measurements, are uncommon due to the clinical effort required (Carayon & Gurses 2005; Chase et al. 2008a).

2.3 Continuous glucose monitoring in critically ill patients

CGM devices, with their 1-5 minute measurement interval (Girardin et al. 2009), have the potential to eliminate the trade-off in TGC between performance/safety and nurse workload. Originally designed to help people with diabetes manage their condition, the most widely used CGM devices consist of a subcutaneous sensor which measures glucose in the interstitial fluid for up to 7 days, and a monitor which stores/displays data (Girardin et al. 2009). Figure 2.1 shows the elements of a commonly available and widely used CGM system.

In general, there are two families of CGM devices. First, there are retrospective devices, which store data in a blinded manner until the end of monitoring. Second, there are real-time devices, which display glucose values throughout the monitoring period. Both types of device require multiple (3-6) independent BG measurements per day for calibration, but, all else equal, can deliver different results due to their different calibration and processing of the sensor signal.

Figure 2.1 The Medtronic Guardian Real-Time CGM - a commonly available and widely used CGM device.

Although well studied in people with diabetes (Gandhi et al. 2011; Hoeks et al. 2011b), there have been relatively few investigations of CGM devices in critically ill patients, despite their obvious potential benefits with TGC. An early study of automated closed loop BG control using CGM in critically ill patients failed to show improvements over manual control, in part due to high glucose sensor glucose (SG) error (Chee et al. 2003). More recently, several studies reported CGM to be clinically useful in critical care units, particularly for early detection of hypoglycaemia (Holzinger et al. 2010; Westhoff et al. 2010). One larger study, analysing SG data from 174 patients, found CGM device accuracy to be "very good" and regarded the device to be safe to guide insulin therapy (Brunner et al. 2011).

However, despite a positive report, there are still concerns regarding the reliability of CGM devices in critical care units (Rabiee et al. 2009; Vlkova et al. 2009), as certain conditions/illnesses including sepsis, septic shock, and peripheral oedema, could potentially affect CGM device performance (Lorencio

et al. 2012). In addition, certain medications/therapies commonly used in the ICU, such as paracetamol, can influence CGM device performance (Moser et al. 2010). Interestingly, vasoactive medications, such as dopamine, norepinephrine, ketanserin, enoximone and nitro-glycerine, are reported to have little or no affect on the accuracy of one type of glucose sensor (Holzinger et al. 2009; Westhoff et al. 2010).

Overall, regardless of specific results, a common factor among all of these studies is the recognition of the potential benefits of CGM in critically ill patients, which include:

1. Improve BG control with a well designed TGC algorithm
2. Improve hypoglycaemia detection, a benefit that is already being realised.
3. Decrease nurse burden by reducing the number of manual BG measurements required to 3-4 per day for CGM device calibration.

However, further investigations are required before CGM devices can be implemented in critical care units to guide clinical insulin delivery to manage hyperglycaemia. The following chapters investigate using CGM devices in an adult ICU, with the overall goal of improving BG control and increasing patient safety, while reducing nurse workload.

Chapter 3. CGM error models, filtering and alarms

Prior to using CGM devices in the critical care unit, a search of the literature was undertaken to determine whether enough information could be gathered to allow an in-silico feasibility study to be completed. The elements required to complete the feasibility study, in addition to a clinically validated patient simulation package that was already available (Chase et al. 2010), were a CGM error model, an appropriate filtering algorithm and an alarming algorithm. This chapter describes those elements and the sources of the information.

3.1 CGM error models

Despite significant outpatient use and promise, the literature contains very few reports of error models derived from clinical SG data. Two studies in particular provide sufficient details of CGM device error characteristics to allow models to be created or reproduced for use in-silico (Breton & Kovatchev 2008; Goldberg et al. 2004). Although the models described in this section were developed by other researchers, the details are presented due to their extensive use throughout this thesis.

3.1.1 Gaussian based error model

The Gaussian based CGM error model was created using data from a study published by Goldberg et al. (Goldberg et al. 2004). This study investigated the accuracy of CGM devices in critically ill patients admitted to the 14-bed medical ICU in Yale New Haven Hospital. Patients admitted to the ICU that were considered at risk of hyperglycaemia were eligible to take part in the study, but those expected to stay for <24 hours were excluded. Risk factors for hyperglycaemia included a BG measurement ≥11.1mmol/L, a clinical history of Diabetes Mellitus (DM), or the use of clinical interventions associated with hyperglycaemia such as corticosteroids, vasopressors and/or enteral/parenteral nutrition.

This study used Medtronic Continuous Glucose Monitoring System (CGMS - Medtronic Minimed, Northridge, CA) monitors, receiving information from Medtronic SOF sensors (Medtronic Minimed, Northridge, CA). Sensors were inserted in the lateral abdomen, or upper, outer buttocks. At the time of this specific study, SOF sensors were validated for up to 72 hours use, after which time they were removed or replaced. Monitoring continued until discharge from the ICU. During monitoring, the CGM device was calibrated at least 4 times per day matching manufacturer guidelines (Mastrototaro 2000), with capillary BG measurements from a Surestep Flexx (Lifescan, Johnson & Johnson, Milpitas, CA) glucose meter. At the end of the monitoring period, data were downloaded using Minimed software Solutions V3.0 (Medtronic Minimed, Northridge, CA).

The study enrolled 22 patients and each patient had between 1 and 5 sensors inserted sequentially, resulting in a total of 41 CGM data sets. The SG data from CGM were compared to capillary BG measurements to assess accuracy and a total of 546 paired BG-SG measurements were available for analysis. Accuracy of the CGMS device was reported based on absolute glucose level, which was defined as the mean of the BG-SG paired measurements. The relevant accuracy results reported are shown in Table 3.1.

Table 3.1 Error characteristics of the CGMS in ICU patients, by absolute BG level (Goldberg et al. 2004)

Mean SG/BG (mg/dL)	Number of paired measurements	MAPE (%)
<100	43	20.0
100-149	219	13.5
150-199	185	11.3
200-249	66	11.4
≥250	33	9.8

The mean absolute percent error (MAPE) accuracy data from Table 3.1 were used to create a mathematical model of SG error (Chase et al. 2006; Gschwendtner 2007). The model consists of five independent Gaussian distributions, one for each of the BG-SG bins reported in Table 3.1. The

probability density function of a zero mean Gaussian distribution with scale parameter s_G, can be described:

$$f_n(x) = \frac{1}{\sqrt{2\pi}s_G} e^{-\frac{x^2}{2{s_G}^2}} \quad 3.1$$

Using the Box-Muller transform, Equation 3.1 takes the form:

$$f_n(x) = s_G\sqrt{-2ln\,(a_1)}\,cos(2\pi a_2) \quad 3.2$$

With uniformly distributed variates $a_1, a_2 \in (0,1]$.

Equation 3.2 represents the basis of the model and for each of the 5 BG-SG bins, the scale parameter is adjusted so the MAPE of the Gaussian distribution is the same as the corresponding MAPE reported in Table 3.1. Furthermore, each scale parameter can be approximated:

$$s_G = \frac{\sqrt{2\pi}}{2} \cdot MAPE_{req} \approx 1.25 \cdot MAPE_{req} \quad 3.3$$

The scale factors used to obtain the appropriate MAPE values from Table 3.1 are shown in Table 3.2.

Table 3.2: Scale factors (S_G) used in Gaussian CGMS error model

BG Range (mg/dL)	Paired BG-SG measurements	MAPE (%)	S_G
<100	43	20	0.25
100-149	219	13.5	0.17
150-199	185	11.3	0.14
200-249	66	11.4	0.14
≥250	33	9.8	0.12

With the error distributions for each BG range defined, SG data is generated by adding error from the model to known, underlying BG values. Ideally, underlying BG data will be produced at 5-minute intervals to replicate the sampling rate of

real CGMS data. Modelled sensor error is added to each individual BG value, as defined:

$$G_{noise,i} = G_{true,i} + f_n(i) \cdot G_{true,i} \qquad 3.4$$

Where:

$G_{true,i}$	True blood glucose concentration (mmol/L)
$G_{noise,i}$	CGMS blood glucose approximation (mmol/L)
$f_n(i)$	Gaussian distributed error factor (where n depends on $G_{true,i}$)

As there are no reported limits on maximum/minimum error for the CGMS sensor in the literature, bounds were placed on the output value $G_{noise,i}$ at 2.2mmol/L and 22.2mmol/L. These values are consistent with the limits on output SG data specified by the manufacturer of the CGMS (Mastrototaro et al. 2002).

The model was validated using a virtually generated cohort that has similar BG characteristics to those described by Goldberg et al. A total of 546 BG measurements, with the same number of BG measurements in each glycaemic interval as Goldberg et al., were used in the validation. Randomly sampled CGM error was added to each BG measurement to give virtual SG measurements, which were analysed and compared to the results reported by Goldberg et al. The metrics used to validate the model were the $r_{Pearsons}$ correlation coefficient, the overall MAPE and the Clarke Error Grid (Clarke et al. 1987). Table 3.3 shows the results of this model validation and in comparing the results with those reported by Goldberg et al., it is clear the model replicates the CGMS error characteristics well.

Table 3.3: Validation of the error model against results reported by Goldberg et al.

			Clarke Error Grid		
	$r_{Pearsons}$	**MAPE ± SD**	**Zone A (%)**	**Zone B (%)**	**Zone C-E (%)**
Goldberg	0.88	12.8 ± 11.9	78.4	20.3	1.3
Guassian model	0.91	12.8 ± 10.0	81.3	18.2	0.5

It should be noted that although model was designed to match Goldberg et al.'s reported results on a similar data set, the overall $r_{Pearson}$, MAPE and Clarke Error Grid results can change depending on the underlying BG data used as $G_{true,i}$ with the model. Also, the model described in this section assumes each SG value is independent of the previous values. However, without any quantification in the literature of the time dependent aspects of CGMS error, such as drift, they could not be included in this model.

3.1.2 Autoregressive based error model

The second CGM error model was created using data from a study published by Breton et al. in 2008 (Breton & Kovatchev 2008). In their paper, researchers proposed three important aspects of CGM that should be quantified in the model:

1. Physiological lag between BG and IG concentrations
2. Sensor lag between IG and SG at the sensor electrode
3. Residual sensor error or noise.

Two data sets containing BG and Freestyle Navigator (Abbott Diabetes Care, Alameda, CA) SG data were used to create the model. The first data set was from 136 patients who used the CGM device at home, calibrating via self-monitoring at irregular intervals. A total of 4000 days of 10-minute sampled SG data and 40,745 BG measurements were collected. After pre-analysis data processing, data set 1 contained 20,660 paired BG-SG paired measurements that could be used for modelling.

The second data set was from 28 patients who were monitored in a clinical environment. The CGM device in these patients produced a value every minute, and BG concentration was measured every 15 minutes using a YSI device (YSI Life Sciences, Yellow Springs, OH). A total of 56 sensors were used, two per patient, and after pre-analysis data processing, data set 2 contained approximately 7,000 paired BG-SG paired measurements

Data set 1 was used to address points 1 and 2 above, specifically, to approximate the overall lag between glucose in the blood reaching the sensor electrode. First,

SG data were recalibrated using a linear regression method, similar to the methods used by sensor manufacturers. Second, data were stratified into bins based on the linearly regressed rate of change of data in a 20 minute sliding window. Third, kernel density estimation was used to approximate the distributions of the sensor readings across a range of different reference glucose values.

The premise behind their methodology was that if SG measurements are delayed due to interstitial lag and/or signal processing, then sensor error will be dependent on the rate of change of glucose. Thus, the location of the kernels for each bin could be assessed to quantify the delay. The main observations made from the kernel density data were:

- At negative rate of change, the sensors tend to read high
- At positive rate of change, the sensor tends to read low
- The extent to which the sensor systematically reads high or low is correlated to the amplitude of the rate of change.

The researchers found that a correlation exists between glucose rate of change and average BG-SG discrepancy, irrespective of reference glucose concentration. The average discrepancy was linearly related to the rate of change and the slope of this relationship gave an estimation of the overall delay as 17 minutes.

The delay was modelled by a first-order diffusion process:

$$\frac{\partial G}{\partial t} = \frac{1}{\tau}(SG - BG) \qquad 3.5$$

Where SG represents glucose at the sensor and τ represents the blood to sensor time lag. Empirical estimations gave a time lag of 5mins, producing a delay of around 15 minutes, which is comparable with the 17 minute delay that was detected in data set 1.

Data set 2 was used to quantify the residual error in SG measurements and their interdependence over time. SG data was synchronised and recalibrated to BG data using a first-order diffusion model, which minimised delay and residual error. The time dependent aspect of SG errors was determined using classical time-series techniques. Specifically, the autocorrelation function and the partial autocorrelation function. The analysis suggested that sensor error was best predicted by a linear combination of the sensor error 15 minutes earlier and a random white noise term.

The distribution of residual SG errors had a mean of 0.04mmol/L and a standard deviation of 0.61mmol/L, but its skewness and kurtosis showed it to be non-Gaussian. However, an unbounded Johnson distribution, with parameters ξ = -5.471, λ = 15.96, γ = -0.5444 and δ = 1.6898, fit the empirical data well. This distribution was coupled with an autoregressive moving average process to model the non-white, non-Gaussian SG errors:

$$\begin{aligned} e_1 &= v_1 \\ e_n &= 0.7(e_{n-1} + v_n) \\ \varepsilon_n &= \xi + \lambda \cdot sinh(\frac{e_n - \gamma}{\delta}) \end{aligned} \qquad 3.6$$

where $v_n \sim \Phi(0,1)$ is IID, ξ = -5.471, λ = 15.96, γ = -0.5444 and δ = 1.6898.

Throughout the remainder of this thesis, where applicable, the implementation of this autoregressive based model was achieved in three steps. First, clinically validated metabolic system models were used to generate a continuous time series of glucose values from clinical BG measurements, insulin treatments and parenteral/enteral feed data. Second, an overall lag was applied to the data using the diffusion model shown in Equation 3.5, to account for physiological and sensor lags. Finally, Equation 3.6 was used to add time-dependent sensor error to the delayed glucose data, creating virtual SG data. Virtual SG data produced using this model have been previously validated and shown to be consistent with empirical data (Breton & Kovatchev 2008).

3.2 Filtering CGM data

Digital filtering techniques have been used to reduce the impact of seemingly random fluctuations in SG data on the overall trends in the time series (Bequette 2010). This research investigates common digital filters, such as finite impulse response (FIR) filters and infinite impulse response (IIR) filters, as well as a novel median least mean square filter. All filters were tested in silico using virtual SG data, generated using the Gaussian CGM error model described in Section 3.1.1.

3.2.1 Patients and methods

A 20 patient benchmark cohort (Chase et al. 2008b) was used to test the filter designs and optimise filter performance. Each patient in the cohort was on the SPRINT TGC protocol during their ICU stay (Chase et al. 2008c). While on SPRINT, BG was measured every 1-2 hours to guide therapy. Demographics for the patients are shown in Table 3.4. The use of these patient records fell under existing ethics approval granted by the Upper South Regional Ethics Committee, New Zealand.

Table 3.4: Demographics for the 20 patient benchmark cohort

Patients (N)	20
Age (years)	68 [60 - 73]
Gender (M/F)	12/8
Length of data (h)	276 [173 - 472]
BG measurements (N)	170 [130 - 260]
APACHE II	19 [16 - 24]

Clinically validated metabolic system models were used to produce a continuous glucose time series from the BG measurements, insulin treatments and parenteral/enteral feed data. This continuous glucose data was used as the foundation for generating virtual SG data. Randomly sampled error values from the Gaussian based error model described in Section 3.1.1 were added to the underlying BG trace, producing virtual SG data. The Gaussian model was used as it was derived from CGM in critically ill patients, which is the application of this research.

3.2.2 Filter designs

3.2.2.1 Finite Impulse Response (FIR) filters

FIR filters are a type of signal processing filter whose response to an external input of finite duration will decay to zero in a finite period of time. The general form of an Nth order FIR filter can be described:

$$y(n) = b_0 x(n) + b_1 x(n-1) + \cdots + b_N x(n-N) \tag{3.7}$$

where $y(n)$ is the filter output, $x(n)$ are the filter inputs, b_N are the filter coefficients, and N is the filter length/order.

There are a several methods available for determining optimal filter coefficients for a FIR filter, such as the window design method, frequency sampling method and weighted least squares design. In this research, a window design method was used and the filter was implemented in MATLAB (The Mathworks, Natick, Massachusetts). Finally, several filter length and cut-off frequency parameters were tested to optimise performance.

3.2.2.2 Infinite Impulse Response (IIR) filters

Unlike FIR filters, an IIR filter's response to an external input will remain non-zero for an infinite duration of time. This behaviour is due to the feedback terms in the general form of an Nth order IIR filter, which can be described:

$$\begin{aligned} y(n) = \frac{1}{a_0}(b_0 x(n) + b_1 x(n-1) + \cdots + b_P x(n-P) - a_1 y(n-1) \\ - a_2 y(n-2) - \cdots - a_Q x(n-Q)) \end{aligned} \tag{3.8}$$

where $y(n)$ is the filter output, $x(n)$ are the filter inputs, b_P are the feedforwared coefficients and a_Q are the feedback coefficients. P is the feed forward filter length and Q is the feedback filter length.

Four different types of IIR filters were tested in this research, specifically, Butterworth, Yulewalk, Chebyshev type 1 and Chebyshev type 2. Briefly, Butterworth filters are designed to minimise ripple in the passband and stopband. Chebyshev filters are designed to minimise the error between ideal and actual filter performance at the expense of some passband (type 1) or stopband (type 2) ripple. Yulewalk filters are recursively designed filters that utilise a least squares fit to a specified frequency band. Each filter was implemented in MATLAB.

3.2.2.3 Median Least Mean Square filter

A novel median least mean square (LMS) filter was also tested in this study (Pretty et al. 2010). The filter uses a composite trailing 3 point and 7 point median value together with a linear least squares approximation to filter data in real-time. The fundamental steps to implement the median filter at any given time, t=x are:

1. Calculate the median value, M_3, of the current and two prior SG readings (three samples, 10-minute window)
2. Calculate the median value, M_7, of the current and six prior SG readings (seven samples, 30-minute window)
3. Determine average of M_3 and $M_7 = M_A$
4. Fit LMS first-order polynomial line to the current M_A value and 12 prior M_A values (1-hour window of median-filtered values)
5. Output value at time $t = x$ is the value of this fitted line at $t = x$

Similar to the FIR and IIR filters, this algorithm was implemented in MATLAB to assess filter performance.

3.2.3 Analysis

Due the random nature of SG data generation, Monte Carlo (MC) simulations consisting of 50 runs per patient were used to assess filter performance. The overall MAPE and Median absolute percent error (APE) of the unfiltered virtual SG data were compared to results for the filtered data. For the FIR and IIR filter variants, five cut-off frequencies (ω_c) and six filter lengths (N_F) were tested, but

only the results from the optimal combination of parameters are presented. The median filter was implemented as described in Section 3.2.2.3.

3.2.4 Filter results

The filter performance results in Table 3.5 show that all filter types reduced the overall MAPE and Median APE, compared to the unfiltered virtual SG data. The Yulewalk IIR filter and Median LMS filter have the biggest impact, reducing the MAPE by approximately ~5.8% and the Median APE by 4.8%.

Table 3.5: Filter results from 50-run simulations of 20 virtual SG data sets

	Unfiltered	Windowed FIR	Butterworth IIR	Yulewalk IIR	Chebyshev I IIR	Chebyshev II IIR	Median LMS
		$\omega_c = 0.01$, $N_F = 3$	$\omega_c = 0.03$, $N_F = 3$	$\omega_c = 0.03$, $N_F = 3$	$\omega_c = 0.02$, $N_F = 3$	$\omega_c = 0.05$, $N_F = 25$	
Median APE (%)	13.5	9.7	9.5	8.9	10.1	13.6	9
MAPE (%)	16.8	11.9	11.4	11	11.9	16.4	10.9

Based on these results, the Median LMS filter was selected for the remainder of the in-silico studies in this research. Furthermore, for retrospective filtering, median filters can be centred on the time point of interest to eliminate filter lag. More advanced filtering methods, such as Kalman filtering, could be investigated in future with a larger cohort of clinical SG data. These results simply provide an initial quantification of the effectiveness of basic digital filtering methods.

3.3 Hypoglycaemic alarms

Hypoglycaemic alarms can be used to predict the onset of hypoglycaemia, alerting the user before the BG level drops to a dangerous level. There are a wide range of algorithms that can be used, such as threshold alarms, integral alarms and Kalman or model-based alarms (Bequette 2010). This section investigates the performance of two types of hypoglycaemia alarms: 1) a simple threshold alarm, and, 2) an integral alarm (Pretty et al. 2010; Signal et al. 2010).

3.3.1 Patients and methods

This research was conducted as a retrospective study using records from seven patients admitted to the Christchurch Hospital ICU between 2005 and 2007.

Patients were included if they had one or more severe hypoglycaemic episodes, defined as BG < 2.2mmol/L, while on the SPRINT glycaemic control protocol (Chase et al. 2008c). Patients were excluded if the hypoglycaemic episode appeared to be due to sensor failure or a recording error, based on surrounding data and expert clinical opinion. Details of the cohort are shown in Table 3.6. The use of these patient records fell under existing ethics approval granted by the Upper South Regional Ethics Committee, New Zealand.

Table 3.6: Demographics for the 7 patients who had severe hypoglycaemia

Patients (N)	7
Age (years)	63 (37 - 81)
Gender (M/F)	4/3
Hypoglyceamia BG level (mg/dL)	2.1 (1.7 - 2.2)
APACHE II	25 (12 - 30)

Clinically validated metabolic system models were used to produce a continuous glucose time series from the clinical BG measurements, insulin treatments and parenteral/enteral feed data. Similar to the process used for testing the digital filters, randomly selected errors from the Gaussian based error model were added to the underlying BG trace, producing in a virtual SG data set. These virtual SG data sets consisted of 5-minuted SG measurements. To simulate the real-time use of CGM, an algorithm was implemented that stepped through the sequence of virtual SG data without knowledge of "future" values. All virtual SG data were filtered using the real-time median filter described in Section 3.2.2.3. Thus, the combination of real-time data streaming with real-time filtering replicated the clinical situation in this study of hypoglycaemia alarming.

3.3.2 Alarm designs

Two hypoglycaemia alarms were tested in this research, a threshold alarm and an integral alarm.

3.3.2.1 Threshold alarm

In its most simple form, a threshold hypoglycaemia alarm will trigger when a single glucose measurement is below a pre-defined threshold. Further

conditions, such as multiple measurements below a threshold or an average rate of change less than zero can be added to improve the robustness of the alarm. This study investigated a threshold alarm with the following alarm conditions:

- 3 consecutive virtual SG measurements < threshold
- Average SG gradient < 0mmol/L.min

Three threshold levels were tested, at 2.8mmol/L, 3.3mmol/L and 3.9mmol/L.

3.3.2.2 Integral alarm

Integral alarms use the area under the (filtered) SG time series to determine when to trigger the alarm. Specifically, in this investigation, the area between the virtual SG data and a specified level was calculated within a window of prior samples, as shown in Figure 3.1. When this integral became less than a preselected threshold value, an alarm was triggered, indicating an impending hypoglycaemic episode. In this study, 3 integration window lengths of 5,7 or 13 SG measurements were tested. Furthermore, for each integration window length, the 3 specified threshold values were tested.

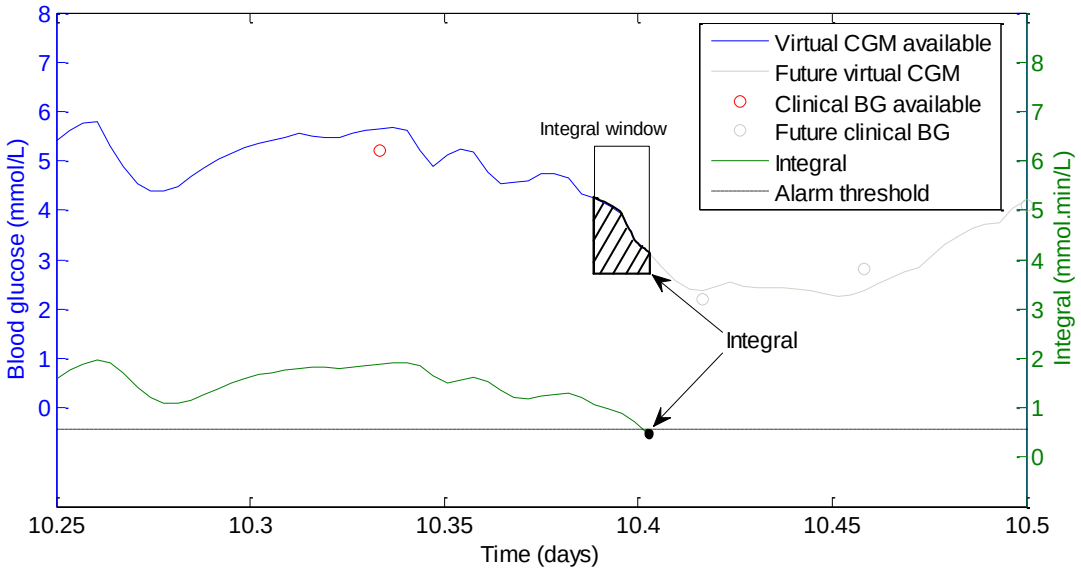


Figure 3.1 Integral alarm method. The area below the SG data within the integral window is used to trigger the alarm. The integral window can only use data prior to the time point of interest, to simulate a realistic clinical scenario.

3.3.3 Analysis

The performance of each alarm variant was quantified by lead time of the alarm warning and the number of false alarms. Lead time was defined as the time difference between the alarm triggering and the clinical BG measurement that diagnosed hypoglycaemia. This value essentially measured the time to intervene and prevent severe hypoglycaemia. The number of false alarms was also recorded to ensure that the method was accurate. An alarm was considered false if both of the following conditions held:

- There was more than one alarm for each actual hypoglycaemic episode per patient
- For two clinical blood glucose measurements either side of the alarm, no value was less than or equal to 2.2mmol/L.

To get meaningful results with the random SG error, MC simulations were used. Each patient's model-based, BG profile was passed through the Gaussian CGM error generator 100 times, creating 100 different virtual SG sequences per patient. Results of these 700 trials with the alarm algorithms were analyzed using nonparametric statistics. Negative time values indicate that an alarm was triggered before a hypoglycaemic event was measured.

3.3.4 Alarm results

3.3.4.1 Threshold alarm performance

A representative example of hypoglycaemia detection using the threshold alarm is shown in Figure 2.1. In this figure, multiple virtual SG values cross the alarm threshold with a negative gradient, satisfying the conditions for alarming. Thus, clinical staff would receive warning of hypoglycaemia, before the clinical BG measurement that actually diagnosed hypoglycaemia.

The results of the MC simulations are shown in Table 3.7. These results show an important trade-off when determining the threshold to be used with a particular threshold alarm design. First, raising the threshold from 2.8mmol/L to 3.3mmol/L and subsequently to 3.9mmol/L increases both the lead time of detection and the BG level at alarm. However, these increases come at the

expense of alarm sensitivity and specificity. For example, with a threshold of 3.9mmol/L the alarm gave a median lead time of 75 minutes, but the median number of false alarms was 7. Conversely, with a threshold of 2.8mmol/L, the lead time is reduced to a median of 15 minutes, but the number of false alarms is also significantly reduced.

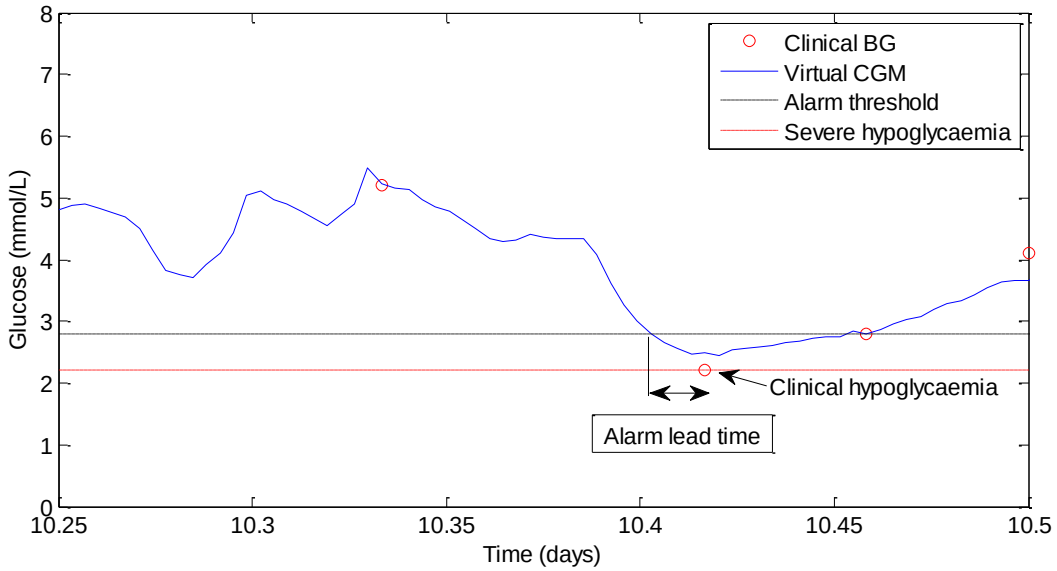


Figure 3.2 A representative example of a threshold alarm giving warning of impending hypoglycaemia.

Table 3.7: Threshold alarm results over all 700 MC runs

Point below threshold	3		
Average gradient (mmol/L.min)	<0		
Alarm threshold (mmol/L)	2.8	3.3	3.9
Hypoglycaemia warning (mins)	-15 [-25 - -5]	-40 [-75 - -25]	-75 [-110 - -45]
BG level at alarm (mmol/L)	2.4 [2.3 - 2.6]	2.8 [2.4 - 2.9]	3.1 [2.9 - 3.4]
False alarms	-1 [-1 - 1]	-1 [-1 - 6]	7 [-1 - 13]

These results highlight the importance of tuning threshold alarms appropriately for their intended application. In the critical care setting, where clinical staff are extremely busy and time is extremely valuable, the number of false alarms is potentially more critical than the lead time. A high rate of false alarms will likely waste a significant amount nurse time. However, in terms of lead time, a nurse at

the bedside should be able to confirm a diagnosis of hypoglycaemia and provide treatment within 15 minutes, should the alarm sound. Thus, an alarm for critical care might be optimised for sensitivity and specificity at the expense of lead time. The opposite could also be true for an otherwise healthy individual using the alarm for daily BG control. The results shown here suggest a threshold of ~3mmol/L would be well suited for critical care applications.

3.3.4.2 Integral alarm performance

Figure 3.3 shows a representative example of the integral alarm triggering at the predicted onset of hypoglycaemia. This example uses the same clinical BG measurements as Figure 2.1, but the SG data appears slightly different due to the randomly selected CGM error. The integral alarm detects the hypoglycaemia well in advance of the clinical BG measurement.

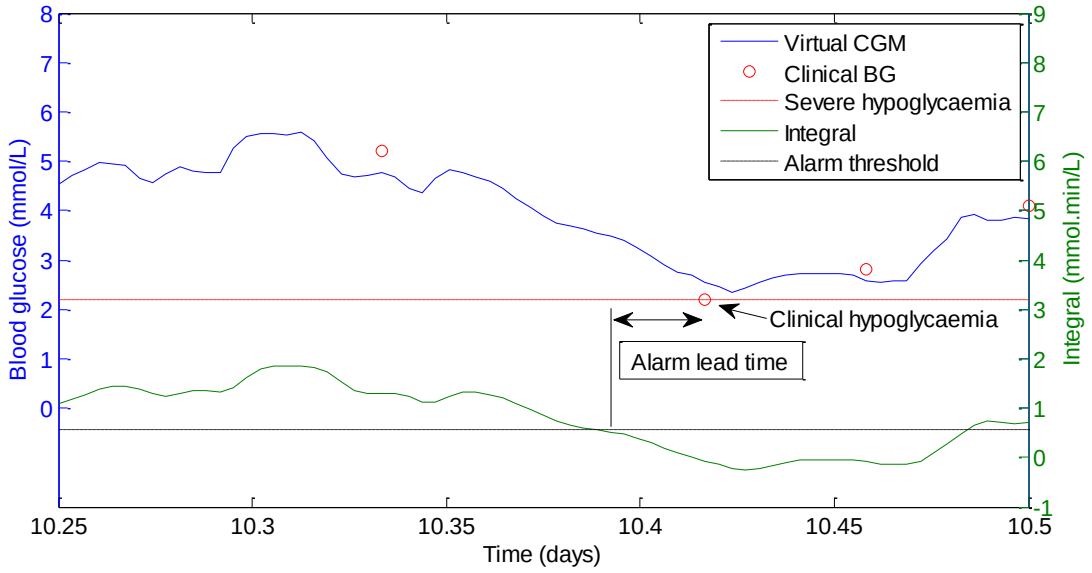


Figure 3.3 A representative example of an integral alarm giving warning of impending hypoglycaemia.

Table 3.8 shows the results from the MC simulations assessing the different integral alarm configurations. Two important trends are evident in the results. First, increasing the integral window length increased reliability of the alarm by providing a higher degree of filtering, which reduced the number of false alarms. For example, increasing the integration window length from 5 to 13, while

holding the integral threshold at 0.28 min.mmol/L, reduced the number of false alarms from 1 [0 - 2] to 0 [0 - 1]. However, the reduction in false alarms came at the expense of lead time, which was likely reduced by the additional filter lag, delaying the integral threshold crossing.

Second, increasing the integral threshold increased the lead time, up to a median of 145 minutes for the 5 integral filter with a threshold of 1.12. This trend was consistent across all integration window lengths and the increase in lead time was similar for each variant. The trade-off of increasing the threshold to increase lead times was the increase in false alarms. Most notably, for the integral alarm of length 7, the number of false alarms increased from 1 [0 - 1] to 2 [1 - 3] when the threshold was increased from 0.28 to 1.12 min.mmol/L.

Table 3.8: Integral alarm results for all 700 MC runs

Integration window (n)	5		
Integral threshold (min.mmol/L)	0.28	0.56	1.12
Hypoglycaemia warning (mins)	-65 [-110 - -35]	-95 [-120 - -55]	-145 [-180 - -120]
BG level at alarm (mg/dL)	3.2 [3.1 - 3.3]	3.6 [3.4 - 3.7]	4.3 [4.1 - 4.4]
False alarms	1 [0 - 2]	1 [0 - 2]	2 [1 - 3]
Integration window (n)	7		
Integral threshold (min.mmol/L)	0.28	0.56	1.12
Hypoglycaemia warning (mins)	-55 [-95 - -30]	-85 [-120 - -50]	-130 [-180 - -115]
BG level at alarm (mg/dL)	3.1 [2.9 - 3.2]	3.4 [3.3 - 3.6]	4.2 [3.9 - 4.3]
False alarms	1 [0 - 1]	1 [0 - 2]	2 [1 - 3]
Integration window (n)	13		
Integral threshold (min.mmol/L)	0.28	0.56	1.12
Hypoglycaemia warning (mins)	-35 [-73 - -15]	-55 [-95 - -25]	-120 [-160 - -85]
BG level at alarm (mg/dL)	2.9 [2.7 - 3.0]	3.1 [2.8 - 3.3]	4.0 [3.4 - 4.2]
False alarms	0 [0 - 1]	1 [0 - 1]	1 [1 - 2]

Overall, there was a clear improvement in the time of detection of hypoglycaemic episodes using SG data compared to standard intermittent BG measurements. However, trade-offs were exposed when selecting integration window length and integral threshold. Ultimately, the optimal values for window length and alarm threshold are likely to depend upon the application of the alarm. In the critical care setting lead time is arguably less important than high sensitivity.

Thus, a longer integration window might be coupled with a lower integral threshold to reduce false alarms at the expense of lead time.

3.3.5 Limitations

Overall, the low number of data sets available for this investigation may mean that the results are not representative of the overall population behaviour. For this reason, extensive optimisation of alarm algorithms was not justified. However, the results shown here do give a good indication of how different alarm parameters affect lead time, sensitivity and the rate of false alarms. Repeating the study with a larger cohort would provide a more statistically powerful result. However, in the Christchurch ICU, all patients receiving insulin are on the SPRINT glycaemic control protocol and only ~4% of patients experience hypoglycaemia, making it difficult to recruit a large cohort. Thus, data from a different study at a different centre might be needed to obtain a large cohort containing clinical hypoglycaemia.

3.4 Summary

The research presented in this chapter investigated 3 key elements required to complete in-silico trials of CGM in critically ill patients. First, CGM error models based on available data from the literature were explored. Two models were described and can be used for future work: 1) A Gaussian based CGM error model, created using ICU specific SG data, and, 2) an autoregressive based CGM error model, derived using SG data from people with diabetes. Second, a range of digital filters including FIR, IIR and median LMS filters were tested. The median filter was the most effective at removing SG error, reducing the overall MAPE from 16.8% pre-filtering to 10.9% post-filtering. Third, threshold and integral hypoglycaemia alarms were investigated and lead times of 15-145 minutes were achieved, depending on alarm parameters. However, the trade-off between alarm lead time and sensitivity to detecting hypoglycaemia reinforces the importance of tuning alarm parameters for the intended application. Overall, this study shows that simple filtering and alarm methods can provide a highly effective warning system to alert at the onset of predicted abnormal glycaemia.

Chapter 4. In-silico TGC with SG

This chapter describes an in silico feasibility study conducted to determine whether CGM devices could be effectively utilised in the Christchurch Hospital ICU, in conjunction with existing TGC protocols. The main goals of this study were to show that nurse workload could be significantly reduced, BG control could be maintained, and hypoglycaemia could be reduced, with the implementation of CGM devices in a successful, effective and existing TGC protocol.

4.1 Introduction

Many different TGC protocols have been trialled in critical care units around the world, from sliding scale protocols (Gonzalez-Michaca et al. 2002; Krinsley 2004; Zimmerman et al. 2004) to advanced model-based control protocols (Chase et al. 2007; Evans et al. 2012; Fisk et al. 2012; Plank et al. 2006). The Christchurch Hospital ICU has been using the SPecialised Relative Insulin Nutrition Tables (SPRINT) protocol since 2005, when the protocol was implemented as a clinical practice change (Chase et al. 2008c).

The SPRINT protocol is a nurse driven model-based protocol that was developed using simulation methods that utilise clinically validated mathematical models of the glucose-insulin regulatory system (Lonergan et al. 2006; Wong et al. 2006). Most TGC protocols only adjust insulin in response to BG levels (Meijering et al. 2006), but SPRINT modulates both insulin and nutrition to obtain maximum performance and safety. Thus, SPRINT was introduced as two 'paper wheels' shown in Figure 4.1

Start up criteria for SPRINT is 2 BG measurements > 8mmol/L, although occasionally a patient will start on SPRINT at the discretion of the attending physician. Once on SPRINT, BG is measured at the bedside every 1-2 hours by the nurse, depending on the stability of BG. The current BG measurement, previous insulin bolus and previous nutrition rates are selected on the corresponding

paper-wheel to determine the new treatment. All insulin doses are given as boluses to reduce the risk of hypoglycaemia that can occur if an insulin infusion is left unattended.

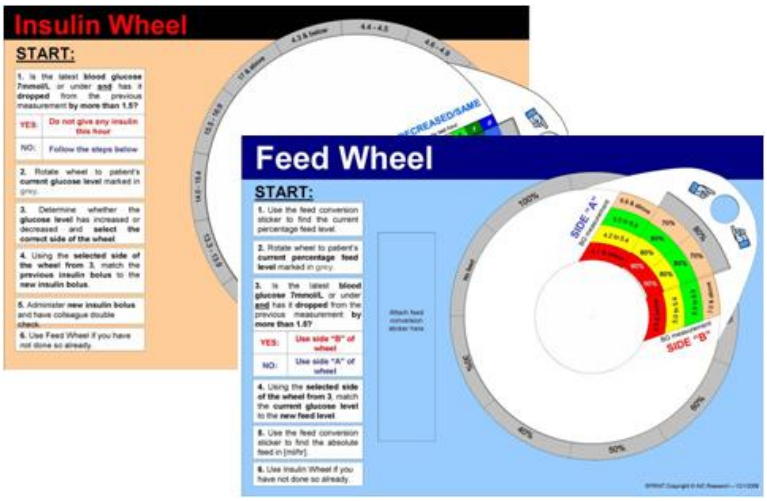


Figure 4.1 SPecialised Relative Insulin Nutrition (SPRINT) protocol insulin and nutrition wheels.

An analysis of the performance of SPRINT across 371 patients showed excellent results. Specifically, more BG measurements were in the desirable glycaemic band (4.4-6.1mmol/L) and, although still present on rare occasions, the instances of severe hypoglycaemia (BG < 2.2mmol/L) were significantly reduced compared to treatment methods prior to SPRINT (Chase et al. 2008c). The only major limitation of the new protocol was the number of BG measurements required per day, which averaged around 16 per patient per day. At approximately 1.5-4 minutes per BG measurement, nurses could easily spend a total of 1 hour or more measuring BG, which is clinically significant.

CGM devices with their 1-5 minute measurement interval offer far more frequent BG measurements than standard intermittent BG monitoring, potentially enabling better BG control and further reduction in hypoglycaemia through early warning alarms/interventions. Furthermore, most CGM devices only require 3-4 manual BG measurements per day for device calibration. Thus, nurse workload in measuring BG could also potentially be reduced by up to ~75%. However, added sensor noise or error is a trade off in some cases of the CGM devices far

higher, automated sampling rate and any excess noise must be effectively managed. Fortunately, these sensor and algorithm technologies are also constantly evolving with every new generation offering improvements.

This study examines how patients might behave when controlled by the SPRINT TGC protocol driven by simulated SG measurements, rather than the normal 1-2 hourly BG measurements taken by the nurse. Hypoglycaemia detection and prevention are also tested, as CGM devices offer the ability to potentially detect and avert these events much earlier than with current methods. This *in-silico* study thus aims to demonstrate that CGM devices coupled with an effective TGC protocol are capable of reducing nursing workload, while maintaining safe and effective glycaemic control.

4.2 Subjects and methods

4.2.1 Subjects

This study uses patient data from the benchmark cohort, proposed by Chase et al. (Chase et al. 2008b). It includes 20 patients who were admitted to the Christchurch hospital ICU during the SPRINT protocol clinical practice change (Chase et al. 2008c). Each patient spent 5 days or longer on SPRINT, during which time some patients experienced breaks in treatment, typically due to surgery or other clinical changes. Patients who experienced breaks from SPRINT had their BG data segmented into 2 or more uninterrupted treatment sections, resulting in 33 individual continuous data sets. Each treatment section was then considered an individual 'patient' for the purposes of this study. Details of the patient cohort are shown in Table 4.1. The use of these patient records falls under existing ethics approval granted by the Upper South Regional Ethics Committee, New Zealand.

Table 4.1 Cohort details, presented as Median [inter-quartile range (IQR)] where applicable.

Number of Patients	33
Age	68 [59.5 - 73]
Gender	60% Male
Length of SPRINT (hours)	167 [81 - 241]
Number of Measurements	102 [57 - 169]
APACHE II	19 [16 - 24]
APACHE II risk of death	33.6 [13.5 - 51.2]

4.2.2 Simulated SG data, filtering and hypoglycaemia alarm design

These methods were implemented using MATLAB. BG profiles were generated incrementally for each patient at 5 minute intervals using clinically validated metabolic system models (Lonergan et al. 2006; Wong et al. 2006). These BG profiles were sampled *in-silico* and CGM sensor errors were added creating virtual SG data.

This study used the Gaussian and Autoregressive based CGM error models described in Chapter 3. Two levels of the Gaussian error model were investigated. First, the model was implemented as described in Chapter 3, to produce equivalent simulated SG errors on a similar cohort to those reported by a study of CGM in a medical ICU in 2004 (Goldberg et al. 2004). Second, the Gaussian model was implemented at a reduced magnitude, where SG errors generated by the model were halved to account for any technological advances in these devices since 2004. Overall, SG errors were randomly sampled from the appropriate Gaussian distribution and added to underlying BG data, creating virtual SG data. The top and middle plots in Figure 4.2 show examples of the SG data generated using the Gaussian CGM error model. Figure 4.3 compares the probability density functions for the CGM error models.

The autoregressive noise model was based on data from a 2008 study of CGM in ambulatory type 1 diabetics monitored using the FreeStyle Navigator CGM device (Breton & Kovatchev 2008). This model was implemented as described in Chapter 3. Briefly, an overall lag was applied to BG data using a first order

diffusion model to account for physiological and sensor lags. Next, time-dependent sensor errors were added to the delayed glucose data, creating virtual SG data. The bottom plot in Figure 4.2 shows an example of SG data generated using the autoregressive CGM error model. The probability density function (PDF) for the autoregressive model in Figure 4.3 is similar to the reduced Gaussian model distribution, but the time dependent aspect of the autoregressive model is not portrayed.

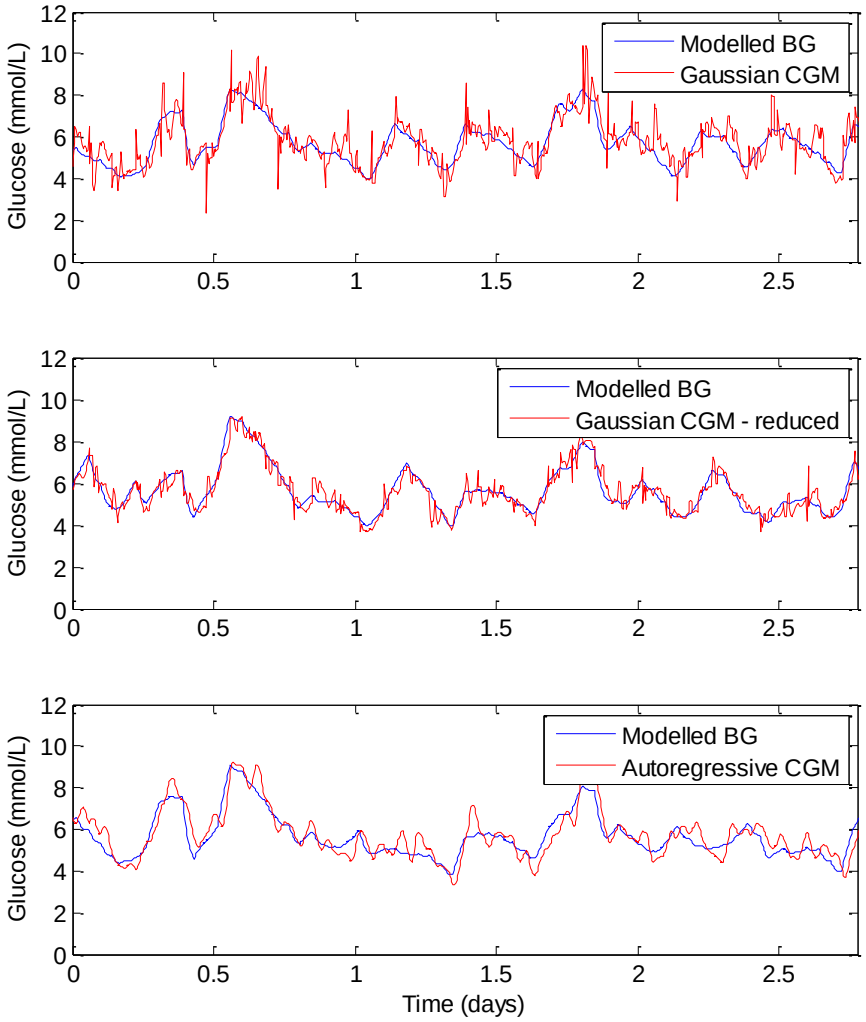


Figure 4.2 Example of simulated SG data overlaying modelled BG data. Top - SG data created using the Gaussian model, Middle - SG data created using the Gaussian model with reduced magnitude, and, Bottom - SG data created using the autoregressive model.

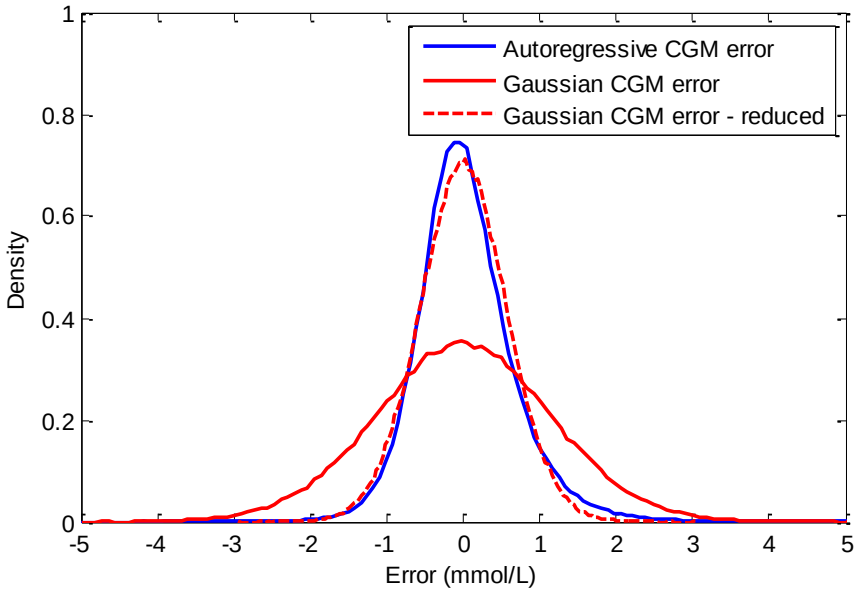


Figure 4.3 Model derived probability density functions for the 3 variations of SG error over ~250,000 data points

In addition to a random or quasi-random error, CGM sensor error often contains a bias or drift aspect (Boyne et al. 2003; Castle & Ward 2010). Drift and bias, which can occur due to sensor degradation or bio-fouling over time, were not considered in this study. Although mentioned in the literature, these factors have not been reported in detail to the author's knowledge and thus could not be implemented in the sensor error models.

Virtual SG data were filtered incrementally, simulating the process that would be encountered in a real-time clinical setting. The novel median filter described in Chapter 3 was used in this study. Briefly, the filter consists of a composite 3 point and 7 point median filter, coupled with a LMS estimator. This filter is effective at removing unwanted and potentially un-physiological noise from SG data (Pretty et al. 2010).

A simple threshold algorithm was used to trigger an alarm and dextrose intervention when a hypoglycaemia appeared imminent. Specifically, if two consecutive filtered SG measurements were below 3mmol/L, an intravenous dextrose intervention was triggered. Four different sized dextrose interventions were tested in this study (3, 6, 12.5, and 25 grams), although these were held

constant for any given simulation. The intervention was given over 3 minutes to approximate the clinical situation and two interventions could not be given within 15 minutes, to allow time for a BG change before the second intervention.

4.2.3 Simulating SPRINT with SG

A single measurement from the filtered virtual SG data was taken every 1-2 hours and used to guide the SPRINT protocol, with no other modifications to the algorithm. The SPRINT algorithm and its intervention frequency were not changed to enable a direct comparison to clinical results and interventions. While the SPRINT protocol only used 1-2 hourly values from the filtered virtual SG data, the full set of SG data was used for hypoglycaemia alarming and intervention.

4.2.4 Analysis

MC methods were used due to the random nature of SG error, and each of the 33 patients were simulated over 10 runs. A total of 7 simulations were run to test the various CGM error models and hypoglycaemia interventions. First, 3 simulations were run to test each CGM error model without hypoglycaemia intervention, and these were compared to baseline results determined using the clinical BG data. Second, 4 simulations were run using the Gaussian CGM error model to test the effect of each hypoglycaemia intervention.

Time in a 4.4 - 6.1mmol/L glucose band and number of hypoglycaemic events were considered the most important metrics of TGC performance for this investigation. Time in band was calculated using only the 1-2 hourly filtered SG measurements that were used to drive SPRINT, ensuring an equal comparison to the baseline results. Hypoglycaemia was defined as a BG < 2.2mmol/L in the baseline case, or a filtered virtual SG measurement < 2.2mmol/L in the simulation runs.

Nutrition and insulin interventions prescribed by SPRINT were also examined. Differences in these interventions compared to the baseline SPRINT results were

analysed to determine the impact of CGM sensor noise on operation of this protocol.

4.3 Results

Table 4.2 shows how glycaemic control was affected when the CGM error models were implemented with no hypoglycaemia interventions. The median and IQR BG results show no clinically significant difference between the baseline, reduced Gaussian error, full Gaussian error, and autoregressive error simulations. The time in band is reduced by up to 4.5% with the addition of SG error and consequently the <4.4mmol/L, <4mmol/L and <2.2mmol/L bands are slightly increased. Insulin and glucose interventions are unchanged, but this may also be partially due the discretisation of these interventions in SPRINT (Lonergan et al. 2006). Per-patient results show similar overall trends.

Table 4.3 shows the results from the 4 simulations of Gaussian error when a dextrose bolus intervention was administered at the predicted onset of hypoglycaemia. The median and IQR BG results show no clinically significant difference between any of the interventions and the baseline. Interventions did not increase time in any band when compared to results with no intervention in Table 4.2. However, the number of interventions and hypoglycaemic events decreases with larger interventions, up to a limit of 12.5g. With an intervention of 25g the number of hypoglycaemic events increases. This potentially counter-intuitive trend continues with 40g interventions (results not shown). Overall, the interventions have reduced the number of hypoglycaemic events from 1 in 33 patient episodes to 0, which is clinically significant.

Table 4.2 Summary statistics from the MC simulations with different noise models for both the whole cohort and per-patient results. Lognormal and non-parametric statistics are used with lognormal as indicated. For summary statistics IQR is presented across results from all 10 MC runs and 33 episodes. For per-patient results IQR is across median per-patient results.

Summary Statistics	**Clinical Baseline**	**Half noise, no intervention**	**Full noise, no intervention**	**Breton noise, no intervention**
Num episodes:	33.0	33.0	33.0	33
Total hours:	6881.0	6881.0	6881.0	6881.0
Num 1-2 hourly BG measurements [IQR]:	3953.0	3978 [3969 - 3984]	4111 [4108 - 4127]	4080 [4067 - 4090]
BG median [IQR] (mg/dL):	105.5 [92.2 - 119.5]	105 [91.4 - 119.4]	106.7 [91.6 - 122.0]	105.3 [91.0 - 121.6]
BG mean (geometric) [IQR] (mg/dL):	105.9	105.16 [104.98 - 105.20]	105.98 [105.81 - 106.06]	105.15 [105.13 - 105.26]
BG StDev (geometric) (mg/dL):	20.2	20.22 [20.21 - 20.24]	20.32 [20.32 - 20.32]	20.34 [20.33 - 20.34]
% BG within 80-110 mg/dL [IQR]	49.6	50.3	45.2	47.2
% BG < 80 mg/dL [IQR]	9.0	9.90 [9.73 - 10.19]	11.37 [11.25 - 11.55]	11.24 [10.93 - 11.38]
% BG < 72 mg/dL [IQR]	3.9	4.46 [4.29 - 4.53]	5.64 [5.48 - 5.78]	5.47 [5.43 - 5.58]
% BG < 40 mg/dL [IQR]	0.0	0.04 [0.03 - 0.08]	0.05 [0.03 - 0.07]	0.10 [0.10 - 0.12]
Num patients < 40 mg/dL [IQR]	1.0	2 [1 - 2]	2 [1 - 2]	3 [2 - 4]
Median insulin rate [IQR] (U/hr):	3.0 [2.0 - 3.0]	3.0 [2.0 - 3.0]	3.0 [2.0 - 3.0]	3.0 [2.0 - 3.0]
Median glucose rate [IQR] (g/hr):	4.6 [3.3 - 5.9]	4.6 [3.3 - 5.9]	4.6 [3.3 - 5.9]	3.9 [3.3 - 5.2]
number of interventions (where applicable)	0.0	0.0	0.0	0.0
Per-patient statistics				
Hours of control [IQR]:	166.0 [71.3 - 252.7]	166.0 [71.3 - 252.2]	166.0 [70.5 - 253.0]	166.0 [71.3 - 252.2]
Num BG measurements [IQR]:	101.0 [43.8 - 149.5]	98.0 [44.5 - 150.7]	102.0 [43.8 - 157.2]	102.0 [43.8 - 160.5]
BG median [IQR] (mg/dL)	106.8 [102.0 - 113.4]	107.1 [102.0 - 112.0]	108.5 [104.0 - 114.0]	107.3 [102.4 - 113.4]
BG mean (lognormal) (mg/dL):	109.7 [102.0 - 116.6]	108.9 [100.3 - 116.4]	109.5 [103.3 - 115.5]	108.8 [101.5 - 115.3]
BG StDev (lognormal) (mg/dL):	20.0 [19.6 - 20.3]	20.1 [19.5 - 20.4]	20.2 [19.8 - 20.6]	20.1 [19.8 - 20.5]
Time to < 140 mg/dL [IQR] (hours):	0.0 [0.0 - 4.0]	0.0 [0.0 - 4.0]	0.0 [0.0 - 3.2]	0.0 [0.0 - 4.0]
% patients to < 140 mg/dL [IQR]:	100.0	100 [100 - 100]	100 [100 - 100]	100 [100 - 100]
Time to < 110 mg/dL [IQR] (hours):	5.0 [2.0 - 12.0]	6.0 [2.0 - 11.2]	5.0 [2.0 - 10.2]	5.0 [1.8 - 11.2]
% patients to < 110 mg/dL [IQR]:	100.0	100 [100 - 100]	100 [100 - 100]	100 [100 - 100]
Median insulin rate [IQR] (U/hr):	3.0 [3.0 - 3.0]	3.0 [3.0 - 3.0]	3.0 [3.0 - 3.0]	3.0 [3.0 - 3.0]
Median dextrose rate [IQR] (g/hr):	4.6 [3.3 - 5.2]	4.6 [3.3 - 5.2]	3.9 [3.3 - 5.2]	3.9 [3.3 - 5.2]

Table 4.3 Summary statistics from MC simulations with intervention. Lognormal or non-parametric statistics are used with lognormal as indicated.

Summary Statistics	**Intervention = 3 grams**	**Intervention = 6 grams**	**Intervention = 12.5 grams**	**Intervention = 25 grams**
Num episodes:	33.0	33.0	33.0	33.0
Total hours:	6881.0	6881.0	6881.0	6881.0
Num 1-2 hourly BG measurements [IQR]:	4114.5 [4095 - 4118]	4095 [4081.5 - 4117.5]	4102 [4091 - 4104]	4086 [4076 - 4100]
BG median [IQR] (mg/dL):	106.2 [91.3 - 121.2]	106.0 [91.4 - 120.7]	106.3 [91.8 - 121.3]	106.3 [91.7 - 121.3]
BG mean (geometric) [IQR] (mg/dL):	105.34 [105.28 - 105.51]	105.58 [105.51 - 105.70]	105.87 [105.74 - 105.95]	106.09 [105.99 - 106.17]
BG StDev (geometric) (mg/dL):	20.3	20.3	20.3	20.3
% BG within 80-110 mg/dL [IQR]	45.37 [45.03 - 45.53]	45.82 [45.66 - 46.42]	45.5 [45.33 - 45.85]	45.60 [45.31 - 46.22]
% BG < 80 mg/dL [IQR]	12.17 [11.80 - 12.32]	11.69 [11.56 - 11.82]	11.48 [11.38 - 11.52]	11.38 [11.21 - 11.45]
% BG < 72 mg/dL [IQR]	5.98 [5.73 - 6.02]	5.57 [5.37 - 5.70]	5.34 [5.27 - 5.53]	5.01 [4.89 - 5.22]
% BG < 40 mg/dL [IQR]	0.01 [0 - 0.05]	0.01 [0 - 0.02]	0 [0 - 0]	0.04 [0.02 - 0.05]
Num patients < 40 mg/dL [IQR]	**1 [0 - 2]**	**1 [0 - 1]**	**0 [0 - 0]**	**1 [1 - 2]**
Median insulin rate [IQR] (U/hr):	3.0 [2.0 - 3.0]	3.0 [2.0 - 3.0]	3.0 [2.0 - 3.0]	3.0 [2.0 - 3.0]
Median glucose rate [IQR] (g/hr):	4.6 [3.3 - 5.9]	4.6 [3.3 - 5.9]	4.6 [3.3 - 5.9]	4.6 [3.3 - 5.9]
number of interventions (where applicable)	64.5	55.8	51.0	55.0
Per-patient statistics				
Hours of control [IQR]:	166.0 [70.5 - 252.7]	166.0 [71.3 - 252.2]	166.0 [70.5 - 252.0]	166.0 [71.3 - 252.7]
Num BG measurements [IQR]:	102.0 [43.5 - 159.0]	102.0 [42.3 - 160.7]	102.0 [42.5 - 158.7]	100.0 [43.0 - 157.0]
BG median [IQR] (mg/dL)	107.7 [104.3 - 112.8]	107.9 [102.4 - 114.8]	108.1 [103.7 - 114.8]	108.2 [103.9 - 112.5]
BG mean (lognormal) (mg/dL):	109.1 [101.3 - 115.3]	109.1 [101.8 - 115.1]	109.5 [101.7 - 114.6]	109.4 [101.2 - 117.8]
BG StDev (lognormal) (mg/dL):	20.2 [19.8 - 20.4]	20.1 [19.7 - 20.5]	20.2 [19.8 - 20.4]	20.2 [19.7 - 20.6]
Time to < 140 mg/dL [IQR] (hours):	0.0 [0.0 - 3.2]	0.0 [0.0 - 3.2]	0.0 [0.0 - 3.2]	0.0 [0.0 - 3.2]
% patients to < 140 mg/dL [IQR]:	100 [100 - 100]	100 [100 - 100]	100 [100 - 100]	100 [100 - 100]
Time to < 110 mg/dL [IQR] (hours):	5.0 [1.8 - 11.2]	5.0 [2.0 - 11.2]	5.0 [1.8 - 11.2]	5.0 [1.8 - 11.2]
% patients to < 110 mg/dL [IQR]:	100 [100 - 100]	100 [100 - 100]	100 [100 - 100]	100 [100 - 100]
Median insulin rate [IQR] (U/hr):	3.0 [3.0 - 3.0]	3.0 [3.0 - 3.0]	3.0 [3.0 - 3.0]	3.0 [3.0 - 3.0]
Median dextrose rate [IQR] (g/hr):	3.9 [3.3 - 5.2]	3.9 [3.3 - 5.2]	3.9 [3.1 - 5.2]	3.9 [3.3 - 5.2]

Figures 4.4-4.6 illustrate the affects of using 3, 12.5 or 25 gram dextrose boluses to prevent the onset of predicted hypoglycaemia. The shaded band in the upper plot represents the target BG band (4.4-6.1mmol/L). Filtered virtual SG measurements are shown as red crosses and the underlying modelled BG data is show by the solid red line. The interventions are identified by vertical red bars in the lower sub-plot of each figure, labelled PN dextrose.

The effect of each intervention size on glycaemia can be very different, which highlights the importance of sizing the dextrose bolus appropriately. Figure 4.5 shows the response to a 12.5 gram dextrose bolus when a patient is approaching hypoglycaemia. The intervention successfully lifted BG to a level just above the target band, preventing hypoglycaemia. Two additional dextrose boluses were deemed necessary by the intervention algorithm, but these occurred within 15 minutes of a SPRINT BG measurement, so they were not administered. At hour 19, the SPRINT protocol administered no insulin, which was enough to raise BG back into the desired band without the need for the dextrose bolus.

Figure 4.4 shows that the 3 gram bolus had little effect and a second dose was given shortly after the first bolus. Later in the monitoring period, there were further interventions given and in each case, the 3 gram bolus was not enough to lift BG into the desired band. Conversely, Figure 4.6 shows that the 25 gram bolus was too large, causing overcorrection of low BG and a loss of control. Following a 25 gram treatment, BG rises significantly and the TGC protocol overcompensates with insulin at the next 1-2 hourly intervention, causing a sudden drop. Hence, an oversized dextrose intervention can increase BG variability and increase hypoglycaemia due to the response of the controller.

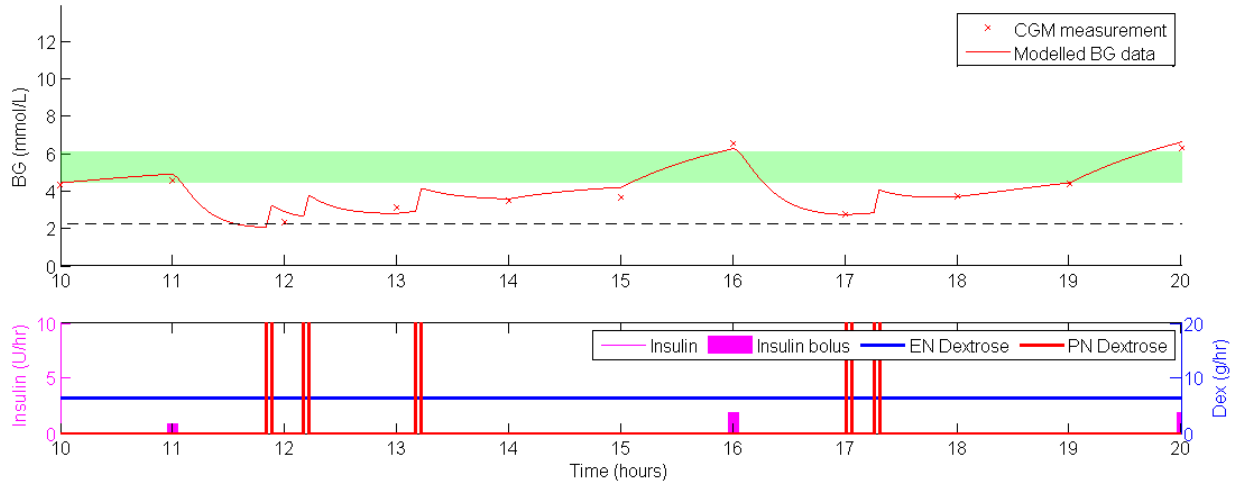


Figure 4.4 The effect on BG of a 3 grams intervention of glucose at the predicted onset of hypoglycaemia.

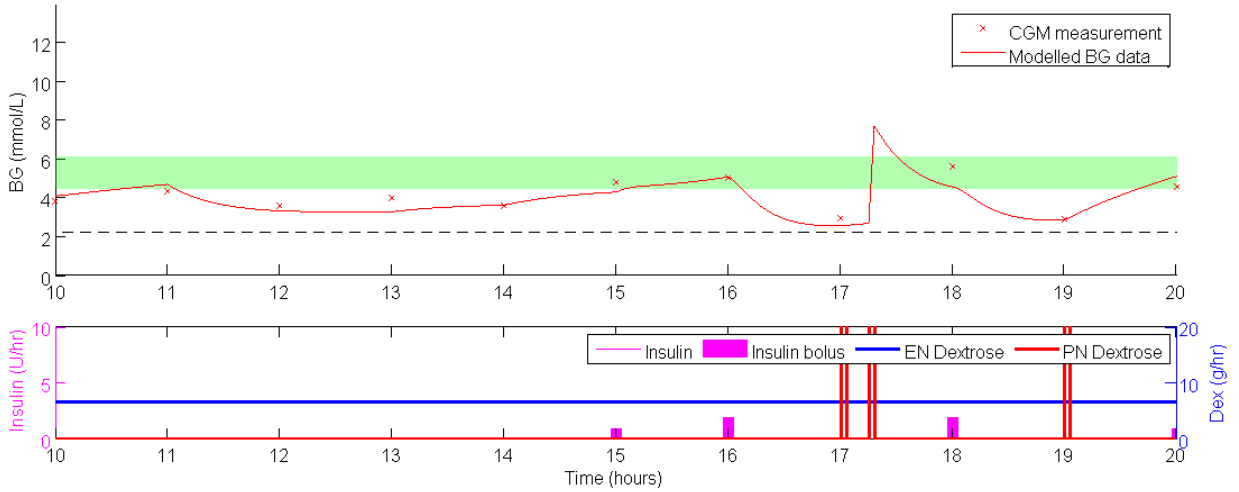


Figure 4.5 The effect on BG of a 12.5 grams intervention of glucose at the predicted onset of hypoglycaemia

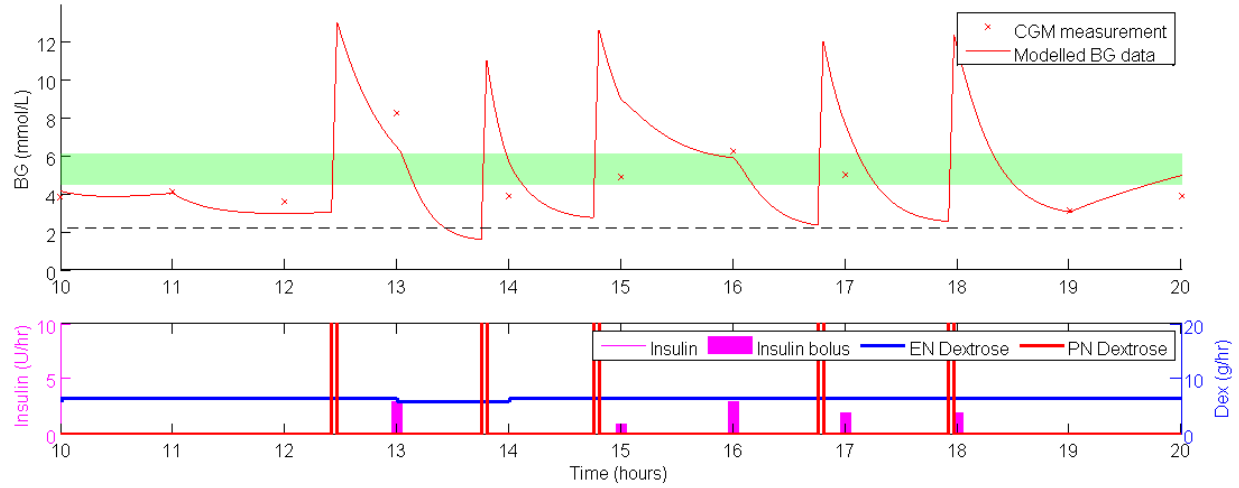


Figure 4.6 The effect on BG of a 25 grams intervention of glucose at the predicted onset of hypoglycaemia

4.4 Discussion

4.4.1 Performance

The aims of this in-silico study were to show that nursing workload could be reduced, the level of glycaemic control maintained, and early hypoglycaemic detection implemented using CGM sensors in the ICU, despite potentially high levels of sensor noise in individual measurements. The performance of the CGM sensors in this role was evaluated in-silico using MC simulations with stochastic CGM error models and ICU patient data. The in-silico model has been extensively clinically validated on independent patient data (Chase et al. 2010).

There was no clinically significant reduction in the BG control metrics with the addition of SG error, as seen in Table 4.2. The ability to maintain good BG control suggest that the additional error introduced by the CGM sensors has little effect, at least in the context of the SPRINT protocol. This outcome may be because SPRINT is model-derived and the inputs/outputs are discretised, so small sensor errors do not generally have a large effect on the insulin and nutrition rates selected by the protocol.

The number of patients experiencing hypoglycaemia and the percentage of total BG measurements < 2.2mmol/L increased with the addition of SG error. The increase in hypoglycaemia shown in Table 4.2 could be due to added SG error causing an underlying normoglycaemic level to falsely appear low, or equally, SG error adversely affecting the control protocol. Regardless of the cause, the increase in hypoglycaemia might be considered clinically unacceptable. However, clinically, if CGM devices were to be used in the hospital setting to guide therapy, they would need to be trusted and acted upon. Thus, this study investigated also using the SG data to trigger intravenous dextrose boluses to prevent hypoglycaemia.

The results in Table 4.3 show that intervening with a dextrose bolus at the onset of potential hypoglycaemia had little effect on overall BG control. The percent BG between 4.4mmol/L and 6.1mmol/L for all four intervention sizes were ~45%. However, the number of patients experiencing hypoglycaemia varied

substantially depending on the size of the intervention, showing the importance of sizing this rescue intervention appropriately.

The 3 gram dextrose intervention had little or no effect on the number of patients experiencing hypoglycaemia. Figure 4.4 shows a representative example of the effects of a 3 gram dextrose bolus, given at the onset of predicted hypoglycaemia. The initial dextrose bolus failed to restore BG adequately, and 15-20 minutes later a second intervention was given. Furthermore, approximately 1 hour later a third intervention was administered, suggesting the 3 gram intervention is too small. Multiple undersized interventions within a short time frame can lead to an unnecessary increase to nurse workload and clinical burden.

The number of patients experiencing hypoglycaemia reduced as the intervention size increased, to a limit of 12.5 grams. Figure 4.5 shows a representative response in BG when a 12.5 g dextrose bolus is given at the predicted onset of hypoglycaemia. The intervention causes a rise in BG to ~7.5mmol/L, preventing hypoglycaemia and restoring BG to the desired band. Interestingly, in this example the following BG at hour 18 has a large positive SG error. When this BG is input to the SPRINT controller, an insulin bolus is prescribed and the BG continues to fall up until hour 19. This outcome is not a reflection on the size of intervention, it simply reinforces the importance of designing the controller appropriately if SG measurements are to be used.

Interestingly, when the intervention is increased to 25 grams the results show a substantial increase in the number of patients experiencing hypoglycaemia. A representative example of a 25 gram intervention is shown in Figure 4.6. The intervention at 12.5 hours causes the BG level to rise to over 12mmol/L. At hour 13, a BG of ~8mmol/L is entered into the SPRINT controller and an insulin bolus is administered. This extra insulin, when the patient is likely already quite sensitive or still has a lot of insulin on-board, causes BG to drop to below 2.2mmol/L and a second dextrose intervention is given. This potentially counter-intuitive trend of rebound hypoglycaemia continues to occur over the following

hours as the controller and intervention combination struggle to control BG. This indicates that the size of the bolus needs to be carefully selected, maybe on a patient-specific or protocol-specific basis, or extra information should be provided to the control protocol so it can factor in the effect of the bolus at the next intervention.

A final significant result is the potential reduction in clinical burden with the use of CGM device. Typically, SPRINT requires 16 measurements per day on average at 1.5 - 4 minutes per measurement for all tasks (Lin et al. 2008). Assuming 4 calibration measurements per day with CGM, the total measurement burden could be reduced by 75%. Thus, nurses could potentially gain 18-48 minutes per patient per day with the use of CGM, which is clinically significant in a busy ICU environment. It is possible that some of this reduction in clinical burden may be offset by hypoglycaemia alarms and interventions. However, enhanced alarm algorithms and constantly evolving CGM technology may resolve or reduce this issue.

4.4.2 Limitations

Although the virtual patient simulation methods used here are clinically well validated (Chase et al. 2007; Chase et al. 2010; Lonergan et al. 2006), results from this in-silico study may differ from actual clinical results. In this study, the BG sequences used to drive SPRINT were model-based, not real SG output data. However, this analysis provides a strong initial proof of concept and the results justify a pilot clinical investigation. Equally, the cross validation study using independent data for this model in Chase et al. (2010) shows these errors would likely be small.

The CGM error models used in this study may not completely describe SG behaviour in critically ill patients. The Gaussian error model was based on ICU data, however the model was generated using sensor error statistics and the time-series information was not available. Thus, SG errors were assumed to be independent. The autoregressive CGM error model is more complete in terms of modelling the interdependence of errors and the delay in glucose diffusion from

blood to the sensor, but the model was generated using data from people with type 1 diabetes. Thus, this autoregressive model may not necessarily provide the best model of CGM behaviour in ICU patients. It is anticipated that true SG error characteristics in critically ill patients fall somewhere between these simulated models. However, to date, there is no publically available SG-BG paired data for ICU patients to create better models.

Calibration drift due to sensor degradation or bio-fouling over time was not considered in this study. Without correction, calibration drift is characterised by a glucose bias that increases over time. This drift could potentially reduce control and increase the rate of false positives/negatives in hypoglycaemia alarms. However, the magnitude of the bias due to drift is often a function of the frequency and quality of calibration measurements, which can be controlled more readily in a critical care setting. This issue has also not been quantified, to date, in the literature for CGM devices in the critical care setting.

4.5 Summary

This chapter has analysed in-silico the use of CGM sensors with simple filters and hypoglycaemia alarms to provide input to the SPRINT TGC protocol. The main results and conclusions include:

- CGM devices, even with significant error, had little impact on TGC performance and time in a euglycaemic range was between 45-50%, compared to 50% in the baseline case.
- Threshold hypoglycaemia alarms have the potential to reduce the rate of hypoglycaemia to zero with minimal false positives.
- The glucose bolus intervention size is critical. This study found ~12.5g to be relatively optimal and that levels of 25-40g can counter-intuitively increase hypoglycaemia. Thus, this intervention is also control protocol specific.
- The use of CGM devices with SPRINT should reduce average nurse burden for BG measurements by up to 75%, potentially saving 18-48 minutes per patient, per day, which is clinically significant in the very busy ICU.

All of these results justify clinical testing for validation and highlight some of the main issues in using CGM devices for TGC in critical care.

Chapter 5. Analysing CGM trend accuracy

The feasibility study in the previous chapter showed that CGM devices have the potential to significantly improve safety and BG control when using TGC protocols with critically ill patients. Most notably, trend information provided by CGM allows algorithms to alarm at the onset of predicted hypoglycaemia, which is not practical or useful with intermittent BG monitoring. With trend information available from the CGM, new metrics and methods of analyses are required to determine the level of trend accuracy. This chapter describes a new tool designed to assess trend accuracy of SG data, both for use in control, as well as for (primarily) assessing device performance.

5.1 Introduction

Unlike traditional self monitoring BG devices, which offer a 'snapshot' of glucose concentration at the time of testing, CGM devices give additional information about the rate of change of BG over short periods of time (minutes). This information is particularly useful for revealing and predicting abnormal glycaemia, such as hypoglycaemia or hyperglycaemia, and deciding on the appropriate course of treatment (Beck et al. 2009; Bequette 2010; Hoeks et al. 2011a; Pretty et al. 2010; Signal et al. 2012a). However, to make good treatment decisions it is important to have good trend accuracy, not just good point-to-point accuracy.

Trend accuracy refers to the ability of a CGM device to accurately capture the true 'shape' of glycaemia over time, whereas, point-to-point accuracy assesses the discrepancy between a CGM and reference BG measurement at a single point in time. One important area where trend information is used is closed loop glycaemic control, where CGM devices are coupled with insulin pumps and an appropriate control algorithm to provide automatic glycaemic control. Several pilot studies have investigated closed loop control in people with diabetes (Breton et al. 2012; Clarke et al. 2009; Weinzimer et al. 2008). However, these methods are still being developed and it is not used as a standard therapy.

Another area where trend accuracy is particularly important is hypoglycaemia alarms, which often inherently use trends to predict the onset of hypoglycaemia (Bequette 2010; Dassau et al. 2010; Eren-Oruklu et al. 2010; McGarraugh 2010; Pretty et al. 2010). In this case, poor trend accuracy can result in a high rate of false alarms, or worse, missed hypoglycaemic events.

In these applications, trend accuracy is particularly important because even though good trend accuracy doesn't guarantee success, poor trend accuracy is likely to cause failure. As trend dependent applications/features, such as closed loop control or hypoglycaemic alarms, become more common in CGM devices the need for good trend accuracy increases.

Many users of CGM devices are likely to be unaware of the level of trend accuracy of their particular device. Furthermore, studies in the literature that use CGM devices or investigate CGM performance often report point-to-point accuracy, but rarely quantify trend accuracy (Corstjens et al. 2006; Djakoure-Platonoff et al. 2003). This lack of clarity could be because there are many methods or metrics available for assessing point-to-point accuracy, such as mean absolute difference (MAD), mean absolute relative difference (MARD), the Bland-Altman plot (Bland & Altman 1999), and Clarke error grid (Clarke 2005). However, there are very few metrics to assess trend accuracy (Kovatchev et al. 2004), particularly independent of BG level and BG accuracy.

One method that does assess CGM trend accuracy is the continuous glucose error grid analysis (CG-EGA) (Kovatchev et al. 2004). CG-EGA evaluates the accuracy of continuous glucose monitoring sensors in terms of both point-to-point accuracy and trend (rate) accuracy. Results from the CG-EGA are presented in a table, showing the proportion of paired BG/SG measurements that fall into clinically acceptable, unacceptable and benign zones. While the results produced by CG-EGA have been reported to be difficult to interpret (Wentholt et al. 2006), the method certainly represents a step in the right direction in terms of assessing both aspects of sensor accuracy.

Hence, there is a major need for independent trend metrics as increasing numbers of CGM devices make their way into the market. Both regulatory bodies and end users need to be confident that CGM devices have good trend accuracy, as well as good point-to-point accuracy, especially if they feature predictive hypoglycaemic alarms based on trends. The aim of this study was to develop a metric or tool that could quantify trend accuracy and present the results in an intuitive plot that is easy to interpret for any user. This tool is intended to be used in conjunction with traditional point-to-point accuracy methods to provide a fuller assessment of sensor accuracy and clinical utility. This chapter describes that tool: the Trend Compass.

5.2 Subjects and Methods

This chapter introduces a novel trend metric that can be used to assess the trend accuracy of SG measurements from a continuous glucose monitoring device, with reference to BG reference measurements determined using a gold standard measurement device, such as a YSI chemistry analyser.

5.2.1 Quantifying trend

Trend accuracy can be defined as the level of agreement between the rates-of-change of two independent devices measuring a single time series, over the same time period. An effective and simple way to quantify trend accuracy is derived from the geometric interpretation of the dot product. The dot product assesses the similarity of 2 vectors *A* and *B*:

$$A \cdot B = \|A\|\|B\|cos\theta \qquad 5.1$$

Where $A = [a_1 , a_2]$ represents two measurements from a BG reference and $B = [b_2, b_2]$ represent two SG measurements at the same time points. Rearranging Equation 5.1 to make θ the subject gives a normalized measure of similarity between *A* and *B*:

$$\theta = cos^{-1}\left(\frac{A \cdot B}{\|A\|\|B\|}\right) \qquad 5.2$$

The output of Equation 5.2 provides the angle between the two vectors, *A* and *B*, where a smaller angle is indicative of better trend accuracy. Thus, Equation 5.2 can be used with clinical data to quantify the level of trend accuracy between paired sets of BG/SG measurements, independent of the point-to-point measurement accuracy which is often referred to simply as error.

The value of θ is dependent on the time interval between BG/SG samples, which should be held constant when comparing different pairs of vectors. This study uses a 1 hour time interval between consecutive samples of BG and SG. More frequent sampling, such as 15 minutes, can be analysed with the Trend Compass by using a 1 hour window, sliding at 15 minute increments. The sensitivity of the Trend Compass to timing errors in the sampling frequency has not been investigated as this chapter was written to present the overall method, which can be refined by consensus or future studies in due course.

5.2.2 Trend Compass plot

Overall trend accuracy from Equation 5.2 can be conveyed visually using the Trend Compass shown in Figure 5.1. A polar coordinate system is used. The angular coordinate depicts the trend accuracy (θ degrees from top or bottom vertical) and the radial coordinate shows the reference BG level (See *Appendix A* for a step-by-step guide to using the Trend Compass). Trend accuracy is plotted against reference BG level to show how it changes over the range of glucose values, because very good trend accuracy is more crucial during hypoglycaemia or hyperglycaemia where important treatment choices are potentially affected. For example, a mismatch in trend at 8mmol/L would likely lead to less severe complications than the same mismatch in trend at 3.2mmol/L.

The top hemisphere of the Trend Compass shows trend accuracy when the reference BG rate of change is ≥ 0 (BG is rising - examples F,A,B in Figure 5.2) and the bottom hemisphere shows trend accuracy when the reference BG rate of

change is < 0 (BG is falling - examples C,D,E in Figure 5.2). Furthermore, the hemispheres are divided into two quadrants, which each give information about the relative rate-of-change between the reference BG and the SG. For example 'B' in Figure 5.2 shows a BG change from 7mmol/L to 8.2mmol/L and an SG change from 7mmol/L to 7.5mmol/L, so the top-right quadrant is used. Alternatively, if SG is rising more steeply than BG like 'F' in Figure 5.2, then the top-left quadrant is used. Note examples 'A' and 'D' in Figure 5.2 have perfect trend accuracy, even though there is a significant offset between SG and BG in 'A', so they are plotted on the vertical line between quadrants. Importantly, these examples are shown to reinforce that this Trend Compass assesses trend accuracy independent of point-to-point error.

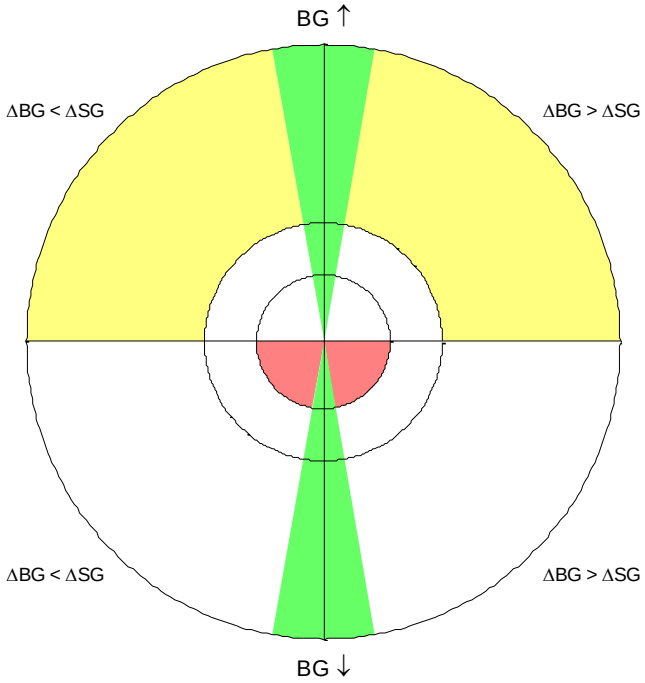


Figure 5.1 The Trend Compass, used to assess the trend accuracy of a set of measurements relative to a corresponding reference set of measurements. Green zones show areas of good trending, and yellow and red zones show areas of moderate to severe clinical risk, respectively

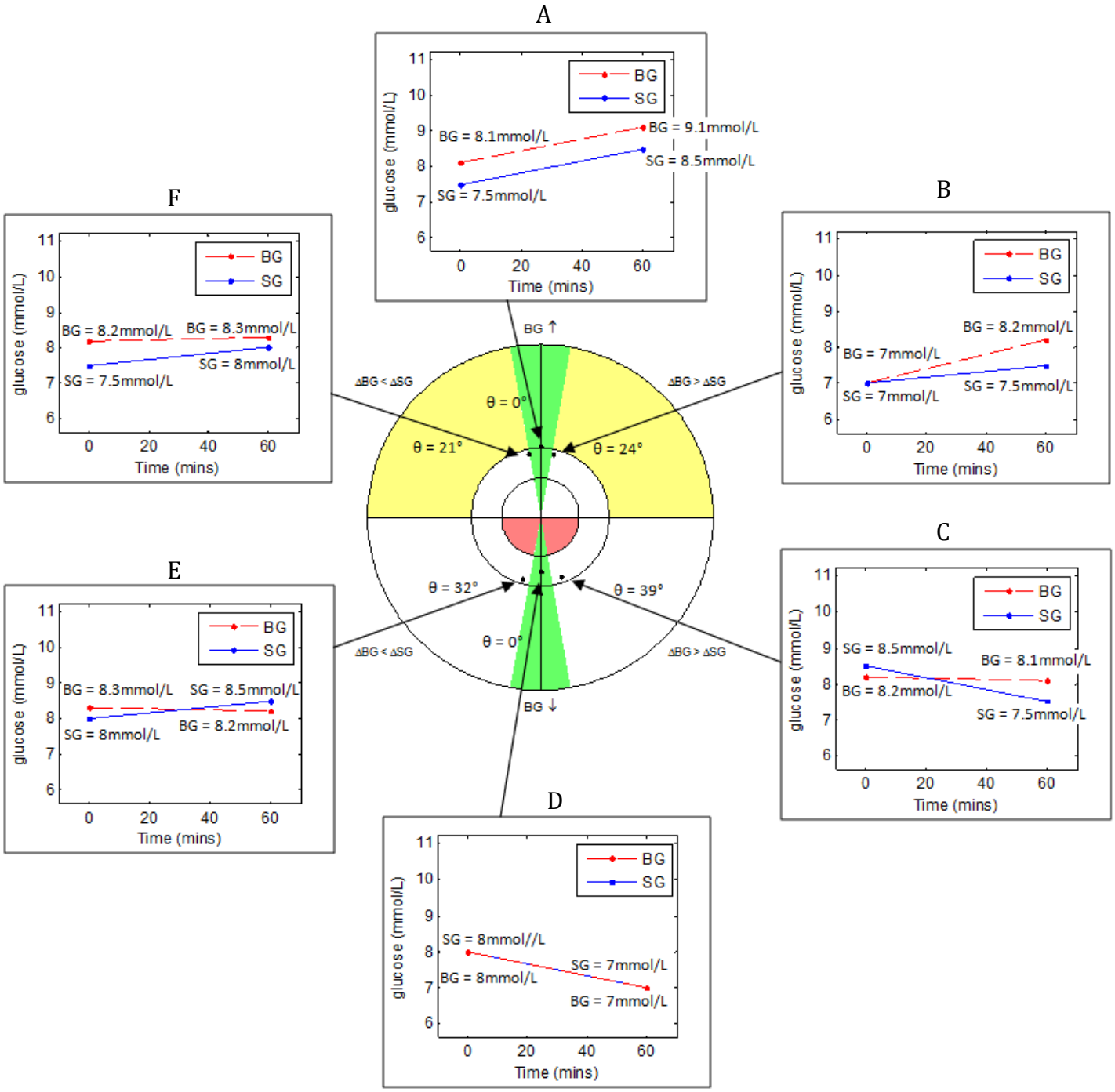


Figure 5.2 Six examples of SG and BG paired measurements with their corresponding point on the Trend Compass. Note: comparing 'A' to 'D' shows that the constant bias has no effect on how trending is displayed on the Trend Compass (Both examples have perfect trend accuracy so ϑ=0°).

In addition to separating the Trend Compass into four quadrants, two green zones around the vertical axis were added to show 'good' trend accuracy. To present the method, the size of the green zones were set at ±10° on the plot,

which captured mismatches in trend of up to 20° (see note at the bottom of Appendix A). The size of the green zones was set with conservative acceptability in mind to present the method and may be changed by future users as desired, so long as it is held constant when comparing the trend accuracy of multiple devices. A few survey inputs from physicians suggest these limits are reasonable, although this survey was not comprehensively done and a large survey might be required for consensus on proper zone settings.

In the radial direction, the Trend Compass has been separated into three zones to reflect the clinically significant glycaemic zones: 1) hypoglycaemia; 2) euglycaemia; and 3) hyperglycaemia. The boundaries presented in this paper are 0 to 5mmol/L for hypoglycaemia, 5 to 8.9mmol/L for normoglycaemia, and greater than 8.9mmol/L for hyperglycaemia. These zones are similar to what is widely accepted and published, but, again, may be changed by the user as desired.

Finally, four regions of the Trend Compass are coloured to highlight clinically significant zones where trend accuracy is most important. The yellow regions show areas where reference BG is above 8.9mmol/L and rising with poor trend accuracy. Hence, moderate caution should be applied. The red regions highlight areas where the consequences of poor trending could be far more significant, such as when reference BG is below 5mmol/L and falling. In both cases, treatment decisions based on poor trending in SG data could increase the risk of adverse outcomes.

5.2.3 Accompanying numerical trend metrics

The Trend Compass was intended to be a visual tool that is fast and easy to interpret. The use of vector agreement as the basis of the Trend Compass is a novel means for assessing trend in this context, and it allows direct, objective numerical comparison between devices. For this reason, a simple evaluation table can also be created for direct analysis, comparison and/or regulatory processes. Table 5.1 represents a simple choice to present the concept and it could easily be augmented as desired for analysis or regulatory purposes.

Table 5.1 A table of metrics to accompany the Trend Compass plot.

Overall trend accuracy				
Percent in green				
Percent in yellow				
Percent in red				
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green				
Percent outside green				
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green				
Percent outside green				

Furthermore, analogous to MAD, a numerical metric frequently used to quantify point-to-point accuracy (Clarke & Kovatchev 2007; Klonoff 2005a; Kovatchev et al. 2008) the user could present trend accuracy using a Trend Index (TI), defined:

$$TI = \frac{1}{n}\sum_{i=1}^{n}|\theta(i)| \qquad 5.3$$

TI describes the average overall trend accuracy and a lower *TI* is indicative of better global trend accuracy.

5.2.4 Simulated data

To validate the Trend Compass in-silico, artificial SG and BG data sets were created in MATLAB. A glucose trace was created using a random walk model and normally distributed error was added to give hourly paired measurements. The paired measurement sets are used to illustrate the use of the Trend Compass. The data sets simulated four typical (limit case) scenarios that might be encountered during real-world use:

- Low glucose variability patient with low sensor error
- Low glucose variability patient with high sensor error
- High glucose variability patient with low sensor error
- High glucose variability patient with high sensor error

5.2.5 Clinical data

Guardian Real-Time (Medtronic; Northridge, CA) Continuous glucose monitoring data and YSI 2300 reference BG measurements from 2 patients were used to show the Trend Compass in use with clinical data. Each patient was monitored for ~3 days, during which time the SG was recorded every 5 minutes and BG was determined approximately every 60 minutes. BG measurements were paired with the SG measurement that was sampled closest to the time of BG sampling. Overall, the median [IQR] sampling interval between BG measurements was 60 [55 - 62] minutes. Finally, this data was used to show the independence of this trend metric to point-to-point error.

5.3 Results

5.3.1 Simulated data

The Trend Compass was first tested using simulated paired SG and BG measurements, sampled at 1 hour intervals. Figures 5.3-5.6 and Tables 5.2-5.5 show the trend accuracy results for the four simulated data sets described in 5.2.4. In all four figures, the blue line represents simulated SG data and the red circles represent BG data.

Figure 5.3 shows a low glucose variability patient, with low sensor error and the corresponding Trend Compass plot. The Trend Compass plot shows very good trend accuracy, with most of the points lying close to the vertical lines at the cardinal North (N) and South (S) position (TI = 11.3°). In the radial direction, the Trend Compass plot depicts the patient as a low glucose variability patient, as all of the points are contained within the normoglycaemic band. Table 5.2 also shows good trend accuracy results for this patient with 91.3% of measurements falling within the green zones.

Figure 5.4 shows a patient with the same glucose trace characteristics as in Figure 5.3, but with a higher level of sensor error. The increase in error has resulted in a Trend Compass plot with visibly more points outside the green

zones (TI = 28.1°). This result is also reflected in Table 5.3, which reports 39.1% of points in the green zones and 8.7% of points in the red zone.

Figure 5.5 shows an example of a patient with high glucose variability, coupled with low sensor error. The Trend Compass plot for this patient appears similar to Figure 5.3 with majority of points near the N and S positions, and a similar TI of 11.1°. However, in the radial direction the points are far more spread out due to the large range of glucose values in the data set. Table 5.4 shows trend accuracy to be good, with 82.6% of data points in the green zones and 0% in the yellow or red zones.

Finally, Figure 5.6 shows an example patient with high glucose variability and high sensor error. The resulting Trend Compass shows a reduction in trend accuracy when compared to Figure 5.5, with a much wider angular spread of results (TI = 27.0°). Table 5.5 shows that this data set has 52% in the green zones, 8.7% in the yellow zones and 4.2% in the red zone.

5.3.2 Clinical data

Using clinical SG data with paired BG measurements, the Trend Compass can quantify the level of trend accuracy allowing sensor performance to be evaluated and compared. The solid blue line in Figure 5.7 shows CGM sensor data and the red circles represent BG data from the same patient. Overall, the trend accuracy is very good and ~70% of the points lie in the green areas. Furthermore, only 11.6% of points are in the yellow zone and only 1.4% in the red zone. The TI for this data set is 18.2° and this value for TI is potentially slightly skewed by a few outlier BG data points seen in the trace. Figure 5.8 uses the same data set as Figure 5.7, but with a 4mmol/L constant bias applied to CGM sensor data. The Trend Compass plot, performance metric table and TI are remain unchanged with the offset SG data, further illustrating the independence of this method from point-to-point accuracy, which is unique to this approach.

Figure 5.9 shows the SG data from Figures 5.7 and 5.8, coupled with BG measurements from a different data set. The trend accuracy is expected to be

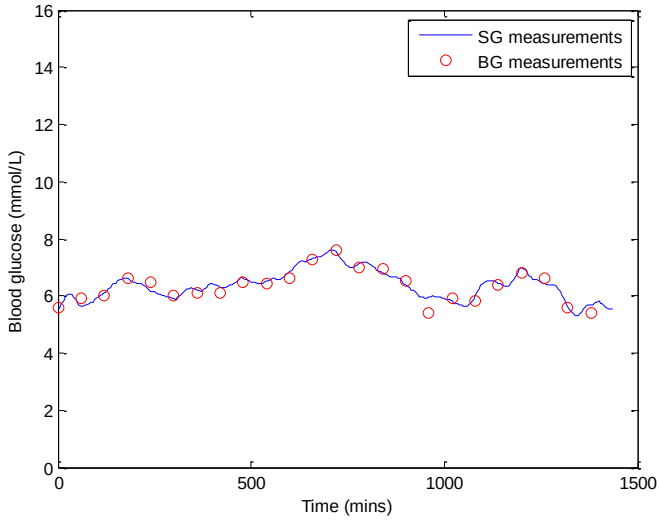
marginal because the SG/BG are sampled from different individuals and are independent. The Trend Compass shows a wide spread of points indicating poor trend accuracy and this is reinforced by the TI of 37. 2°. Sections A and C in Figure 5.9 show periods of good trending and Section B shows a period of poor trend accuracy. The trend metrics shown in Table 5.8 report 34% of points are in the green zones, 13.6% in yellow and 3% in the red zones.

5.4 Discussion

The aim of this study was to develop a novel tool that could quantify trend accuracy, independent of point-to-point accuracy. The results present an intuitive plot that gives a quick visual assessment of relative CGM trend accuracy, and allows detailed quantified results for in-depth comparison. The Trend Compass is described in this manuscript with reference to SG data from a continuous glucose monitoring system that is compared to paired BG measurements from a reference method.

With the introduction of CGM devices, trend accuracy has become very important due to increased investigation of closed loop glycaemic control and the increased use of hypo/hyperglycaemia alarm algorithms, which all inherently use trend patterns. The Trend Compass was not designed to replace conventional accuracy metrics, such as MARD or the Bland Altman plot. In fact, it is intended to be used in conjunction with traditional measures of point-to-point sensor accuracy.

Using error metrics alongside the Trend Compass gives the user much more useful information about the overall performance of a sensor. Equally, as an objective measurement of trend accuracy, the Trend Compass could potentially be useful for regulatory bodies when assessing sensor performance prior to approval. Thus, it is an added metric not an overlapping one, which is a result of its design that it is independent of point-to-point accuracy.



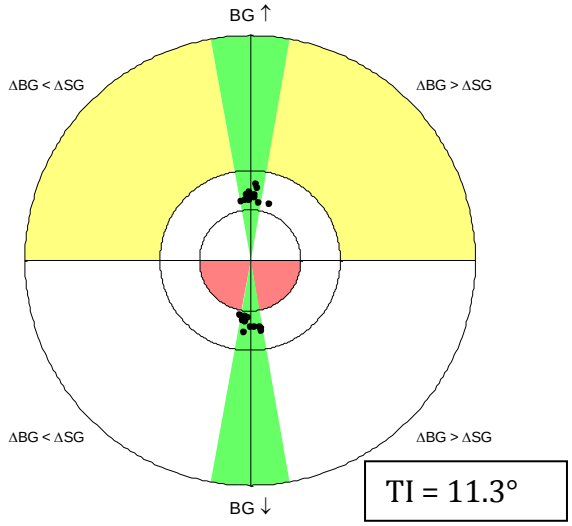
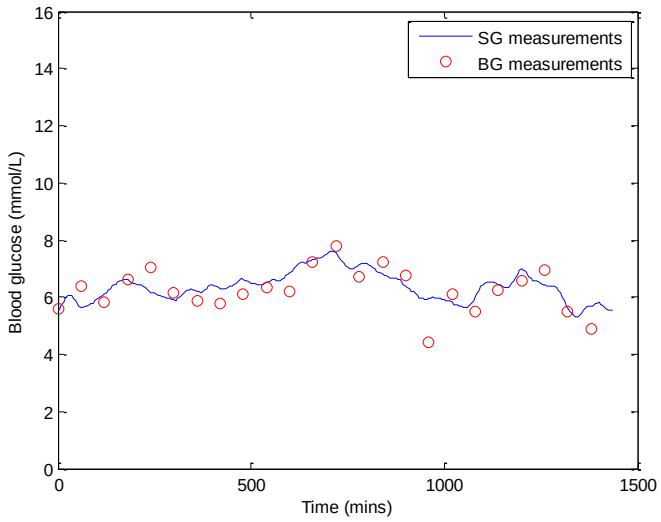


Figure 5.3 (left) BG and SG measurements for a stable patient with low sensor error. (right) Trend Compass plot for this data set with TI metric.

Table 5.2 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	91.3			
Percent in yellow	0			
Percent in red	0			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	47.8	0	47.8
Percent outside green	0	4.3	0	4.3
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	43.5	0	43.5
Percent outside green	0	4.3	0	4.3



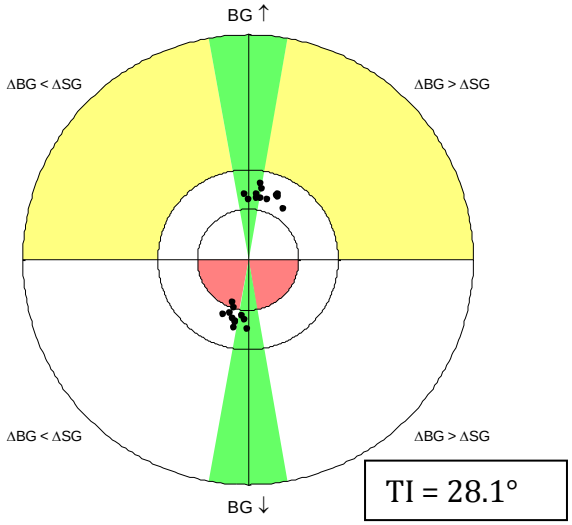
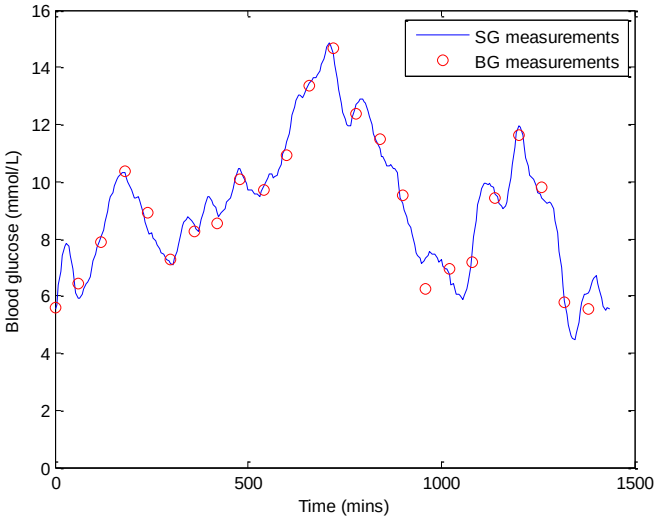


Figure 5.4 (left) BG and SG measurements for a stable patient with high sensor error. (right) Trend Compass plot for this data set with TI metric.

Table 5.3 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	39.1			
Percent in yellow	0			
Percent in red	8.7			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	26.1	0	26.1
Percent outside green	0	26.1	0	26.1
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	13	0	13
Percent outside green	8.7	26.1	0	34.8



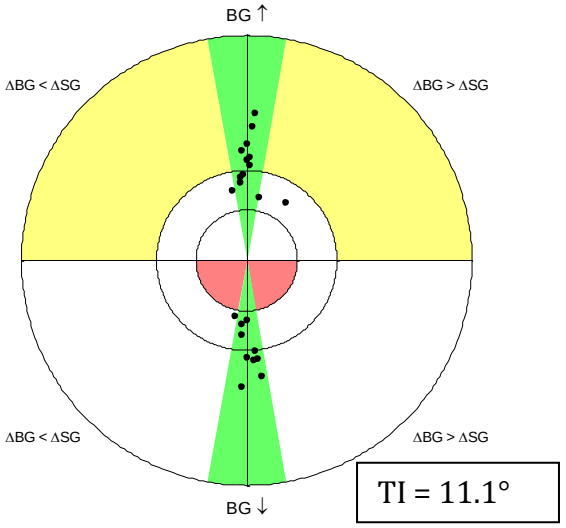
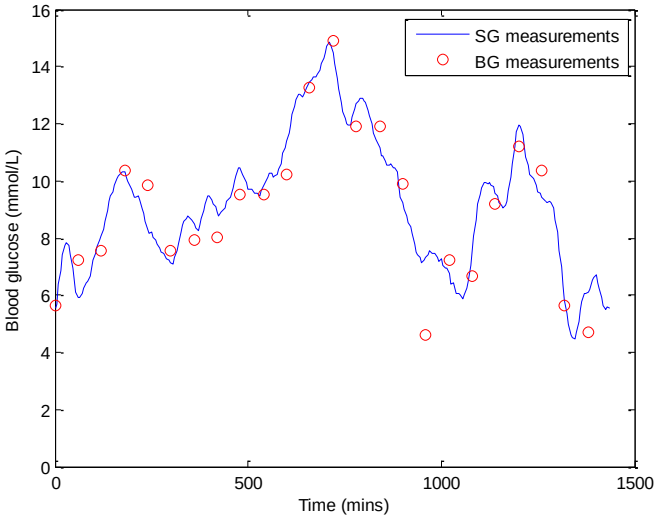


Figure 5.5 (left) BG and SG measurements for a variable patient with low sensor error. (right) Trend Compass plot for this data set with TI metric.

Table 5.4 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	82.6			
Percent in yellow	0			
Percent in red	0			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	13	30.4	43.5
Percent outside green	0	13	0	13
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	13	26.1	39.1
Percent outside green	0	4.3	0	4.3



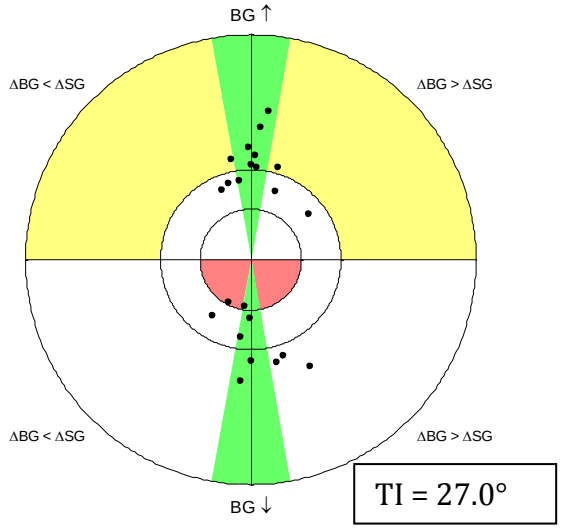
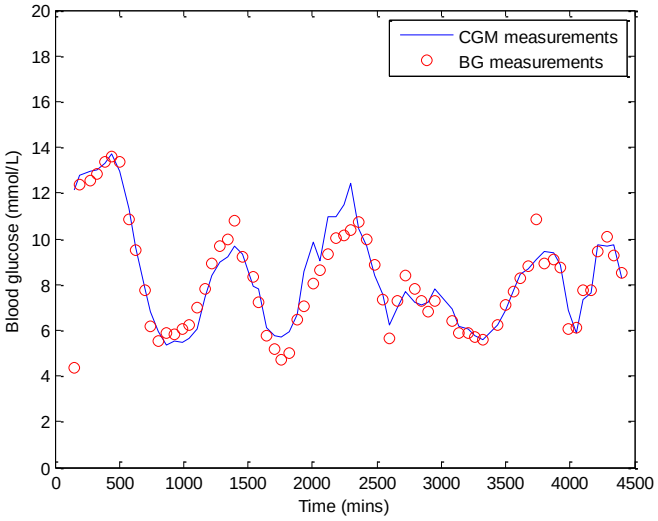


Figure 5.6 (left) BG and SG measurements for a variable patient with high sensor error. (right) Trend Compass plot for this data set with TI metric.

Table 5.5 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	52.2			
Percent in yellow	8.7			
Percent in red	4.3			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	4.3	26.1	30.4
Percent outside green	0	17.4	8.7	26.1
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	4.3	8.7	8.7	21.7
Percent outside green	4.3	4.3	13	21.7



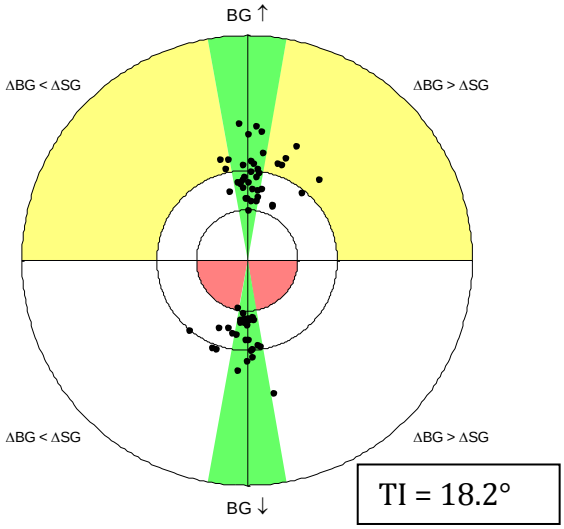
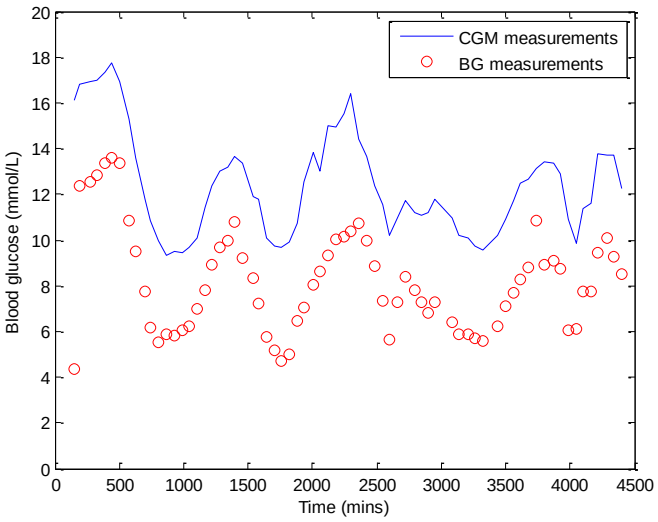


Figure 5.7 (left) Clinical CGM data and BG measurements from the same subject. (right) Trend Compass plot for this data set with TI metric.

Table 5.6 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	69.6			
Percent in yellow	11.6			
Percent in red	1.4			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	24.6	14.5	39.1
Percent outside green	0	7.2	11.6	18.8
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	26.1	4.3	30.4
Percent outside green	1.4	4.3	5.8	11.6



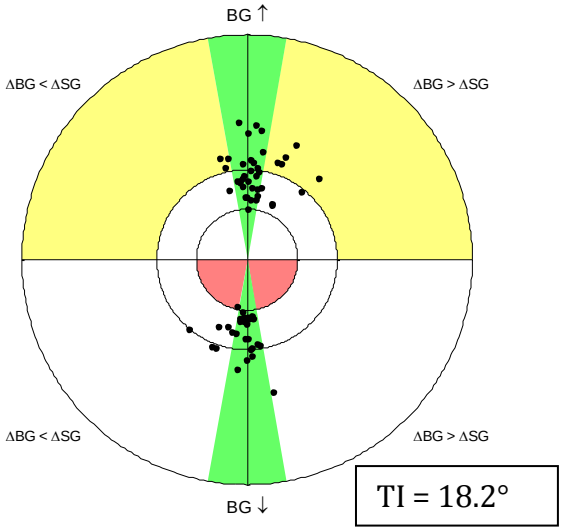
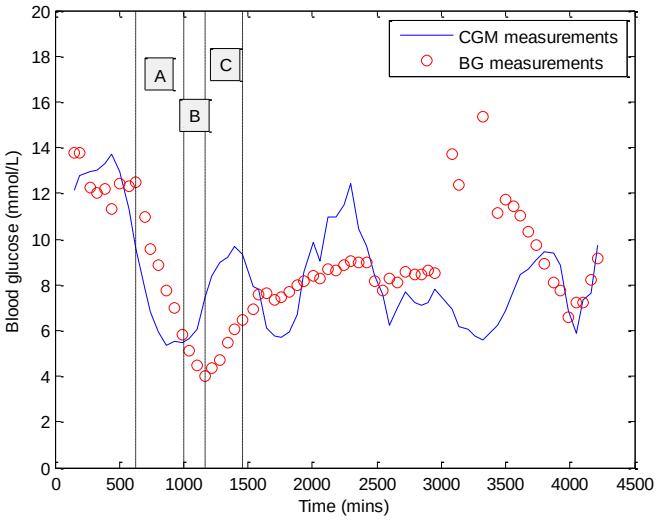


Figure 5.8 (left) Clinical CGM data with a 4mmol/L bias and BG measurements from the same subject. (right) Trend Compass plot for this data set with TI metric.

Table 5.7 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	69.6			
Percent in yellow	11.6			
Percent in red	1.4			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	24.6	14.5	39.1
Percent outside green	0	7.2	11.6	18.8
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	26.1	4.3	30.4
Percent outside green	1.4	4.3	5.8	11.6



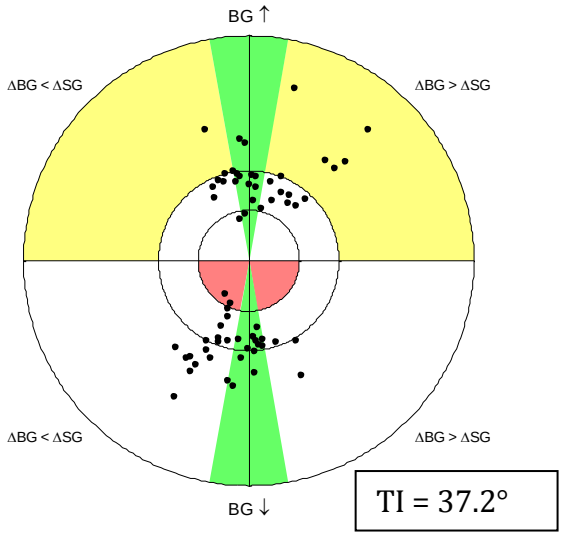


Figure 5.9 (left) Clinical CGM data and BG measurements from two different subjects. (right) Trend Compass plot for this data set with TI metric.

Table 5.8 Performance table showing trend accuracy for this data set

Overall trend accuracy				
Percent in green	34.8			
Percent in yellow	13.6			
Percent in red	3			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	1.5	12.1	3	16.7
Percent outside green	1.5	18.2	13.6	33.3
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	0	13.6	4.5	18.2
Percent outside green	3	10.6	18.2	31.8

The results for the simulated data show how the Trend Compass can effectively differentiate between good trend accuracy and poor trend accuracy. Figures 5.3 and 5.4 assess the trend accuracy for a stable, with low glucose variability patient with different levels of sensor noise. Comparing the plot of SG-BG data for each patient it is difficult to determine which data has better trend accuracy, although it is obvious that the data in Figure 5.3 has a lower sensor error. It is important that the Trend Compass is able to differentiate between the trend accuracy of the two devices in a robust way. In this case with a simulated stable patient the Trend Compass clearly shows that the data in Figure 5.3 has better trend accuracy. This outcome extends to Figures 5.5 and 5.6, which show different sensor error levels for the same high glucose variability simulated patient. Again, the Trend compass is clearly able to show which sensor has better trend accuracy, in this case the data plotted in Figure 5.5.

Another aspect that needed to be robust is the impact of patient variability. Comparing Figure 5.3 to Figure 5.5, and Figure 5.4 to Figure 5.6, it is clear that the patient variability doesn't impair the ability of the Trend Compass to reliably assess trend accuracy. This aspect is very important as different patients or cohorts can have very different glycaemic dynamics, so the assessment of trend accuracy must be robust to these differences. Figures 5.3 and 5.5 show two different patient dynamics, but with similar levels of sensor error. The Trend Compass effectively conveys that in both cases the trend accuracy is very good. This is further reinforced with the accompanying performance table and TI metric. Figures 5.4 and 5.6 also show two different patient dynamics, but this time for a higher level of sensor error. Again, the Trend Compass is consistent in showing both data sets with moderate to poor trend accuracy.

When using the Trend Compass with clinical data the usefulness of the method is immediately clear, as shown by comparing Figures 5.7, 5.8 and 5.9. Figure 5.7, which contains well correlated data collected from one patient, shows good trending compared to Figure 5.9, which contains uncorrelated data collected from two different individuals. The discrepancies between the trend accuracy of the two data sets can be easily interpreted from visual inspection of the Trend Compass plots alone. Interestingly, the Trend Compass for the uncorrelated data set still has ~35% of points in the green zones. This result is likely due to the sections marked 'A' and 'C' in Figure 5.9, which both show relatively good trending between SG and BG by chance.

Figures 5.7 and 5.8 show how the Trend Compass can classify trend accuracy independent of point-to-point measurement error. The blue SG trace in Figure 5.8 is the same data as shown in Figure 5.7, but with a positive 4mmol/L offset. This offset significantly increases the point-to-point error, but the Trend Compass remains unchanged. This lack of change occurs because the relative slope between SG and BG has not changed with the offset, and that relative slope is the fundamental mechanism used to quantify trend with this method. This example further reinforces the intended use of the Trend Compass to assess solely trend accuracy in conjunction with traditional point-to-point

measurement error metrics, creating a more complete assessment of sensor performance.

The importance of assessing trend accuracy as a function of BG level is made clear by the paired BG-SG measurements in section 'B' in Figure 5.9. The trend accuracy in section B falls within the red zone of the Trend Compass, because the BG is falling while the SG is reporting a rise in glucose at a substantially different rate. The implications of a drop in glucose being reported as a rise by a CGM device could be very dangerous, potentially leading to missed treatment of hypoglycaemia. Furthermore, alarm algorithms that use trend information may not alert the user at the onset of these events.

5.5 Summary

The Trend Compass is a tool that can quantify trend accuracy between two devices measuring a single time series, such as a CGM device and a reference BG. It is robust when used with different patient cohorts (different dynamics), as well as different levels of sensor error. The resulting trend accuracy is easily interpreted on the Trend Compass plot, and if required, accompanying performance table and TI metric. Importantly, it assesses trend accuracy independent of BG level and point-to-point accuracy. It is possible for a device to have very poor point-to-point accuracy, but excellent trend accuracy. Thus, assessing trend accuracy is as important as assessing point-to-point measurement error as CGM devices become more widely used. A tool such as the Trend Compass provides an easy to interpret, reliable method to do so.

Chapter 6. Pilot clinical trial with CGM devices

This Chapter presents the results and findings of an observational pilot study using standard off-the-shelf CGM devices in the Christchurch Hospital ICU. The goal of the trial was to assess the reliability of CGM in critically ill patients and to determine whether sensor site or CGM device type could have a significant effect on SG data. Overall performance reliability results are presented and discussed along with several interesting observations, which are presented as case studies.

6.1 Introduction

The in-silico feasibility study in Chapter 4 showed that CGM devices offered several potential benefits over conventional intermittent BG monitoring in the ICU. The study showed that CGM devices coupled with a well designed TGC protocol and hypoglycaemia intervention protocol, have the potential to improve BG control, reduce/eliminate hypoglycaemia and reduce nurse burden. However, these findings were based on simulation results using modelled CGM behaviour, not real clinical SG data.

Prior to implementing any new device as a clinical practice change, a pilot study should be completed to assess the device's true behaviour in its intended clinical environment. This necessity is particularly important for CGM devices in the ICU, where in the limited literature available, there is some debate as to whether CGM devices perform well enough to guide clinical therapy (Brunner et al. 2011; Rabiee et al. 2009; Vlkova et al. 2009). Furthermore, the types of patients admitted, severity of illnesses encountered, and, treatments and therapies used can all vary between critical care centres, reinforcing the importance of testing these devices in the particular ICU that intends to use them.

This chapter presents results from an observational pilot study using CGM devices in Christchurch Hospital ICU. Specifically, the reliability of CGM devices in critically ill patients, and the impact of both sensor calibration and sensor

location were tested. The overall goal was to better understand the variability induced by these factors, and their potential clinical impact in use for TGC.

6.2 Subjects and Methods

6.2.1 Subjects

This study uses data from an observational pilot study of CGM in patients admitted to the Christchurch Hospital ICU. This analysis uses CGM and BG data from 10 patients who were recruited into the study. All patients were recruited by a physician in the ICU and informed written consent was obtained from the next of kin if the patient was unable to consent. Inclusion criteria were two consecutive BG measurements greater than 8mmol/L, indicating the need for insulin therapy using the STAR protocol (Evans et al. 2012). Exclusion criteria were an anticipated ICU admission period of less than 3 days. This study and use of data was approved by the Upper South A Regional Ethics Committee, New Zealand. Table 6.1 shows the patient demographics.

Table 6.1 Patient demographics

Patients	10
Age (years)	51 [39 - 64]
Sex (M/F)	5/5
APACHE II	24 [17 - 27]
APACHE III	85 [52 - 99]
SAPS II	52 [30 - 59]
ICU admission (days)	20 [10 - 33]
Outcome (L/D)	6/4
Diabetes (None/T1/T2)	10/0/0

6.2.2 Continuous Glucose Monitoring

Each participant in the study was monitored concurrently using 3 independent CGM devices for a period of up to 6 days. All CGM devices and sensors used in this study were un-modified, off the shelf devices, designed for people with diabetes outside of the hospital. Thus, these CGM devices and sensors were being assessed off label. Two different CGM devices were tested in this study. However, both devices used the same sensor hardware technology. Figure 6.1 shows the

devices, the Medtronic Guardian Real-Time CGM device and the Medtronic iPro2 CGM device (Medtronic Diabetes, Northridge, CA).

Figure 6.1 (Left) Medtronic Guardian Real-Time CGM device. (Right) Medtronic iPro2 CGM device.

The components that make up each model of CGM are slightly different. The Guardian device uses a wireless transmitter to transfer data to the monitor, where it can be viewed in real time, and, the iPro2 simply stores data internally for download at a later date. The Medtronic Enlite sensor (Medtronic Diabetes, Northridge, CA) used with each device is shown in Figure 6.2.

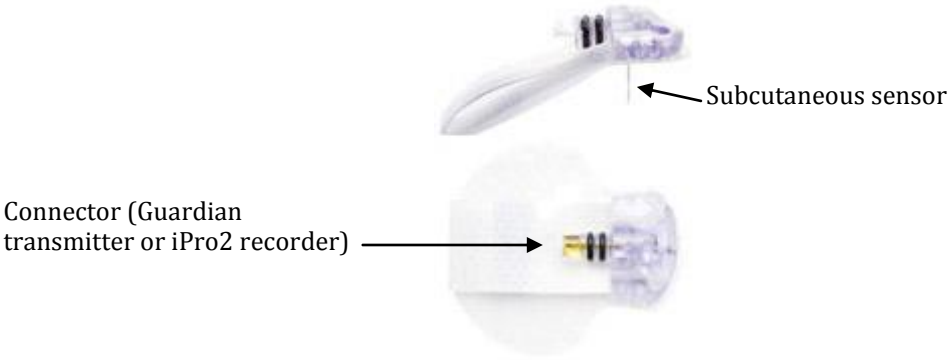


Figure 6.2 Side and top view of the Medtronic Enlite subcutaneous glucose sensor, which is compatible with both the Guardian Real-Time and the iPro2 CGM devices.

For each patient, two sensors were located on the abdomen, one of which was connected to a Guardian Real-Time CGM device and the other connected to an iPro2 CGM device (*both devices:* Medtronic Diabetes, Northridge, CA). A third sensor was located on the patient's thigh and was connected to a second iPro2

recorder. This configuration allowed comparison of results across different devices and sensor locations within each subject. Thus, enabling inter-site and inter-device variability to be quantified

One significant difference between the two CGM devices is the calibration algorithm. Calibration algorithms use independent calibration BG measurements to convert the raw sensor current (I_{SIG}) into a series of SG values for the user. Here, all else equal, different calibration and processing can yield different results.

The iPro2 devices store the sensor signal information internally and these measurements are retrospectively calibrated after the device is removed. Retrospective calibration allows the calibration algorithm to use information both before and after the time point of interest to obtain an optimal calibration to each reference point. In contrast, the Guardian CGM device displays a glucose value in real time and the calibration algorithm can only use prior data for calibration, but it thus enables real-time glycaemic management. Thus, both devices see the data differently when they process it. However, for real time control only a real time device is feasible.

Calibration BG measurements were obtained by specifically trained ICU nurses at least 3 times per day as recommended by the device manufacturer (Minimed 2006). A blood sample was drawn from the patient's arterial line and a BGA (ABL90 FLEX, Radiometer, Copenhagen) was used to determine the glucose concentration. The value from the BGA was immediately entered into the Guardian Real-Time device and then recorded for retrospective calibration of the iPro2 devices.

6.2.3 Intermittent BG monitoring

In addition to BG measurements used for calibration of SG data, each patient had intermittent BG monitoring every few hours. The STAR protocol requires, on average, 12-14 BG measurements per day to guide insulin/nutrition therapy (Fisk et al. 2012). In this study, BG measurements were determined using Super

Glucocard II (Arkray, Japan) glucose meters, by the ICU nurse working at the bedside. Several patients had additional BG measurements determined using Nova Statstrip (Nova Biomedical, Waltham, MA, USA) and/or Roche Accu-chek Inform II (F. Hoffmann-La Roche Ltd, Basle, Switzerland) hospital grade glucose meters, which both actively measured and adjusted for haematocrit level. Occasionally, a single blood sample was distributed across multiple BG meters to assess measurement precision and in those cases the median value was used as the 'true' BG. All meter BG measurements collected were distinctly separate from BG measurements used for CGM device calibration, providing an independent comparator for SG data.

6.2.4 Analysis

CGM data were stratified into 3 subsets by CGM device type and sensor location to allow comparison between the three combinations: abdomen Guardian, abdomen iPro2 and thigh iPro2. Overall accuracy of SG data in each subset was quantified using MARD and Bland Altman plots were produced to show how SG errors were associated with glucose level. Cumulative distribution functions (CDFs) were used to show the overall inter-site and inter-device discrepancies in SG data.

Trend agreement between SG data sets was assessed using TI, which is described in detail in Chapter 5. The TI was calculated for the two abdomen SG data sets to assess inter-device trending, and the two iPro2 SG data sets to assess inter-site trending. However, in this study the true trend accuracy of each sensor relative to a reference set of independent BG measurements could not be assessed due to clinical constraints on the frequency of manual BG measurements. Thus, the trend agreement between two concurrent SG data sets is presented in its place.

Finally, 3 sets of BG, SG and trend data are presented as case studies, showing some of the interesting aspects of CGM behaviour in critically ill patients. These case studies highlight unanticipated aspects of CGM behaviour that potentially need further investigation before CGM devices are permanently implemented in the ICU clinical environment.

6.3 Results

6.3.1 Overall cohort

The overall results from the analysis of BG and SG data are shown in Table 6.2. The BG results show that intermittent BG measurements were taken frequently in this study and the median time between consecutive measurements was 1.5 hours. The median [IQR] BG levels were 6.9 [6.2 - 7.6] mmol/L. Overall, these results show that STAR protocol successfully controlled BG to a normal level.

Table 6.2 Blood glucose and SG data results

Blood glucose results			
Number of patients	10		
Time between BG (hours)	1.5 [0.9 - 2.3]		
Median [IQR] BG (mmol/L)	6.9 [6.2 - 7.6]		
CGM results	*Guardian - Abdomen*	*iPro2 - Abdomen*	*iPro2 - Thigh*
Number SG Data sets	10	10	10
Duration of CGM (days)	4.8 [3.0 - 6.0]	4.8 [2.8 - 6.0]	5.3 [3.0 - 6.0]
Time between calibration (hours)	7.5 [5.1 - 8.2]	7.5 [3.6 - 9.0]	6.3 [3.0 - 8.1]
Time between reference BG (hours)	1.8 [1.0 - 2.8]	1.7 [1.0 - 2.7]	1.8 [1.0 - 2.8]
Median [IQR] SG (mmol/L)	6.9 [5.9 - 8.1]	6.7 [6.0 - 7.4]	6.7 [6.1 - 7.3]
MARD (%)	24.0	11.8	12.4
CGM Trend results			
Inter-device Trend Index (degrees)*	29.0 [20.1 - 34.4]		
Inter-site Trend Index (degrees)*		19.4 [17.8 - 23.4]	

*See Chapter 5 for a thorough description of Trend Index

The lower section of Table 6.2 show results for each combination of CGM device and sensor location assessed in this study. All of these subsets have good CGM duration, with most data sets containing more than 3 days of data. For a majority of patients, SG data was calibrated at least every 8 hours and between calibration BG measurements, reference BG measurements were taken every ~1.8 hours. Overall, the median [IQR] results reported by SG data are very similar to those results reported by BG data. Assessing the overall accuracy SG data, the MARD for the Guardian device in the abdomen was 24%, compared to ~12% for the two iPro2 data sets. The trend results show that trending between the two iPro2 devices located in different sites was better than trending between the iPro2 and Guardian Real-Time devices, both located in the abdomen. The median TI for

inter-site agreement of SG data was 19.4 degrees and the median TI for inter-device agreement was 29 degrees.

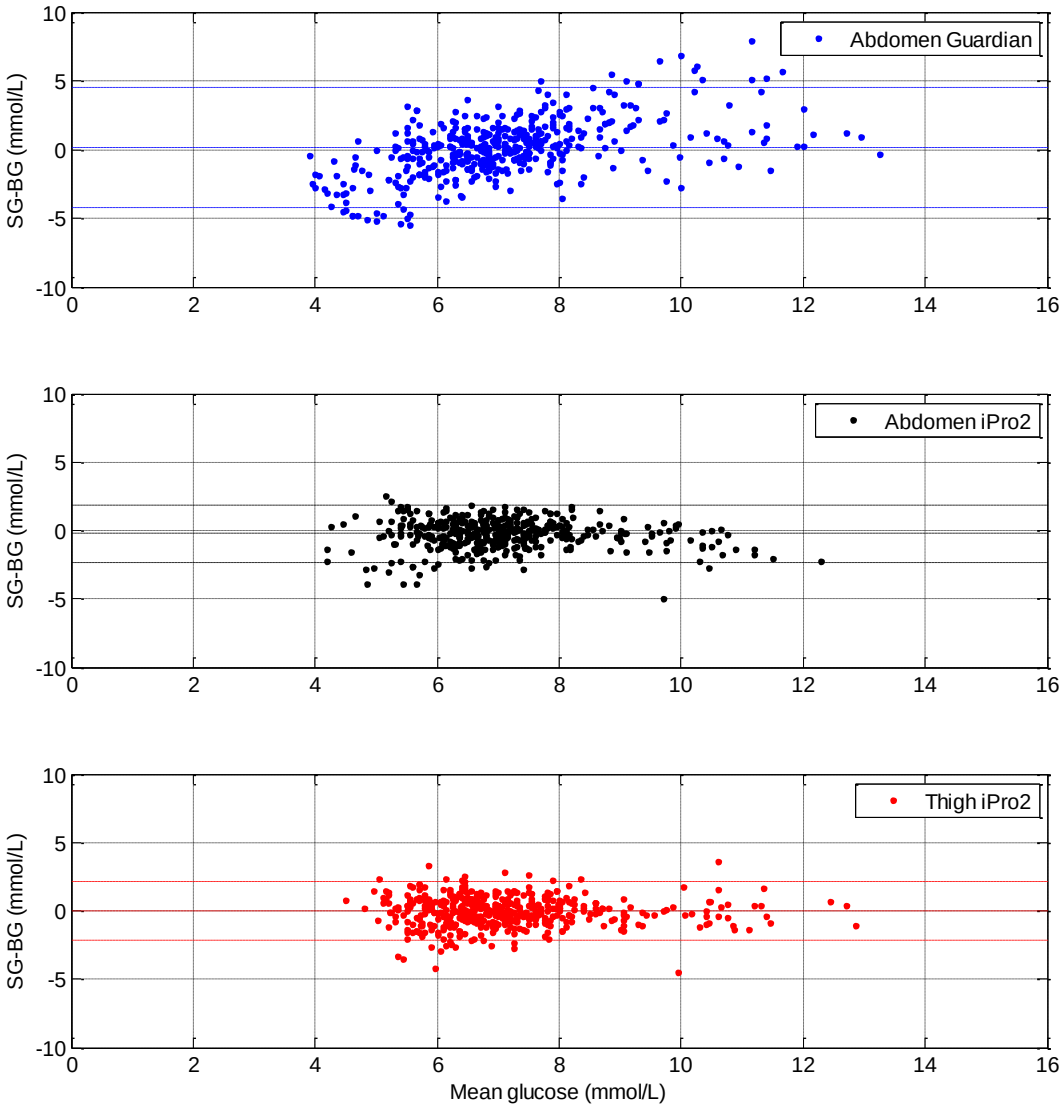


Figure 6.3 Bland Altman plot for the three different CGM device and sensor location combinations

The Bland Altman plots in Figure 6.3 show how SG error changes with glucose level. The top plot shows data from the abdomen Guardian CGM device. The overall mean error is only 0.2mmol/L, but the 95% confidence bounds are at -4.2 and 4.4mmol/L, suggesting error can be relatively large for this device when monitoring critically ill patients. At lower BG levels, the Guardian CGM had a tendency to read low and at high BG levels it had a tendency to read high, as shown by the positive slope in the scattered data. The middle and bottom Bland

Altman plots show data from the abdomen iPro2 and thigh iPro2 CGM devices, respectively. Both of these SG data sets have an overall mean error close to zero and the 95% confidence intervals are much tighter than that of the abdomen Guardian, where the difference is thus between real time and retrospective calibration. There appears to be little or no association between SG error and glycaemic level in either iPro2 data set.

Figure 6.4 shows two cumulative distribution functions (CDFs), one for inter-site discrepancies in SG data and one for inter-device discrepancies in SG data. At the time of every 5-minute SG measurement, inter-site discrepancy was calculated as thigh iPro2 SG - abdomen iPro2 SG and inter-device discrepancy was calculated as abdomen Guardian SG - abdomen iPro2 SG. The inter-site CDF is steep and narrow suggesting good agreement between the two CGM devices, irrespective of sensor location. Conversely, the inter-device CDF is flatter and wider, suggesting the type of CGM device, or calibration method, has a larger impact on SG data. The 5th to 95th percentile interval for inter-site is 3.2mmol/L, compared to nearly double that value at 6.1mmol/L for inter-device comparison, reinforcing that CGM device and/or calibration type has a substantially larger impact.

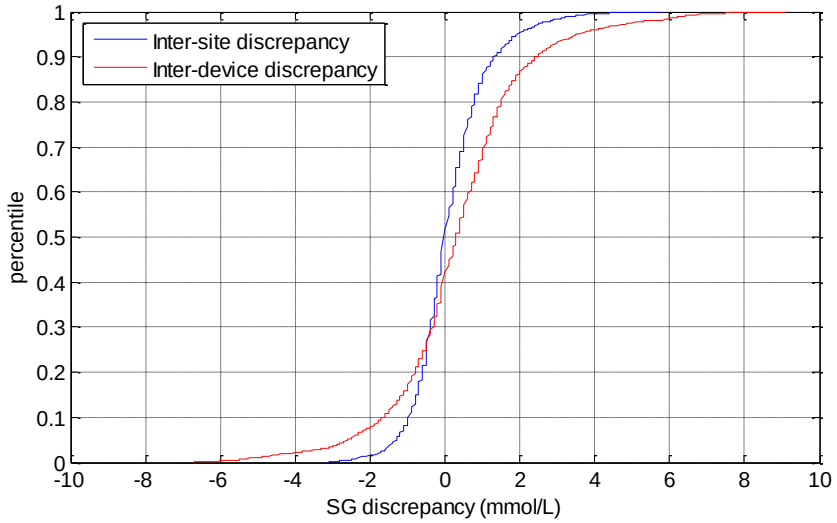


Figure 6.4 Inter-site and inter-device discrepancies between SG data. Inter-site discrepancies were calculated as thigh iPro2 SG - abdomen iPro2 SG and inter-device discrepancies were calculated as abdomen Guardian SG - abdomen iPro2 SG

6.3.2 Case studies

Due to the combinations of CGM devices and sensor locations, this study had the unique opportunity to observe several interesting characteristics of CGM behaviour in critically ill patients. This section presents 3 of those observations as individual case studies. First, Figure 6.5 shows an example of 3 CGM devices working very well in the critical care setting. During the first ~24 hours of monitoring there was some mismatch between each devices SG data, but for the remainder of monitoring the SG traces were almost overlapped. This particular patient was an otherwise healthy spinal injury patient with little or no oedema and no signs of sepsis. The CGM devices tracked glycaemic trends well and the Guardian Real-Time device would have provided useful data at the bedside if nurses were able to use the data.

Figure 6.6 shows a Trend Compass plot, assessing the trend agreement between the two iPro2 devices monitoring this patient. Overall the trend agreement between the devices is good, with the majority of points in the green wedges. This result is reinforced by the numerical results in Table 6.3, which show that 63.1% of points fell in the green zones, and only 4.5% fell in the red zone. Furthermore, no points were in the yellow zone.

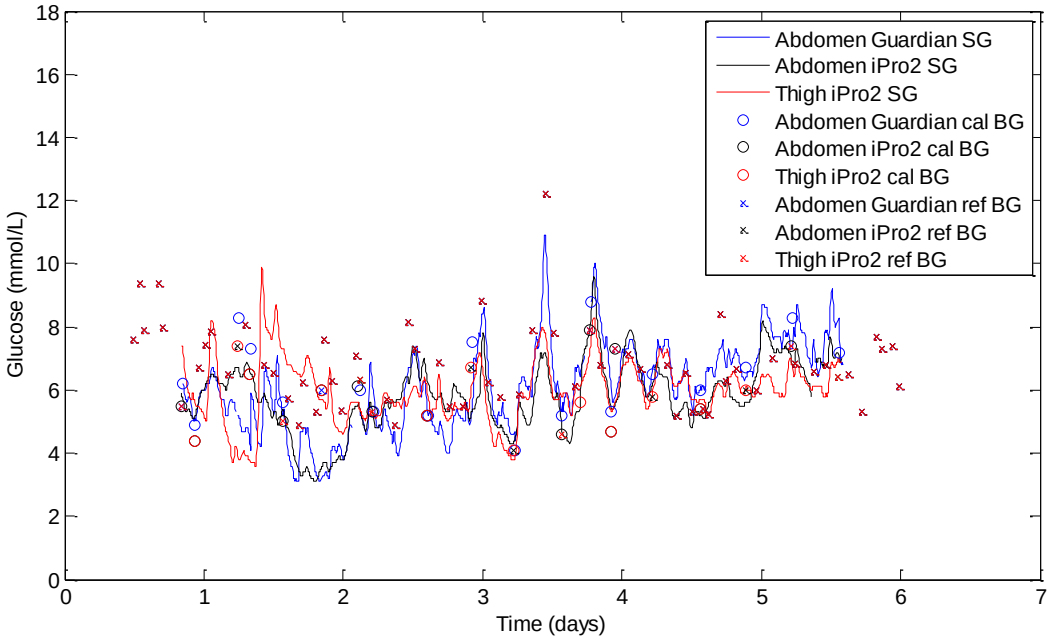


Figure 6.5 Three CGM devices monitoring a patient with good inter-device/site agreement

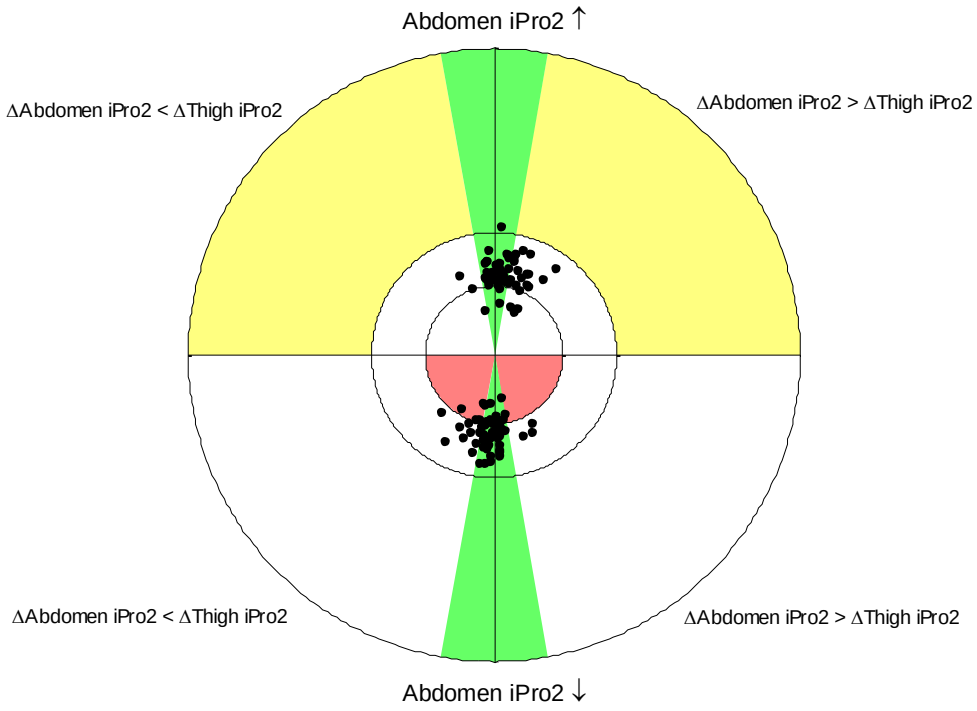


Figure 6.6 Trend Compass plot showing inter-site agreement between abdominal iPro2 SG data and Thigh iPro2 SG data

Table 6.3 Performance Table corresponding to the Trend Compass in Figure 6.6

Overall trend accuracy				
Percent in green	63.1			
Percent in yellow	0			
Percent in red	4.5			
When BG is rising	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	1.8	25.2	0.9	27.9
Percent outside green	3.6	17.1	0	20.7
When BG is falling	BG < 5mmol/L	5mmol/L < BG < 8.9mmol/L	BG > 8.9mmol/L	overall
Percent in green	9	26.1	0	35.1
Percent outside green	4.5	11.7	0	16.2

Figure 6.7 shows SG and BG data collected from a patient with severe oedema. This patient had an estimated 18 litres of additional fluid onboard during the first few days of monitoring, with most of it in the abdominal region. Due to the additional fluid, the simple process of inserting each sensor was made difficult and the first sensor inserted in the abdomen failed to adhere to the skin, due to fluid constantly seeping from the insertion site. This sensor was replaced and the

other two sensors, one in the abdomen and one in the thigh, were inserted successfully. However, after 2-3 hours, one of the abdominal sensors failed and had to be removed early. Thus, Figure 6.7 contains only two complete SG data sets. It should be noted that this patient was the only patient in the study to have a sensor adhesion failure.

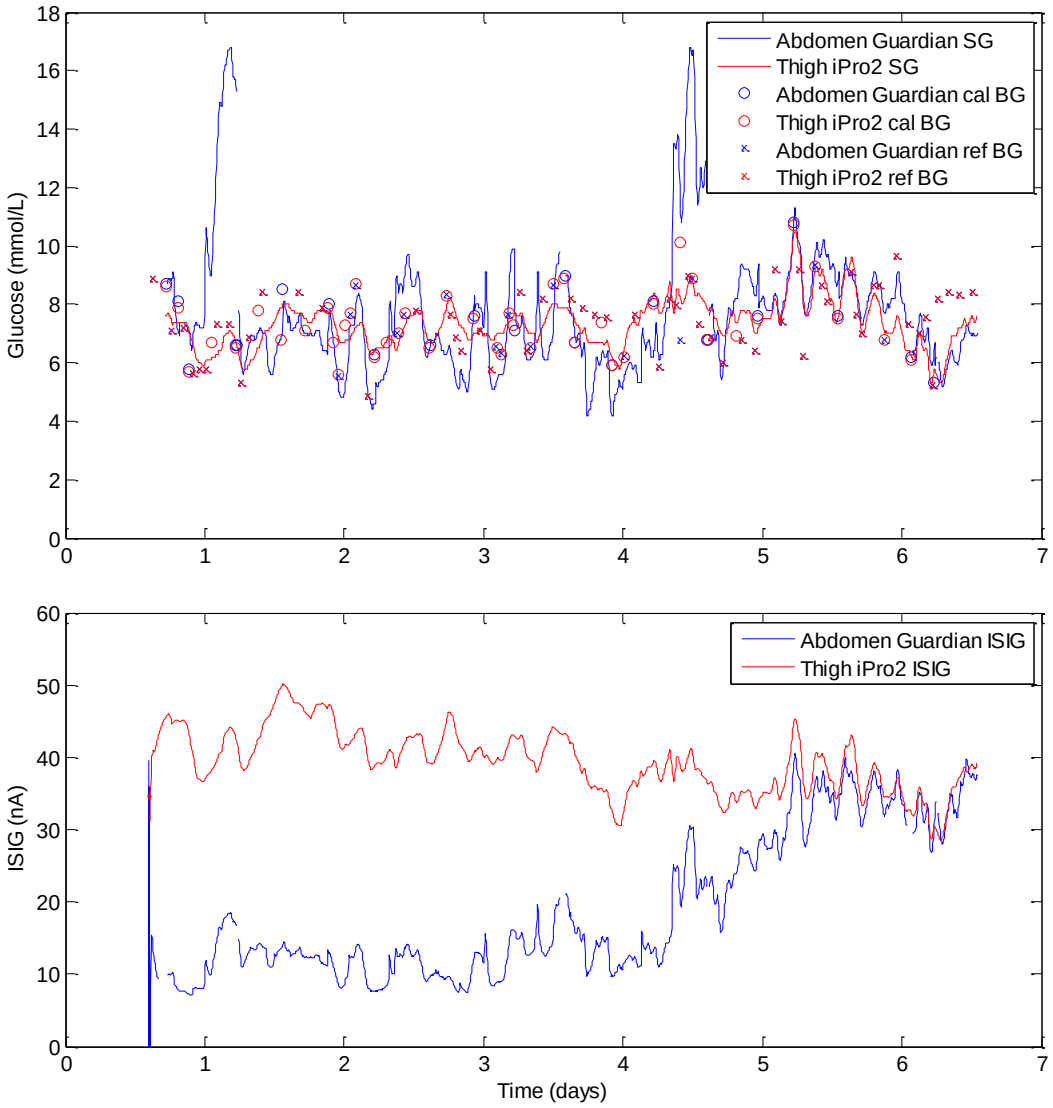


Figure 6.7 SG and I_{SIG} data from two CGM devices monitoring a patient with severe oedema

The top plot of Figure 6.7 shows SG and BG data collected throughout the monitoring period. Abdominal Guardian SG data is much more variable than the Thigh iPro2 SG data during days 1-4. After day 5, both SG traces reported similar trends in glycaemia. The bottom plot shows the raw sensor signal, or I_{SIG}, for each

CGM device over the monitoring period. The I_{SIG} data shows a clear separation in sensor sensitivity over the first 4 days of monitoring. The abdominal sensor current was typically in the region of 10-20nA, compared to the thigh sensor which was between 30-50nA. However, during day 5, the abdomen I_{SIG} rose to the level of the thigh I_{SIG} and for the remainder of the monitoring period both sensors reported similar dynamics in the region of 30-40nA.

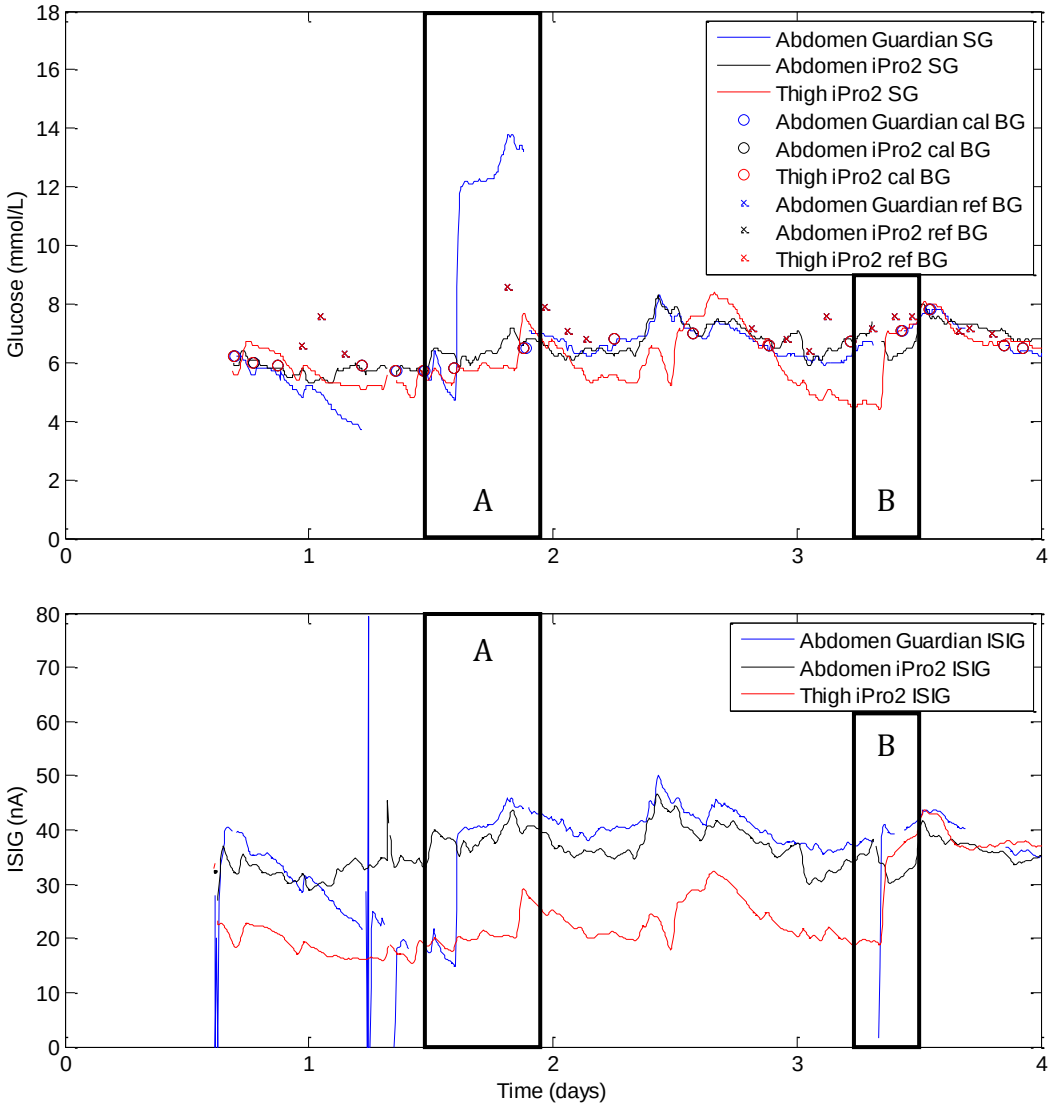


Figure 6.8 Example of un-physiological 'step' increases in I_{SIG} and their appearance in SG

Figure 6.8 shows two examples of large step increases in sensor current and how they appear in the SG data after calibration. Similar to Figure 6.7, the top plot shows BG and SG data, while the bottom plot shows I_{SIG} data. The box labelled 'A'

in Figure 6.8 illustrates a step increase in the abdominal Guardian I_{SIG} data from ~20nA to ~40nA. In this instance, the real-time calibration algorithm amplified the step causing the SG data to rise from 4.7mmol/L to above 12.2mmol/L. At around 1.9 days when the Guardian device was next calibrated, the algorithm detected the SG was too high and adjusted it to the correct level. The box labelled 'B' shows a similar rise in the thigh iPro2 I_{SIG} from ~20nA to ~35nA. However, the retrospectively calibrated thigh iPro2 SG only increased from 4.5mmol/L to 7mmol/L and no further adjustments were made at the next calibration at ~3.4 days.

6.4 Discussion

This chapter presents the results of a pilot study investigating the reliability of CGM, used off label, in critically ill patients. The results from study give significant insight into the performance of CGM devices in this cohort and highlight some important aspects that require further investigation, prior to clinical implementation

6.4.1 Overall cohort

The results in Table 6.2 show that glycaemia was monitored closely during the study, by intermittent BG measurements and CGM, with both methods producing similar overall glycaemic results. In terms of CGM accuracy, the MARD for the Guardian SG data was approximately twice the MARD of iPro2 SG data, irrespective of whether the iPro2's sensor was located in the abdomen or thigh. This result strongly suggests that the accuracy of the device, in terms of MARD, is dependent on device type, or, more likely, the resulting calibration algorithm.

The Guardian uses a real-time calibration algorithm that must adjust the 'calibration factor' using only prior data. Thus, any disturbances that might induce error at the time of calibration could result in substantial inaccuracies in SG data until the time of next calibration. Conversely, the iPro2 is calibrated retrospectively, so one would expect the same disturbances to have less impact on the overall accuracy, as future calibration BG values are known and used to

obtain the final SG trace. Overall, the MARD values presented are all in the region of reported MARD for CGM in outpatients, where the devices were designed to be used (Keenan et al. 2012; Kovatchev et al. 2008; Larson & Pinsker 2013), indicating that they otherwise performed as might be expected despite the different cohort.

Despite the relatively high MARD of the Guardian SG data, the clinical impact of those errors was determined to be minor. A Clarke Error Grid analysis of paired SG and reference BG data (results not shown) showed that 99.1% of the points fell in zones A and B, which would not lead to inappropriate treatment in cohorts for which this error grid was designed (Clarke 2005). Furthermore, no points fell in the clinically dangerous regions D and E. The same analysis of iPro2 SG data showed similar clinical results, with 100% of points in zones A and B. However, compared to the Guardian analysis, there were a higher proportion of points in zone A for both iPro2 data sets.

The Bland Altman plots in Figure 6.3 highlight another interesting difference in the error characteristics of each CGM device type. The Guardian SG error appears to be associated with BG level, but the iPro2 error appears independent of BG level. At low BG levels the Guardian tended to report under the true value and at high BG levels it tended to over report glycaemia. Interestingly, there have been reports of associations between CGM error and BG level, but they typically show the CGM to report high at lower BG levels and low at higher BG levels (Beardsall et al. 2013; Voskanyan et al. 2007). A larger data set is required from patients that cover a wider range of glycaemia before the association observed in this study can be confirmed. Furthermore, patients on the STAR protocol are generally very well controlled (89% of BG measurements in 4.4-8.0mmol/L and 0.9% < 4.0mmol/L) so it is thus difficult to obtain enough data to conclusively assess error characteristics during hypoglycaemia or hyperglycaemia.

To determine the impact of inter-site and inter-device variations on the data produced by CGM devices, SG data from one device were compared directly to SG data from another, over the entire monitoring period. Figure 6.4 shows that the

discrepancy in SG data between two different devices is generally much more significant than the discrepancy in SG data due to sensor location. Again, this outcome is likely due to the calibration scheme used, as, in this study, the sensor technology for all three CGM devices was the same. Furthermore, without an accurate reference BG measurement every 5 minutes, it is impossible to determine the true underlying BG concentration. However, the true BG is likely some combination of the SG data produced by all three CGM devices.

6.4.2 Individual patient case studies

Three case studies are presented, to show some of the interesting observations in data collected from the 10 patients in this study. First, as shown in Figure 6.5, CGM devices are capable of working very well in critically ill patients. However, there are still several questions that need to be answered before they are implemented as normal clinical practice, such as:

- Which patients stand to benefit from CGM?
- What conditions/drugs/therapies (if any) have a negative effect on sensor performance?
- What are the performance characteristics of CGM in the ICU and how can data be utilized?

Fortunately, several researchers have already started asking these questions, among others, in an effort to improve healthcare for critically ill patients (Adolfsson et al. 2012; Bridges et al. 2010; Kovatchev et al. 2008; Roberts et al. 2012).

Second, as shown in Figure 6.7, it is possible that conditions, such as severe oedema, could have an effect on sensor performance. While the data in Figure 6.7 represents anecdotal evidence from a single patient in this study, it still presented an interesting case study on a topic that has not been thoroughly investigated. The main focus for this specific discussion is thus the I_{SIG} produced by each sensor.

In particular, the I_{SIG} produced by the abdominal sensor is much lower than the I_{SIG} produced by the sensor in the thigh, where there is much less excess fluid.

Interestingly, after a few days of monitoring, as the patient's condition improved and excess fluid was removed, the I_{SIG} from the abdominal sensor rose to the level of the iPro2 I_{SIG}. In addition, for the remainder of monitoring the I_{SIG} from both sensors tracked each other well. However, these observations could have been due to other factors such as the sensor itself, drugs/therapies, and these other factors cannot be ruled out by this study. Further investigation with a larger cohort containing patients with severe oedema is required to determine whether or not it has an effect on sensor performance.

Third, as shown in Figure 6.8, it is possible for spurious, non-physiological changes in the I_{SIG} data to occur without warning. The way that these 'step changes' appear in SG data is dependent on the calibration algorithm used. In box 'A' of Figure 6.8, the 20nA increase in I_{SIG} caused a ~7.8mmol/L increase in SG data, whereas in box 'B', a 15nA increase in I_{SIG} only caused a ~2.5mmol/L increase in SG data. In box 'A', the rise occurred in the sensor that was monitored by the Guardian real-time CGM. As previously mentioned, the Guardian algorithm could only prior data to estimate the calibration factor, which converted I_{SIG} data into SG data. Consequently, the calibration factor remained fairly constant at approximately 5.5, both before and after the rise in I_{SIG}. It was not until 6 hours after the rise I_{SIG}, when the next calibration BG was entered, that the calibration factor was reduced to 3.

Conversely, the rise in I_{SIG} shown in box 'B' of Figure 6.8 occurred in a sensor connected to an iPro2 with a retrospective calibration scheme. Therefore, at the time of the rise in I_{SIG}, the calibration algorithm used future data to determine that the calibration factor should be reduced from 4.3 to 3.3. This adjustment prevented the SG data from rising significantly above reference BG measurements. These two examples clearly illustrate one of the major tradeoffs between real-time and retrospective calibration of SG data, and thus between these types of devices.

Finally, in Figure 6.8, there are several drop outs of I_{SIG} to ~0nA in the sensor connected to the Guardian CGM device. These drop outs are frequently observed

at the start of monitoring when a voltage is first applied to the sensor. It is possible that a loose connection between the sensor and transmitter could have caused the dropouts in I_{SIG} later in the monitoring period. Fortunately, the calibration algorithms recognize that these dropouts are unusual and simply leave gaps in the SG data.

6.4.3 Limitations

There are three main limitations to this study that need to be addressed. First, this study uses BG and SG data from a relatively small proof-of-concept pilot trial cohort of 10 critically ill patients. These patients are broadly representative of the patients admitted to Christchurch ICU, but a larger study is required to provide conclusive evidence regarding the results presented here. Second, patients on the STAR TGC protocol tend to remain in the 4.4-8mmol/L glycaemic band, and, consequently, a wide range of BG levels are not included in this cohort. Again, a larger study with a broader population, potentially from multiple centres, would likely provide the data required to assess CGM characteristics in hypoglycaemia and hyperglycaemia. Third, it was not possible in this study to have a high accuracy reference BG measurement taken every 5-15 minutes, as this study was a pilot study done in the unit as observed. Thus, we cannot conclude whether all large rises/falls in SG were due to glycaemia or sensor artefacts. However, most major rises/falls in SG could likely be identified and classified using a stochastic detection method, similar to the work presented in Chapter 12.

6.5 Summary

This study used CGM and BG data from 10 patients to assess the reliability of CGM in critically ill patients. Overall cohort results and three case studies were used to show several important findings from this study to date. First, CGM devices can monitor certain patients with a high degree of accuracy, but some illnesses, drugs and therapies might affect sensor performance. Second, severe oedema could potentially affect sensor performance, but further investigation is required to confirm this. Third, CGM device type can have a significant effect on

the accuracy of SG data, but sensor location tends to have less impact. These findings will help direct further studies of CGM devices in critically ill patients in an effort to get them implemented as a clinical practice change.

Chapter 7. Additional uses for CGM in ICU

The use of CGM devices in critical care units is not limited to improving TGC methods. The high density data provided by CGM devices also allows new investigations of glycaemia that were not possible with intermittent BG monitoring. This chapter investigates the association between glucose complexity and mortality in critically ill patients. Specifically, the effects of sensor location and device type on complexity analysis is determined and discussed.

7.1 Introduction

The 2001 landmark study by Van den Berghe et al. was the first data showing lower BG levels in critically ill patients were associated with improved outcomes (Van den Berghe et al. 2006). Since then, it has been determined that that glycaemic variability also plays a very important role (Egi et al. 2006; Hermanides et al. 2010b; Krinsley 2008). More recently, CGM devices have allowed researchers to test the latest hypothesis, that glucose complexity could also be associated with mortality in critically ill patients (Brunner et al. 2012; Lundelin et al. 2010). Glucose complexity, in a very simplistic view, is a measure of the 'fuzziness' of a glucose trace.

The current hypothesis is that a healthy glucose regulatory system will be highly reactive to disturbances by making many small adjustments to keep glucose concentration within a normal range (Lundelin et al. 2010). Conversely, a failing glucose regulatory system will be ‘sluggish’ and the glucose profile should appear smooth with very few high frequency adjustments. To date, there have been two studies that have investigated glucose complexity in critically ill patients (Brunner et al. 2012; Lundelin et al. 2010). Both studies used Detrended Fluctuation Analysis (DFA) to quantify glucose complexity, reporting an association between the results of DFA and overall patient outcome mortality.

Outside of glucose complexity, DFA has already been widely used to quantify the scaling and correlation properties of other non-stationary, physiological time series. For example, DFA has been applied to inter-breath-interval of human respiration, inter-beat-interval of human heartbeat and inter-stride-interval of human stride to differentiate between healthy and pathological conditions (Eke et al. 2002; Goldberger et al. 2002; Hausdorff et al. 1996; Lee et al. 2004; Peng et al. 1995; Peng et al. 2002; Penzel et al. 2003). The aim of this study was to extend the knowledge of glucose complexity in critically ill adults by investigating the effects of CGM device type/calibration and CGM sensor location on DFA results. Furthermore, this study introduces multifractal DFA, an extension to the widely used monofractal DFA, to determine, based on the requirements of the underlying mathematics, whether it is a more appropriate method for analysing glucose complexity in SG data.

7.2 Subjects and methods

7.2.1 Patients

This study uses data from 10 patients recruited into a pilot study of CGM devices in the Christchurch Hospital ICU, as also presented in Chapter 6. Inclusion criteria for the study were an expected duration of ICU admission longer than 5 days and the need for glycaemic control using STAR protocol (Evans et al. 2012). Consent was obtained from the patient or next of kin if the patient was unable to consent. This study and use of data was approved by the Upper South A Regional Ethics Committee, New Zealand. Table 7.1 shows the cohort demographics.

Table 7.1 Cohort demographics. Displayed as Median [IQR] where applicable

Patients	10
Age (years)	51 [39 - 64]
Sex (M/F)	5/5
APACHE II	24 [17 - 27]
APACHE III	85 [52 - 99]
SAPS II	52 [30 - 59]
LOS (days)	20 [10 - 33]
Outcome (L/D)	6/4
Diabetes (None/T1/T2)	10/0/0

7.2.2 Continuous glucose monitoring

As noted in Chapter 6, each participant in the study was monitored using 3 CGM devices for a period of up to 6 days. Two sensors were located on the patient's abdomen, one of which was connected to a Medtronic Guardian Real-Time monitor and the other connected to a Medtronic iPro2 recorder. The third sensor was located on the patient's thigh and was connected to a second Medtronic iPro2 recorder. This configuration allowed inter-site and inter-device variations to be assessed.

One significant difference between the two CGM devices is the calibration algorithm. Calibration algorithms use independent calibration BG measurements to convert the I_{SIG} into a series of SG values for the user. The iPro2 device stores the sensor signal information internally and is retrospectively calibrated after the device has been removed from the patient. Retrospective calibration allows the calibration algorithm to use information both before and after the time point of interest to obtain an optimal calibration to each reference point. However, the Guardian CGM device displays a glucose value in real time and the calibration algorithm can only use prior data for calibration. Otherwise, all sensor technology was identical.

Calibration BG measurements were obtained by specifically trained ICU nurses at least 3 times per day per manufacturer guidelines. A blood sample was drawn from the patient's arterial line and a BGA was used to determine the glucose concentration. The value from the BGA was immediately entered into the Guardian Real-Time device and then recorded for retrospective calibration of the iPro2 devices.

7.2.3 Glucose complexity

Glucose complexity was quantified in this study using DFA, which has been widely used to quantify the scaling and correlation properties of non-stationary, physiological time series. For a self-similar time series (X), the scale invariant structure can be described by $X(ct) = c^H X(t)$, where the power law exponent (H) describes the scaling. In some cases, the scaling properties of a time series

are associated with changes in physiology and may be useful to help better understand certain illnesses or medical conditions (Goldberger et al. 2002; Peng et al. 2002). In terms of glucose complexity, the current hypothesis is that a healthy glucose regulatory system will make many small adjustments to keep glucose concentration within a healthy range and the high complexity is characterised by a low value of *H*. In contrast, a failing glucose regulatory system will be 'sluggish' and appear smooth with low complexity, characterised by high values of *H*.

Two sub-categories of DFA are available depending on the properties of the signal or time series. Monofractal DFA is used when the scaling properties of the time series can be quantified by a single power law exponent, which is independent of time and space. However, the limitation of monofractal DFA is that real world physiological signals often do not exhibit simple monofractal scaling behaviour over the entire time period (Goldberger et al. 2002; Ihlen 2012). In these cases, multiple scaling exponents are required to fully characterise the correlation properties of the signal and multifractal DFA should be employed.

The results of a multifractal DFA are typically displayed in a multifractal spectrum. The width, shape and location of the multifractal spectrum can all be used to give important information about the relationship between the time series and the physiological phenomenon being studied. This approach has been illustrated in previous studies that have used multifractal DFA to help differentiate between healthy and pathological conditions, such as heart disease (Goldberger et al. 2002).

An example of two spectrums produced using multifractal DFA are shown in Figure 7.1. The spectrum plotted with red dots was produced from a signal with monofractal scaling properties and the spectrum plotted with blue crosses was produced from a signal with multifractal scaling properties, data for both of which were provided by Ihlen (Ihlen 2012). Note the spectrum for the monofractal signal is very narrow, indicating the scaling exponents are almost

independent of time and space. Thus, a single H from monofractal DFA is sufficient to characterise the scaling and correlation properties of the signal.

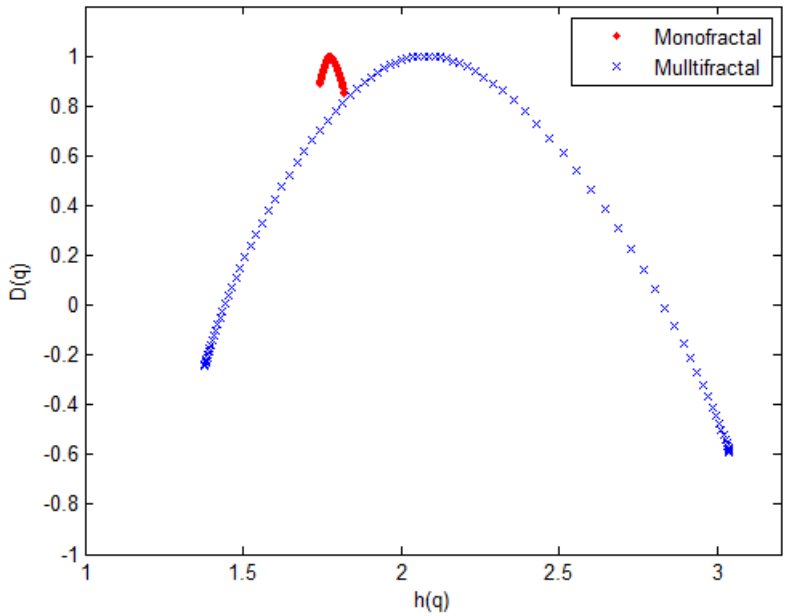


Figure 7.1 Example of multifractal spectrum that is produced from multifractal DFA. Note the monofractal signal produces a very narrow spectrum, indicating monofractal scaling is present and monofractal DFA is sufficient to characterise the scaling and correlation properties of the signal.

This study uses both monofractal DFA and multifractal DFA implemented in MATLAB, based on the descriptions provided in (Ihlen 2012; Kantelhardt et al. 2002). A thorough discussion of both methods can be found elsewhere for both monofractal DFA (Eke et al. 2002) and multifractal DFA (Kantelhardt et al. 2002).

7.2.4 Implementation of DFA

The implementation of DFA on a time series x, can be summarised as follows. First, the time series is summed and mean subtracted:

$$Y(i) = \sum_{k=1}^{i} [x_k - \langle x \rangle], \qquad i = 1, \dots, N\,G \tag{7.1}$$

Second, the profile $Y(i)$ is divided into N_s non-overlapping segments of equal length s and the trend for each segment is approximated using a least-square fit, as shown in Figure 7.2 for segment sizes of 32, 64 and 128 SG measurements.

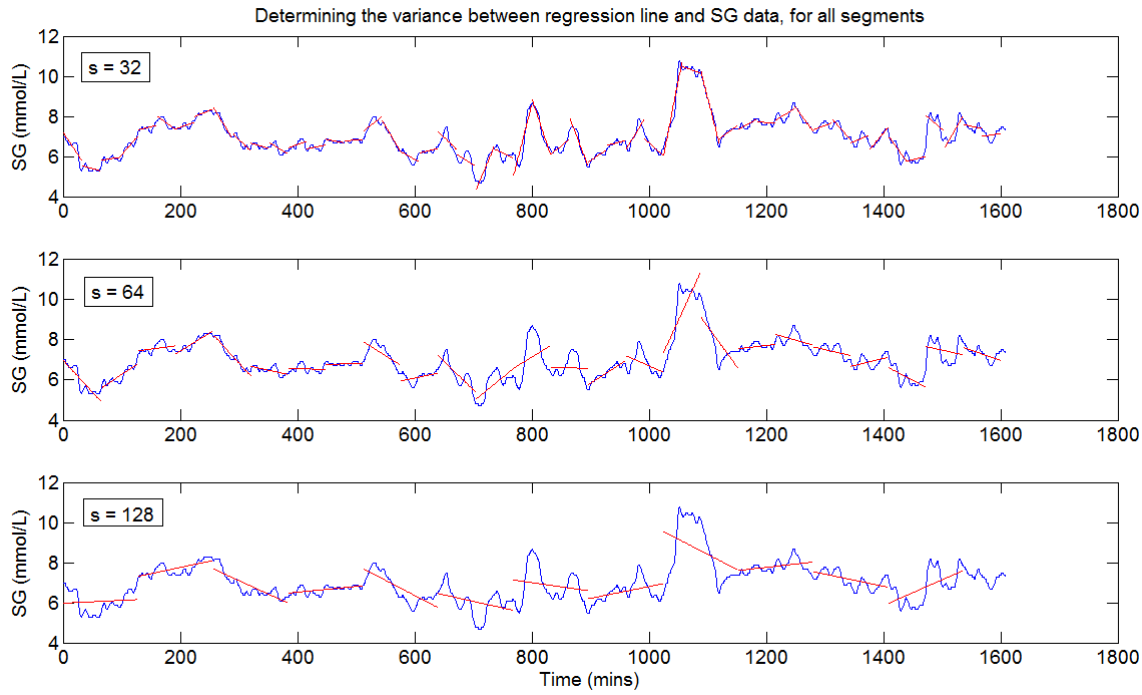


Figure 7.2 Three examples showing segmented SG data with linear regression lines. The segment size increases from top to bottom. The variance between each regression line and the corresponding SG data in that segment is calculated using Equation 2.

The variance between the time series and the least-square fit for each segment v is calculated:

$$F^2(v,s) = \frac{1}{s}\sum_{i=1}^{s}(Y[(v-1)s+i] - y_v(i))^2 \qquad 7.2$$

Third, the qth order fluctuation function is calculated using Equation 7.3, where a monofractal DFA is obtained by holding q=2. Essentially, Equation 7.3 gives a qth order RMS of the variances calculated using Equation 7.2. For positive q, large deviations from the fitted trend will have more influence on the fluctuation function than small deviations, and, for negative values of q small deviations will

have a larger influence on the fluctuation function. The degree to which the fluctuation function is influenced by q is determined by the magnitude of q.

$$F_q(s) = \left\{ \frac{1}{N_s} \sum_{v=1}^{N_s} [F^2(v,s)]^{q/2} \right\}^{1/q} \qquad 7.3$$

The scaling behaviour of the fluctuation functions is illustrated by analysing a plot of $log(F_q(s))$ versus $log(s)$ for each q. For a scale invariant series, the relationship is linear and the slope represents the power law exponent, H, as in Figure 7.3 (left). For a multifractal time series, the scaling exponent will change for different values of q, as shown in Figure 7.3 (right).

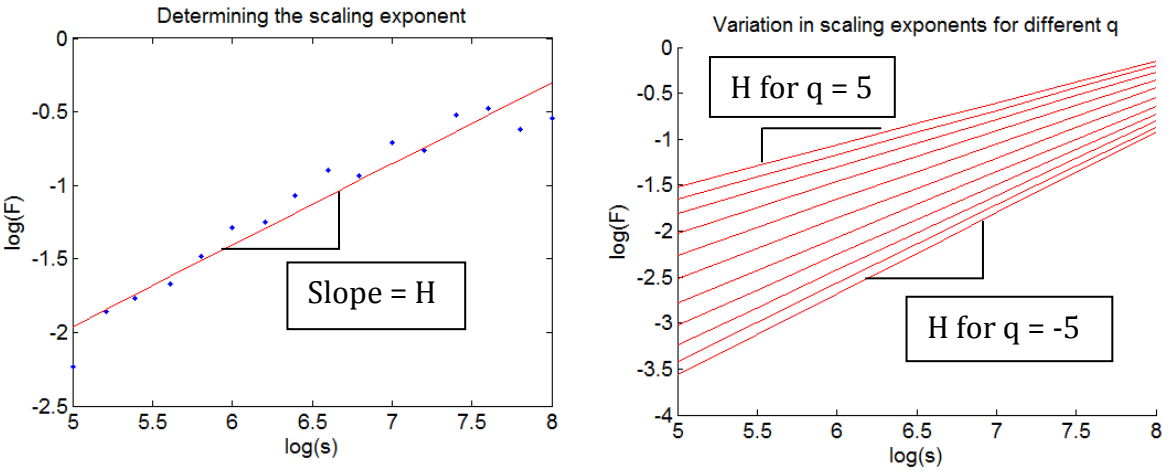


Figure 7.3 (left) Example plot of log(F) versus log(s) for a single value of q, where the slope of the linear regression line is the scaling exponent, H. (right) Example of how H changes for different values of q, for a multifractal time series

For the case of q=0, Equation 7.3 cannot be employed so a logarithmic averaging procedure is used instead:

$$F_0(s) = exp\left\{ \frac{1}{2N_s} \sum_{v=1}^{N_s} \log[F^2(v,s)] \right\} \qquad 7.4$$

Multifractal DFA performs best when the signal being analysed is noise-like, rather than a random walk. To determine the type of signal, a monofractal DFA can be run prior to multifractal DFA. If the power law exponent, H, is between 0.2 – 0.8 then the time series is noise-like and multifractal DFA can be run without modification. However, if H is between 1.2 – 1.8 then the time series is more like a random walk. For random walk signals, the integrating process shown in Equation 7.1 can be skipped, and +1 is added to the power law exponents determined in the multifractal analysis (Ihlen 2012).

To aid the interpretation of the multifractal scaling properties of a time series, the mass exponent ($\tau(q)$), q-order singularity exponent ($h(q)$) and q-order singularity dimension ($D(q)$) are calculated:

$$\tau(q) = q.H(q) - 1 \qquad 7.5$$

$$h(q) = \tau'(q) \qquad 7.6$$

$$D(q) = q.h(q) - \tau(q) \qquad 7.7$$

A plot of $\tau(q)$ vs. q (Figure 7.4 - left) or $h(q)$ vs. q (Figure 7.4 - right) can be used to determine the degree of multifractal scaling in a time series. A plot of $D(q)$ vs. $h(q)$ displays the mulitfractal spectrum. The width, shape and location of the multifractal spectrum can all be used to give important information about the relationship between the time series and the physiological phenomenon being studied.

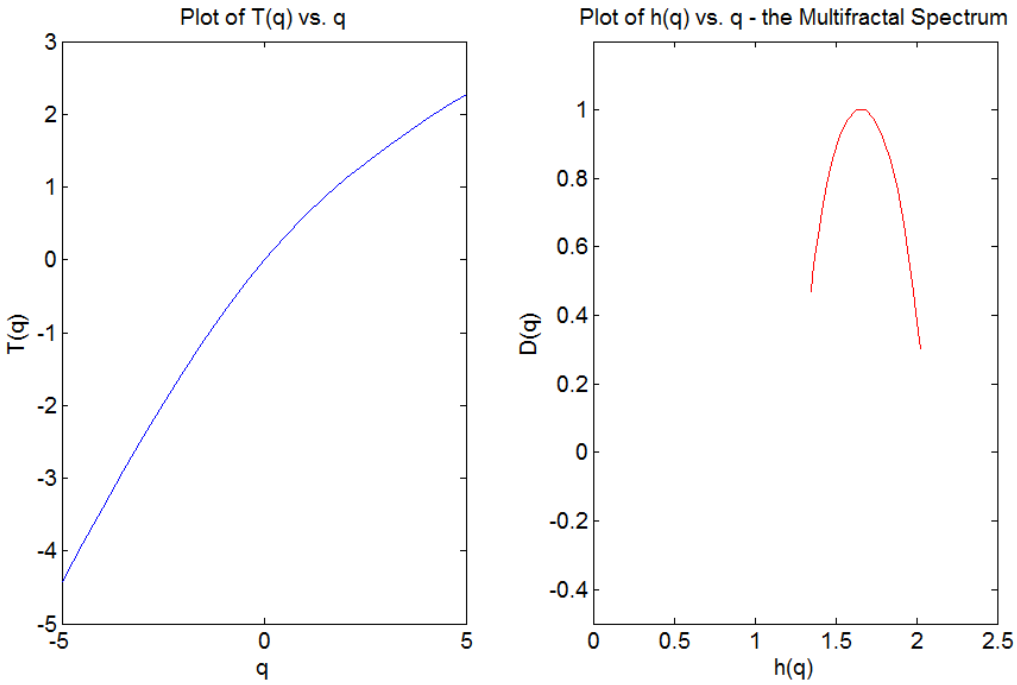


Figure 7.4 (left) plot of $\tau(q)$ vs. q and (right) $h(q)$ vs. q, either of which can be used to interpret the scaling properties of a multifractal time series.

7.2.5 Analysis

This study uses DFA to investigate the glucose complexity of 10 critically ill patients who were monitored by 3 simultaneous CGM devices during their ICU stay. Specifically, it investigates:

1. Whether CGM device type/calibration or CGM sensor location affects DFA results.
2. Whether monofractal or multifractal DFA is more appropriate for CGM signals, given the use of monofractal DFA as a discriminator in (Brunner et al. 2012; Lundelin et al. 2010).
3. Whether DFA results are associated with mortality.

Each patient enrolled in the study had 3 CGM devices monitoring glucose levels for up to 6 days. The warm-up period for these devices is 1-2 hours (Minimed 2006,2010), but due to off label use here the first 12 hours of SG data were excluded to ensure the devices were performing properly during the period of interest.

Data sets with less than 500 SG measurements were excluded and SG data sets with less than 1000 SG measurements were analysed with increased care to ensure robust results, as this value is a recommended minimum (Ihlen 2012). Finally, the SG data were analysed using both monofractal DFA and multifractal DFA to determine whether the analysis method has an impact on results.

In particular:

- Data from the two iPro2 CGM devices, one on the thigh and one on the abdomen, were compared to assess sensor location effects independent of technology.
- Data from the iPro2 on the abdomen was compared to data from the Guardian Real-Time, located on the opposite side of the patient's abdomen to assess the impact of CGM device type/calibration on DFA results.
- Analyses were repeated using the I_{SIG}, which removed any effects induced by device calibration.

Monofractal DFA results are presented in a table, as the result from each analysis is a single scaling exponent, H. Results from the multifractal DFA are presented in figures containing a plot of the multifractal spectrum. When comparing results, similar values of H from the monofractal analysis and spectrums of similar shape/position from the multifractal analysis indicate little or no difference in the scaling properties of the time series. In the two case study examples shown, plots of the CGM glucose trace and a plot of the scaling exponent as a function of q-order statistical moments are also shown. Numerical results are presented as median [IQR] where applicable.

7.3 Results

Of the 30 CGM sensors inserted into the 10 participants during the study, 26 sensors from 9 patients produced data suitable for analysis. The 4 excluded data sets contained less than 500 measurements and were considered too short for analysis. Three of the excluded data sets were from a short stay patient eliminating all their CGM data, and the other data set was from a sensor with adhesion issues that was dislodged early in the monitoring period.

7.3.1 Monofractal DFA

The results from the monofractal DFA are shown in Table 7.2. The top half shows the results of DFA when analysing SG data from the CGM devices. It is important to reiterate that the Guardian and iPro2 devices use different calibration algorithms with the same sensor technology. Thus, these results include any effects of calibration on DFA results.

The *CGM device type* section shows the results when comparing CGM device type. Across the cohort, the retrospectively calibrated iPro2 reported significantly higher scaling exponents of 1.56 [1.46 – 1.60] compared to 1.43 [1.37 - 1.48] for the Guardian Real-Time device ($p = 0.03$). Furthermore, when comparing the two different abdominal devices monitoring the same individual, the H values for iPro2 data were 0.10 [0.03 – 0.11] higher than the H values for Guardian data ($p = 0.08$).

The *Sensor location* section shows the results comparing a sensor inserted in the abdomen to a sensor inserted in the thigh of the same CGM device type (iPro2). There is no significant difference in the H values for different sensor locations ($p > 0.1$).

The *Outcome mortality* compares scaling exponents based on ICU mortality. Results show no significant difference in the scaling exponent for patients who lived compared to patients who died.

The bottom half of Table 7.2 shows results from the same analysis of the I_{SIG}, which removes the effects of CGM device calibration. The data sets are stratified into the same groups as the top half of Table 7.2 based on CGM type, sensor location and mortality. When the effects of calibration are removed, the results show no significant differences between the groups, in any of the sub analyses.

Table 7.2 Results from monofractal DFA of SG and I_{SIG} data over cohort

Analysing calibrated SG data			
CGM device type (both in abdomen)			
	Guardian	**iPro2**	**P value**
Number of data sets	9	8	
Scaling exponent (H)	1.43 [1.37 - 1.48]	1.56 [1.46 - 1.60]	0.03*
Difference in H (iPro2 - Guardian)	0.10 [0.03 - 0.20]		0.08**
Sensor location (both iPro2)			
	Abdomen	**Thigh**	**P value**
Number of data sets	8	9	
Scaling exponent (H)	1.56 [1.46 - 1.60]	1.52 [1.50 - 1.61]	0.90*
Difference in H (Thigh - Abdomen)	0.04 [-0.06 - 0.11]		0.64**
Outcome mortality			
	Lived	**Died**	**P value**
Number of data sets	17	9	
Scaling exponent (H)	1.51 [1.46 - 1.57]	1.47 [1.39 - 1.59]	0.50*
Analysing pre-calibration ISIG data			
CGM device type (both in abdomen)			
	Guardian	**iPro2**	**P value**
Number of data sets	9	8	
Scaling exponent (H)	1.42 [1.34 - 1.52]	1.54 [1.37 - 1.60]	0.33*
Difference in H (iPro2 - Guardian)	0.06 [-0.04 - 0.10]		0.53**
Sensor location (both iPro2)			
	Abdomen	**Thigh**	**P value**
Number of data sets	8	9	
Scaling exponent (H)	1.54 [1.37 - 1.60]	1.51 [1.47 - 1.60]	0.73*
Difference in H (Thigh - Abdomen)	0.04 [-0.03 - 0.09]		0.38**
Outcome mortality			
	Lived	**Died**	**P value**
Number of data sets	17	9	
Scaling exponent (H)	1.50 [1.35 - 1.55]	1.51 [1.42 - 1.60]	0.36*

*** Wilcoxon Rank-Sum test, ** Wilcoxon signed Rank test**

7.3.2 Multifractal DFA

Figure 7.5 shows the multifractal spectrums for all SG and I_{SIG} data sets used in the analyses. The three subplots on the left side were created using SG data, comparing CGM device type (top), sensor location (middle), and outcome mortality (bottom). Three equivalent plots on the right side of Figure 7.5 were created using I_{SIG} data. There is no significant association between any of the CGM parameters tested and the shape, width or location of the multifractal spectrums.

Figure 7.6 shows four examples of how monofractal and multifractal DFA can give contradicting results and why the results must be interpreted with care. In all four cases, monofractal DFA of the two SG data sets resulted in the same scaling exponent. However, the multifractal DFA resulted in significantly different spectrums for each of the paired data sets. Furthermore, the width of the spectrums suggests the scaling properties of the time series are multifractal and that multifractal DFA is a more appropriate analysis technique (Goldberger et al. 2002; Ihlen 2012).

Figure 7.8 shows SG data and multifractal DFA results for two example patient cases. Example 'A' shows SG data from three sensors that report very similar glucose traces, with little discrepancy between them. However, the multifractal DFA of each of these SG data sets results in three quite different multifractal spectrums, indicating that the scaling properties of each signal are unique. In contrast, example 'B' shows SG data from two sensors that do not track each other very well. Despite the large difference in variability, the multifractal DFA resulted in two spectrums that are almost identical. These two examples clearly highlight the difference between typical variability metrics and complexity analyses that assess very different aspects of the signal.

7.4 Discussion

The aim of this study was to investigate the effects of CGM device type/calibration and CGM sensor location on DFA results, as well as whether DFA results are associated with mortality, in critically ill patients. Due to the configuration of CGM devices and sensor locations in this study, there was a unique opportunity to study the effects of these parameters on DFA results.

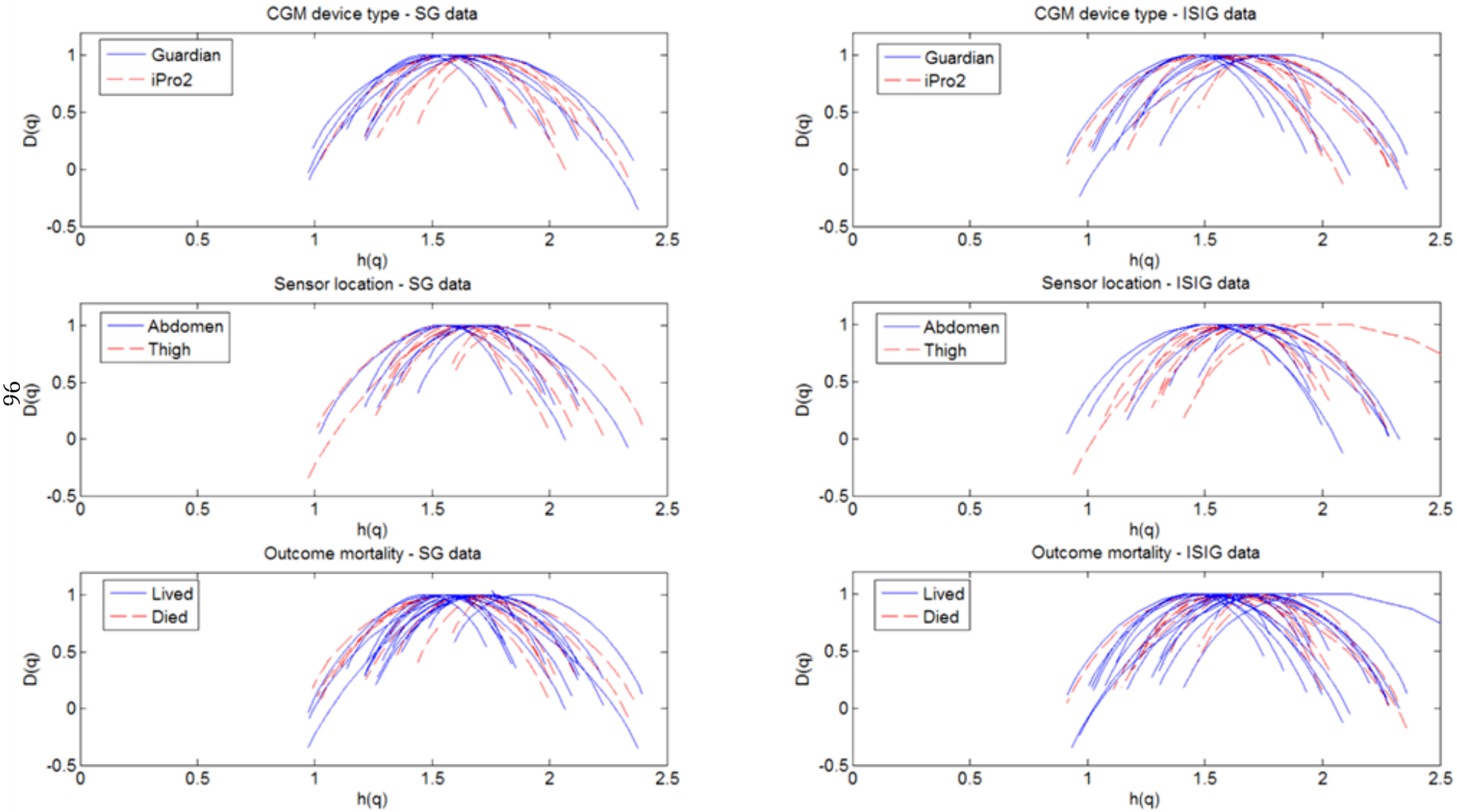


Figure 7.5 Multifractal spectrums comparing CGM device types, sensor locations and outcome mortality. The plots on the left were created using SG data and the plots on the right were created using I_{SIG} data.

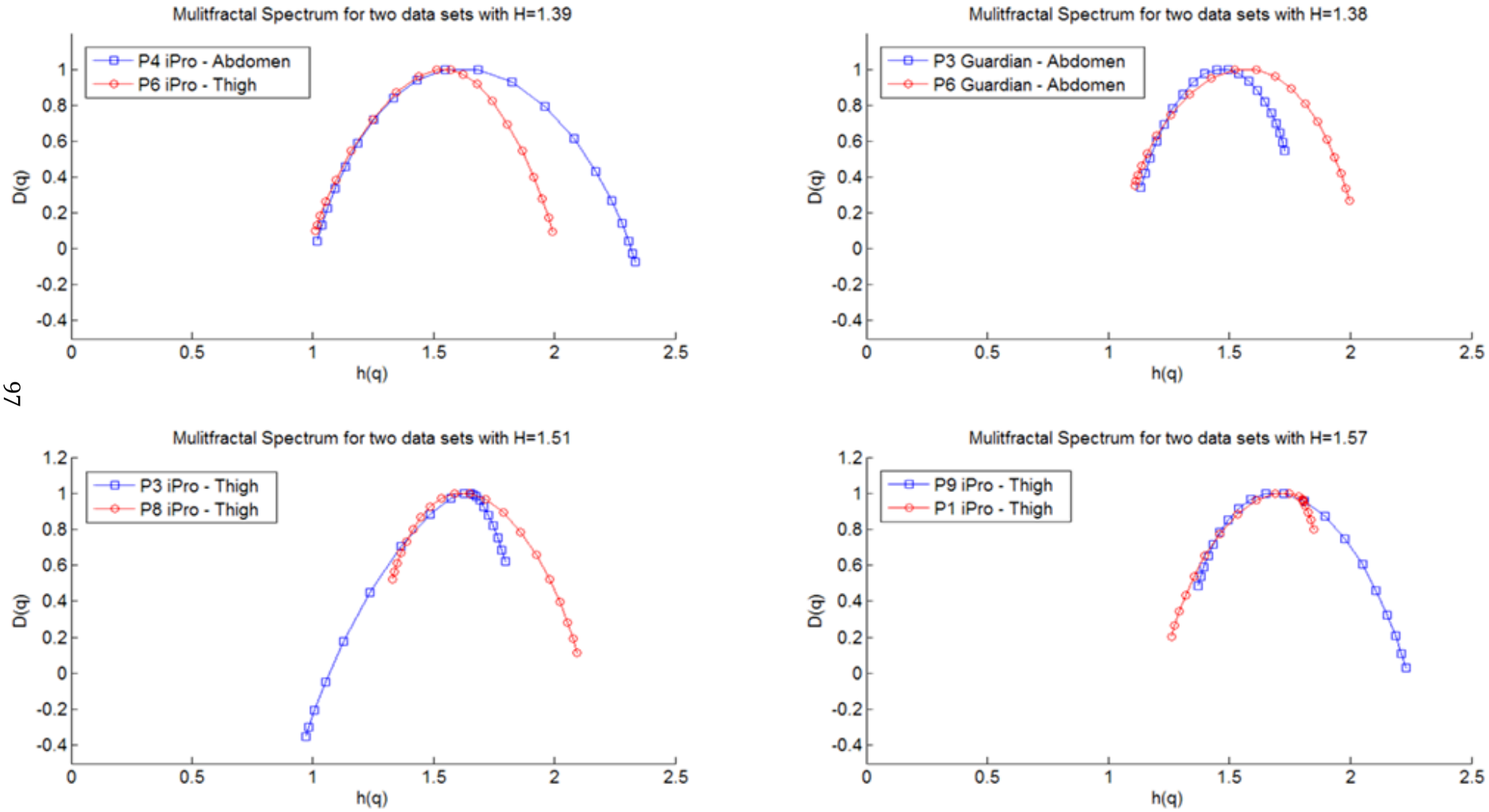


Figure 7.6 Multifractal Spectrum comparison for data sets that had the same scaling exponent from monofractal DFA.

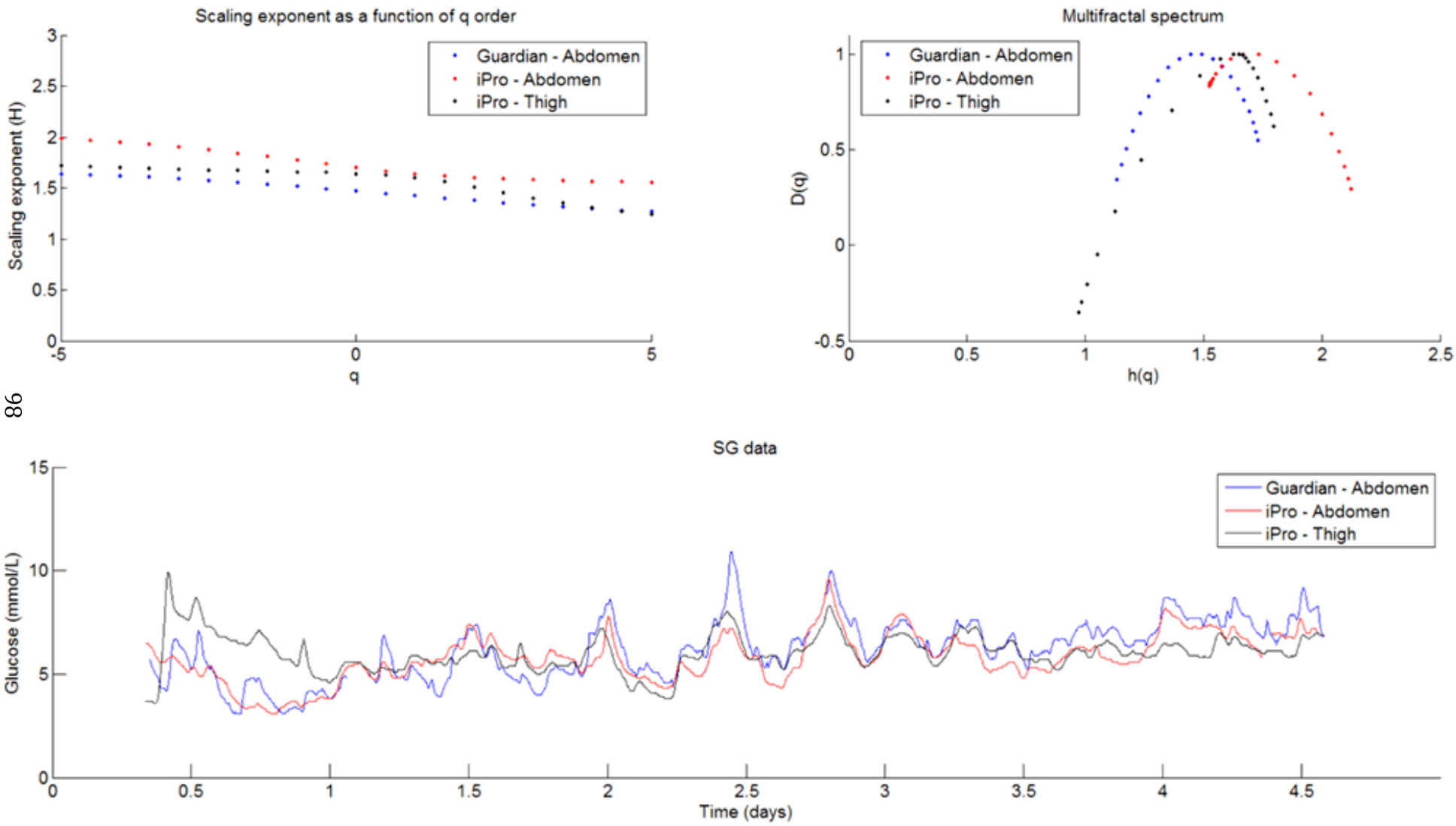


Figure 7.7 This example shows good agreement between SG data for each of the three CGM devices, but the multifractal spectrums for each data set are quite different.

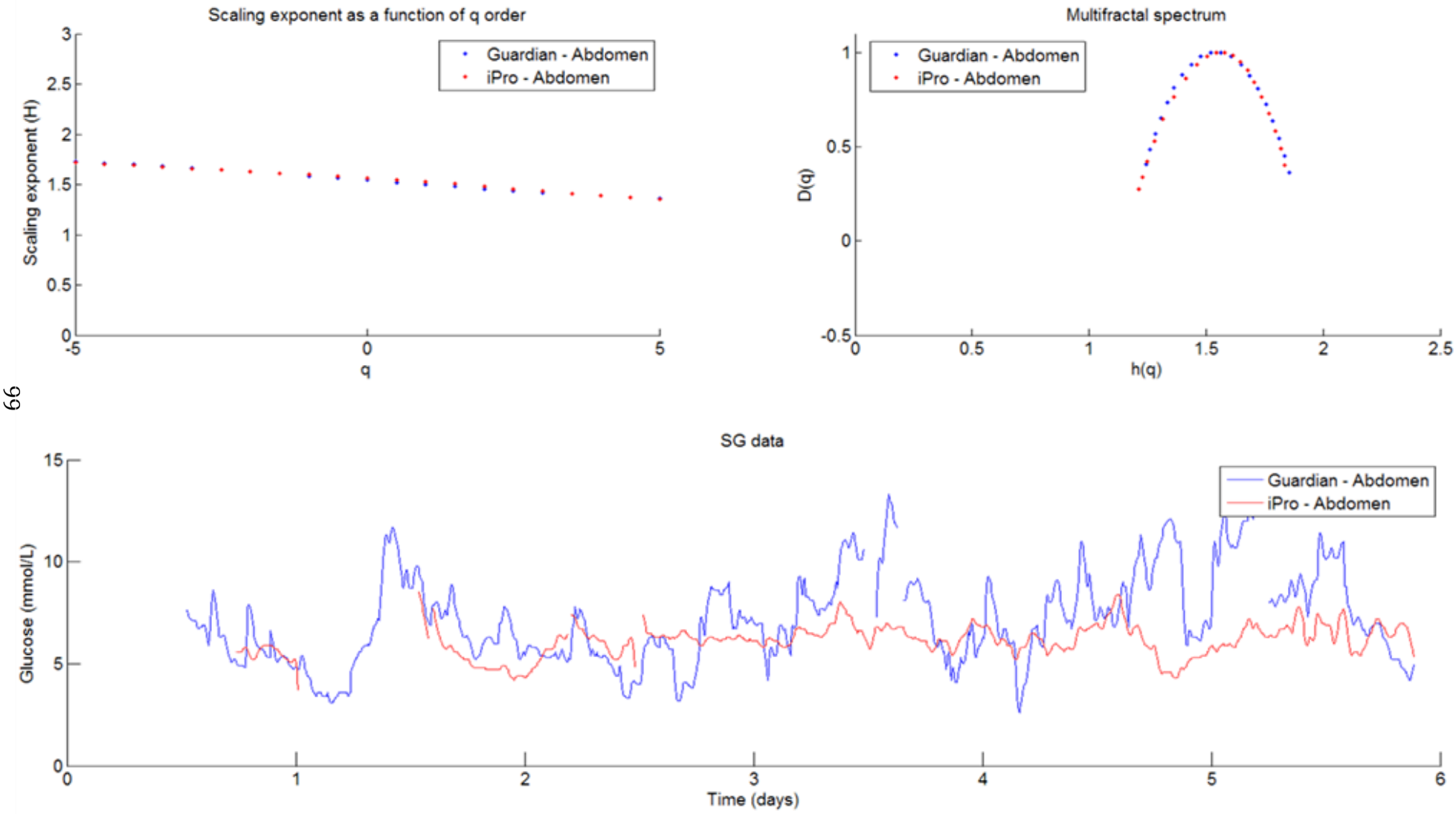


Figure 7.8 This example shows average agreement between SG data for two CGM devices, but the multifractal spectrums for each data set overlap.

7.4.1 Monofractal DFA

The effects of CGM device type on monofractal DFA results revealed an important trend. The two sensors located in the patient's abdomen were identical, but the H values from the iPro2 data were significantly higher than H values from the Guardian device. This outcome is most likely due to the calibration algorithms used to process the raw sensor data. The retrospective calibration utilised by the iPro2 device potentially has a higher degree of smoothing/filtering that leads to lower complexity of the output SG trace.

Due to privacy restrictions placed on the calibration algorithms, this difference could not be fully confirmed by comparing the algorithms directly. However, an analysis of the raw pre-calibration I_{SIG} data from the sensors showed no significant difference between the H values from each device. Thus, the significant difference in H observed when assessing SG data can be attributed to the calibration algorithm and/or potentially the CGM hardware used to store and transmit data, rather than to the specific sensing mechanism.

Sensor location (thigh iPro2 data vs. abdomen iPro2 data) had little effect on the scaling exponent determined through monofractal DFA and the range of H values for each sensor location were similar. Furthermore, comparing the results from two sensors monitoring the same individual showed no consistent trend of one sensor having a higher H value than the other. Repeating the analysis using the raw sensor I_{SIG} to eliminate any calibration effects gave similar results and the scaling exponent was not dependent on location. These results suggest that CGM sensor location, at least thigh vs. abdomen, should not have a significant effect on the results of a study using monofractal DFA.

Monofractal DFA of SG data revealed important findings, but there were no detectable associations between DFA results and mortality. This result was unexpected because two previous studies by Lundelin et al. and Brunner et al. both reported a strong association and relatively clear discrimination between glucose complexity (H from DFA) and mortality. The authors acknowledge that due to low sample size it was unlikely a statistically significant association would

be detected. However, the results in Table 7.2 clearly show heavily overlapped values that appear unlikely to separate, often with different results for the same patient. Similar results are seen in the I_{SIG} data, which eliminates any effects of calibration.

7.4.2 Multifractal DFA

Analyses of SG data using multifractal DFA were unable to associate the shape, width or position of the multifractal spectrum with CGM device type, sensor location or mortality for this data set. Furthermore, overlaying all of the multifractal spectrums from all sensors in a single figure, similar to Figure 7.5, showed no clear trends to differentiate between CGM device types, sensor locations or mortality outcomes. The I_{SIG} results showed essentially the same results. Again, the mortality outcome could be attributed to the relatively small sample size and associations could potentially show up in a similar study of a larger cohort, although there were no initial results in Figure 7.5 suggesting they might arise.

However, one very important finding from the multifractal DFA was the width of the multifractal spectrums. Each spectrum was spread across a wide range of scaling exponents, which suggests the scaling of SG data is multifractal in all cases here, and not monofractal, by nature. Thus, future studies that investigate glucose complexity using SG data should test for multifractal scaling and consider multifractal DFA in their analyses.

Another interesting finding from the multifractal analysis is depicted in the subplots of Figure 7.6. In these four cases, monofractal DFA of both data sets produced the same H value, but their multifractal spectrums were clearly and significantly different. These differences between monofractal and multifractal results are important as they could skew the interpretation of results and lead to incorrect conclusions from studies of this nature. These results reiterate the need for a thorough analysis and highlight why care must be taken when interpreting results.

Figures 7.7 and 7.8 clearly illustrate the independence of glucose complexity and glucose variability. Figure 7.7 has three SG traces that track each other well and have similar variability, but their multifractal spectrums are very different, indicating the SG traces have different complexities. Conversely, Figure 7.8 has two SG traces that don't track each other very well and the blue Guardian SG data has higher variability than the iPro2 SG data. However, the multifractal spectrums for each sensor show very similar multifractal scaling properties for each time series, indicating the signal complexities are similar. Therefore, complexity analyses should be included in addition to variability analyses to determine associations in future studies of glucose dynamics.

7.4.3 Limitations

There are several limitations of this study that need to be addressed. First, the data set used here contains 26 sets of CGM data from a relatively small cohort of critically ill patients. Thus, a statistically significant and testable association between DFA and mortality was unlikely to be detected. Despite this limitation, the analysis did highlight several important aspects that should be considered in future studies including the effects of sensor location and CGM device/calibration on DFA results. Equally, the total number of sensors available for assessing the impact of calibration and location ensure novel results that are robust.

Second, the sensors and devices used in this study were different to those used by Lundelin et al. and Brunner et al. (Brunner et al. 2012; Lundelin et al. 2010). Therefore, a direct comparison between all three studies is not appropriate. The devices in those studies were no longer available when this study data was collected and a newer generation of technology was used. Thus, the study conditions that would allow an exact comparison could not be replicated. However, any potential differences due to advancing technology would, again, indicate the impact of technology on DFA assessment of physiology.

7.5 Summary

The results of this study revealed or reiterated 5 important aspects that should be considered when analysing glucose complexity in critically ill adults:

1. Monofractal DFA results were sensitive to the type of CGM device used to collect the glucose data, and the calibration in particular. Data from the iPro2 CGM device gave consistently higher DFA results compared to data from the Guardian Real-Time CGM device.
2. Sensor location (abdomen vs. thigh) had no significant effect on DFA results.
3. Multifractal DFA results were not always consistent with monofractal DFA results. Furthermore, the width of the multifractal spectrums suggests that multifractal DFA is more appropriate for this type of data and should be considered in future glucose complexity studies.
4. Glucose complexity is not the same as glucose variability, and DFA could be used in addition to glucose variability analyses to better characterise a glucose time series.
5. An association between DFA results and mortality could not be detected in this limited data set.

Further investigations of glucose complexity are required before solid conclusions can be drawn. However, this study clearly highlights where care should be taken in the design of those studies and the analysis of the results.

Chapter 8. Summary of ICU

The first half of this thesis has investigated CGM in critically ill patients. The overall goal of this work was to assess whether these devices, which were designed for use outside of the hospital, could perform reliability in the critical care environment. Specifically, to determine whether CGM devices could be integrated with existing TGC protocols to improve control and safety while reducing nurse workload.

Prior to testing CGM devices clinically, a literature search revealed the details of several elements that would allow an in-silico study of TGC using CGM devices to be completed. Two mathematical models of CGM error characteristics were created based on data published in two independent studies. The first model was created using accuracy data from a test of CGM devices in critically ill patients (Goldberg et al. 2004). This model was comprised of a series of Gaussian distributions that captured the point-to-point error characteristics of SG data, but assumed consecutive errors were independent. The second model incorporated a time-dependent aspect into the model and was based on an autoregressive model (Breton & Kovatchev 2008). While more complete in terms of a CGM error model, it was created using data from otherwise healthy diabetics, not critically ill patients, so there is no assurance it is applicable in critical care.

In addition to modelling CGM error characteristics, several filtering methods and hypoglycaemia alarms were tested to determine which algorithms perform well. Finite impulse response filters, infinite impulse response filters and a novel median LMS filter were tested. The median LMS filter performed particularly well and was selected to be used in the in-silico study. In addition to filtering methods, several alarm algorithms were also tested. A major risk with TGC protocols is hypoglycaemia, which can negatively affect outcomes in critically ill patients. Thus, two alarming algorithms were tested to assess their reliability in detecting hypoglycaemia, and the amount of lead time that could be given before an episode. Both the integral based and threshold based alarms performed well.

The CGM error models, median LMS filtering algorithm and a variant of the threshold alarm algorithm were coupled with existing patient simulation methods, to enable an in-silico clinical trial to be completed. Briefly, the clinically validated patient simulation methods use a set of clinically validated mathematical models to determine a patient's response to insulin and nutrition, which are prescribed by a TGC protocol. The in-silico study replaced the BG measurements required by the TCG protocol, with filtered measurements created using the CGM error models. The results showed that the SPRINT TGC protocol performed well with filtered SG measurements and nurse workload could potentially be reduced by up to 75%. Furthermore, dextrose interventions at the predicted onset of hypoglycaemia could effectively eliminate the risk of hypoglycaemia with TGC if the proper dose is used. Overall, the results justified a clinical trial of CGM devices in the ICU.

An important difference between intermittent BG measurements and the data provided by CGM devices is the availability of trend information. For example, intermittent BG measurements offer a snapshot of the BG concentration at a single point in time, whereas CGM devices give additional information about glucose trends. Although there are a wide number of metrics and methods available for assessing point-to-point accuracy, there are very few available for assessing trend accuracy. Even more importantly, none of those few existing trend assessments is independent of point-to-point sensor accuracy. A new method of quantifying trend accuracy was developed and presented in an easy to interpret graphical interface called the Trend Compass. A performance table and TI metric can accompany the table to give further information if required.

The Trend Compass was tested in-silico, where it successfully differentiated between good and poor trending in a robust manner, irrespective of underlying glycaemia. Real clinical data was also used to show the effectiveness of the Trend Compass at quantifying trend accuracy. Furthermore, the clinical data was used to highlight the difference between trend accuracy and point-to-point accuracy. Overall, the Trend Compass, and accompanying metric table (if required),

provide an easy to interpret method of assessing trend accuracy in SG data independent to point accuracy.

With positive in-silico study results, a pilot clinical trial of CGM devices in the ICU was undertaken. The main goal of the trial was to assess the reliability of off-the-shelf CGM devices in monitoring critically ill patients, with a specific interest in the effects of CGM device type and sensor site on SG data. Thus, each of the 10 participants in the study was monitored using three CGM devices, a Medtronic Guardian Real-Time and a Medtronic iPro2 placed on their abdomen, and a second Medtronic iPro2 placed on their thigh. All CGM devices used the latest Medtronic Enlite sensing technology, allowing a fair assessment of inter-site and inter-device variability in SG data. In addition to CGM, each patient had 12-14 BG measurements per day for TGC, which provided a further range of comparator data to SG data.

Overall, the accuracy of the abdomen Guardian SG data relative to reference BG measurements was lower than both the abdomen and thigh iPro2 SG data. Furthermore, an analysis of inter-site and inter-device discrepancies in SG data suggested that the CGM device type has a larger impact on observed performance than sensor site. Several unanticipated findings were also investigated and presented as case studies, as they highlighted important factors that require further investigation or may help direct future studies of CGM devices in the ICU.

The pilot study showed that CGM devices are capable of monitoring patients with a high degree of accuracy. However, certain illnesses, drugs and therapies can potentially affect sensor performance. One case study showed that severe oedema may have affected sensor performance, but further investigation is required to confirm this initial result. The effects of different calibration algorithms were also investigated and, in particular, the way they respond to artefacts in the sensor data. Overall, the limited data set and preliminary analysis indicated that CGM devices have the potential to improve TGC in critically ill patients. Such improvements include using SG measurements to drive insulin

therapy, and/or, using SG data for hypoglycaemia detection and alarming as guardrails. However, further understanding of the clinical factors that affect CGM performance is needed, based on larger trials, before improvements to TGC can be realized.

The CGM data collected in the pilot study was also used to investigate a new hypothesis, that glucose complexity is associated with mortality in critically ill patients. More specifically, the unique configuration of devices used in the pilot study allowed inter-site and inter-device effects on complexity analysis to be assessed. Briefly, it has been suggested that a healthy regulatory system will respond quickly to changes in glycaemia, resulting in a complex glucose trace that is characterised by frequent, small fluctuations. Conversely, a failing regulatory system should appear sluggish, with a smoother texture and longer, slower dynamics. The level of complexity in SG data can be quantified using a signal processing technique called DFA.

There have been two prior studies that have investigated glucose complexity and reported statistically significant associations with mortality. Both studies used older, and in one of the studies mixed, CGM technology and a mono-fractal version of DFA. The results from our data set suggest that glucose complexity is far more strongly and statistically significantly associated with the type of CGM device used, rather than mortality. Our data failed to show any significant association between glucose complexity and mortality. Furthermore, a multi-fractal DFA analysis suggested that the SG data in our study, and potentially all SG data, is multi-fractal by nature, contradicting the assumptions of the previous studies. Thus, future studies of glucose complexity should consider a multi-fractal approach, at least in addition to mono-fractal analyses, and care should be taken when interpreting the results.

Overall, the work presented in Chapters 2-7 contributes valuable information and insight into the use of CGM devices in critically ill patients. This work not only considers the engineering aspects, but also the clinical perspective, which is paramount if CGM devices are going to be implemented permanently. Finally, the

positive results discussed here provide a significant step toward achieving the ultimate goal of using CGM devices in hospital environments.

PART II - CGM IN NEONATAL ICU

Chapter 9. Background - Neonatal ICU

Similar to the adult ICU, patients in the NICU could also benefit from CGM if it proves to be reliable and is used effectively. Neonatal hypoglycaemia is a common metabolic problem in newborn babies that may cause seizures and permanent brain injury in newborns (Stanley & Baker 1999). However, there remains significant controversy regarding the definition of hypoglycaemia, and consequently, the effect it can have on the child's later development (Cornblath et al. 2000; Koh et al. 1988). Diagnosis is by BG measurements, typically taken several hours apart, and episodes of hypoglycaemia that occur between BG measurements can go undetected (Harris et al. 2010). Thus, CGM devices with their nearly continuous estimate of BG concentration have the potential to improve the detection and diagnosis of hypoglycaemia in newborn infants.

9.1 Blood glucose physiology

Glucose is the principle substrate for cerebral metabolism and maintaining glucose homeostasis is critical for an infant's healthy development. Prior to birth, neonates receive glucose from their mother through the placenta, via a relatively complex set of mechanisms that tend to keep its metabolism relatively constant (Hay 2006). At birth, the umbilical cord is severed, removing the maternal supply of glucose and forcing the infant to self-regulate its BG concentration (Halamek et al. 1997).

Self regulation of BG concentration occurs via a negative feedback system, where the endocrine cells of the pancreas act as the primary controller. The endocrine cells, or Islets of Langerhans, make up around 1-2% of the pancreas and include four types of hormone-secreting cells:

1. Alpha cells: Account for about 18-20% of islet cells and secrete the hormone glucagon.
2. Beta cells: Constitute about 73-75% of islet cells and secrete the hormone insulin.
3. Delta cells: Make up only about 4-6% of islet cells and secrete the hormone somatostatin.

4. F cells: Only around 1% of islet cells are F cells, which secrete the hormone pancreatic polypeptide.

The interactions of these hormones are complex and not completely understood, but, in terms of BG regulation, the hormones of primary interest are glucagon and insulin. Figure 9.1 is a schematic showing the fundamental aspects of endogenous BG self regulation, specifically, the action of the hormones glucagon and insulin.

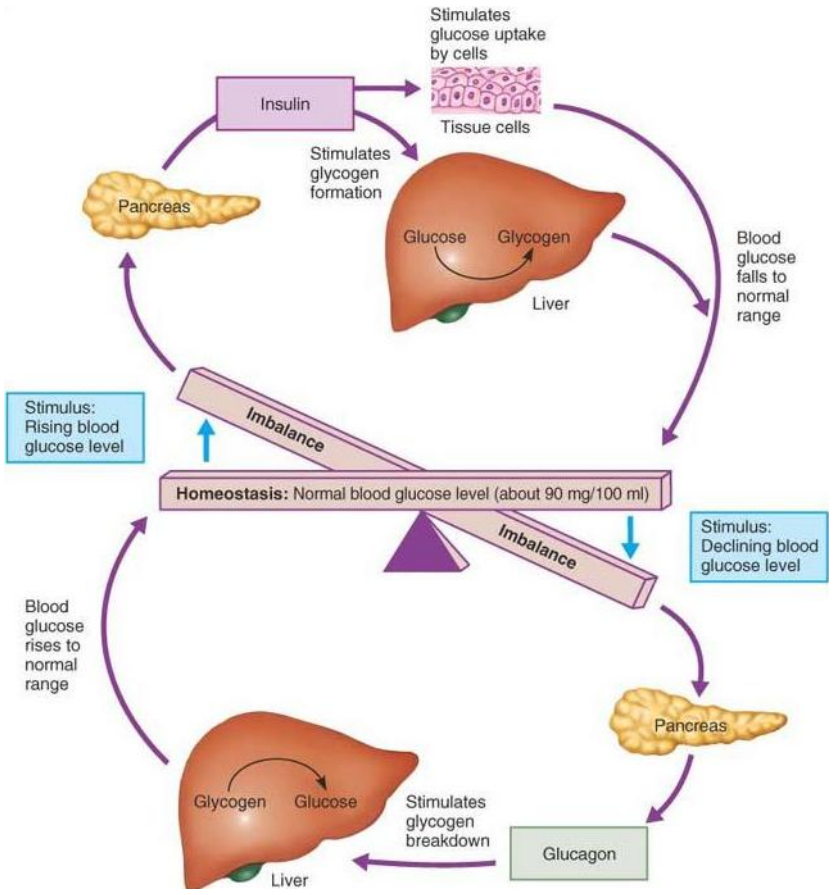


Figure 9.1 Schematic showing how BG concentration is controlled via a negative feedback system (Marieb & Hoehn 2005)

Low BG levels stimulate the alpha cells in the pancreas to release glucagon, which leads to an increase in BG concentration. This occurs predominantly in the liver where glucagon stimulates hepatic glucose output by accelerating glycogenolysis, and halting glycolysis to allow the reverse process gluconeogenesis. Briefly, glycogenolysis is the breakdown of glycogen, into glucose, and, gluconeogenesis is the formation of glucose from lactic acid and

certain amino acids. Importantly, both of these metabolic pathways result in an increase in plasma glucose concentration.

Conversely, high BG levels stimulate the beta cells in the pancreas to secrete insulin, which leads to a decrease in BG concentration. Unlike glucagon, which acts primarily in the liver, the great majority (80-90%) of insulin mediated glucose disposal occurs in muscle cells and adipose tissue. Insulin in the plasma binds to receptors on the plasma membrane of the cells, and through a series of protein activation cascades, the glucose transporter GLUT-4 is translocated to the plasma membrane. The presence of GLUT-4 on the plasma membrane facilitates the active transport of glucose from the plasma into muscle cells and adipose tissue, thus reducing BG concentration.

In addition to glucose disposal in muscle cells and adipose tissue, insulin also alters several other processes, which effect BG concentration. Insulin increases the rate at which the liver synthesizes glucose into glycogen for storage (glycogenesis), as well as slowing glycogenolysis and gluconeogenesis. Furthermore, insulin stimulates lipogenesis, a storage pathway where glucose is synthesised into fatty acids, which are eventually stored in adipose tissue.

Any abnormalities or deficiencies in these glucose regulatory processes at birth can cause newborns to suffer from metabolic problems, such as hypoglycaemia.

9.2 Neonatal hypoglycaemia

Neonatal hypoglycaemia is defined as low glucose levels in a newborn baby. Causes of neonatal hypoglycaemia include, but are not limited to: low glycogen stores, increased glucose use, decreased glycogenolysis or gluconeogenesis, or, depleted glycogen stores (Stanley & Baker 1999). Clinically, infants considered to be at risk of hypoglycaemia are those born preterm, small for gestational age (SGA), large for gestational age (LGA), infant of a diabetic mother (IDM), or sick (e.g. Sepsis, asphyxia) (Harris et al. 2010). In at risk babies it is recommended that BG concentration be monitored closely, as hypoglycaemia can present both

symptomatically (symptoms include jitteriness, irritability, lethargy, seizures, apnoea, grunting and sweating (Halamek et al. 1997)) and asymptomatically. Currently, a widely used definition of neonatal hypoglycaemia is a BG measurement less than 2.6mmol/L, which was deduced from two studies, published in 1988 (Koh et al. 1988; Lucas et al. 1988).

The first of these studies measured neurophysiological function via evoked potentials during periods of low BG concentration (Koh et al. 1988). The researchers sought to determine whether there was a measurable difference in evoked potentials between children who are 'symptomatic' and 'asymptomatic'. They recruited 17 subjects between 1 day and 16 years of age, but only 5 were younger than 4 months old. Each patient experienced either spontaneous hypoglycaemia, or, provoked hypoglycaemia by fasting or insulin administration. In each subject, either brainstem auditory or somatosensory evoked potentials were recorded immediately after every BG measurement.

The results showed 10 out of the 17 subjects had abnormal changes in evoked potentials, recorded at BG concentrations ranging 0.7-2.5mmol/L. However, in these same subjects, normal evoked potentials were recorded for BG concentrations ranging 1.9-5.6mmol/L. Furthermore, in the remaining 7 patients no changes in the evoked potentials were observed, despite 1 patient dropping to 1.9mmol/L. A conclusion of this study was that although it cannot be known whether transient abnormalities in the evoked potentials are associated with permanent brain injury, BG should be maintained above 2.6mmol/L to prevent these abnormalities. Other studies have not been able to demonstrate the same findings with evoked potentials and hypoglycaemia (Greisen & Pryds 1989).

The second study investigated the occurrence and persistence of hypoglycaemia and how it related to the neurodevelopmental outcome at 18 months of age (Lucas et al. 1988). This study recruited 661 newborn infants with a mean birth weight of 1337g and a mean gestational age of 30.5 weeks. These infants had weekly plasma BG measurements and more frequent plasma measurements when finger-stick BG measurements (done 6-hourly for the first 48-72 hours)

showed hypoglycaemia (1 fingerstick BG measurement below 1.5mmol/L, or, 2+ below 2.5mmol/L). A total of 543 infants were followed up at 18 months, at which time Bayley motor and mental development scores (Bayley 1965) were determined to quantify neurodevelopment.

The authors reported a strong association between the frequency of low blood glucose concentrations and poor Bayley scores at 18 months. In particular, babies who had five or more BG measurements below 2.6mmol/L on separate days during the first few months of life had lower Bayley scores. Thus, this study concluded that BG concentrations below 2.6mmol/L were associated with negative outcomes in these infants. However, when this cohort of children was assessed at ages 7.5 to 8 years the strong association between low blood glucose concentrations and neurological outcome was not apparent, casting doubt on earlier findings. Nevertheless, the hypoglycaemic threshold of 2.6mmol/L is still widely used for these infants (Burns et al. 2008; Duvanel et al. 1999).

The idea of using a 'one size fits all' threshold for hypoglycaemia is a highly contentious subject. The definition of hypoglycaemia almost certainly cannot be defined by a single number (Cornblath et al. 2000). It is highly likely that a safe glucose concentration at one point in time for a particular patient will be unsafe at a different point in time or for a different patient (Halamek et al. 1997). Furthermore, the wide range of inter-patient variability suggests that the definition should be patient-specific, depending on factors, such as the infants physiological maturity and the influence of pathology (Cornblath et al. 2000). However, due to the methodological limitations of measuring BG, no data currently exists that define the concentration of plasma glucose or its duration that causes damage. Thus, a BG measurement below 2.6mmol/L remains a widely used definition of hypoglycaemia in infants.

A major limitation of traditional BG monitoring is the frequency of BG measurements, which is often restricted due to blood draw constraints or the pain and discomfort it causes the infant. Typically, BG is measured every 1-4 hours in at risk patients and more frequently only if hypoglycaemia has

previously been diagnosed. In addition, each BG measurement represents a discrete value in time and little can be inferred about the BG concentration between measurements, as depicted schematically by Figure 9.2. Thus, it is possible for hypoglycaemia to occur between BG measurements and remain undetected.

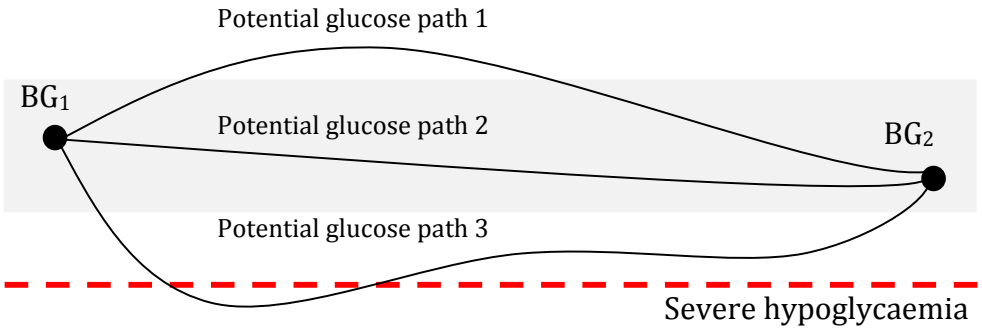


Figure 9.2 Individual BG measurements represent a discrete value in time and give little information about the BG concentration between measurements (3 possibilities shown), especially measurements taken several hours apart.

This limitation on the frequency of BG measurements has been especially problematic for clinical studies that have attempted to find a causal relationship between neonatal hypoglycaemia and a child's later development. However, the advancement of continuous glucose monitoring technology in the 1990's offered a potential alternative to traditional BG monitoring. In particular, the ability to define the BG trajectory between measurements in situations like that shown in Figure 9.2.

9.3 The role of CGM

The first commercially available CGM device was the MiniMed CGMS, which was FDA (Food and Drug Administration) approved in 1999 (Mastrototaro 2000). A later version, the CGMS System Gold (Medtronic Minimed, Northridge, CA) is shown in Figure 9.3 and was used throughout this neonatal application. Briefly, the CGMS System Gold consists of a small pager-like monitor that receives information from a sensor inserted into the subcutaneous layer, just beneath the skin (Klonoff 2005b). The SOF sensors used in this research utilise a glucose

oxidase reaction to produce a small electrical current that is proportional to the glucose concentration. The monitor provides a value every 5 minutes, but must be manually calibrated at least 3 times per day using independent BG measurements (Minimed 2003). Importantly, the CGM described here is a retrospective device and users remain blinded to glucose levels until after the monitoring period.

Figure 9.3: Medtronic Minimed CGMS System Gold used in this study

The first reported use of CGM in newborn babies during the neonatal period was published in 2005 (Beardsall et al. 2005). The aim of the study was to assess the validity of using CGM devices in very low birth weight (VLBW) infants and to obtain more data on the prevalence of abnormalities in glycaemia. Sixteen babies were recruited into this study with a median gestational age of 26.5 weeks and a median birth weight of 706 grams. Monitoring with CGMS began in the first 24 hours after birth and lasted for up to 7 days which resulted in a total of 1302 hours of SG data. Across the cohort, the reported median (range) percent duration of hypoglycaemia was 0.4% (0-11%) and all episodes were asymptomatic. Three babies (19%) had hypoglycaemia that was not detected clinically and one baby (6%) had a hypoglycaemic episode that lasted 2 hours 45 minutes before it was detected by BG measurement. Accuracy of CGM during periods of hypoglycaemia could not be assessed due to a low number of BG measurements at that level, but the study concluded CGMS is practical for

glucose monitoring in VLBW infants and hypoglycaemia is prevalent in this population.

In 2009, a second study using CGM in VLBW infants was published (Iglesias Platas et al. 2009). Thirty-eight VLBW infants were recruited in the first 24 hours after birth and monitoring continued for up to 14 days. The cohort had a median gestational age of 27.6 weeks and a median birth weight of 958 grams. Results of this study showed that 14 patients (37%) experienced hypoglycaemia and the average percentage of time in the register with hypoglycaemia was 0.55%. Furthermore, an analysis of each day showed that episodes of hypoglycaemia appeared evenly throughout the study period with 2.6-13.2% of patients affected per day, after the first day of life. Similar to the study in 2005, this study concluded that CGMS is a feasible way of studying glucose levels in VLBW infants and, when applied, it reveals "abnormal glycaemia" at a much higher rate than standard intermittent BG measurements.

These two studies of VLBW infants were followed by an investigation of CGM in 102 larger babies considered at risk of hypoglycaemia (Harris et al. 2010). The cohort had a median birth weight of 2327 grams and the median gestational age of 35 weeks. From the 102 babies recruited, hypoglycaemia was detected in 45 babies using CGM and in 32 babies using intermittent BG measurements. CGM detected 265 episodes of hypoglycaemia lasting between five and 475 minutes, with 107 episodes lasting 30 minutes or more. Astonishingly, only 19% of hypoglycaemic episodes were detected by standard, clinically specified BG measurements, further highlighting the potential benefits of CGM in neonates. Recurrent hypoglycaemia presented in 21 babies who had repeated episodes lasting longer than 30 minutes. Similar to previous studies, CGM was well tolerated and remained in use for up to 7 days.

All three studies reported that CGM was well tolerated by the infants and didn't interfere with nursing care. Abnormalities were detected at a much higher rate compared to intermittent BG monitoring and SG values provided additional information about the duration and severity of hypoglycaemic episodes.

However, there were concerns about the accuracy of SG values and whether they could be relied on at low concentrations of blood glucose (Beardsall et al. 2005). Overall, the agreement between CGMS and glucometer BG readings is acceptable, but CGM tends to provide a higher BG value at low BG levels and accuracy is lowest on day 1 of monitoring (Beardsall et al. 2013).

To draw sound conclusions from studies utilising CGM, it is not only important to have a thorough understanding of the physiology, but also the CGM technology used to collect data. Furthermore, signal processing and mathematical techniques can be used to enhance the quality of data or aid in the interpretation of results. The following chapters investigate CGM in critically ill newborn babies, with a focus on hypoglycaemia detection and quantification.

This research is part of a wider study, The CHYLD Study (Children with HYpoglycaemia and their Later Development), which is a large multi-disciplinary prospective study investigating the development of young children who were at risk of developing hypoglycaemia in the early neonatal period. The children are followed up at 2 and 4.5 years of age at which point their growth, cognitive and neurodevelopment are assessed in relation to the duration, severity, and frequency of hypoglycaemia in the first days of life. Ultimately, this research aims to improve the definition, diagnosis and management of neonatal hypoglycaemia, including the use of CGM.

Chapter 10. CGM device calibration

If CGM devices are to be used effectively in the clinical environment, it is important to understand how the sensors/monitors work and any limitations associated with their use. Certain aspects of CGM, such as the algorithm used to calibrate SG data, can be adjusted or modified to better suit a specific application. This chapter focuses on how SG outputs are determined, and specifically, the effect calibration algorithms can have on these output measurements. The emphasis here is on neonatal hypoglycaemia, so this study investigated how device calibration affected metrics used to quantify hypoglycaemia in neonates.

10.1 Introduction

The CGM system used in this research consisted of a Medtronic Minimed CGMS System Gold monitor coupled with a Medtronic Minimed SOF sensor. The SOF sensors are inserted using a specifically designed device that ensures the process takes only a few seconds and is nearly pain free. Once in place, SOF sensors indirectly measure glucose in the interstitial fluid using amperometric enzyme electrodes based on glucose oxidase.

This particular sensor uses a 3-electrode configuration, consisting of a platinum working electrode where oxidation occurs, a counter electrode where reduction occurs, and a reference electrode (Rossetti et al. 2010). Glucose oxidase (GO_x) is immobilised at the working electrode and catalyses the oxidation of interstitial glucose into gluconolactone, per the reaction defined:

$$Glucose + GO_x(FAD) \rightarrow Gluconolactone + GO_x(FADH_2) \qquad 10.1$$

The cofactor Flavin Adenine Dinucleotide (FAD) acts as electron acceptor, reducing to $FADH_2$. The $FADH_2$ is re-oxidised in the presence of oxygen to produce hydrogen peroxide (H_2O_2), in a reaction defined:

$$GO_x(FADH_2) + O_2 \rightarrow GO_x(FAD) + H_2O_2 \quad 10.2$$

The physical makeup of these sensors also includes several permeable polymer membranes surrounding the electrodes. These membranes were designed to address two significant issues common with subcutaneous glucose sensors (McGarraugh 2009):

1. Oxygen deficit: The transport of glucose is restricted to maintain a proper glucose-oxygen ratio, as oxygen concentrations are typically 1 order of magnitude lower than glucose concentrations.
2. Interference from other molecules: Molecules such as uric acid and absorbic acid may interfere with the process, so their access is restricted.

The H_2O_2 permeates through the membranes to the platinum working electrode, which is poised at 0.535 V relative to the reference electrode (Gross et al. 2000). Oxidation of H_2O_2 at the working electrode produces an electrical current (I_{SIG}), which is in the order of nano amps, defined by the $2e^-$ term in the reaction:

$$H_2O_2 \rightarrow 2H^+ + O_2 + 2e^- \quad 10.3$$

The CGMS System Gold monitor samples the I_{SIG} every 10 seconds. Over a period of 1 minute, the highest and lowest values are discarded and the remaining 4 measurements are averaged to create interval values (Mastrototaro et al. 2002). Over a period of 5 minutes, the highest and lowest interval values are excluded, and the remaining 3 interval values are averaged and stored internally by the monitor (Mastrototaro et al. 2002). An empirically derived 10 minute time delay is added to stored values to account for delays between capillary BG concentration and interstitial fluid glucose concentration. Several possible sources of delay have been reported including a physiological lag of ~3-12 minutes, electrochemical sensor lag of 1-2 minutes, and additional lags due to filtering (Keenan et al. 2009). The time-shifted I_{SIG} values are considered *Valid* I_{SIG} and form the basis of the SG output. The CGMS System Gold is a retrospective device and no glucose data is displayed in real time.

To obtain meaningful glucose levels from the 5-minute *Valid* I_{SIG} values, independent reference measurements of BG are required for calibration. Calibration BG measurements are typically obtained using a finger-stick glucose meter, similar to those used by patients with diabetes. Finger-stick BG devices are reported to have errors in the range of 2-10% (Abbott 2010; Roche 2007,2008; Solnica et al. 2003), and often have reduced performance in critically ill patients due to varying levels of haematocrit, medication and other effects (Hoedemaekers et al. 2008; Kanji et al. 2005; Karon et al. 2007). Further complicating the matter, it is not uncommon for term newborn babies to have elevated haematocrit levels during the first few days of life, reaching 65%+ in some cases (Jopling et al. 2009).

The manufacturer specified a calibration procedure that involves an initial calibration approximately 1 hour after sensor insertion and then at least 4 times per day (Minimed 2003). Calibration BG measurements are manually entered into the device by the user and the value is stored until the end of the monitoring period. At the completion of monitoring, data from the CGMS System Gold is downloaded to a personal computer (PC) using CGMS system solutions software V3.0C. The retrospective calibration algorithm uses the manually entered calibration measurements to determine calibration parameters, which convert the *Valid* I_{SIG} data into glucose data.

Exact details of the calibration algorithms are not disclosed in the literature due to their commercial sensitivity, but several sources suggest the retrospective calibration algorithm used by the CGMS Systems Solutions Software employs linear regression over a series of calibration measurements (Klonoff 2000; Mastrototaro et al. 2002). Linear regression calibration methods may contribute to the CGMS System Gold reporting higher values during hypoglycaemia and reporting lower values during hyperglycaemia (Beardsall et al. 2013; Voskanyan et al. 2007). However, it may also balance the impact of errors in finger stick glucose meters by not over-relying on any single calibration measurement. Thus, important clinical observations, such as excursions from normal BG levels will be

directly affected by the calibration algorithm used and the quality of calibration BG measurements.

Several previous studies have investigated CGM calibration schemes, both retrospective and real-time, in adults and children (Bequette 2010; Buckingham et al. 2006; King et al. 2007). However, in this study, blood gas determinations of BG concentrations were available for the dataset and are assumed to be a "gold-standard" assessment due to their very low error. Alternative calibration algorithms can be applied to the CGM readings, utilising the high accuracy BG measurements, and compared to the data calibrated using the manufacturer algorithm. This chapter explores and quantifies the impact of calibration measurement quality, calibration method, and non-linear filtering on metrics of hypoglycaemia in neonates using CGM devices.

10.2 Subjects and Methods

10.2.1 Subjects

This study used CGM data from 50 babies ≥ 32 weeks gestation that were at risk of hypoglycaemia and admitted to the Waikato Hospital NICU. Exclusion criteria included serious congenital abnormities or a skin condition preventing continuous glucose monitoring. Informed consent was obtained from parents of all babies in the study. The study was approved by the Northern Y Regional Ethics Committee (New Zealand). Table 10.1 shows the demographics for patients enrolled into the study.

Table 10.1 Cohort demographics

Cohort Demographics	
Number of CGM traces	50
Sex (M/F)	26/24
Gestational Age (weeks)	34 [33 - 37]
Birthweight (g)	2172 [1880 - 2990]
Primary Risk (# infants):	
Diabetes	15
Premature	19
Small or Large for gestational age	14
Other	2

10.2.2 Continuous Glucose Monitoring

All patients had interstitial glucose monitoring using CGMS System Gold devices and SOF sensors. Sensors were located on the infant's lateral thigh and were connected to the monitor via the supplied cable. Monitoring began on admission to the NICU and finished after 7 days or when the baby was no longer considered at risk of hypoglycaemia. During the monitoring period nurses recorded all BG concentrations, feeding and medication for the management of hypoglycaemia. Nurses remained blinded to SG measurements and all calibration BG measurements were entered per the manufacturer's recommendations. Data were downloaded to a PC using CGMS system solutions software version 3.0C, which calibrated the SG data retrospectively.

10.2.3 Calibration Measurements

Blood samples were taken by nursing staff via heel-pricking at one hour of age, then two to four hourly (before feeds) for 12 hours. In babies receiving intravenous dextrose, BG was measured 4 hourly for 12 hours, and then less frequently as clinically indicated. The median [IQR] interval between calibration BG measurements was 4.8 [3.5 – 6.4] hours.

All BG calibration measurements were made using a BGA (Radiometer, ABL800Flex, Copenhagen). This device has a reading range of 0.0 to 60.0mmol/L and a C.V. of 1.4-2.2% (Cembrowski et al. 2010; Watkinson et al. 2012). Furthermore, a study of a device using the same glucose electrode reported a coefficient of variation of 2.1% in ICU patients with little or no decrease in performance due to haematocrit, pH or PaO_2 (Watkinson et al. 2012). Due to the location of the BGA in this study, a short time delay (estimated < 15mins maximum) was possible between taking the blood sample and obtaining the glucose concentration.

10.2.4 Calibration Algorithms

The factory calibration algorithm used by the CGMS is based on linear regression (Chee et al. 2002; Mastrototaro et al. 2002). The underlying processes used in the factory calibration are described in Figure 10.1 (Mastrototaro et al. 2002).

Specifically, when only 1 calibration BG is available the algorithm will perform a single point calibration (Figure 10.1a). When multiple calibration BG are available the algorithm will perform a linear regression calibration (Figure 10.1b).

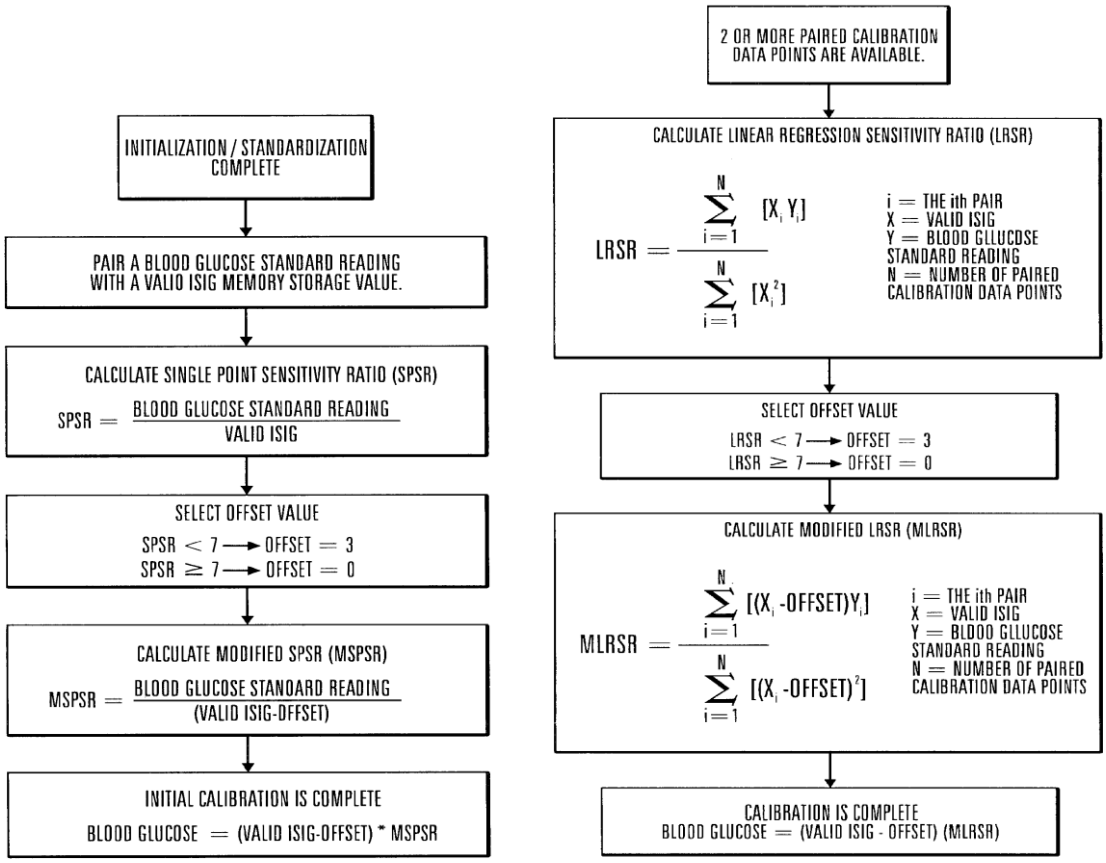


Figure 10.1 a (left) - Underlying manufacturer calibration process used when a single calibration BG is available. b (right) - Underlying manufacturer calibration process used when multiple calibration BG values are available. (Mastrototaro et al. 2002)

With multiple calibration BG measurements available, the algorithm uses linear regression to calculate a sensitivity ratio. The sensitivity ratio quantifies the expected electrical current from the sensor in response to an external glucose stimulus. The magnitude of the sensitivity ratio is used to determine the *Offset*, which accounts for any non-linear behaviour in the sensor. It is important to note that the *Offset* values in the dataset used in this study contain other values in addition to the 0 or 3 stated in Figure 10.1, suggesting the algorithm has evolved since the filing of the patent (Mastrototaro et al. 2002), but that a similar overall approach is used.

A modified linear regression sensitivity ratio is calculated using the adjusted *Valid* I_{SIG} and calibration BG pairs. This sensitivity, or *Slope*, is then used as the gain/multiplier to give a BG estimate from the *Valid* I_{SIG} values that are not paired with calibration BG measurements. The algorithm described in (Mastrototaro et al. 2002) suggests the calibration parameters were held constant for 24 hours, starting and finishing at midnight. However, the dataset used in this study has a *Slope* function that gradually changes over time, as shown in Figure 10.2, suggesting the version of algorithm used in this study incorporates additional smoothing and/or an adjustable calibration window. However, the manufacturer calibration method used with the CGMS System Gold is still based on linear regression, which is the aspect of interest for this work.

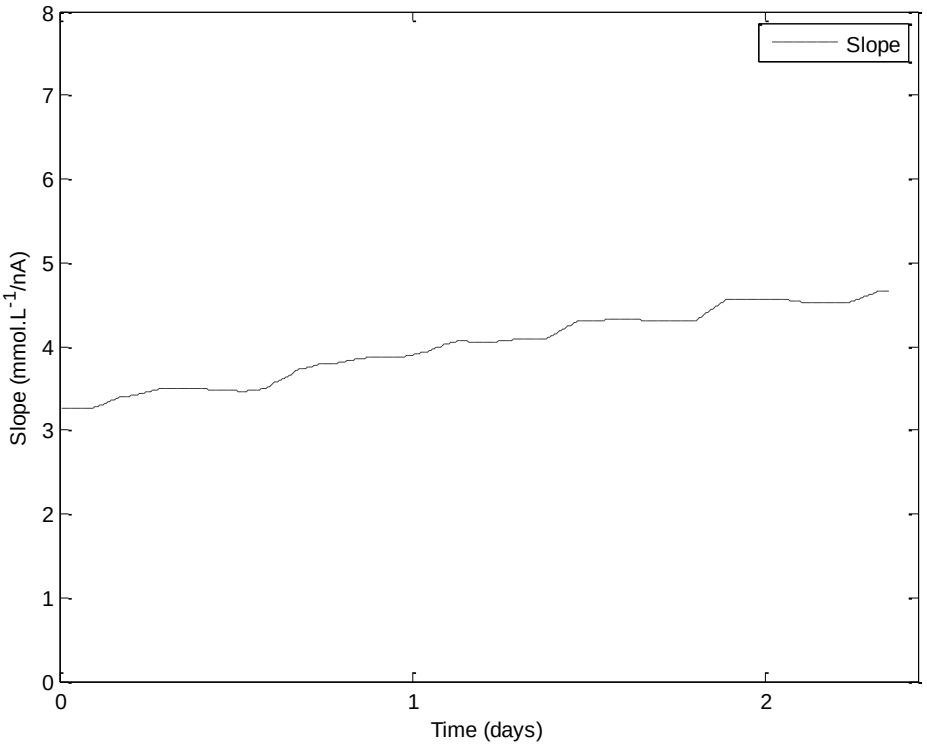


Figure 10.2 Representative *Slope* function from a patient monitored in this study

The linear regression algorithm is aimed primarily at ambulatory individuals with type 1 diabetes who use the CGM device to help manage BG levels. One reason to utilise linear regression is that this population typically uses a finger stick glucometer for calibration, which analyses capillary BG and typically has up to 10% error (Abbott 2010; Roche 2007,2008; Solnica et al. 2003). The use of

linear regression implicitly balances errors in calibration BG measurements and *Valid* I_{SIG} data. Hence, SG outputs do not necessarily exactly correspond to BG measurements, as shown in Figure 10.3.

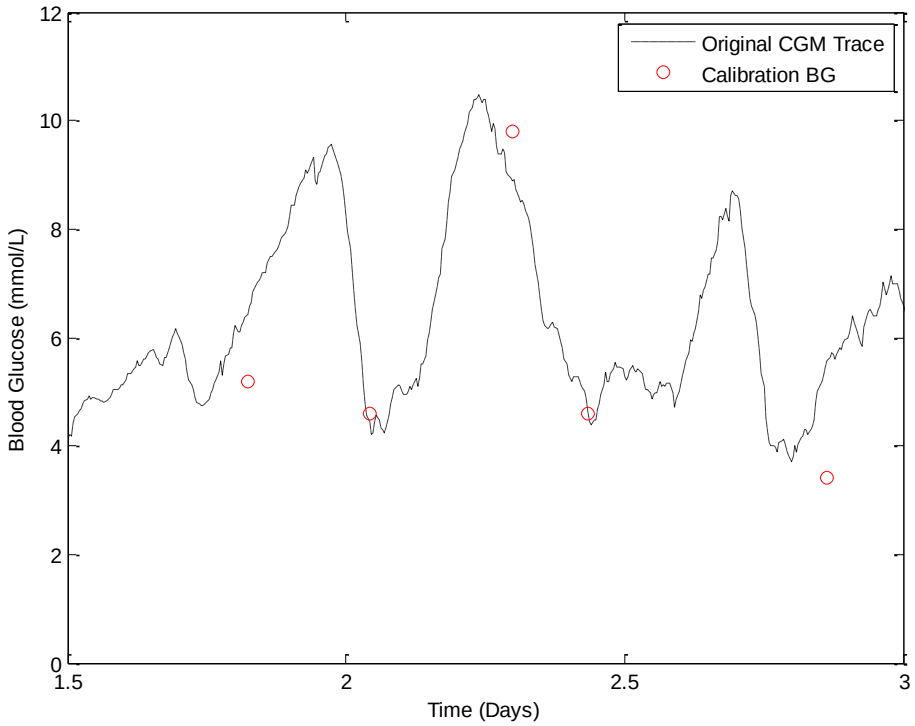


Figure 10.3 Representative section of manufacturer calibrated SG data showing discrepancies between SG data and calibration BG measurements.

The calibration BG measurements in this study were determined using a BGA with a reported level of error between 1.4-2.2% (Cembrowski et al. 2010; Watkinson et al. 2012), which is significantly more accurate than typical glucometers. Thus, data were recalibrated to make better use of the accurate calibration measurements by forcing the SG output through all calibration BG measurements. It should be noted there are many ways that the data could be recalibrated and the algorithm used in this study represents just one example based as directly as possible upon the details available for the current method (Bequette 2010; Mastrototaro et al. 2002; Minimed 2003).

The CGM blood glucose estimation is determined using:

$$BG_{CGM} = Slope\ (Valid\_I_{sig} - Offset) \quad 10.4$$

Where:

- BG_{CGM} – Blood glucose estimate by the CGM (mmol.L^{-1})
- *Slope* – Calibration parameter found using linear regression (mmol.L^{-1}.nA^{-1}). Note the parameter *Slope* is referred to as MLRSR in Figure 10.1.
- *Valid* I_{SIG} – Electrical current detected by the monitor from the sensor (nA)
- *Offset* – Calibration parameter to compensate for non-linear sensor behaviour, typically observed at high glucose levels.

To recalibrate, Equation 10.4 is rearranged to:

$$Reqd\ Slope(i) = \frac{BG_{Cal}(i)}{(Valid\ I_{sig}(i) - Offset(i))} \quad 10.5$$

Where:

- BG_{Cal} – Blood glucose level for calibration (BGA measurement) (mmol.L^{-1})
- *Reqd Slope* – *Slope* that forces BG_{CGM} through calibration measurements (mmol.L^{-1}/nA)

The recalibration algorithm adjusts the *Slope* parameter and forces the output SG trace to pass through all calibration BG measurements, while preserving the raw sensor current, *Valid* I_{SIG}, and *Offset* parameter. At each calibration measurement a value of *Reqd Slope* is calculated using Equation 10.5. Linear interpolation of *Reqd Slope* gives the new, continuous S*lope* function. Figure 10.4 shows a manufacturer calibration *Slope* function compared to the recalibrated *Slope* function.

The recalibrated *Slope* function is inserted back into Equation 10.4 with the unmodified *Valid* I_{SIG} and *Offset* parameters, to give recalibrated SG values. Recalibrated SG data provides a comparator to assess the impact of calibration on outcome SG trace. A representative example of the effect of calibration on SG data is shown in Figure 10.5.

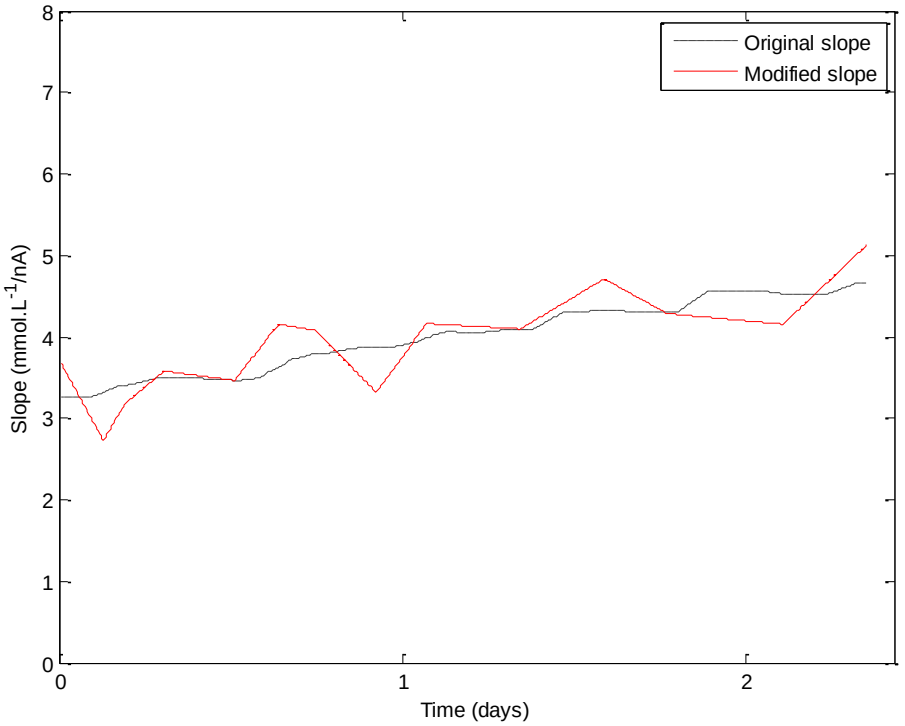


Figure 10.4 Comparison of a *Slope* function generated by the manufacturer's calibration algorithm (black dashed line) and the re-calibrated *Slope* parameter (red line).

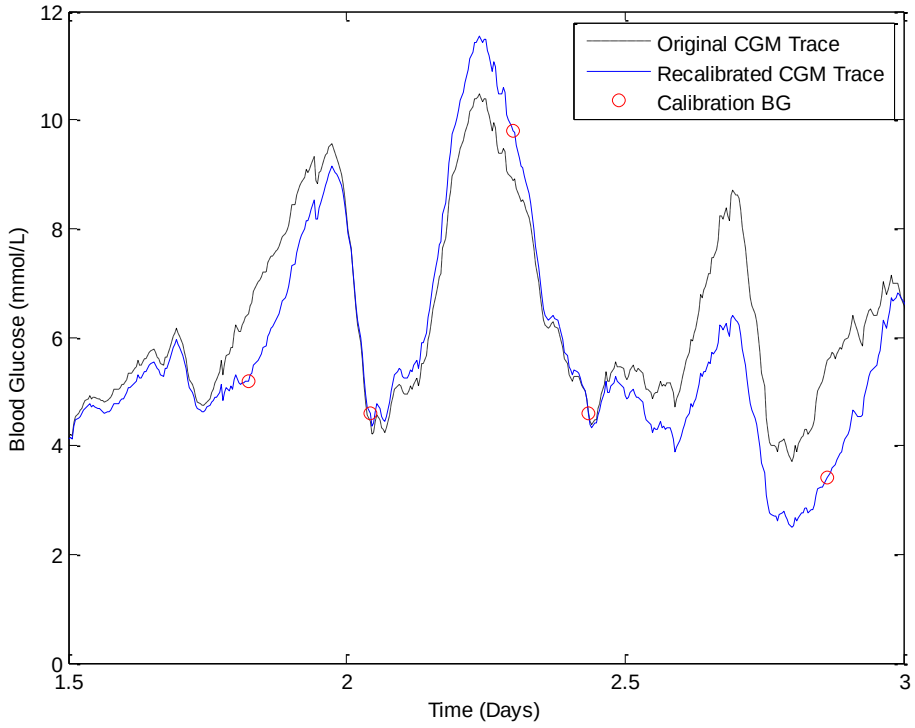


Figure 10.5 Recalibrated SG data (blue solid line) passes through all calibration BG measurements, compared to the same sensor data calibrated using the manufacturer's algorithm.

10.2.5 Non-linear Filtering

Non-linear filters can be used to remove unwanted and potentially un-physiological high-frequency noise from the CGM signal. In particular, median filters have proven to be a simple and effective method of removing this noise and smoothing SG traces (Bequette 2010; Pretty et al. 2010). Several median filters of different lengths were tested, including composite median filters. Figure 10.6 shows a recalibrated SG trace and the effects of 3 different median filters.

A retrospective composite median filter was used in this study because it allows faster and slower glucose dynamics to be captured more effectively. The filter averages a 3 point median and a 7 point median both centred about the time point of interest (Pretty et al. 2010). The filter was implemented both prior to recalibration (on the *Valid* I_{SIG}) and post calibration on the SG output to assess any differences due to where in the process filtering occurs.

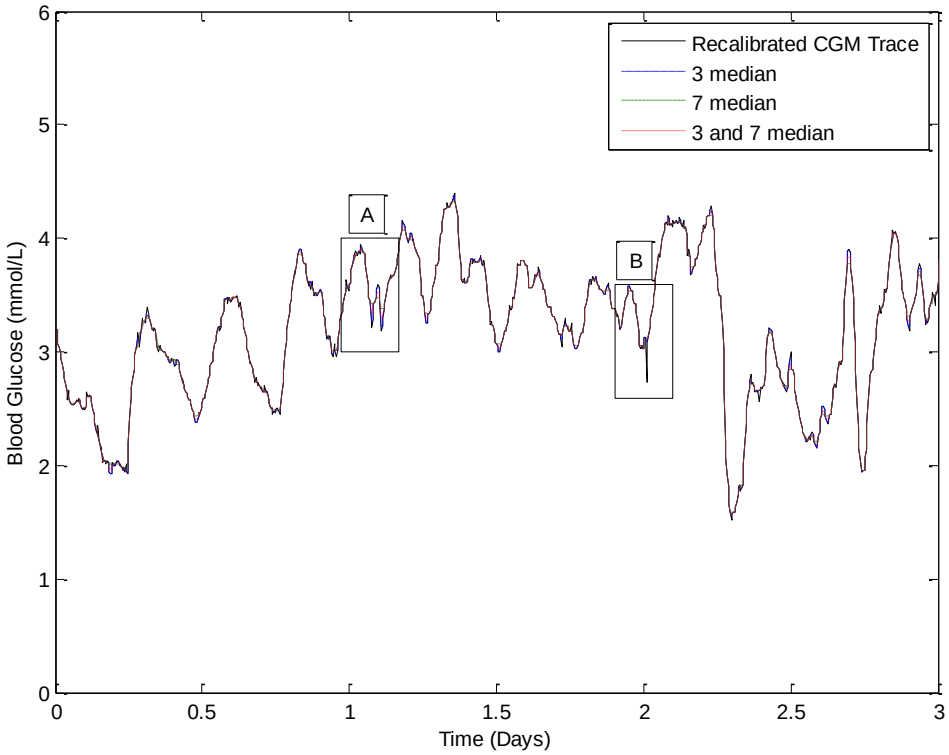


Figure 10.6 Representative SG data with non-linear median filtering from 3 different median filters. Sections 'A' and 'B' highlight the effects of each individual filter on CGM dynamics that were often observed in this data set.

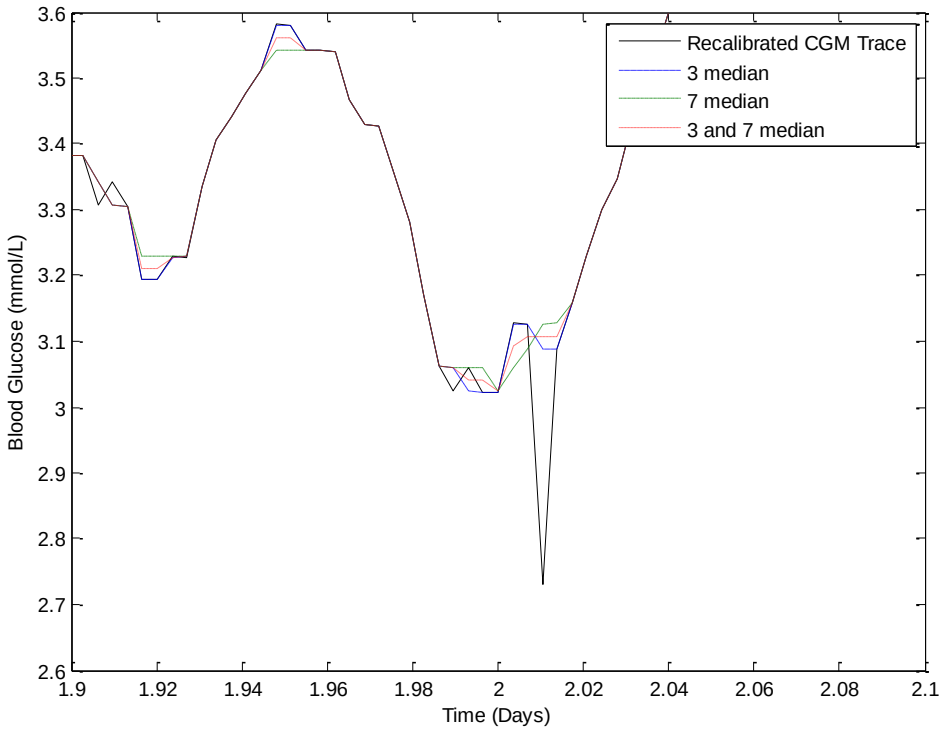


Figure 10.7 Enlarged area 'A' in Figure 10.6 showing a sudden, potentially un-physiological, drop in the SG data that has been removed by the filter. All three filters have effectively removed the artefact.

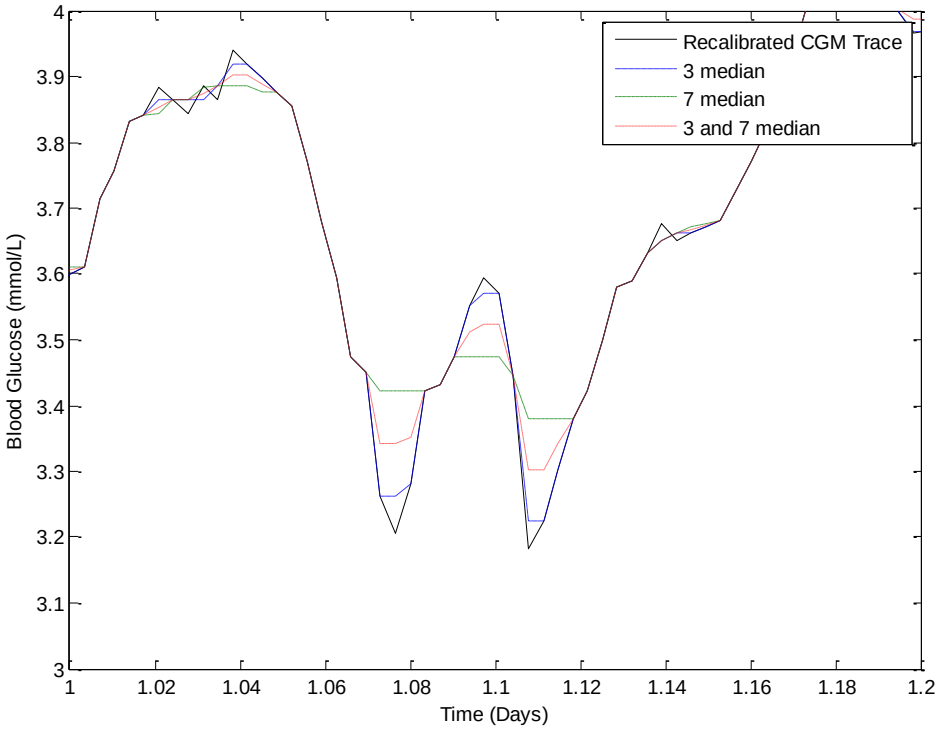


Figure 10.8 Enlarged area 'B' in Figure 10.6 showing the effect of filter length on the degree of filtering. The composite 3 and 7 length median filter removes high frequency changes, while capturing the overall trend of the SG data.

10.2.6 Analyses

10.2.6.1 Overall cohort

Four analyses of the CGMS data from the 50 babies were performed in this study:

1. Original SG output
2. Recalibrated SG output
3. Recalibrated and median filtered SG output
4. Filtered *Valid* I_{SIG} then recalibrated SG output.

Each of the recalibrated variations, with and without filtering, is compared to the original SG output to see the effect of recalibrating/filtering on clinical measures of hypoglycaemia. Hypoglycaemia was defined as one or more consecutive SG measurement(s) below 2.6mmol/L, surrounded by SG measurements greater than or equal to 2.6mmol/L. The metrics used to quantify hypoglycaemia were:

- **Number:** Number of independent hypoglycaemic events
- **Duration:** Percent of SG record below 2.6mmol/L
- **Severity:** Lowest measurement of hypoglycaemic event.
- **Hypoglycaemic index:** Similar in concept to Hyperglycaemic index (Vogelzang et al. 2004) and defined as the area between the 2.6mmol/L threshold and the SG trace (for SG trace < 2.6mmol/L) summed over the entire length of stay, normalised by the length of data record. Note: the units used in this study are μmol/L, not mmol/L.

10.2.6.2 Per-patient hypoglycaemic index by day

A subset of 43 infants with at least 3 days of SG data were analysed for hypoglycaemic index on a day-by-day basis. This analysis looked at the re-occurrence rate of hypoglycaemia in each individual. Seven patients were excluded from the analysis due to having less than 3 days of SG data. For each day, using the original SG output, patients were ranked by hypoglycaemic index. The ranks were preserved for the three recalibrated and filtered CGM analyses to highlight changes in hypoglycaemic index for individual patients. For example, a patient with a high hypoglycaemic index using default manufacturer calibration may have a lower hypoglycaemic index when recalibrated. A further integral index metric was used to represent the total hypoglycaemic index across the 43 patients on that day, quantifying this per-patient result to a single value over the cohort for easy comparison.

10.3 Results

10.3.1 Overall cohort

Figure 10.9 shows the distributions of CGM errors, calculated as CGM minus BG, at the time of calibrations. Specifically, these discrepancies represent the residual difference between SG values and calibration BG values, after calibrating using the manufacturer's algorithm. Thus, the accuracy of the device cannot be deduced from these plots, as all BG measurements were used for calibration and the data set contained no independent BG values that could be used as references. The top plot shows the distribution for all paired measurements (1074 pairs). The middle plot shows the distribution of errors when either the CGM or BG measurement was less than or equal to 3mmol/L (145 pairs). The bottom plot shows the distribution of errors when either the CGM or BG measurement was greater than 7mmol/L (62 pairs). In all plots the dashed vertical lines represent the median, and the solid vertical lines represent the IQR.

In terms of capturing hypoglycaemia, the calibration BG measurements detected 53 hypoglycaemic events during the study. Of these, 16 were at a time when the CGM was also hypoglycaemic. However, the other 37 hypoglycaemic calibration BG measurements occurred at a time when the baseline CGM was above 2.6mmol/L.

Figure 10.10 shows a section of SG trace comparing original (black dashed line), recalibrated (solid blue line), Recalibrated and filtered (green dotted line), and Filtered *Valid* I_{SIG} then recalibrated variations (purple dash-dot line). In this example, overall trends in calibration parameters and SG output are preserved. However it is clear in Figure 10.10 that metrics of hypoglycaemia will vary depending on the method of signal processing used.

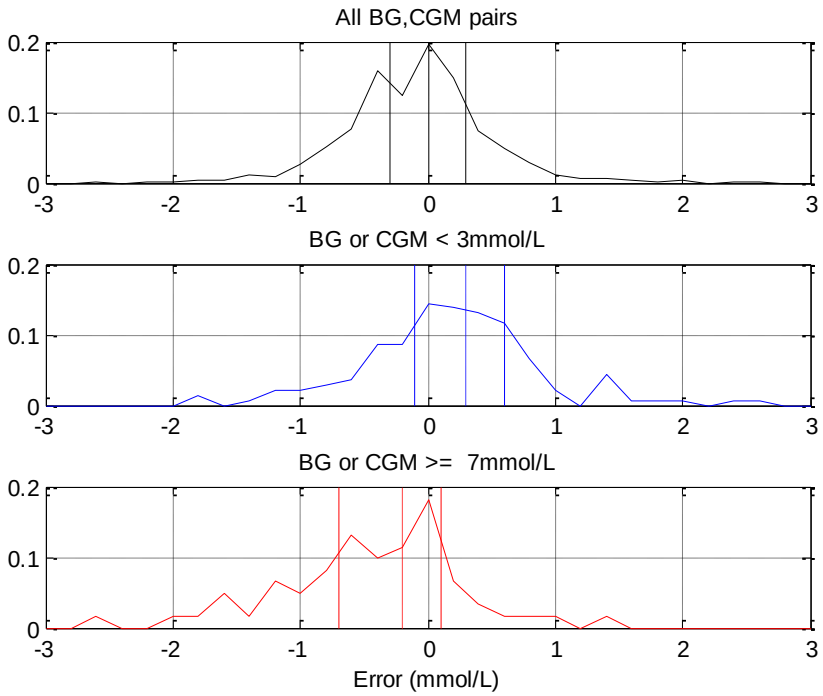


Figure 10.9 Distribution of errors between CGM and BG measurements with median (dashed vertical line) and inter-quartile range (solid vertical lines), for different glucose levels.

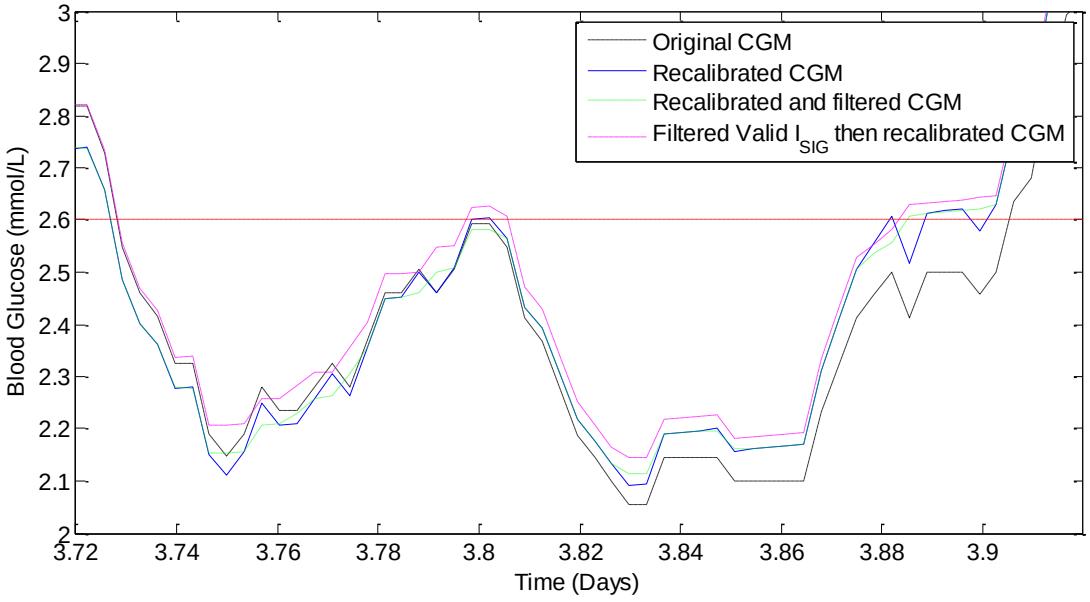


Figure 10.10 Comparison of a section of SG trace containing hypoglycaemia for Original CGM, Recalibrated CGM, Recalibrated and filtered CGM, and Filtered *Valid* I_{SIG} then recalibrated CGM. Hypoglycaemia is defined as one or more consecutive SG measurement(s) below 2.6mmol/L, surrounded by one or more SG measurement(s) above 2.6mmol/L. Note: recalibrating increases the number of hypoglycaemic events from 1 to 4, then filtering reduces it back to 1 in this example.

Figure 10.11 shows the SG trace in this study that had the largest change in hypoglycaemia metrics after recalibration. The original SG trace (black dashed line) suffered from particularly bad factory calibration for the first 12 hours of monitoring, during which time there was a significant period of hypoglycaemia. However, all four calibration BG measurements in the first 12 hours were above 2.6mmol/L, and consequently the recalibrated SG trace (solid blue line) had no hypoglycaemia.

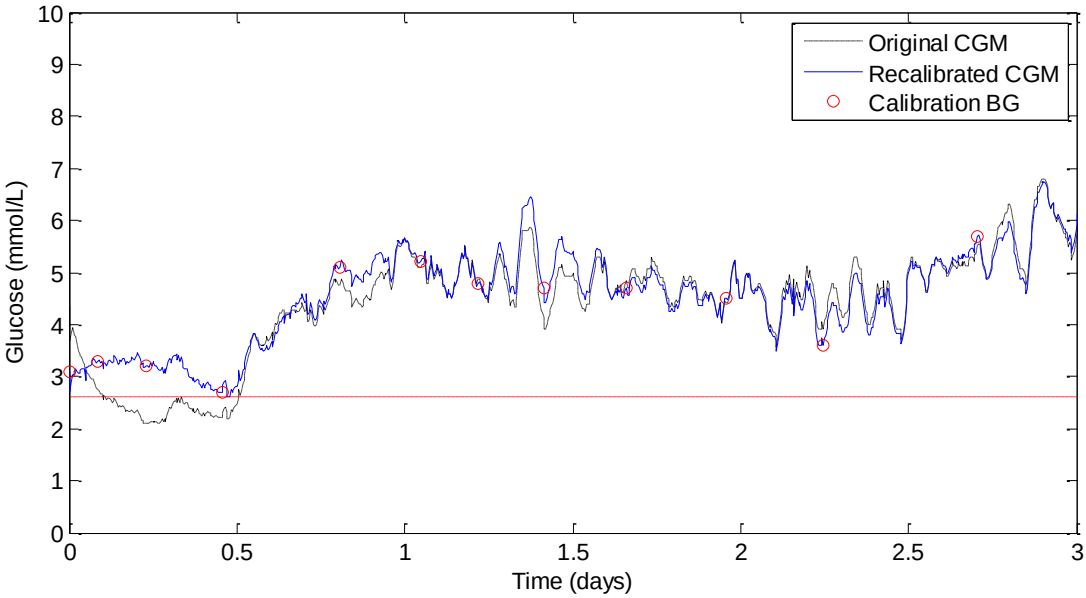


Figure 10.11 The SG trace that had the largest change in hypoglycaemia metrics after recalibration. The original SG trace (black dashed line) contains a long period of hypoglycaemia in the first 12 hours of monitoring. However, all four calibration BG measurements in this period were above 2.6mmol/L and consequently the recalibrated SG trace (blue solid line) had no hypoglycaemia.

The top section of Table 10.2 compares the number, duration and severity of hypoglycaemia events, as well as the hypoglycaemic index for each variation of the CGM calibration. The results are presented for the overall cohort and per-patient to show any potential skewed results from individual patients. The bottom section of Table 10.2 shows the effect of recalibration and/or filtering on the hypoglycaemic state of patients over the monitoring period. For example, the top left cell in bottom section of Table 10.2 ("24") indicates that 24 patients who had hypoglycaemia in the original SG data still had hypoglycaemia in the

recalibrated SG data. Conversely, the second row in column one ("1") indicates that 1 patient who had hypoglycaemia in the original SG data had no hypoglycaemia in the recalibrated SG data.

10.3.2 Per-patient hypoglycaemic index by day

Figure 10.12 shows the effect of recalibration and filtering on the hypoglycaemic index of individual patients on a day-by-day basis. The top row of Figure 10.12 shows the hypoglycaemic index for each patient over the first 24 hours of monitoring. The second row of Figure 10.12 shows the second 24 hour period and the third row shows the third 24 hours. From left to right in each row, the change in hypoglycaemic index for each patient due to calibration and/or filtering can be seen (the patient order is preserved across these plots). For example, the infant with most hypoglycaemia by original CGM calibration on day 1, shown in Figure 10.11, had no hypoglycaemia with any of the recalibration methods. Finally, the integral index summarising these results is also indicated on each panel of Figure 10.12.

10.4 Discussion

The aim of this study was to investigate how calibration and nonlinear filtering of SG data affect metrics of hypoglycaemia in at risk preterm babies. This knowledge is important for accurately relating hypoglycaemia to long term outcomes.

10.4.1 Overall cohort

The top section of Table 10.2 shows all metrics of hypoglycaemia increased after recalibration compared to original CGM results, which can potentially be explained by skew in the distribution of BG vs. CGM readings at low BG concentrations. Figure 10.9 shows the distribution of errors between the CGM-BG paired measurements. The data set contains 1074 paired BG-SG measurements of which 51% have a BG measurement higher than the CGM and 49% have a BG measurement lower than the CGM, and these data are, overall, relatively centred, as expected from the regression aspect of calibration.

Table 10.2 Effect of recalibration and filtering on recorded CGM hypoglycaemia for the entire cohort and per-patient. Results are presented as median [Inter-quartile range] where applicable.

Overall cohort results	**Original CGM**	**Recalibrated CGM**	**Recalibrated and Filtered**	**Filtered Valid ISIG then Recalibrated**
Number of Hypoglycaemic events	161	193	131	146
Duration (% of CGM record < 2.6mmol/L)	2.2	2.6	2.5	2.6
Hypoglycaemic index (µmol/L)	4.9	7.1	6.9	6.8
Hypoglycaemia events between 2.4-2.6mmol/L	87	87	51	61
Hypoglycaemia events between 2.2-2.4mmol/L	35	40	35	34
Hypoglycaemia events between 2.0-2.2mmol/L	18	38	23	30
Hypoglycaemia events less than 2.0mmol/L	21	28	22	21
Number of patients with no hypoglycamia	25	19	21	19
Per-patient results				
Number of Hypoglycaemic events	1 [0 - 3]	1 [0 - 4]	1 [0 - 3]	1 [0 - 3]
Duration (% of data hypoglycaemic)	0.1 [0.0 - 1.0]	0.5 [0.0 - 2.4]	0.5 [0.0 - 2.3]	0.6 [0.0 - 2.4]
Hypoglycaemic index (µmol/L)	0.0 [0.0 - 2.1]	0.8 [0.0 - 7.1]	0.5 [0.0 - 6.0]	0.4 [0.0 - 5.6]

Hypoglycaemic State	**Recalibrated CGM**	**Recalibrated and Filtered**	**Filtered Valid ISIG then Recalibrated**
Original hypoglycaemia --> Hypoglycaemia (# patients)	24	22	23
Original hypoglycaemia --> No hypoglycaemia (# patients)	1	3	2
Originally no hypoglycaemia --> Hypoglycaemia (# patients)	7	7	8
Originally no hypoglycaemia --> No hypoglycaemia (# patients)	18	18	17

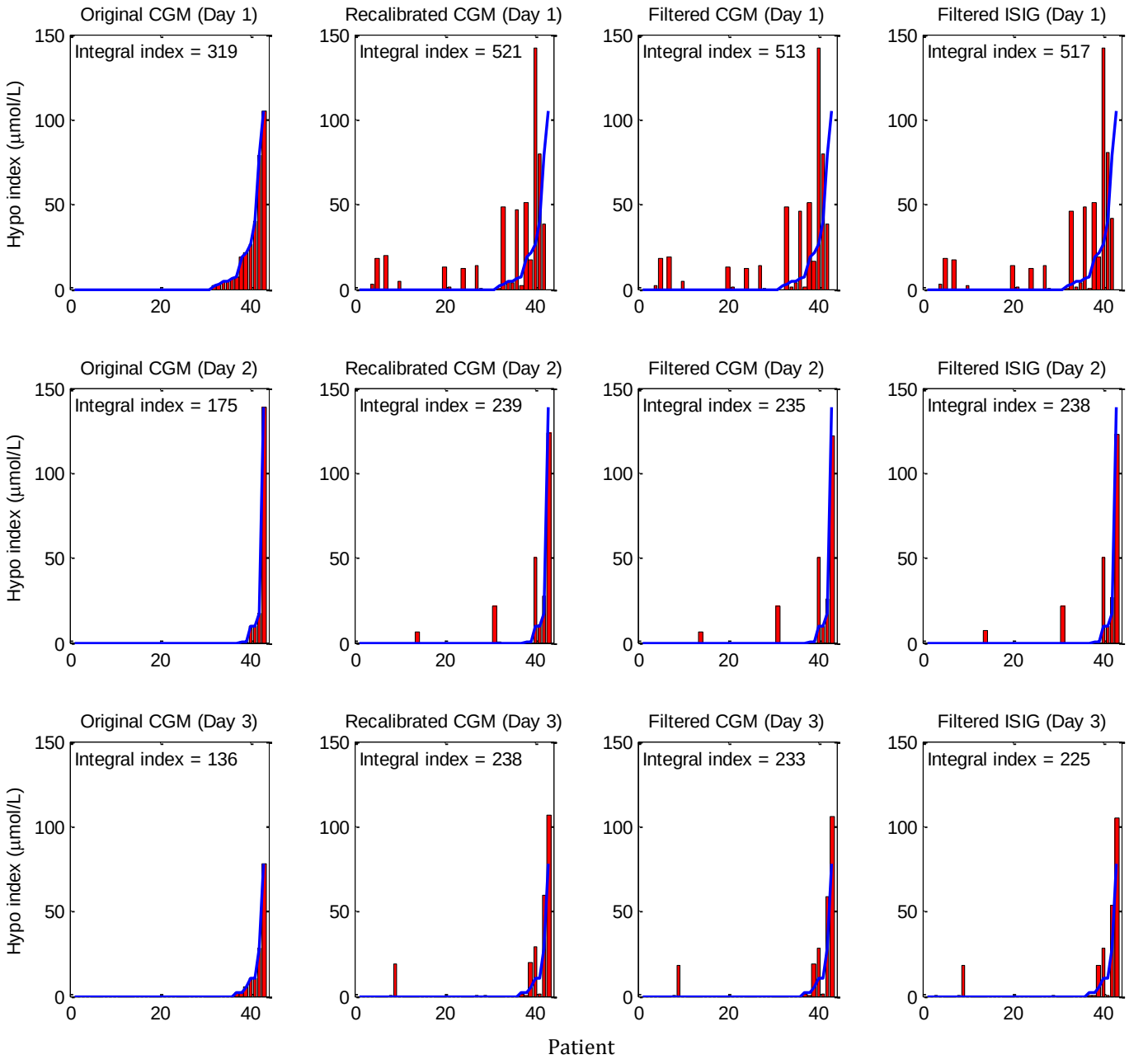


Figure 10.12 Comparison of ranked Hypoglycaemic Index for 43 patients (3 days) for Original CGM, Recalibrated CGM, Recalibrated and Filtered CGM, and Filtered I_{SIG} then Recalibrated. The curved solid line repeats the ranked distribution as determined by original SG data. The integral index captures the overall area of each panel for a single comparator value.

More importantly, the second plot in Figure 10.9 shows a definite positive shift in the median when only considering low glycaemic levels. Of the 145 pairs containing either a CGM or BG measurement below 3mmol/L, 63% have a BG measurement lower than the CGM. These lower measurements pull the SG trace down to the more accurate BGA value when recalibrating and cause the hypoglycaemia metrics to increase. Additionally, the bottom plot in Figure 10.9 shows the opposite is also true for high BG concentrations. At least for this dataset, CGM readings have a greater tendency to be lower than their BG counterparts when the concentrations are above 7mmol/L.

Furthermore, of the 53 calibration BG measurements that were hypoglycaemic, only 16 occurred at a time when the CGM was below 2.6mmol/L. The remaining 37 hypoglycaemic BG measurements were at a time when the CGM was reporting values greater than 2.6mmol/L. Thus, recalibrating the SG data to capture the 'missed' events is further justified, especially when using the SG data to classify hypoglycaemia.

When comparing recalibrated SG data to recalibrated and filtered SG data, the large reduction in the number of hypoglycaemic events in Table 10.2 (193 to 131) with little change in hypoglycaemic duration (2.6% to 2.5%) can be explained with reference to Figure 10.10. Two different phenomena occur that reduce the number of events from 4 to 1 in this exemplar case. First, at 3.8 days the peak in the SG trace is trimmed by the filter (filter after recalibration) preventing it from crossing the normoglycaemic threshold and reducing the number of hypoglycaemic events. The opposite can also occur, where a hypoglycaemic event observed in the recalibrated SG trace is removed by filtering, where, in this case, a trough would be trimmed.

The second phenomenon is seen at ~3.88 to 3.9 days, where high frequency fluctuations in SG measurements are smoothed by the filter. Smoothing these fluctuations around the threshold is likely to be the major influence on the reduced number of hypoglycaemic events observed. The variations in the number of hypoglycaemic events observed over the 4 analyses suggests that this

metric alone (number of events) may not be reliable when classifying hypoglycaemia using CGM due to this variability when near the threshold value.

The bottom section of Table 10.2 shows the number of patients that gained, lost or stayed with/without hypoglycaemia when recalibrated and filtered, compared to original SG data. Of the 25 babies who had hypoglycaemia in the original data set 22 to 24 had hypoglycaemia in the modified data sets, and, 17 to 18 out of 25 of babies who had no hypoglycaemia in the original data set still had no hypoglycaemia. These results suggest that over the duration of monitoring, the CGM should be consistent ~80% of the time about which patients had experienced hypoglycaemia, independent of calibration method.

10.4.2 Per-patient hypoglycaemic index by day

Figure 10.12 shows three important trends. First, the hypoglycaemic index, a measure of duration and severity, for a patient on any given day can change significantly after recalibration. For example, the highest ranked patient on day one using the original SG output had a hypoglycaemic index of ~105μmol/L, which was reduced to 0μmol/L after recalibration. Conversely, several patients have no hypoglycaemic events on day 1 (original CGM), but, after recalibration all have a non-zero hypoglycaemic index. For days 2 and 3 there is far more agreement in which patients had hypoglycaemia, although in some cases the index still changes with recalibration. This outcome suggests that after day one, regardless of calibration method, patients who experienced hypoglycaemia could be identified with a higher level of confidence.

The next trend shown by Figure 10.12 is that hypoglycaemia is most prevalent during day 1. Comparing the plots in row 1 and 2 of Figure 10.12 shows the decrease in integral index from day 1 to 2 is in the range of 45% - 54%, suggesting hypoglycaemia is more prevalent in the overall cohort on day one, by any calibration method. Interestingly, further decreases were not observed from day two to three in either the original SG data or the recalibrated/filtered SG data of Figure 10.12. This result reinforces the importance of capturing the first 24 hours of SG data and thus of proper placement and initial device calibration.

These first two trends suggest the effects of calibration scheme and/or filtering on hypoglycaemia detection are most significant on day 1. This could be due to two reasons: First, the error between calibration BG measurements and SG measurements tends to be larger on day 1, so recalibrating has more affect. Second, these infants are at risk by definition and hypoglycaemia can be more prevalent on this first day of life (Bhat et al. 2000; Hawdon et al. 1992). Neither effect can be ruled out in this analysis.

The third trend is an increase in integral index with recalibration, regardless of the day, which can potentially be explained by the previously mentioned tendency of the original CGM device to deliver a higher value at low glucose levels due to the linear regression based calibration approach. Conversely, there is a reduction in integral index with the addition of filtering, which is likely explained by the 'rounding' or clipping of troughs, in this analysis.

10.4.3 Limitations

The main limitation of this study is the number of BG measurements available. Due to the pain and discomfort of blood sampling in neonates, typically by heel prick, it was unethical to measure BG more frequently than ~4 hourly. Ideally, a reference measurement would be sampled for nearly every SG measurement, in addition to the ~4 calibration measurements per day. The reference measurements would allow the impact of calibration on a 'true' level of hypoglycaemia to be assessed more thoroughly and conclusively.

The limited number of BG measurements also restricts the recalibration and filtering strategies. For example, the delays between BG and interstitial glucose could not be optimised so the overall delay was left constant at the manufacturer set 10 minutes. Studies in the literature report delays typically in the range of 5-15minutes (Bailey et al. 2009; Boyne et al. 2003; Garg et al. 2010; Kamath et al. 2009; Kovatchev et al. 2009). However, without sufficient reference BG measurements the delay cannot be determined in this study, for this group of patients. Finally, the filter implemented in this study is effective, but represents one of a wide variety of available filters. Other more advanced filtering options

are available, but without several reference BG measurements it is not possible to determine the optimal filtering strategy.

10.5 Summary

The aim of this study was to investigate how recalibrating and filtering SG data affects metrics of hypoglycaemia in preterm infants. The results suggest that conventional hypoglycaemia metrics are heavily dependent on both the error in calibration BG measurements and the calibration algorithm used. All metrics of hypoglycaemia for our cohort increased after recalibration, confirming that the original, unmodified SG output tended to report high at lower levels. If highly accurate calibration measurements are available it may be more appropriate to recalibrate the data, especially when trying to accurately classify hypoglycaemia or other specific extreme events.

More importantly and generally, calibration BG measurement error and thus calibration algorithms play a significant role in quantifying hypoglycaemia using SG data. In particular, metrics such as number of hypoglycaemic events are particularly sensitive to recalibration effects. While this conclusion may be expected, its potential impact is quantified here, in this case for at risk neonates for whom hypoglycaemia may carry long-term negative consequences.

Chapter 11. Effects of calibration measurement errors

The recalibration procedure described in the previous chapter was specifically designed to utilise calibration BG measurements with minimal measurement error that are entered into the CGM monitor without delay. In the clinical environment, this is not always possible for a number of reasons. Therefore, it is important to have an understanding of how these errors can impact SG data, and specifically in this study, how hypoglycaemia metrics are affected. This chapter investigates the impact of measurement error and calibration time delays on metrics used to quantify hypoglycaemia.

11.1 Introduction

The recalibration algorithm described in the previous chapter was designed under the assumption that the calibration measurements used were very accurate and entered into the device without delay. However, in the busy neonatal care environment it is possible for unexpected time delays to occur between measuring BG and entering the value into the CGM device for calibration. The magnitude of time delays can depend on a number of factors, including the meter or method used to measure the BG concentration and the location of the BG sensors relative to the patient (Aragon 2006; Carayon & Gurses 2005; Ginsberg 2009).

Calibration algorithms cannot currently detect or correct for these time delays. Consequently, any delay could potentially introduce significant error in the output SG trace, especially during periods of rapidly changing BG (Castle & Ward 2010). For example, if the blood glucose level is steadily dropping at 0.08mmolL^{-1}min^{-1}, a perfectly accurate BG measurement entered just 15 minutes late will be 1.2mmol/L higher than the true BG level. Thus, when CGM devices are used in the neonatal unit where time delays are possible, it is important to understand the potential impact of these delays and whether or not they are clinically significant.

Another source of inaccuracy that could impact SG data is calibration BG measurement error. As mentioned in the previous chapter, point-of-care glucometers normally have measurement errors in the range 2-10% and can exhibit reduced performance in critically ill or neonatal care patients, particularly due to medications or fluctuating haematocrit levels (Carayon & Gurses 2005; Hoedemaekers et al. 2008; Kanji et al. 2005). In contrast, BGAs can measure BG concentrations with less than 2% error for a wide range of patient states, with little influence from haematocrit, PH, or PaO_2 (Watkinson et al. 2012). Hence, the choice of calibration measurement sensor could have a significant impact on CGM accuracy, independent of calibration timing.

These errors can add uncertainty to the SG data and consequently to any glycaemic metrics calculated from it. Therefore, studies utilizing CGM devices to quantify glycaemic events, such as hypoglycaemia, should be aware of these potential sources of error and the potential impact on results. In particular, any uncertainty in SG data would limit the detectable resolution of any true clinical changes and should be accounted for in the study design. This chapter quantifies the effect of timing delays and calibration BG measurement errors, both together and separately, on metrics used to classify hypoglycaemia in newborn infants.

11.2 Subjects and methods

11.2.1 Patients

This post-hoc analysis uses CGM device data and BG data from 155 newborn infants admitted to Waikato Hospital NICU during the SUGAR-BABIES trial. Eligibility for the study included babies greater than 35 weeks gestation, less than 48 hours old and at risk of neonatal hypoglycaemia. Primary risk factors for neonatal hypoglycaemia included having a mother with diabetes, prematurity and/or being small or large for gestational age. All patients were monitored using a CGMS System Gold device and SOF sensor, which was inserted in the baby's lateral thigh as soon as was practical after birth.

Patients with less than 24 hours of CGM data were excluded from this study and SG data sets exceeding 72 hours were trimmed. This data was used to create one of two timing error models and for the main analysis. Across all 155 infants, the median [25th - 75th percentile] duration of CGM recordings was 1.8 (1.5 – 2.0) days with 5.90 (5.1 – 6.9) calibrations per day. This study and use of data was approved by the Northern Y Ethics Committee, New Zealand. Table 11.1 summarises the demographics of this cohort.

Table 11.1 Cohort demographics

Cohort Demographics	
Number patients	155
Sex (M/F)	80/75
Gestational Age (weeks)	37 [36 - 38]
Birthweight (g)	2705 [2404 - 3393]
Primary Risk (# infants)	
Diabetes	50
Premature	57
SGA or LGA	41
Other	7

11.2.2 Timing error models

Timing error models are used to assess the impact of delays entering BG measurements into the CGM for calibration. Two models were created for this study, one using data from the Waikato Hospital NICU study and one using data from an on-going CGM study in the Christchurch Hospital ICU. The Christchurch Hospital ICU study and use of data was approved by the Upper South A Regional Ethics Committee, New Zealand.

All calibration BGs in both centres were determined using a BGA (Waikato: Radiometer ABL800Flex, Copenhagen; Christchurch: Radiometer ABL90Flex, Copenhagen). The BGA recorded the time and glucose concentration electronically, but nurses were required to manually enter the BG value into the CGM where it was stored internally and used for calibration. The time discrepancies between measuring BG and entering it into the CGM for calibration form the basis of this model. These delays were calculated as CGM calibration

time minus BGA measurement time and restrict the data set to the positive time domain as it is not possible to enter a calibration BG prior to measuring BG concentration in the BGA. Any additional time delays between obtaining the blood sample and determining the glucose concentration were considered negligible due to the close proximity of the BGA to the patients.

Data from Waikato Hospital included 1947 time delay values that could be included in the model. The vertical blue bars in Figure 11.1A show the distribution of time delay data and the red line shows the exponential model (μ = 12.96) that was used to capture the data. Figure 11.1B shows a similar plot for the 155 time delay values from the Christchurch Hospital study, which were also modelled by an exponential decay model (μ = 8.84).

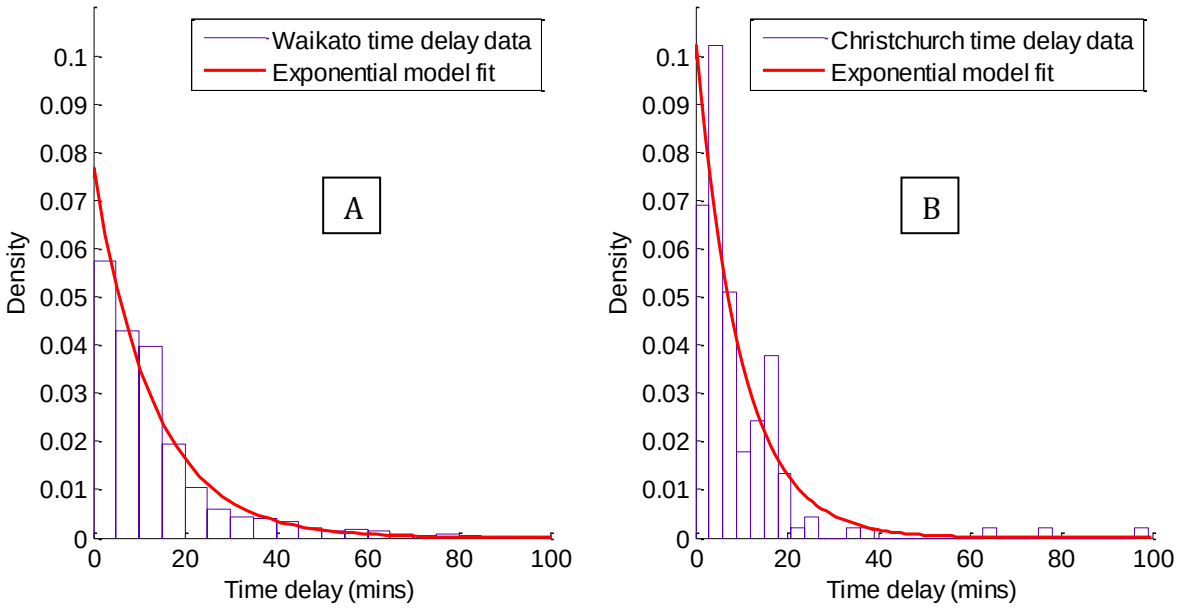


Figure 11.1 A) Raw timing error data and exponential model fit from the Waikato Hospital study. B) Raw timing error data and exponential model fit from the Christchurch Hospital study.

11.2.3 Measurement error models

Measurement error models describe the level of inaccuracy in a calibration BG measurement, commonly parameterized by accuracy and precision. A wide range of error levels for different devices can be found in the literature (Critchell

et al. 2007; Ho et al. 2004; Nuntnarumit et al. 2011; Rao et al. 2005). However, this study focused on three glucose meters, shown in Figure 11.2:

1. Abbott Optimum Xceed (Abbott Diabetes Care, Alameda, CA)
2. Nova Statstrip Glu (Nova Biomedical, Waltham, MA, USA)
3. Roche Accu-chek Inform II (F. Hoffmann-La Roche Ltd, Basle, Switzerland).

Figure 11.2 The three glucose meters used in this study. From left to right: Abbott Optium Xceed, Nova Statstrip Glu, Roche Accu-chek Inform II.

All three statistical error models were developed using data from the Christchurch ICU study. Under this protocol, every blood sample drawn for BG analysis was first analysed by the BGA and then the remaining blood distributed across up to 15 separate BG meters (up to 5 of each model). Not every type of glucometer was available for each patient in the study, so the numbers of paired meter-BGA measurements available was different for each model. All BG measurements were made by trained staff, minimizing potential user associated error (Ginsberg 2009).

All devices used in this study measure glucose concentration in a whole blood sample and display a plasma equivalent glucose concentration as recommended

by the International Federation of Clinical Chemistry and Laboratory Medicine (D'Orazio et al. 2006). Furthermore, BGA's in the same family as the ones used in this study have also been shown to correlate well with lab plasma glucose determinations (Scott et al. 2008), allowing direct comparison between all devices.

For each glucometer measurement, the error between the glucometer and BGA value was calculated as *meter – BGA*, so that a positive error is indicative of a high meter value. Errors were stratified into bins, based on the BGA measurement. Mean and standard deviation of each bin was then used to describe the error distribution for each bin, which was assumed Gaussian around any bias. Positive bias assumes the glucose meter is reading higher than the recorded BGA value, and vice-versa.

The *Abbott Optium Xceed* is an inexpensive and commonly available device that measures the glucose concentration of a whole blood sample, and estimates the plasma equivalent glucose concentration using a constant adjustment factor of 1.12 (Abbott 2010). This factor is derived from the difference between plasma glucose and whole blood glucose concentration for an individual with a normal haematocrit level. The Abbott test strips are validated for a haematocrit range of 20-70 (Abbott 2010). The top section of Table 11.2 shows the error model derived from the experimental Abbott and BGA data. There were a total of 724 paired meter-BGA measurements available for the model.

The *Nova Statstrip GLU* was designed for point of care (POC) testing in the hospital environment and adjusts for haematocrit level when calculating plasma glucose concentration. The Nova test strips used in this study were validated for a haematocrit range of 20-65% (Nova 2011). The middle section of Table 11.2 shows the error model derived from the 229 paired meter-BGA measurements.

The *Roche Accu-chek Inform II* was also designed for POC testing and monitoring in hospitals. This device also adjusts for haematocrit level and test strips were validated for a haematocrit range of 10-65% (Roche 2007). The bottom section

of Table 11.2 shows the error model derived from 344 paired meter-BGA measurements.

Table 11.2 Measurement error data for the Abbott Optium Xceed, Nova Statstrip, and Roche Accu-chek Inform II glucose meters. The Nova and Roche models have a reduced number of bins to avoid skewing due to low measurement numbers.

Abbott Error Model					
Reference BG (mmol/L)	*< 5.9*	*6.0 - 6.9*	*7.0 – 7.9*	*8.0 – 8.9*	*> 9.0*
Number of measurements	141	277	224	42	40
Error mean (mmol/L)	0.5099	0.5433	0.2299	0.1952	0.635
Error std. dev. (mmol/L)	0.4982	0.7519	0.5521	0.8748	0.3965
Nova Error Model					
Reference BG (mmol/L)	*<6.9*		*7.0 – 7.9*	*>8.0*	
Number of measurements	67		141	21	
Error mean (mmol/L)	-0.0134		-0.0823	-0.1905	
Error std. dev. (mmol/L)	0.2564		0.2471	0.3463	
Roche Error Model					
Reference BG (mmol/L)	*<6.9*		*7.0 – 7.9*	*>8.0*	
Number of measurements	174		160	10	
Error mean (mmol/L)	-0.181		-0.4212	-0.27	
Error std. dev. (mmol/L)	0.2615		0.2645	0.0949	

All three measurement error models were validated against the error characteristics described by the manufacturers (Abbott 2010; Nova 2011; Roche 2007). The Roche and Nova measurement error models compared well to the manufacturers data. The Abbott measurement error model has increased bias and variation compared to the manufacturer's data. Finally and equally importantly, the models and specific values described here are comparable to what has been reported in the literature for these devices (Makaya et al. 2012; Nuntnarumit et al. 2011). Thus, the models summarised in Table 11.2 are representative of these devices in use.

11.2.4 Analysis

This analysis used CGM and BG data from 155 babies admitted to the Waikato Hospitals NICU. Timing and measurement errors were randomly sampled from

the models described in Sections 11.2.2/11.2.3 and added to calibration BG measurements, as depicted in Figure 11.3.

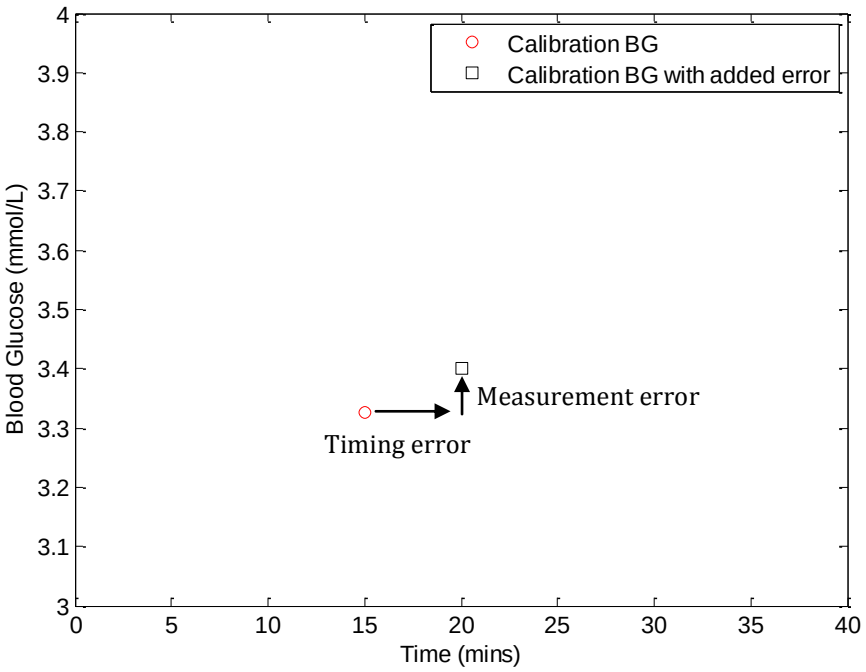


Figure 11.3 Example of the error that was added to calibration BG measurements to simulate measurement error and delays in entering calibration BG measurements into the CGM.

The calibration BG measurements with added error were used to calibrate the SG data and hypoglycaemia in this CGM data was quantified by:

- **Number:** Number of independent hypoglycaemic events
- **Duration:** Percentage of SG recordings below 2.6mmol/L
- **Hypoglycaemic index:** Total area between the 2.6mmol/L threshold and the SG trace (when SG < 2.6mmol/L), divided by the total monitoring duration

MC methods were employed to ensure robust results from the randomly sampled errors. Each simulation consisted of 1000 runs and a total of 11 simulations were completed, one for each individual model (two timing error models and three measurement error models) and one for every possible combinations of timing/measurement error models (a further 6 simulations). The simulation protocol can be summarized as follows:

1. For every simulation run on a each patient, the difference in hypoglycaemia metrics for the 'random error' SG data and the original 'no error/baseline' SG data are determined. These differences are calculated as *hypoglycaemia in SG data with error - hypoglycaemia in baseline SG data*, so a positive result suggests an increase in hypoglycaemia with the error(s) added to calibration BG measurements.
2. The median difference in hypoglycaemia across 1000 MC runs for each patient was recorded
3. The results for the overall cohort are presented as a median [25th -75th percentile] (5th - 95th percentiles) of the values calculated in step 2 for all patients.

Figure 11.4 shows a representative example of the variation in a SG data trace across 1000 runs in the simulation for this subject. The baseline SG data calibrated with no error in calibration BG measurements is shown by the red line in Figure 11.4.

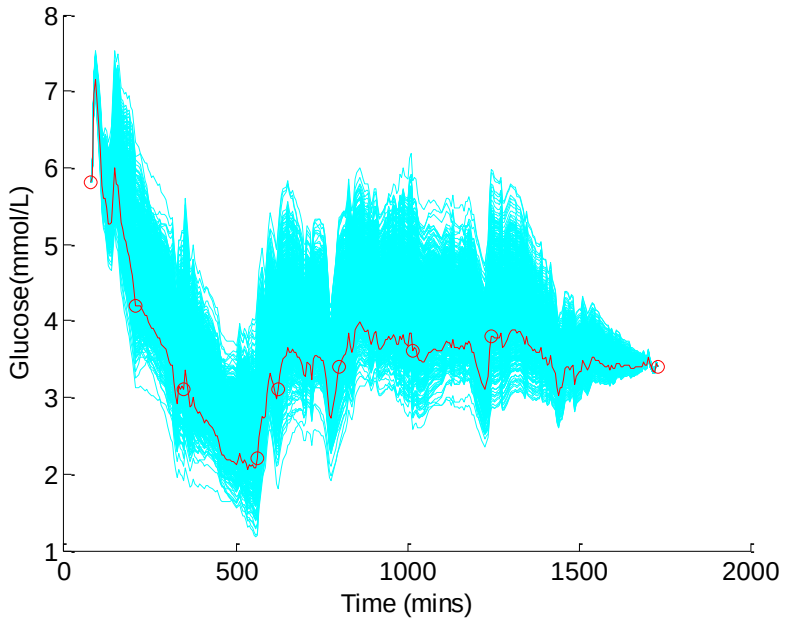


Figure 11.4 Example of the variation in SG data across 1000 runs in a simulation, shown in blue. The red line shows the SG data when calibrated with no additional error.

11.3 Results

11.3.1 Timing error only

Table 11.3 shows timing error tends to have little effect on the number of hypoglycaemic events. The results for either timing error model both have a median difference from the original metrics of 0 and the 5^{th}-95^{th} range variation is only ±2 events. Percent duration and hypoglycaemic index are increased by timing error for both models. The Waikato model caused a median increase of 0.21% to the duration of hypoglycaemia and 0.27μmol/L to hypoglycaemic index. Simulations using the Christchurch model caused a median increase in duration of 0.17% and a median increase in hypoglycaemic index of 0.18μmol/L.

Figure 11.5 shows a data set that has a period where timing error in calibration BG measurements has a larger impact on the SG data that measurement error. Between 500 - 700mins, the timing error results in a wider range of SG values, as shown by the thick green band. However, for the remainder of the SG data the measurement error appears to be dominant.

11.3.2 Measurement error only

The results from simulations using only measurement error are also shown in Table 11.3. The Abbott results show a negative difference across all metrics compared to baseline results. Hence, the Abbott results tended to under-report all hypoglycaemic metrics. The number of hypoglycaemic events recorded had a difference of -1[-3 0](-8 0) events. The median number of hypoglycaemic events reported from simulations using the Nova model were the same as baseline results. However, there was a slight tendency for percent duration and hypoglycaemic index to be over reported, with results of 0.49[0.1 1.6](-0.1 6.7)% and 2.93[0.5 8.2](0 16)μmol/L, respectively. The Roche sensor tended to over-report hypoglycaemia to the highest degree, with median differences of 4.45% and 19.4μmol/L for duration and index, respectively.

Table 11.3 Overall cohort results from each 1000 run MC simulation showing the median [IQR] (5^{th}-95^{th} range) difference in hypoglycaemia metrics from baseline. Metrics calculated *hypoglycaemia in CGM data with error - hypoglycaemia in baseline CGM data*

Baseline Hypoglycaemia				
Number of Hypoglycemic Events	2 [1 5] (0 13)			
Percentage Duration	6.13 [1 13] (0 29)			
Hypoglycemic Index	10.13 [1.3 27] (0 87)			
Results From Monte Carlo Simulations Using Error Models				
Number of Hypoglycemic events	*No measurement error*	*Abbott measurement error*	*Nova measurement error*	*Roche measurement error*
No timing error		-1 [-3 0] (-8 0)	0 [0 1] (-3 2)	0 [0 2] (-3 4)
Waikato timing error	0 [0 1] (-2 2)	-1 [-2 0] (-8 0)	0 [0 1] (-3 2)	1 [0 2] (-3 4)
Christchurch timing error	0 [0 0] (-2 2)	-1 [-2 0] (-8 0)	0 [0 1] (-3 2)	1 [0 2] (-3 4)
Percentage Duration				
No timing error		-4.68 [-9.0 -1.0] (-17 0)	0.49 [0.1 1.6] (-0.1 6.7)	4.45 [1.8 10] (0 23)
Waikato timing error	0.21 [0 1.3] (-1.7 4.1)	-4.25 [-9.0 -1.0] (-15 0)	1.02 [0.1 3.2] (-0.5 8.7)	5.36 [2.1 11] (0 26)
Christchurch timing error	0.17 [0 0.9] (-1.5 3.6)	-4.40 [-8.5 -0.6] (-15 0)	0.84 [0 2.7] (-0.4 8.1)	5.23 [2.2 11] (0 25)
Hyperglycemic Index				
No timing error		-7.64 [-22 0] (-59 0)	2.93 [0.5 8.2] (0 16)	19.4 [4.2 38] (0 70)
Waikato timing error	0.27 [0 3.1] (-3.3 14)	-6.84 [-22 -0.3] (-48 0)	3.84 [0.7 12] (-0.1 27)	20.8 [4.1 42] (0 82)
Christchurch timing error	0.18 [0 2.3] (-2.9 11)	-7.24 [-21 -0.4] (-50 0)	3.77 [0.60 11] (0 23)	21.3 [4.7 42] (0 80)

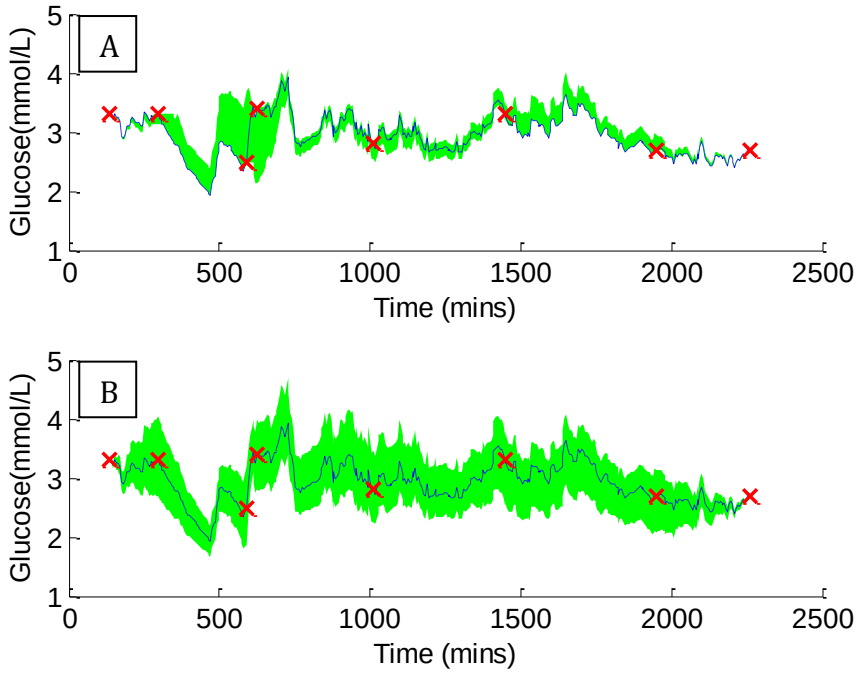


Figure 11.5 A) Example patient where timing error in calibration BG has a large effect on the calibrated SG data. B) The same data set with measurement error added to calibration BG. Note between 500 - 700mins timing error has a larger impact than measurement error.

The tendency of each measurement error model to cause under or over reporting of hypoglycaemia metrics can be seen in Figure 11.6. The top plot of Figure 11.6 shows a representative SG data set after a simulation using the Abbott measurement error model. The middle plot shows the same dataset after simulation using the Nova measurement error model and the lower plot shows the data after simulation using the Roche measurement error model. In each plot, the coloured band shows the 5th -95th percent range of reported measurements at each time point, across 1000 simulation runs. The dark blue line shows the recalibrated SG data with no error and the red crosses show calibration BG with no error.

This specific SG dataset contains a significant hypoglycaemic event that occurs between 430 and 570 minutes. However, when simulating with the Abbott measurement error model (A in Figure 11.6) this hypoglycaemic event only appears in approximately 50% of the 1000 simulation runs. In the Nova (B in

Figure 11.6) and the Roche (C in Figure 11.6) simulations this hypoglycaemic event is detected for 100% of the MC runs. As a consequence of the negative bias, especially in the Roche model, this event tends to drop below the 2.6mmol/L threshold earlier and rise later than the baseline trace, resulting in a concomitant increase in the hypoglycaemia duration and index metrics, as well.

Figure 11.7 shows the baseline (no error) duration of hypoglycaemia (x-axis) plotted against the duration of hypoglycaemia for each run (y-axis), for a 1000 run simulation. The spread in the y direction shows the amount of variation in the hypoglycaemia metric across the 1000 runs. The black, blue and red lines show the 25th, median and 75th percentiles, respectively. The left plot shows results when using the Abbott model and the plot shows a clear tendency to under-report the duration of hypoglycaemia. The right plot in Figure 11.7 is for the Roche model and in this plot the duration metric has been over estimated, shown by the majority of results being above 45° line.

11.3.3 Combined timing and measurement error

The combination of measurement and timing error leads to an increase in both the median difference and variation of these across all metrics. However, the measurement error has a significantly larger impact than the timing error on hypoglycaemia metrics.

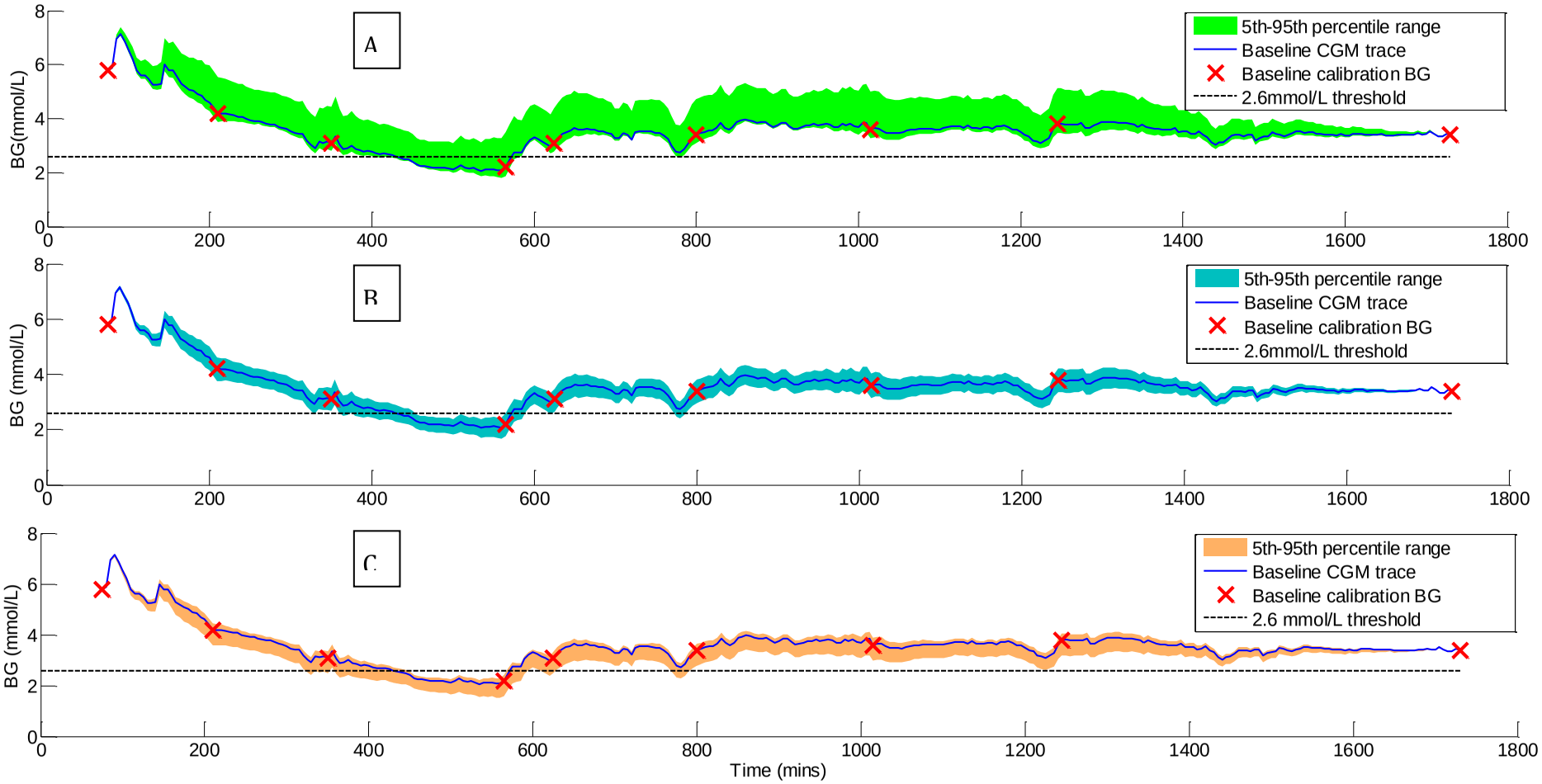


Figure 11.6 Example SG traces showing the effect of A) Abbott measurement error, B) Nova measurement error, and, C) Roche measurement error. The coloured band in each plot shows the 5th-95th percentile variation in the SG trace over 1000 MC simulations

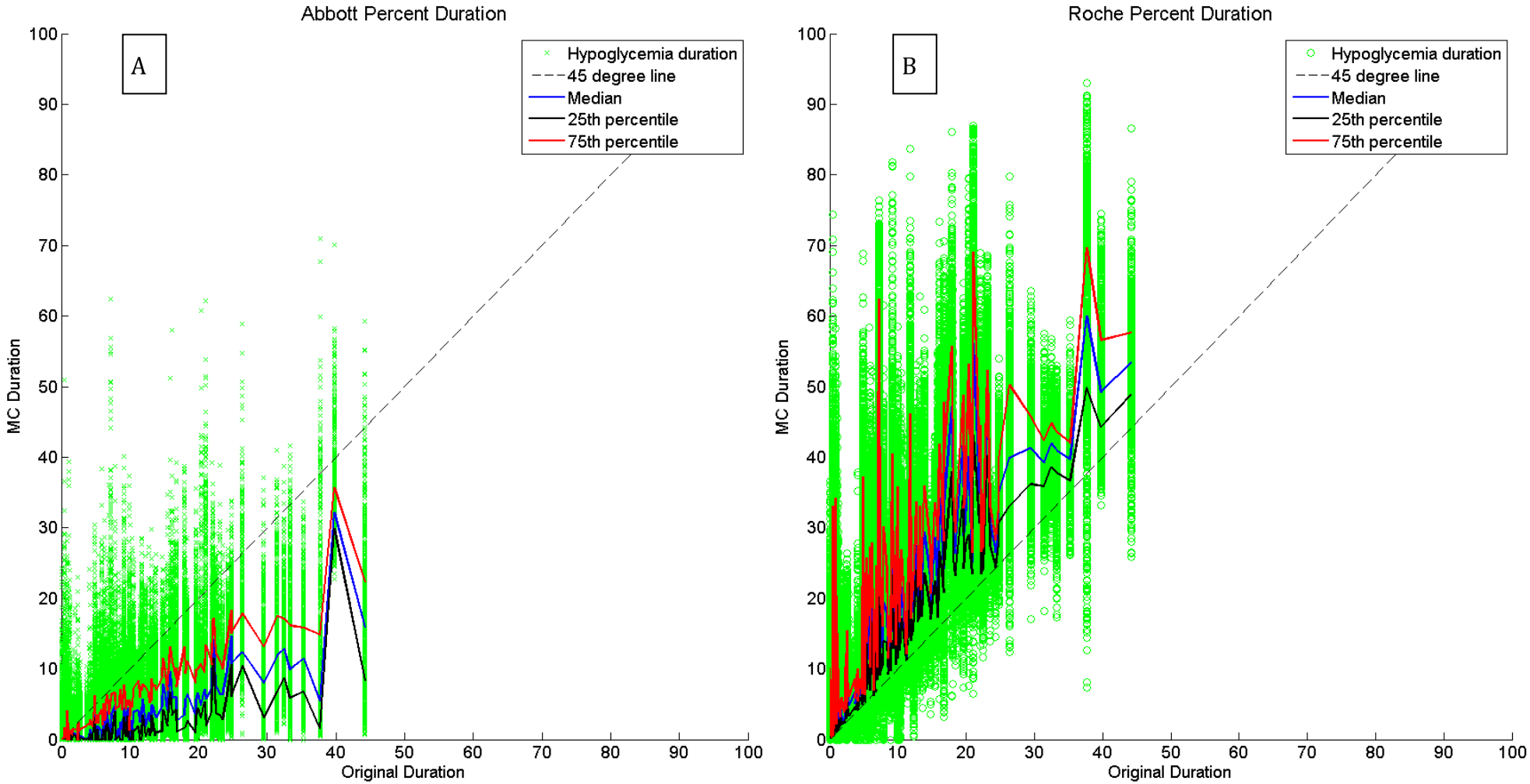


Figure 11.7 Duration of hypoglycaemia in each of 1000 MC runs plotted against baseline duration of hypoglycaemia, for all 155 CGM data sets. Plot 'A' shows simulation results using Abbott measurement error and plot 'B' shows simulation results using Roche measurement error.

11.4 Discussion

11.4.1 Timing error vs. Measurement error

The results suggest that measurement error has a much larger effect than timing error on the metrics used to quantify hypoglycaemia. The Nova measurement error model, which had the least impact of the 3 measurement error models, increases the median duration of hypoglycaemia by more than double that of the either timing error model. Furthermore, the median hypoglycaemic index from the simulation using the Nova measurement error model is an order of magnitude greater than the result from the simulation using the Waikato timing error model. Additionally, when the effects of measurement and timing error are combined, there is little change to the results compared to measurement error alone. These results clearly illustrate the dominance of measurement error over timing error in the overall cohort, and that the observed timing errors have little impact on these hypoglycaemic metrics.

On a per-patient level, it is possible for the impact of timing error to be much more significant than that of measurement error. For example, in Figure 11.5, the calibration BG just after 500 minutes is determined during a period of steep glucose ascent. When timing error alone is randomly added to this calibration BG, as in panel A, we see a wide range of possible SG values for each time point between 500 - 700 minutes. However, in panel B, when simulating measurement error alone the rate of change of glucose has little or no effect and the range of possible SG values remains fairly constant throughout the entire monitoring period.

Clinically, when real-time CGM devices are used to monitor glycaemia, the effects of these errors can be reduced by acknowledging the state of the SG trace at the time of calibration. If glycaemia appears stable then it is likely that impact of timing error is negligible, and the time between measuring BG and entering it in to the CGM is not vital. This approach allows time and care to be taken to ensure the most accurate measurement of blood glucose levels is obtained and the resulting calibration is optimal for the conditions. For example, a BGA might be used instead of a POC glucometer if timing errors are less critical. However, for a

patient with a rapidly changing glucose concentration at the time of calibration, timing error can become more significant than measurement error. This issue is particularly important when monitoring newborn babies glucose levels as they can fluctuate significantly during the first few hours of life. In these specific cases, it is imperative that the calibration measurement is obtained and entered into the CGM quickly, perhaps using a glucometer instead of a BGA to reduce timing error and delay.

11.4.2 The Impact of Bias

The tendency of Abbott to under-report the hypoglycaemic metrics, while Nova and Roche over-report, is caused by the level of sensor measurement bias, which is a property of the device itself. For the Nova and Roche error models, a negative bias causes the trace to be pulled down more frequently during calibration in the simulations. Hence, the number of hypoglycaemic events tends to increase, as well as the time and area below the threshold. In contrast, the positive bias in the Abbott meters causes the trace to be pulled up during calibration causing hypoglycaemia to be under-reported. These effects are clearly shown in Figure 11.6.

The Abbott meter, which is not designed specifically for the ICU environment, is likely more sensitive to clinical factors including varying haematocrit and certain medications. Hence, the relatively large bias and low precision may not reflect the underlying device capacities without these influences. The cohort data used to create the models came from patients with haematocrit levels of 24-36% and while they all remained within the validated range for the test strip, these haematocrit values are significantly lower than the normal haematocrit level of 40-45%. In this study, the low haematocrit levels combined with a constant correction factor of 1.12 would have contributed to the large positive bias seen in the Abbott model.

For example, using a whole blood to plasma conversion equation from the literature (Mahoney & Ellison 2007) and a haematocrit value of 27%, the ‘true’ conversion factor is calculated to be 1.07. This value is 5% lower than the 1.12

used by the Abbott device. Additionally, newborn babies can have haematocrit levels as high as 65% (Jopling et al. 2009) causing the Abbott meter to under-report the true plasma glucose by ~6%. Thus, this factor can have a substantial impact when used in these cohorts.

The clinical impact of these findings is important for future studies using CGM devices to classify hypoglycaemia in neonates. Glucometer device characteristics, such as bias, and clinical factors, such as haematocrit levels, could have a substantial impact on the results of a glycaemia study. Thus, it is imperative they are reported in detail, which is not always the case.

For example, if two separate studies were investigating hypoglycaemia using CGM devices, the study outcomes could differ simply due to the type of glucose meter used to calibrate the CGM. Thus, it is important to select an appropriate method of measuring glucose for calibration when undertaking CGM studies. It is equally important to understand the range of impact or uncertainty that the device may impart to the results. Finally, the use of mixed CGM device types or calibration algorithms should be accounted for.

11.4.3 Variation in hypoglycaemia metrics

The IQR and 5^{th}-95^{th} range in the results table are used to assess the variation in the hypoglycaemia metrics across the 1000 MC runs. The results showed a counter-intuitive trend. Specifically, the Roche results had a wider 5^{th}-95^{th} range (more variation in results), for duration and index, than the results from the lower precision Abbott meter.

This counter-intuitive trend can be explained using Figure 11.7, which shows the duration of hypoglycaemia from each of the 1000 simulation runs as a function of baseline duration, for each SG dataset. In plot A, it is the positive bias in the Abbott model that causes a reduction in hypoglycaemia, which ultimately leads to truncation of the percent duration at zero. This truncation reduces the perceived variation due to the precision of the meter. Conversely, plot B in Figure 11.7 illustrates how the negative bias in the Roche model increases the

prevalence of hypoglycaemia, allowing more variation (due to precision) to be observed in the metric. For example, in plot A, a baseline duration of 10% hypoglycaemia results in a maximum duration of ~30% and ~25% of simulation runs showing no hypoglycaemia. However, in plot B, for baseline duration of 10% the range of values for durations from the 1000 runs spans from 0 to ~80%, making the Roche meter appear worse despite greater precision. Such counter intuitive results further reinforce the need to understand, in detail, the calibration device and measurement value when assessing clinically significant outcomes.

11.4.4 Limitations

The main limitation of this study is the method used to calibrate the CGM after adding timing and measurement error to calibration BGs. As stated previously, the calibration algorithm was designed to be used with very accurate calibration BG measurements, with little or no timing error. Thus, it inherently assumes the calibration BGs are perfect and forces the SG trace to pass through them. This approach makes it particularly sensitive to errors in calibration BG measurements. Inbuilt calibration algorithms based on linear regression would likely reduce the effect of calibration errors on hypoglycaemic metrics, but it is likely that the trends discussed in this study will still remain.

11.5 Summary

Overall, measurement error tends to have a much larger impact on hypoglycaemia metrics than timing error. However, if a patient is particularly variable, then timing error can have a much more measureable impact. The effect of bias in calibration BG measurements was twofold: 1) A negative error bias increases the prevalence of hypoglycaemia; and, 2) it also increases the amount of variation seen in hypoglycaemic metrics. The opposite was also true for positive biases. Finally, if CGM devices are to be used clinically for assessing events such as hypoglycaemia it is important that the investigators are aware of the potential impact that errors in calibration BG measurements can have on SG data.

Chapter 12. Detecting Unusual SG measurements

In addition to the effect of device calibration and errors in calibration BG measurements, the detection of artefacts is another important aspect of CGM that needs further investigation. Artefacts occasionally appear in the SG data and are typically characterised by unusually large step changes in glucose level, often for no apparent reason. These step changes are frequently larger than the commonly observed limits of physiological behaviour. Detecting unusual SG measurements, even without knowing the cause, offers the user valuable feedback on the quality of SG data at any point in time, and can thus better inform decision making. This chapter presents a novel method for detecting unusual SG measurements, using a modified stochastic forecasting method developed for insulin sensitivity prediction.

12.1 Introduction

If CGM devices are going to be used in the clinical setting to monitor, diagnose and potentially aid in the treatment of abnormal glycaemia, clinicians need to know the data are reliable and accurate. In the context of detecting hypoglycaemia in neonatal infants, investigators would benefit from knowing whether a significant drop in glucose is likely 'real' or simply an artefact of the device. Sensor artefacts are known to occur (Bequette 2010; Keenan et al. 2009), but an understanding of the cause of seemingly random jumps or drops in SG data are still being investigated. A thorough review of the literature in 2011 found 10 studies that reported abnormal CGM behaviour due to motion or pressure, but still concluded that this area was poorly understood and needed further research (Helton et al. 2011).

Even with a limited understanding of what causes artefacts in SG data, it is still possible to detect these anomalies in a probabilistic sense. Consider a scenario where SG data are retrospectively analysed to classify hypoglycaemia in neonates, where frequent BG measurements are not available. Three consecutive measurements in a SG trace read 4mmol/L, 2.5mmol/L and 4mmol/L. If

hypoglycaemia was classified as a measurement below 2.6mmol/L, then this sequence would be recorded as a hypoglycaemic event. However, if the rest of the SG trace was very stable with low variability, intuition would suggest this 'event' is potentially a sensor artefact. Figure 12.1 shows an example of a hypoglycaemic drop in glycaemia that could potentially be an artefact of the CGM device rather than a true representation of the underlying glucose concentration.

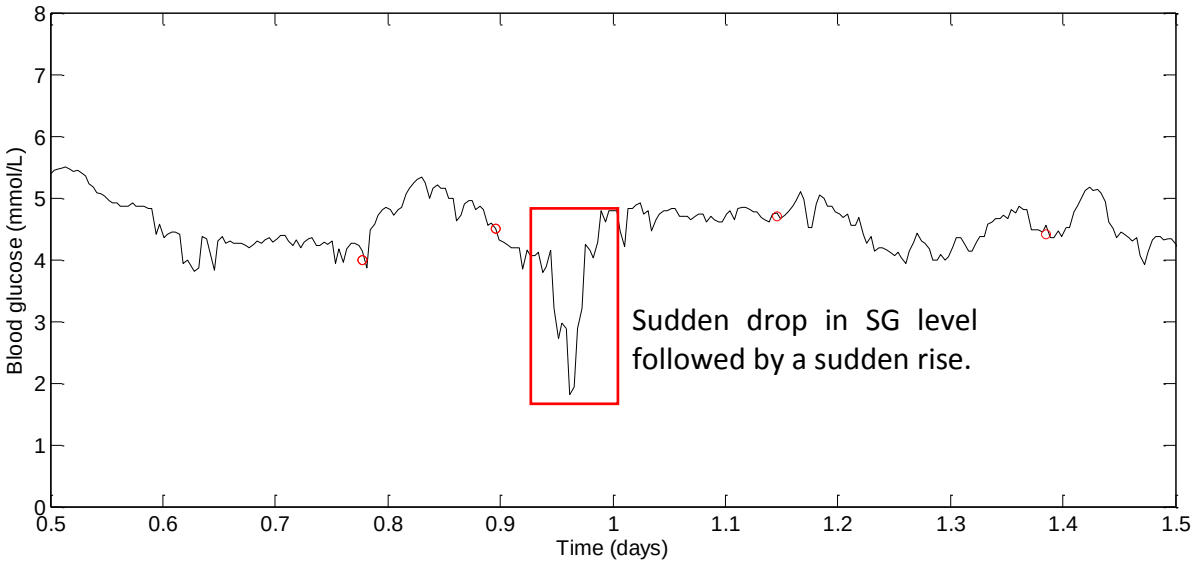


Figure 12.1 Plot of SG data showing a sudden drop to 2mmol/L at approximately 1 day after monitoring began. The drop is followed by a sudden rise, suggesting this feature is potentially an artefact of the CGM device rather than a true representation of the underlying glucose concentration.

The idea of detecting outlier SG data using automated algorithms is not completely new. Previous studies have investigated different methods of detecting outlier measurements including a discrete wavelet transform based online detector (Quan et al. 2010) and a post-processing support vector machine (Leal et al. 2013; Tarin et al. 2010). However, currently, there is no reliable and widely used method of determining whether sudden drop-outs or changes in SG data are likely to be “true” or simply an artefact.

This chapter develops and presents a tool to aid clinicians in identifying unusual CGM behaviour, retrospectively or in real-time, and highlights sections of the CGM glucose trace that potentially need to be interpreted with care. Specifically,

the focus is to identify unusual CGM sensor behaviour, not unusual glycaemic excursions.

12.2 Subjects and methods

12.2.1 Subjects

This study used CGM data from 50 babies ≥ 32 weeks gestation that were at risk of hypoglycaemia and admitted to the Waikato Hospital NICU during the BABIES study. Exclusion criteria included serious congenital abnormities or a skin condition preventing continuous glucose monitoring. Informed consent was obtained from parents of all babies in the study. The study was approved by the Northern Y Regional Ethics Committee, New Zealand. Table 12.1 shows the demographics for patients enrolled into the study.

Table 12.1 Cohort demographics

Cohort Demographics	
Number of CGM traces	50
Sex (M/F)	26/24
Gestational Age (weeks)	34 [33 - 37]
Birthweight (g)	2172 [1880 - 2990]
Primary Risk (# infants):	
Diabetes	15
Premature	19
Small or Large for gestational age	14
Other	2

12.2.2 Continuous glucose monitoring

All patients had interstitial glucose monitoring using the CGMS System Gold and SOF sensors. Sensors were located on the infant's lateral thigh and were connected to the CGMS system monitor via the supplied cable. Monitoring began on admission to the NICU and finished after 7 days or when the baby was no longer considered to be at risk of hypoglycaemia. During the monitoring period nurses were asked to record all BG concentrations, feeding and medication for the management of hypoglycaemia. However, they remained blind to the glucose concentrations determined by the device by design.

The device was calibrated per the manufacturer's recommendations and all of the data entered into the device were checked against clinical records for accuracy. Upon completion of monitoring, data were downloaded to a PC using CGMS system solutions software version 3.0C, which calibrated the CGM readings retrospectively. A total of 234 days of SG data (67438 measurements) were collected and the median [IQR] duration of monitoring per-patient was 4.7 [4.0 - 5.7] days.

12.2.3 Calibration measurements

The CGM device required calibration 2-4 times daily to convert the electrical current produced by the sensor into a meaningful glucose value. Blood samples were taken by nursing staff via heel-pricking and the glucose concentrations were used to calibrate the CGM device. The median [IQR] interval between samples was 4.8 [3.5 – 6.4] hours.

All BG calibration measurements were made using a BGA (Radiometer, ABL800Flex, Copenhagen) that analysed a whole blood sample using the glucose oxidase method. This device has a reading range of 0.0 to 60.0mmol/L and a C.V. of 1.4-2.2% (Cembrowski et al. 2010; Watkinson et al. 2012). Furthermore, the study by Watkinson et al. showed that a device from the same family, using the same glucose electrode, had a coefficient of variation of 2.1% in ICU patients and performance was not affected by haematocrit, pH or PaO_2 (Watkinson et al. 2012). Due to the location of the analyser, a short time delay (estimated < 15mins maximum) was possible between taking the blood sample and introducing the resulting measurement into the device.

12.2.4 Stochastic model

A stochastic model based on the kernel density method was implemented to classify unusual SG measurements using the previous SG measurement and information about the history of CGM behaviour. This lag-1 model is an extension to the methods described by Lin (2007) who developed a stochastic model for predicting variability over time in insulin sensitivity.

$$P(CGM_n = y | CGM_{n-1} = x) = \sum_{i=1}^{n} \omega_i(x) \frac{\phi(y; y_i, {\sigma_{y_i}}^2)}{p_{y_i}} \quad 12.1$$

$$\omega_i(x) = \frac{\phi(x; x_i, {\sigma_{x_i}}^2)/p_{x_i}}{\sum_{j=1}^{n} \phi(x; x_j, {\sigma_{x_j}}^2)/p_{x_j}} \quad 12.2$$

Where: $p_{x_i} = \int_0^\infty \phi(x; x_i, {\sigma_{x_i}}^2)$ and $p_{y_i} = \int_0^\infty \phi(y; y_i, {\sigma_{y_i}}^2)$

The model was generated using Equations 12.1 and 12.2, which define the 2-dimensional kernel density estimation in conditional SG measurement variability. Each $\phi(x; x_i, {\sigma_{x_i}}^2)$ and $\phi(y; y_i, {\sigma_{y_i}}^2)$ is a normal probability density function centred at the corresponding x_i and y_i. To ensure that x and y are positive, p_{x_i} and p_{y_i} normalise each $\phi(x; x_i, {\sigma_{x_i}}^2)$ and $\phi(y; y_i, {\sigma_{y_i}}^2)$ to the positive domain. Specifically, for a given SG measurement CGM_{n-1}, Equation 12.1 provides a continuous, empirical estimate of the conditional probability density function for the next SG measurement, CGM_n. These conditional probability density functions provide the basis for classifying SG measurements and identifying unusual CGM behaviour.

12.2.5 SG measurement classification

Using the stochastic model, a given SG measurement, CGM_n, is classified as follows:

1. The previous measurement, CGM_{n-1}, is used to find the corresponding conditional PDF from the model.
2. CGM_n is located in the PDF and its percentile value in the conditional PDF is determined.
3. The percentile is used to classify CGM_n, where a very high or very low percentile is indicative of an outlier. These outliers are classified as unusual SG measurements.

The measurement-to-measurement sections of the SG trace were colour coded based on the percentile value, to highlight areas of unusual CGM behaviour quickly and effectively. Three ranges were used to specify the colour: within 80% range (10th-90th percentile) was blue, within 90% range (5th-95th percentile) was cyan, within 99% range (0.5th-99.5th percentile) was yellow, and outside 99% range was red. These intervals were chosen based on the data used in this

study and can be customised for different patient groups and/or different CGM sensors. As the scale starts at 80% range, the focus here is on classifying outliers, rather than the full range.

12.2.6 5-fold validation of stochastic model

A 5-fold validation was used to check the fit of the stochastic model. The data set (N=50) was randomly divided into 5 sets of 10 patients. For each 10 patient group, the remaining 40 patients were used to create a stochastic model which was then tested on the group of 10 patients. The model fit was assessed by counting the number of clinical SG measurements (from the 10 patients) captured by the model's 80% range, 90% range and 99% range.

MC methods were used to reduce the effect of randomly selected outliers on overall results. MC methods provide a robust means of estimating the range of possible outcomes for a process involving one or more random variables. In this study, the MC simulation involved repeating the 5-fold validation 25 times, and reporting the Median [IQR] results.

12.3 Results

12.3.1 Clinical SG data and stochastic model generation

Figure 12.2 shows a plot of all the SG data (CGM_{n-1}, CGM_n). The contour lines represent the 5th, 25th, 50th, 75th, and 95th percentiles of the stochastic model surface. Figure 12.3 shows a distribution of the data density by glycaemic level. Figure 12.4 shows a surface plot of the stochastic model. Conditional probability density functions are slices parallel to the CGM_n axis, and each slice has an area under the curve of 1.0. Figure 12.5 shows a comparison of the PDFs obtained from the model versus the PDFs obtained directly from the SG data. Each PDF shows the expected distribution of CGM_n given a previous measurement (CGM_{n-1}) of 2, 4, 6, 8 or 10mmol/L. It should be noted that the PDFs could be generated for any value of CGM_{n-1} within the bounds of the model, and that Figure 12.5 shows just five examples. Table 12.2 shows numerical results from the MC validations which confirm a good model fit.

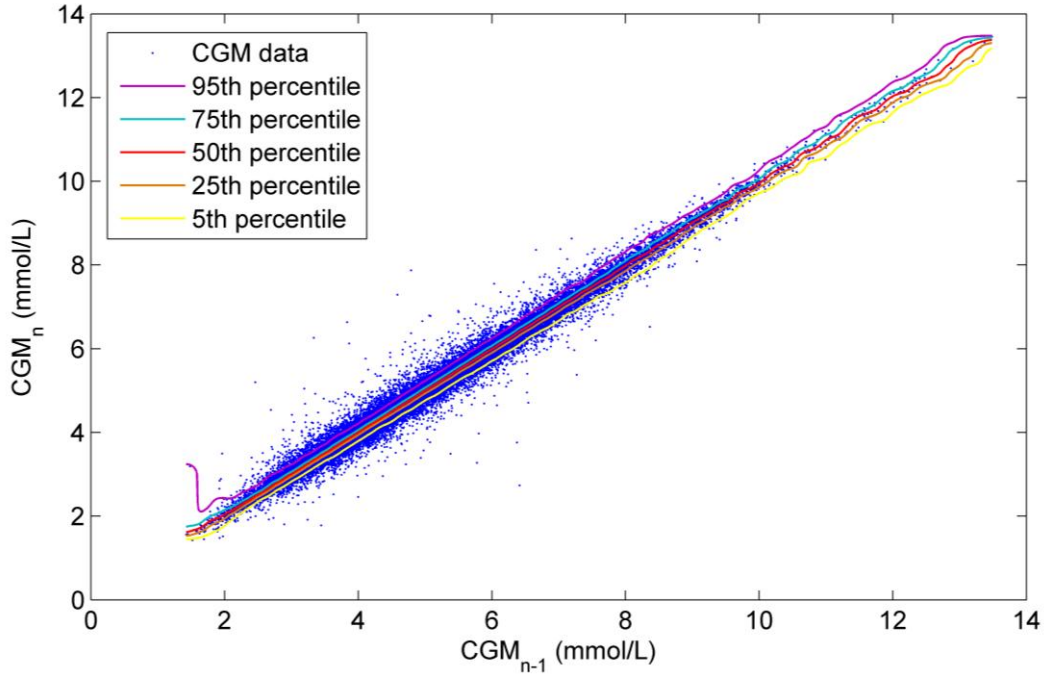


Figure 12.2 Plot of SG measurement pairs (CGM_{n-1}, CGM_n) with contour lines representing the 5th to 95th percentiles, from the bottom of the plot up.

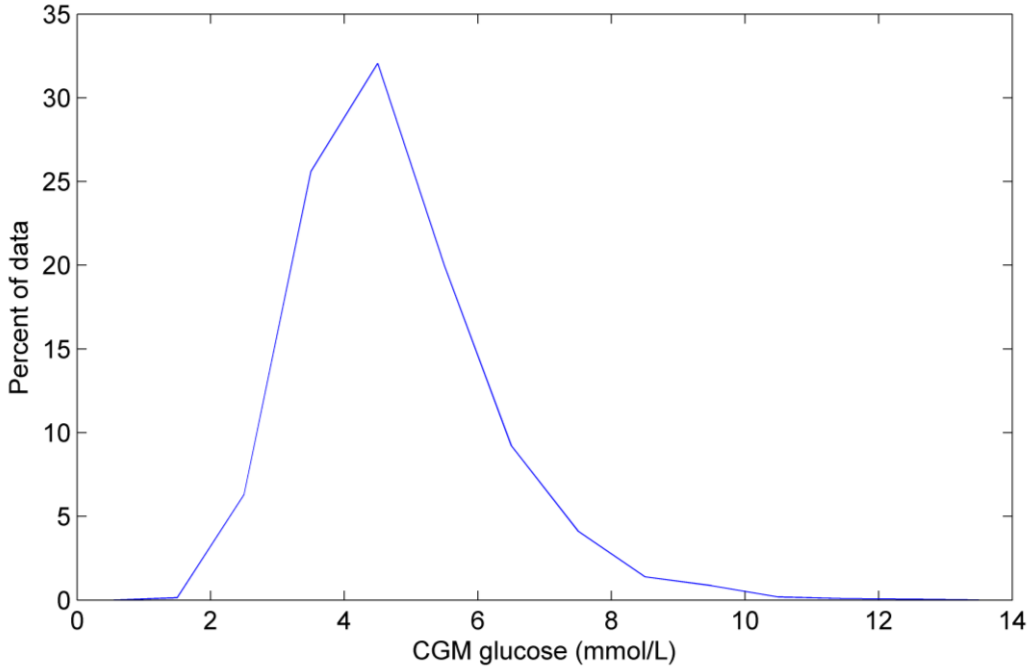


Figure 12.3 Density of the data set by glycaemic level. Density is shown as a percent of the total data set (67,438 measurements).

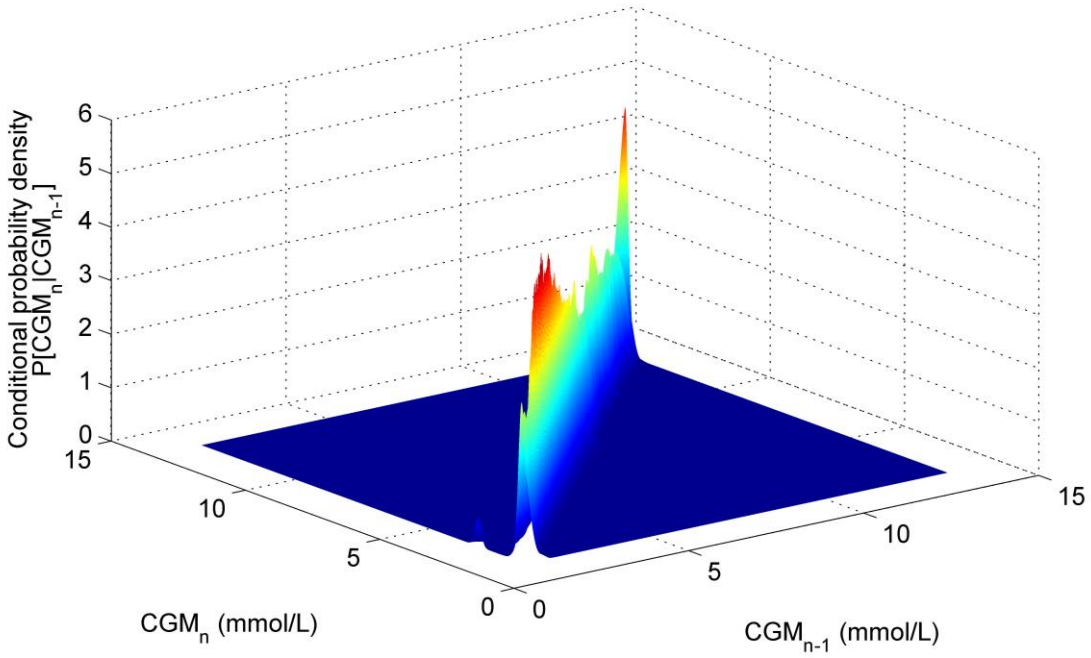


Figure 12.4 Stochastic model surface for this data set. Conditional probability density functions are the surface slices along CGM_{n-1} axis, each slice has an area under the curve summing to 1.0. A colour gradient was used to show the height of the surface.

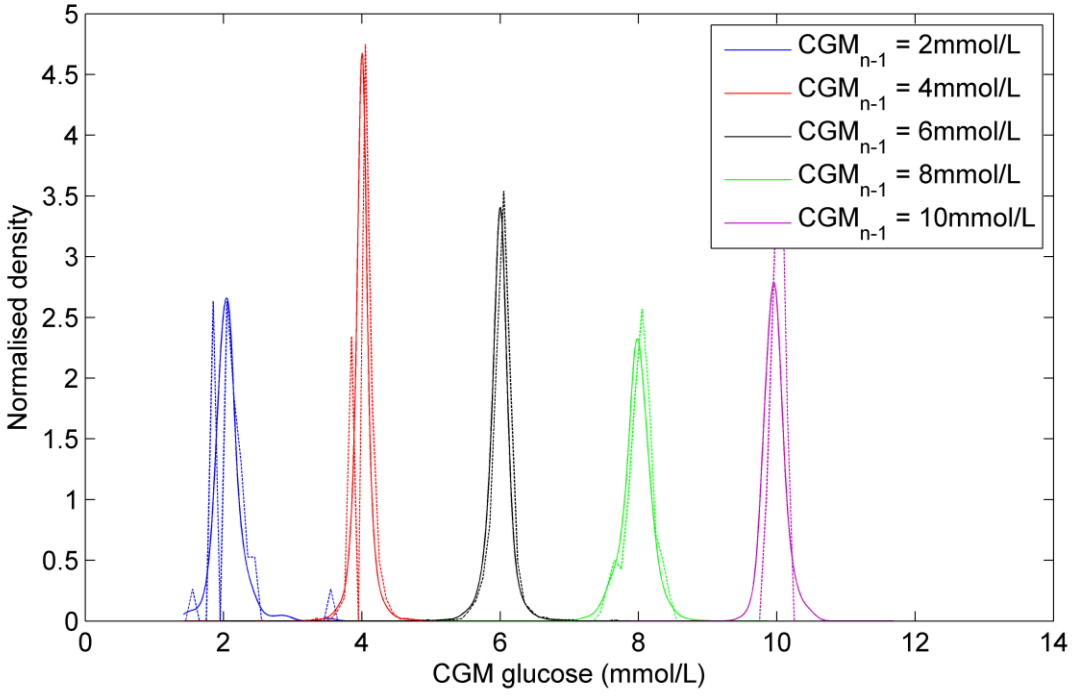


Figure 12.5 Comparison of conditional probability density functions for different levels of CGM_{n-1}. PDFs from the model are solid lines and empirical PDFs from actual SG data are dotted

Table 12.2 Results from a 5-fold validation of the model

5-fold Model Validation (25 MC runs)	80% range	90% range	99% range
Variation across MC runs (Median [IQR])	0.83 [0.79 - 0.86]	0.91 [0.89 - 0.93]	0.99 [0.98 - 0.99]

12.3.2 Classification of representative SG data

Figures 12.6, 12.7 and 12.8 show examples of SG traces that have been coloured using the stochastic classification method. Figure 12.6 shows a stable trace, which is almost entirely dark blue, indicating the measurement-to-measurement change throughout the trace is not unusual. Figure 12.7 shows a trace with several potentially unusual measurements throughout the trace. The hypoglycaemic event that occurs at approximately one day after monitoring began is coloured red and classified as very unusual (>99% range). Figure 12.8 shows a trace with a few potentially unusual measurements for the first three days of monitoring. After day 3 a high proportion of the SG measurements in this figure are classified as very unusual and are coloured red.

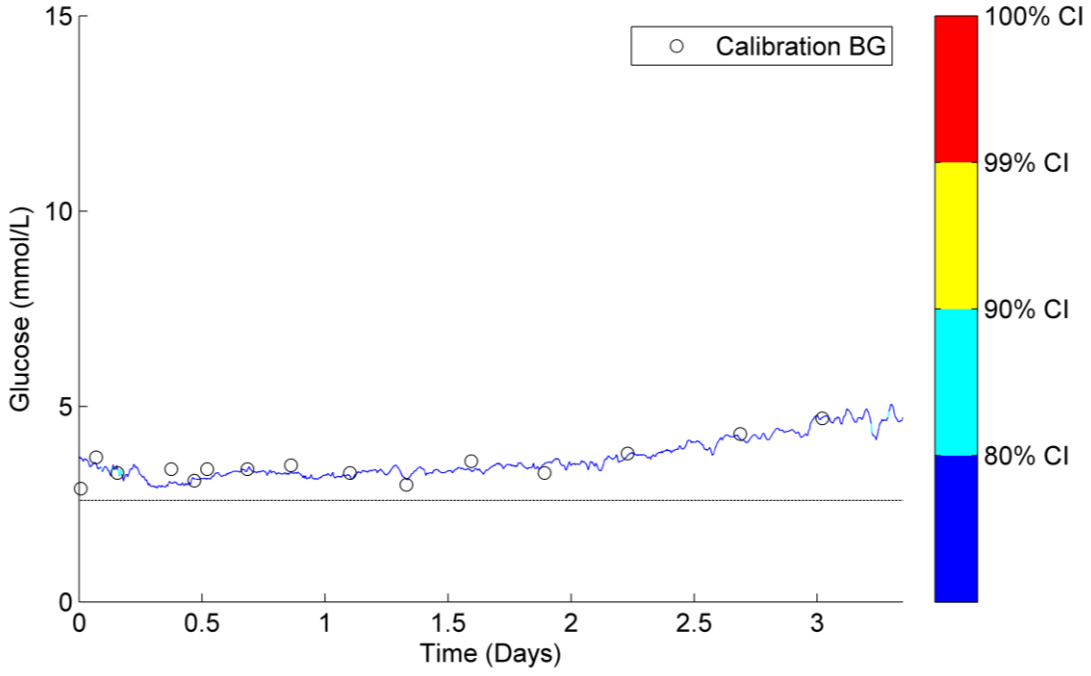


Figure 12.6 Stable SG trace with no yellow or red measurements indicating no SG measurements were classified unusual.

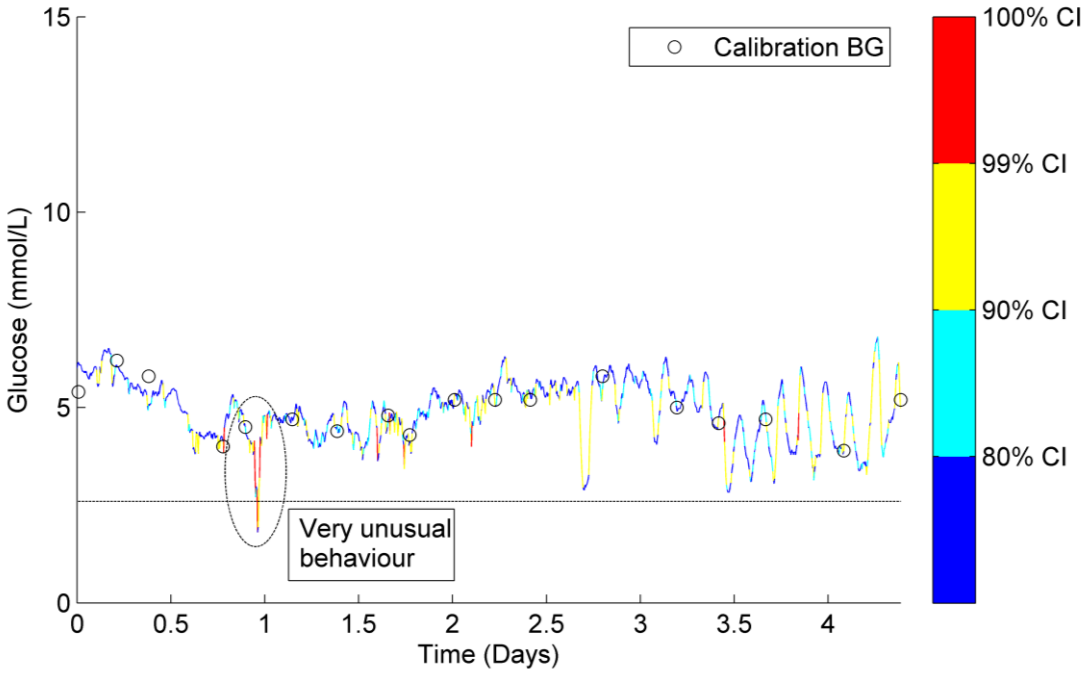


Figure 12.7 SG trace with several measurements classified as mildly unusual. Note the hypoglycaemic event at ~1 day which has been classified as very unusual (red).

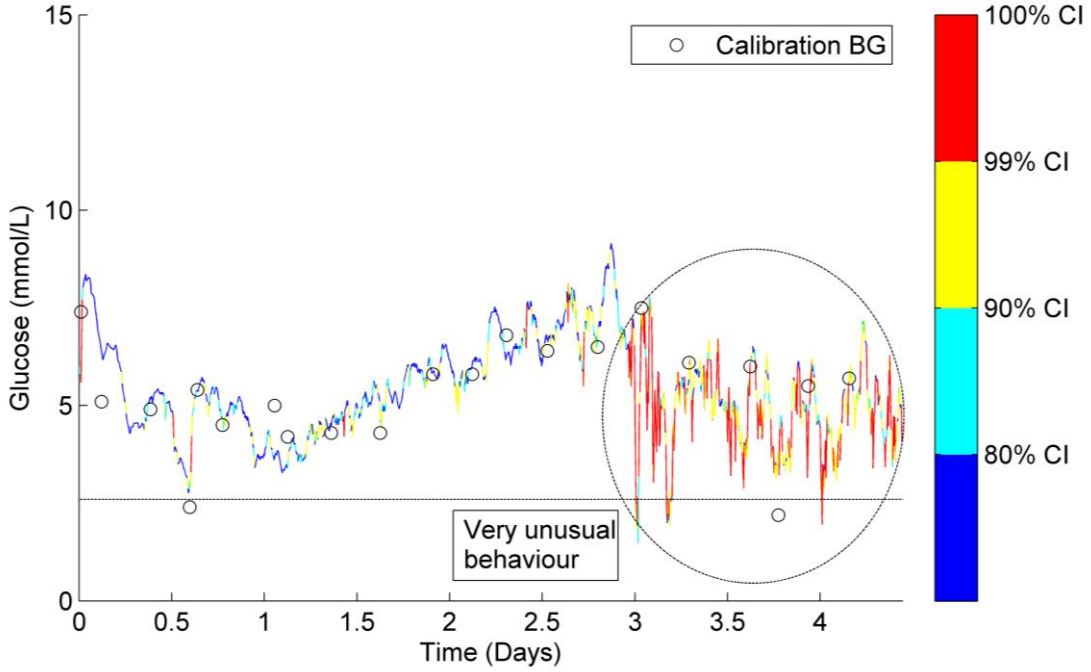


Figure 12.8 SG trace with several measurements classified as mildly unusual. After day 3 the trace is classified as very unusual (red) and could be indicative of sensor malfunction.

12.4 Discussion

12.4.1 Clinical SG data and stochastic model generation

The aim of this study was to design a tool that could aid clinicians in identifying unusual CGM behaviour that should potentially be interpreted with care. Stochastic modelling methods from Lin (2007) and a method of colouring SG traces were used to highlight unusual CGM behaviour clearly and efficiently, in either real-time or retrospectively.

Figure 12.2 and Figure 12.3 give information about the raw data used to create the stochastic model. More than 99% of the data is within 2-10mmol/L range, shown in Figure 12.3. There are several outliers in Figure 12.2 that have a very large change in glycaemia over the 5 minute CGMS Gold measurement interval. The high data density means these outliers have little effect on the model fit, shown by the smooth and tight percentile lines in Figure 12.2. However, below 2mmol/L there are 97 SG measurements and due to the relatively low data density the outliers have more impact on the model fit. This effect is clearly seen in the 95th percentile line of the model, which strays upward at levels below 2mmol/L. Similarly, above 10mmol/L, there are only 232 measurements and the percentile lines all have a wave-like shape, again showing the effect of outliers where data density is low. A greater data density would alleviate these issues without changing the approach.

The quality of data used to create a stochastic model will affect how future SG measurements are classified. If a model is created using low quality data, containing a significant number of outliers, then the model could potentially classify future outlier measurements as ‘usual’ or ‘expected’. Data quality is particularly important when using small data sets, as single outliers can have more effect on the overall model. However, the growing use of databases and electronic records means that collecting large amounts of data that are fully representative will become easier, even for specific sub-cohorts. Thus, this work foreshadows an application with potential for effective and typical clinical use in future, as the method of creating such models is general.

Figure 12.4 shows the surface of the stochastic model. The colour gradient shows how the shape of the model changes in the domain of CGM_{n-1} and that a single, global probability density function is not applicable to this data set. Figure 12.5 further reinforces this conclusion with 5 probability density functions taken from the model at different CGM_{n-1} values, resulting in 5 different shaped density functions. These PDFs are also used to show that the model fits the empirical data well. The model PDFs (solid lines) overlay the empirical data (dotted lines) with only minor discrepancies.

Additionally, the model fit was checked using a 5-fold validation, with results shown in Table 12.2. Across the 25 MC runs, the 80% range typically captured 83% of the data, the 90% range typically captured 91% of the data and 99% range typically captured 99% of the data. These results show the model to capture the SG data well over the expected range of SG values, with little variation in the percentage of data captured in each range.

12.4.2 Classification of representative SG data

Figure 12.6, Figure 12.7 and Figure 12.8 show different SG data sets and how the stochastic model classified the individual SG measurements within them. Figure 12.6 shows a very stable, flat SG trace with only small variations over the 3.5 days of monitoring. The SG trace passes near all calibration measurements and there doesn't appear to be any unusual CGM behaviour. The stochastic model classified almost the entire trace as dark blue indicating no unusual CGM behaviour. The interpretation of this trace would not likely be influenced with the additional information provided by the model.

Figure 12.7 shows a less stable SG trace with a lot more variability. This trace contains a few yellow and red sections that potentially need to be interpreted with care. The focus of this discussion is the 'hypoglycaemic event' that occurs at ~day 1. In the sequence of 5 measurements that lead up to the 1.8mmol/L minimum, there are two drops of ~1mmol/L per 5 minute measurement interval (~12mmol/L/hour). The model has determined these drops are extreme outliers, as would be expected physiologically, and consequently, they have been

coloured red. The trace then rises to above 4mmol/L in 5 measurements, similarly with two rises of ~1mmol/L per 5 minute measurement interval. Although the physiological limits of glucose rate-of-change are still unknown, the level of sensor error that has been reported in previous CGM studies (Breton & Kovatchev 2008; Goldberg et al. 2004) suggest that this hypoglycaemic event could potentially be either glycaemia or sensor error, with the error being the more likely cause.

Interestingly, the bottom of the 'hypoglycaemic event' at ~day 1 contains a dark blue classification; indicating the measurement-to-measurement change immediately after the nadir was not unusual. This outcome was not unexpected, because the lag 1 model classified the change between any two consecutive SG measurements independent of previous classifications. Additionally, the data in Figure 1 is generally centred on the line $CGM_{n-1} = CGM_n$. Consequently, the model surface in Figure 12.4 peaks along or near this line. Thus, for the hypoglycaemic event at day 1 when the CGM reported 1.8mmol/L followed by 1.9mmol/L, $CGM_{n-1} \approx CGM_n$, and the change was classified as expected (blue = within 80% range).

It is important to note that the aim of the stochastic model is not to try and determine the accuracy of the CGM device or the specific cause of a drop in CGM glucose, but rather to highlight the fact it should be interpreted with care. Furthermore, if the stochastic model was implemented in a real-time clinical setting and the downward SG measurements were observed, it would be beneficial for the clinician to know whether or not the sequence of measurements is typical of CGM devices and that patient cohort.

It should also be noted that without an accurate BG measurement at ~1 day, no exact conclusion can be drawn about the whether the hypoglycaemic event in this data was a sensor artefact, or a true glycaemic event. However, this lack of confirmation is often the reality with CGM data. Clinical protocols might use stochastic information to justify an added BG measurement to clarify a potentially significant event. After an event, such traces would yield insight not present at the bedside.

Figure 12.8 shows an example of SG data that becomes increasingly more variable and unstable at approximately day 3 of monitoring. Before day 3, the SG trace is predominantly blue and cyan with only small patches of yellow and occasionally red. However, after day 3 the SG trace is almost entirely red indicating the stochastic model has classified these measurements as very unusual. The sudden apparent degradation of reliable SG measurements could be due to a sensor failure. This conclusion is not unreasonable, given the sensors used in this study were validated for 3 days of continuous monitoring. Again, without more frequent, accurate BG measurements during the period after day 3, no definitive conclusions can be drawn. However, this example represents another potential use of this stochastic model classification method that might be useful to users of CGM devices in determining sensor failures.

Finally, the stochastic model and classification methods were used retrospectively in this study. For real-time use, a stochastic model would be generated using prior SG data, and further SG measurements would be entered into the system in real-time. Classification of paired SG measurements (CGM_{n-1}, CGM_n) takes a fraction of a second, so the corresponding colour coded segment of SG trace would be displayed without significant delay, estimated as less than 1 second. The major limitation in implementing the method in real-time is the ability to stream SG data to a computer or onboard processing in real-time. Although the technology is available, only a limited number of CGM devices currently offer the necessary capabilities.

12.5 Summary

A stochastic model based method was presented and shown to be capable of classifying SG measurements to highlight unusual CGM behaviour. The method uses a colour coded SG trace to convey the information quickly and efficiently and is computationally light enough to be used retrospectively or in real-time.

There are several potential uses for the stochastic classification, which include, but are not limited to, classification of hypoglycaemia and detection of potential

sensor failure. Equally, they can augment alarm methods (Pretty et al. 2010) or be used to time BG measurements more optimally, such as in neonates where blood draws are restricted. Overall, while BG measurements are required to draw definitive conclusions about glycaemic events, the stochastic model provides another level of information to aid users in interpretation and decision making.

Chapter 13. Using CGM to detect neonatal hypoglycaemia

Chapters 9-12 discussed the inner workings of CGM devices and how to utilise them in clinical studies of neonatal hypoglycaemia. Research in this chapter applies that knowledge and some of those techniques to analyse neonatal hypoglycaemia in a large cohort of at risk infants. Furthermore, the advantages/disadvantages of CGM and intermittent BG monitoring are discussed. Finally, a potentially optimal protocol for monitoring glycaemia in newborn babies is presented in the conclusions.

13.1 Introduction

Only a small number of studies have investigated neonatal hypoglycaemia using CGM devices. To the author's knowledge, since the first report of CGM in VLBW infants in 2005 (Beardsall et al. 2005) there have been two other published studies of CGM in newborn infants (Harris et al. 2010; Iglesias Platas et al. 2009). These studies report the prevalence of hypoglycaemia to be ~2.4% of all SG values and the hypoglycaemic time per patient to be between 0-11%. Furthermore, CGM is reported to detect hypoglycaemia at a significantly higher rate than intermittent BG monitoring, due to hypoglycaemia often occurring between consecutive BG measurements. Although there are obvious benefits to CGM, there have also been concerns about the devices accuracy at low glucose concentrations and its ability to detect hypoglycaemia (Ligtenberg et al. 2011).

Correlation between SG measurements and paired meter BG readings in neonates is in the range of 0.87 - 0.96 (Beardsall et al. 2005; Beardsall et al. 2013; Iglesias Platas et al. 2009). Mean error of CGMS values in neonates ranges from 6.9% to 12.7% over 7 days of monitoring, with the highest error occurring on day 1 (Beardsall et al. 2013). In addition, a bias exists that causes the CGM to over read at low glucose levels and under read at high glucose levels (Beardsall et al. 2013). Bias of this nature is likely due to regression in the calibration algorithm, which was discussed in detail in Chapter 10. Overall, CGM accuracy in

this group of patients is acceptable and SG data can provide valuable information about hypoglycaemia that cannot be obtained by intermittent BG monitoring.

This chapter investigates hypoglycaemia in a large cohort of at risk neonates whose glucose levels were monitored using CGM devices. Knowledge and some of the techniques discussed in previous chapters are utilised to ensure robust and reliable results. Ultimately, results from this research will be compared to developmental outcomes at 2 and 4.5 years of age, to improve the definition, diagnosis and management of neonatal hypoglycaemia.

13.2 Subjects and Methods

13.2.1 Patients

This analysis uses CGM device data and BG data from 161 newborn infants born in Waikato Hospital, New Zealand. Babies greater than 35 weeks gestation, less than 48 hours old and at risk of neonatal hypoglycaemia were eligible for the study. Primary risk factors for neonatal hypoglycaemia included having a mother with diabetes, prematurity and/or being small or large for gestational age. This study and use of data was approved by the Northern Y Ethics Committee, New Zealand. Table 13.1 summarises the demographics of this cohort.

Table 13.1 Summary of patient demographics. Data are shown as median [IQR] where appropriate

Cohort Demographics	
Number patients	161
Sex (M/F)	83/78
Gestational Age (weeks)	37 [36 - 38]
Birthweight (g)	2755 [2410 - 3410]
Primary Risk (# infants)	
Diabetes	54
Premature	58
SGA or LGA	42
Other	7

13.2.2 Clinical procedures

All patients had interstitial glucose monitoring using CGMS System Gold devices and SOF sensors. Sensors were located on the infant's lateral thigh and were connected to the monitor via the supplied cable. Monitoring began as soon as practical after birth and finished after 7 days or when the baby was no longer considered at risk of hypoglycaemia. During the monitoring period nurses recorded all BG concentrations, feeding and medication for the management of hypoglycaemia. Nurses remained blinded to SG measurements and all calibration BG measurements were entered per the manufacturer's recommendations. Data were downloaded to a PC using CGMS system solutions software version 3.0C.

Blood glucose concentrations were measured according to current clinical guidelines in Waikato Hospital. Blood samples were obtained by heel lances at one hour after birth, then three to four hourly before feeds for the first 24 hours, then six to eight hourly for the subsequent 24 hours. All blood glucose concentrations were measured using a BGA (Radiometer ABL800Flex, Copenhagen). This BGA has a reading range of 0.0 to 60.0mmol/L and a C.V. of 1.4-2.2% with little or no decrease in performance due to variations in haematocrit, pH or PaO_2 (Cembrowski et al. 2010; Watkinson et al. 2012). All BG measurements collected during the study were used for CGM calibration

Mothers were encouraged to provide skin-to-skin contact and feed the baby within the first hour after birth. Prior to birth many mothers expressed and stored breast milk, and when possible babies who did not breast feed adequately were given expressed breast milk by syringe. Babies who were to be formula fed were offered up to 60 ml/kg·day on day one and 90 ml/kg·day on day two. In addition, some babies were given 200mg/kg dextrose gel as an experimental treatment for hypoglycaemia diagnosed by a BG measurement <2.6mmol/L.

13.2.3 Data processing

All SG data were recalibrated using the algorithm described in Chapter 10, primarily to address 3 main issues with CGMS use in neonates. First, the calibration BG measurements from the BGA have little or no measurement error

and an accurate record of the time of measurement is electronically stored by the BGA. These highly accurate, timely BG measurements address the issues discussed in Chapter 11, specifically the effect of calibration errors on measures of hypoglycaemia. Thus, the effects of timing error and measurement error are eliminated, or at least minimised, in this analysis.

Second, recalibrating removes any bias introduced by linear regression aspects of CGM calibration. An over reading bias has been reported at low glucose levels causing hypoglycaemia to be under reported (Beardsall et al. 2013). Recalibration forces SG data to pass through all calibration BG measurements and removes any bias that might otherwise be present.

Third, there have been concerns about the accuracy of the CGMS System Gold in neonates. MAPE by day is reported to be 13% on day 1 and between 7-8% on days 2 through 7 (Beardsall et al. 2013). Furthermore, on days 2-7 at least 73% of measurements are below the American Diabetes Association's (ADA's) recommended target of 10% error. However, on day 1 only 56% of measurements were below 10% error. Recalibrating with frequent calibration BG measurements addresses the issue of reduced accuracy during the first day of monitoring, while preserving important glucose dynamics captured by CGM.

The first and last hour of SG data were trimmed from all data sets to remove end effects that are often present in this data. After trimming, SG data sets shorter than 1 day were excluded from the analysis. Remaining CGM and BG data were reference to the time of birth, to allow a fair comparison between patients. Finally, median filtering using the composite filter described in Chapter 10 removed unwanted and potentially un-physiological dynamics from SG data.

13.2.4 Analysis

Hypoglycaemia was assessed in both the BG data and the recalibrated/filtered SG data for each patient. For SG data it was defined as one or more consecutive SG measurement(s) below 2.6mmol/L, surrounded by SG measurements greater than or equal to 2.6mmol/L. For BG data a single measurement below

2.6mmol/L was considered an episode of hypoglycaemia. The metrics used to quantify hypoglycaemia were:

- **Number:** Number of independent hypoglycaemic events
- **Duration:** Percent of SG record below 2.6mmol/L
- **Severity:** Lowest SG measurement of hypoglycaemic event.
- **Hypoglycaemic index:** Similar in concept to Hyperglycaemic index (Vogelzang et al. 2004). It is defined as the area between the 2.6mmol/L threshold and the SG trace (for SG < 2.6mmol/L) summed over the entire length of stay, normalised by the length of data record. Note: the units used in this study are μmol/L, not mmol/L.

An additional sub-analysis investigated the evolution of glucose levels over time in these at risk babies. Separate CDF plots were generated using SG data from 0-12, 12-24, 24-36 and 36-48 hours after birth. Each plot contained a CDF per patient and the analysis only included patients with more than 2 hours of SG data. The number, *n*, of patients included is shown in the box in each plot.

13.3 Results

Table 13.2 shows results from the analysis of 161 CGM and BG data sets. Throughout the study period 2271 BG measurements were recorded and 399 of them diagnosed as hypoglycaemic. Of the 399 hypoglycaemic episodes, 198 occurred when the CGM was calibrated and producing data. The CGM detected a total of 337 hypoglycaemic episodes, which suggests that 139 episodes were missed by intermittent BG measurements.

CGM also offered insight into the duration of hypoglycaemia. Across the cohort, 6.5% of SG measurements were hypoglycaemic and the hypoglycaemic index was 18μmol/L. In terms of the severity of hypoglycaemia, 52% of CGM hypoglycaemic episodes were between 2.4 - 2.6mmol/L. It should be noted that it was possible for hypoglycaemia to occur below the factory CGM cut-off of 2.2mmol/L because of the calibration method used. Figure 13.1 shows an example of BG monitoring compared to CGM. Between 200-400 minutes after birth, both methods detect an episode of hypoglycaemia. However, later in the

monitoring period at 1400mins, and again at 2200mins, the CGM reports hypoglycaemia that occurred between consecutive BG measurements

Table 13.2 Hypoglycaemia results from CGM device and BG data

Overall cohort results	
BG results	
Number of BG measurements	2271
hypoglycaemic BG measurements	399
Number of BG measurements during CGM	1741
hypoglycaemic BG during CGM	198
CGM results	
Duration of CGM monitoring (days)	327.8
Number of hypoglycaemic episodes	337
Percent duration of hypoglycaemia (%)	6.5
Hypoglycaemic index (µmol/L)	18
Hypoglycaemia between 2.4 - 2.6mmol/L	174
Hypoglycaemia between 2.2 - 2.4mmol/L	81
Hypoglycaemia between 2.0 - 2.2mmol/L	35
Hypoglycaemia below 2.0mmol/L	47
Per-patient results	
BG results	
Number of BG measurements	13 [10 - 16]
hypoglycaemic BG measurements	2 [1 - 3]
CGM results	
Duration of CGM monitoring (days)	1.9 [1.6 - 2.3]
Number of hypoglycaemic episodes	1 [1 - 3]
Percent duration of hypoglycaemia (%)	4.4 [0.0 - 10.3]
Hypoglycaemic index (µmol/L)	6.3 [0.0 - 21.1]

Per-patient results from the analysis are also shown in Table 13.2. The median number of BG measurements per patient was 13 and the median number of episodes of hypoglycaemia was 2. The results from SG data show a median monitoring period of 1.9 days and a median number of hypoglycaemic episodes of 1. The median per-patient duration of hypoglycaemia was 4.4%, which is slightly less than the 6% reported for the entire cohort, suggesting there were a few patients with long durations of hypoglycaemia that have skewed the results slightly. A similar trend is seen in per-patient hypoglycaemic index results, but the magnitude of skewing is greater.

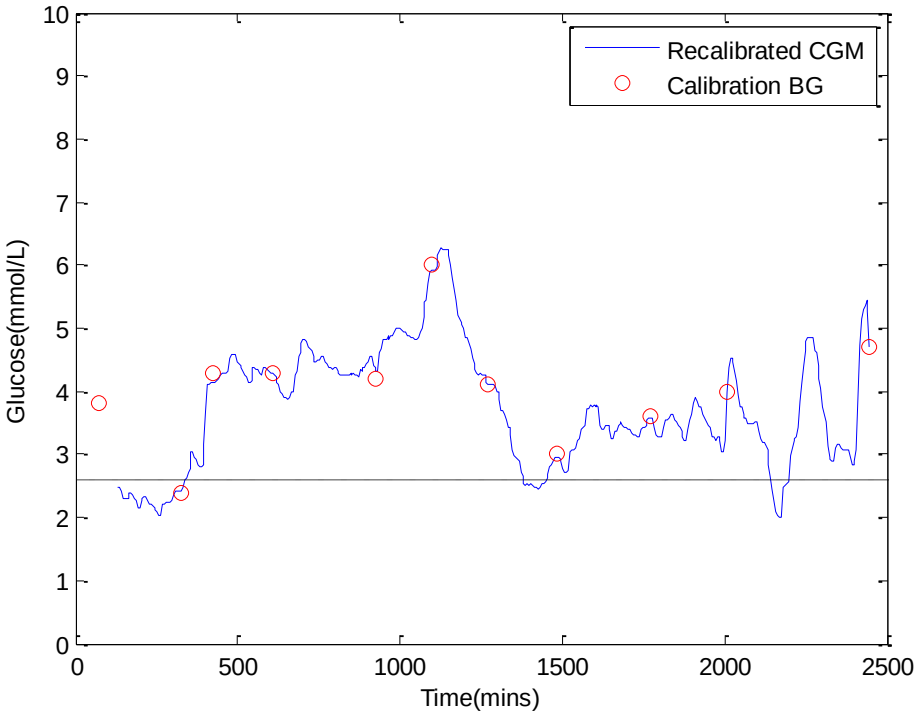


Figure 13.1 SG data and BG data for a representative patient. Note the CGM device and BG measurement between 200-400mins both capture hypoglycaemia, but later in the monitoring period at 1400mins and 2200mins the CGM captures hypoglycaemia missed by the BG measurements.

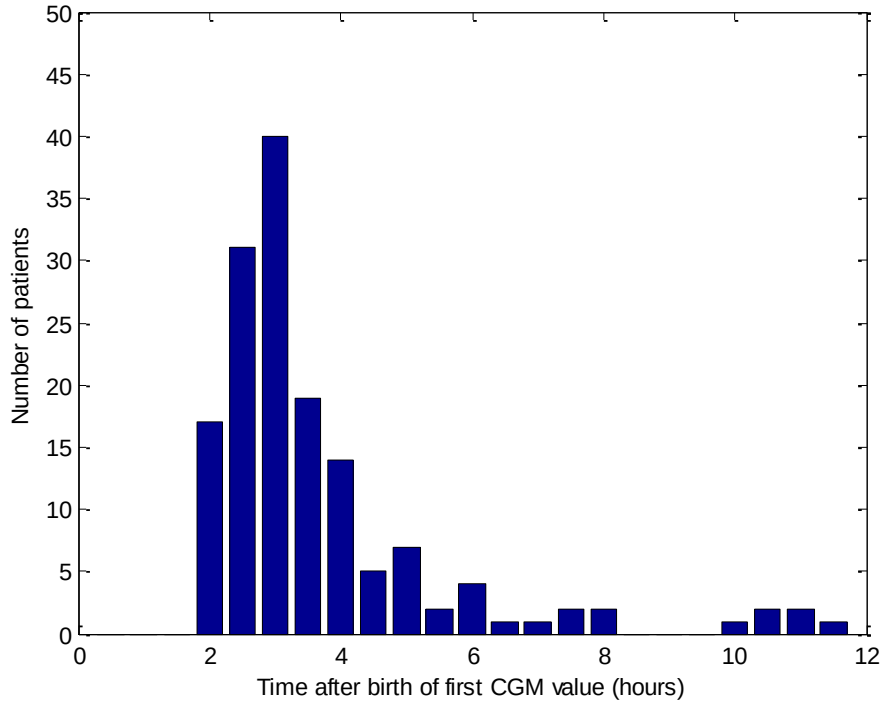


Figure 13.2 Time after birth of first SG measurement.

Figure 13.2 shows the time after birth that the CGM device started outputting data. The median [25th - 75th percentile] time after birth until SG data was available was 206 [174 - 270] minutes. Figure 13.3 shows the number of hypoglycaemic episodes that occur during the first few days after birth, comparing both glucose monitoring techniques. Majority of BG hypoglycaemia occurs during first 12 hours after birth during which time the CGM reports a lower number of episodes. However, after the first 12 hours after birth, the CGM consistently reports more hypoglycaemia than intermittent BG monitoring alone.

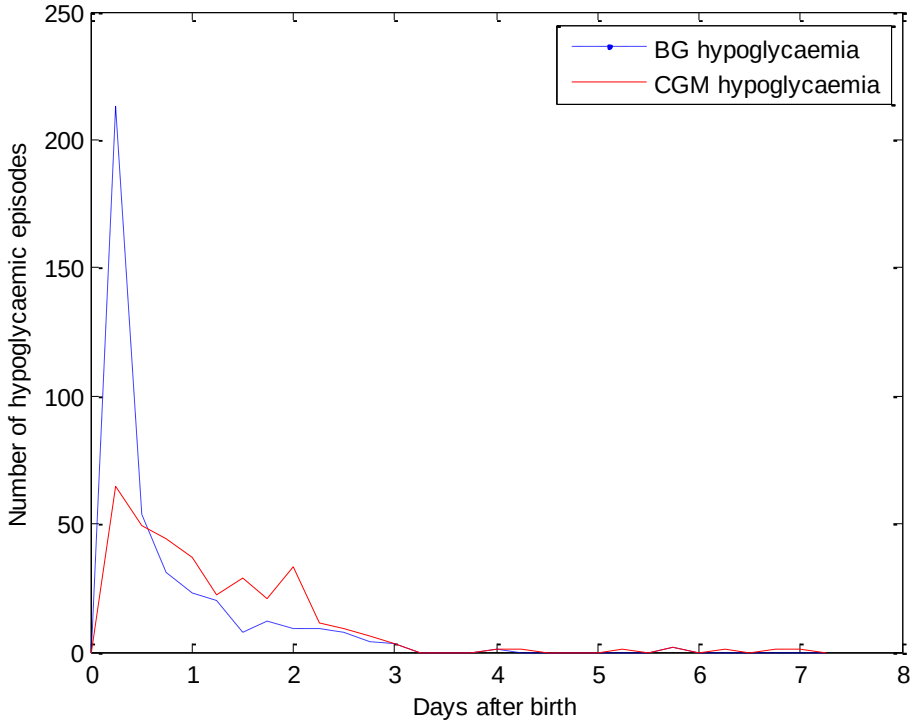


Figure 13.3 Hypoglycaemia as a function of time after birth. The blue dashed line shows hypoglycaemia detected using intermittent BG measurements and the red line shows the number of hypoglycaemic episodes detected using SG data.

Each plot in Figure 13.4 shows the CDF of each patient's SG data for the stated time period after birth. The heavy black vertical line represents the hypoglycaemic threshold to clearly show the level of hypoglycaemia. During the period of 0-12 hours after birth 57% of patients experienced hypoglycaemia, diagnosed using SG data. This reduced during the period 12-24 hours after birth to 39%, and further reduced to 25% and 19% in the subsequent two 12 hour time periods, respectively. There were at least 141 sets of SG data in each

analysis. These results, in conjunction with those seen in Figure 13.3, clearly show that hypoglycaemia is most prevalent in the first day of life in at risk infants.

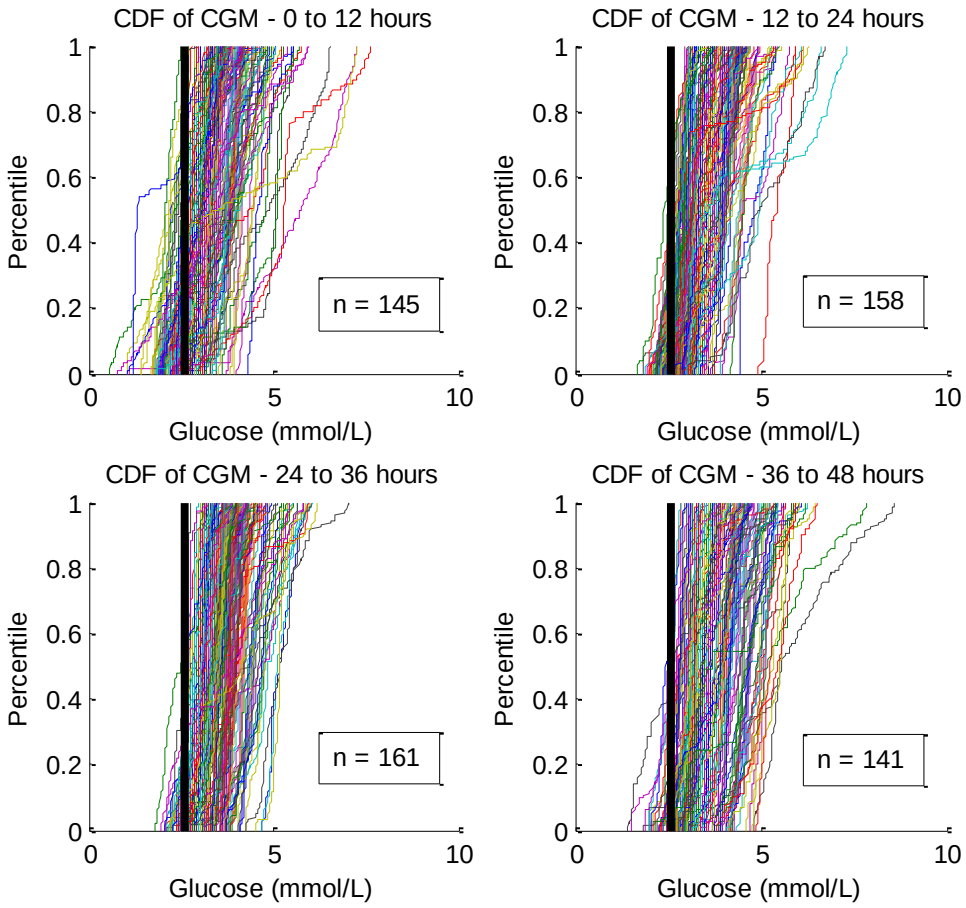


Figure 13.4 CDFs of each patient's SG data for the stated time period after birth. The black vertical line represents the hypoglycaemic threshold and 'n' is the number of patients in each sub-analysis.

13.4 Discussion

The aim of this study was to assess neonatal hypoglycaemia in a large cohort of at risk infants who were monitored using CGM devices. Results from the SG data are compared to results from intermittent BG monitoring.

In this group of at risk neonates, ~17% of BG measurements diagnosed hypoglycaemia. Interestingly, over 50% of hypoglycaemic episodes were

diagnosed in the first 6 hours after birth and over 80% in the first day, as shown in Figure 13.3. In addition, as a percentage of the number of BG measurements per day, hypoglycaemia accounted for 27% of measurements on day 1 and only 7% on day 2. Thus, it is unlikely the results are skewed due to increased sampling on day 1. Nevertheless, these results suggest capturing the first 24 hours of BG data in babies at risk of hypoglycaemia is critical.

Another interesting aspect of Figure 13.3 is the level of hypoglycaemia detected in SG data compared to BG data, throughout the first day of monitoring. The recalibration forces the SG trace to pass through all BG measurements, so it is expected that the number of hypoglycaemic episodes detected by CGM be at least the number detected by BG. However, as Figure 13.2 shows, the median time before the CGM device outputs data is approximately 3 hours 25 minutes and in some rare cases SG data isn't available for 10-12 hours. During the first 6 hours after birth in this study, BG measurements detected 148 hypoglycaemic episodes that occurred before SG data was available. Thus, on day 1 it is essential to have frequent intermittent BG measurements in at risk infants, even if CGM devices are used in parallel.

During days 2 and 3 after birth, when CGM and BG measurements occur concurrently, SG data consistently reports a higher number of hypoglycaemic episodes. This is clearly shown in Figure 13.3, and also in Table 13.2. Considering only BG measurements that were done during CGM monitoring, there were 198 episodes of hypoglycaemia detected by BG and 337 detected by CGM. For example, Figure 13.1 shows episodes of hypoglycaemia at 1400mins and 2200 minutes after birth that were missed by intermittent BG measurements. If these episodes were asymptomatic, it is likely that clinical staff would have never been aware of them without CGM. This example further reinforces the benefit of CGM in these infants.

The prevalence of hypoglycaemia appears to diminish at the end of day 3 and remain almost non-existent during days 4-7. However, SG data sets in this study were relatively short, with a median [25th - 75th percentile] length of 1.9 [1.6 -

2.3] days. There is a not enough data in this study to draw definitive conclusions about the prevalence of hypoglycaemia during the entire first week after birth.

One of the main advantages of CGM in a study of this nature is the ability to classify the duration of hypoglycaemia, which is not possible with infrequent intermittent BG measurements. Across the cohort, 6.5% of SG measurements were below 2.6mmol/L and per patient the median [25th - 75th percentile] were 4.4% [0-10.3%]. These results are marginally higher than other results in the literature, such as the median (range) of 0.4% (0 - 11%) reported by Beardsall and colleagues. However, this outcome is likely due to the cohort used in this study. All infants enrolled in the study were at risk of hypoglycaemia and had at least one diagnosed hypoglycaemic BG measurement. Hence, the results here appear reasonable given the cohort selection.

Overall, the number of patients experiencing hypoglycaemia diminishes over the first 48 hours. Figure 13.4 shows CDFs for each patient during four 12 hour blocks over the first 48 hours. The number of patients with SG values below 2.6mmol/L drops from 57% during 0-12 hours to 19% during 36-48 hours. In addition, each block of 12 hours had at least 141 patients contributing data so the decrease observed is robust to insufficient data. This trend of diminishing hypoglycaemia over the first 48 hours is also supported by the BG results.

This study has several limitations that need to be addressed. First, although feeding protocols are present in all neonatal hospital environments, the use of dextrose gel for treatment of hypoglycaemia was experimental in the study that supplied this data. Researchers involved with this study are still blind to the control/placebo arms so they could not be separated for individual analysis. Second, the SG data sets are relatively short in this study as it was only the first 48 hours that were of interest to the study that supplied this data, so it not possible to draw conclusions after this time period. Finally, the time after birth that SG data was available varied from 2-12 hours. Typically, the CGMS System Gold will start outputting data after 1-2 hours, but it is not always possible to get a sensor inserted in the infant straight after birth.

13.5 Summary

Overall, this study has show that hypoglycaemia in at risk neonates is most prevalent on the first day after birth. An optimal measurement protocol for at risk infants would likely involve CGM for the first week after birth with frequent intermittent BG measurements for the first day. The duration of hypoglycaemia in these patients can vary from 0-10+%, but typically lies in the region of 4-6%. Results from this research will be valuable when assessing long term outcomes of these patients and the effects hypoglycaemia might have had on their neurological development.

Chapter 14. Summary of CGM in NICU

The second half of this thesis has investigated CGM in critically ill infants in neonatal intensive care. The overall goal of this work was to assess whether these devices, which were designed for use outside of the hospital, could perform reliably in the NICU and give additional information to conventional methods of BG monitoring. Specifically, CGM devices were used to monitor newborn infants at risk of hypoglycaemia to improve our understanding of the rate and severity of this common metabolic problem.

Three previous studies that used CGM devices in neonates reported that the devices were well tolerated by infants and SG data revealed trends in glycaemia that were missed by the standard approach using intermittent BG measurements. Hence, the use of CGM devices appeared to show a well-tolerated measurement that provided clinically significant new data. However, there were still concerns about the accuracy of CGM devices in infants, particularly on the first day of monitoring and at low BG levels where calibration may not be robust. With a thorough understanding of the CGM technology, signal processing and mathematical techniques could be used to enhance the quality of the data and/or aid in the interpretation of results.

Calibration algorithms convert raw sensor signals into useful SG data and the effects of CGM device calibration on metrics used to classify hypoglycaemia were investigated. The built-in factory algorithm is based around linear regression and SG data tended to report high at low BG levels. An alternative calibration scheme was tested, that made more optimal use of the high accuracy calibration BG measurements available in the NICU from a BGA. Novel non-linear filtering was also implemented to reduce the impact of any sudden, outlying drops/rises in SG data due to sensor noise or artefacts.

The results showed that all metrics of hypoglycaemia increased after recalibration and the CGM device was confirmed to be reporting higher values at low BG levels. Filtering had a significant effect on the number of hypoglycaemic

episodes, but the duration of hypoglycaemia showed little change. These results highlighted two important considerations for future studies of CGM devices in neonates. First, if high accuracy calibration measurements are available, from a BGA or lab, then it may be more appropriate to recalibrate SG data. Second, when assessing hypoglycaemia in SG data, the duration of hypoglycaemia and/or hypoglycaemic index are more robust metrics when using CGM devices than the standard approach of counting the number of hypoglycaemic episodes.

The sensitivity of these hypoglycaemia metrics to errors in calibration BG measurements was also assessed. Specifically, error in BG measurements and delays in entering calibration BG measurements into the CGM device were investigated in-silico. Three BG error models were created using clinical data from an assessment of glucometer accuracy in ICU patients, and two timing error models were created using data from two investigations of CGM devices in critically ill patients.

Across the cohort, measurement error tends to have a much larger impact on hypoglycaemia metrics than any clinically observed or realistic timing error. However, if a patient is particularly variable, then timing error can have a much more measureable impact due to the high rate-of-change of BG. This specific issue is most notable if calibration occurs during a time of high BG rate-of-change. The effect of bias in calibration BG measurements was twofold: 1) A negative error bias increased the prevalence of hypoglycaemia; and, 2) it also increased the amount of variation seen in hypoglycaemic metrics between MC simulations. The opposite was also true for positive biases. Clinically, if a patient is variable, then calibration BG measurements should be entered into the device with minimal delay, potentially at the expense of a high accuracy measurement from a BGA or similar. However, if the patient's glycaemia is relatively stable, then it is more beneficial to get the most accuracy BG measurement possible within clinical constraints. These outcomes highlight a need to use these devices in concert with observations about patient-specific dynamics that are readily observable in the SG data.

Further, if CGM devices are going to be used in the clinical setting to monitor, diagnose and potentially aid in the treatment of abnormal glycaemia, clinicians need to know the data are reliable and accurate. In the context of detecting hypoglycaemia in neonatal infants, investigators would benefit from knowing whether a significant drop in glucose is likely 'real' or simply an artefact of the device. Stochastic modelling methods, originally developed for insulin sensitivity prediction, were utilised and developed to detect outliers in SG data. In this work the emphasis was on detecting unusual episodes of hypoglycaemia, but the method presented is fully general.

A stochastic classification model, based on kernel density estimation, was used to classify unusual SG data. SG data were displayed on a time series plot and colour coded based on their likelihood of being 'real', so that outlier data could be quickly and easily identified. The stochastic classification successfully identified unusual CGM device behaviour, and in particular several questionable episodes of hypoglycaemia. Overall, while BG measurements are required to draw definitive conclusions about glycaemic events, the stochastic model provides another level of information to aid users in interpretation of the relatively very rich data provided by CGM devices and thus in decision making.

With a series of tools/methods available and a better understanding of how to utilise CGM devices in neonates, hypoglycaemia in a cohort of 161 at risk infants was investigated. All patients had monitoring using the Medtronic CGMS System Gold, which was calibrated multiple times per day using BG measurements from a BGA. Due to having high accuracy calibration measurements, SG data were recalibrated and forced to pass through the calibration BG measurements to eliminate the impact of any error that could have been introduced by the linear regression aspect of factory calibration. Both BG data and SG data were used to quantify hypoglycaemia, and the results were compared.

Results from BG measurements showed that ~17% of BG measurements taken during the study identified hypoglycaemia and over 80% of the episodes occurred in the first day after birth. Hypoglycaemia was detected at a higher rate

by BG measurements than SG measurements on the first day after birth, largely due to delays in inserting the CGM sensor and a short warm-up period. However, with concurrent BG and SG data available, the SG data consistently identified hypoglycaemia at a higher rate suggesting the BG measurements were not capturing some episodes. The SG results showed that the duration of hypoglycaemia in these patients can vary from 0-10+% of the monitoring period, but typically lies in the region of 4-6%. Overall, this work showed that hypoglycaemia in at risk neonates is most prevalent on the first day after birth and an optimal measurement protocol for at risk infants would likely involve CGM for the first week after birth with frequent intermittent BG measurements for the first day.

In summary, the research presented in Chapters 9-13 shows that CGM devices can add value to traditional BG monitoring in the neonatal unit. The findings presented in this thesis will be especially valuable when assessing long term outcomes of these patients, and in particular, the effects hypoglycaemia might have had on their neurological development.

Chapter 15. Future work

The work presented in this thesis shows how CGM devices can be used effectively and robustly in both adult and neonatal ICUs to aid clinical care. In particular, it shows how to use them to maximise their strengths and minimise or manage their weaknesses. Overall, the results are promising and these devices have the potential to work reliably in both clinical departments. However, further research is needed before the maximum benefits of CGM can be realised. This chapter outlines several aspects of future work that could be investigated to advance the knowledge of CGM and ultimately improve care for critically ill patients.

15.1 Adult ICU Future work

The research presented in the first half of this thesis showed that CGM devices have the ability to improve TGC methods in an adult critical care unit. A pilot trial of CGM devices demonstrated that off-the-shelf CGM devices, currently targeted towards outpatient diabetes monitoring, are capable monitoring certain ICU patients with a high level of accuracy. A larger observational trial of CGM devices is required to determine which patient, illness, drug and/or therapy combinations negatively impact CGM performance. With a sound understanding of how these factors influence device performance, prospective patients who would most benefit from using CGM could be identified based on the expected level of CGM device performance. A larger trial of CGM devices should also include the newly released hospital CGM device, Sentrino, which is reported to have excellent reliability and safety in critically ill patients and, importantly, is approved for use in those patients (Kosiborod et al. 2013).

Data collected from a larger trial of CGM device in ICU could also be used for a variety of different in-silico studies. An ICU specific mathematical model of CGM error characteristics could be developed, allowing virtual SG data to be generated that mimics the behaviour of the actual CGM device. Virtual SG data could be used in conjunction with clinically validated mathematical models of the

insulin-nutrition regulatory system to develop TGC protocols. Unlike the in-silico research in this thesis, which replaced BG measurements collected by nurses with virtual SG data, a complete model of CGM dynamics and errors would allow specifically designed TGC protocols to make use of all SG data. Furthermore, alarming algorithms could then be optimally designed to be implemented as guardrails to protect patients from abnormal and dangerous glycaemic excursions, such as hypoglycaemia.

Ultimately, the aim of such future studies should be to develop the technology and protocols that would allow a fully automated closed loop system to operate in the ICU. These computer driven protocols would likely modulate insulin and/or nutrition to ensure excellent control of BG, with minimal user input. With technology constantly evolving, closed loop control in the ICU could potentially be realised in the not too distant future.

15.2 Neonatal ICU Future work

The research presented in the second half of this thesis showed that CGM devices can be used in conjunction with traditional BG monitoring to improve hypoglycaemia detection and classification in neonates. This research was carried out as part of a wider study, The CHYLD Study, which is a large multi-disciplinary prospective study investigating the development of young children who were at risk of developing hypoglycaemia in the early neonatal period.

The infants who had CGM are scheduled to be followed up at 2 and 4.5 years of age at which point their growth, cognitive and neurodevelopment is assessed in relation to the duration, severity, and frequency of hypoglycaemia in the first days of life. These results will not only assess the impact of neonatal hypoglycaemia as it is currently defined (BG < 2.6mmol/L), they will also potentially allow improved definitions of hypoglycaemia to be identified. New definitions or diagnostic guidelines for hypoglycaemia could change clinical practice methods all over the world, and, more importantly, reduce the number of negative complications associated with abnormal glycaemia in the NICU.

Another avenue that could be explored with the addition of CGM devices in the NICU is prophylactic dextrose gel treatments, as a means of preventing dangerous hypoglycaemia. A recent study investigating the effectiveness of dextrose gel in treating hypoglycaemia, identified by intermittent BG measurements, showed that dextrose gel was effective at restoring low BG concentrations to normal levels. Thus, the next logical step is to integrate real-time CGM devices and alarming algorithms in to the NICU so that a preventative gel treatment can be administered at the predicted onset of hypoglycaemia, reducing or even eliminating the risk of that episode.

Overall, it is anticipated that CGM devices could be implemented in a similar way to the ICU, for closed loop control of BG and safeguarding against the negative outcomes associated with abnormal glycaemia. A version of the CG protocol STAR has been developed for, and is currently used in, the Christchurch Hospital NICU. CGM devices could be integrated with this computerised system to improve BG monitoring, and ultimately improve care for infants in the NICU.

References

Abbott 2010, 'Abbott Optium Test-Strip Packet Insert'. Abbott Diabetes Care Ltd., UK.

Adolfsson, P, Ornhagen, H, Eriksson, BM, Cooper, K & Jendle, J 2012, 'Continuous glucose monitoring--a study of the Enlite sensor during hypo- and hyperbaric conditions', *Diabetes Technol Ther,* vol. 14, no. 6, pp. 527-32.

Aragon, D 2006, 'Evaluation of nursing work effort and perceptions about blood glucose testing in tight glycemic control', *Am J Crit Care,* vol. 15, no. 4, pp. 370-7.

Bagshaw, SM, Bellomo, R, Jacka, MJ, Egi, M, Hart, GK & George, C 2009, 'The impact of early hypoglycemia and blood glucose variability on outcome in critical illness', *Crit Care,* vol. 13, no. 3, p. R91.

Bailey, T, Zisser, H & Chang, A 2009, 'New features and performance of a next-generation SEVEN-day continuous glucose monitoring system with short lag time', *Diabetes Technol Ther,* vol. 11, no. 12, pp. 749-55.

Bayley, N 1965, 'Comparisons of Mental and Motor Test Scores for Ages 1-15 Months by Sex, Birth Order, Race, Geographical Location, and Education of Parents', *Child Dev,* vol. 36, pp. 379-411.

Beardsall, K, Ogilvy-Stuart, AL, Ahluwalia, J, Thompson, M & Dunger, DB 2005, 'The continuous glucose monitoring sensor in neonatal intensive care', *Arch Dis Child Fetal Neonatal Ed,* vol. 90, no. 4, pp. F307-10.

Beardsall, K, Vanhaesebrouck, S, Ogilvy-Stuart, AL, Vanhole, C, VanWeissenbruch, M, Midgley, P, Thio, M, Cornette, L, Ossuetta, I, Palmer, CR, Iglesias, I, de Jong, M, Gill, B, de Zegher, F & Dunger, DB 2013, 'Validation of the continuous glucose monitoring sensor in preterm infants', *Arch Dis Child Fetal Neonatal Ed,* vol. 98, no. 2, pp. F136-40.

Beck, RW, Hirsch, IB, Laffel, L, Tamborlane, WV, Bode, BW, Buckingham, B, Chase, P, Clemons, R, Fiallo-Scharer, R, Fox, LA, Gilliam, LK, Huang, ES, Kollman, C, Kowalski, AJ, Lawrence, JM, Lee, J, Mauras, N, O'Grady, M, Ruedy, KJ, Tansey, M, Tsalikian, E, Weinzimer, SA, Wilson, DM, Wolpert, H, Wysocki, T & Xing, D 2009, 'The effect of continuous glucose monitoring in well-controlled type 1 diabetes', *Diabetes Care,* vol. 32, no. 8, pp. 1378-83.

Bequette, BW 2010, 'Continuous glucose monitoring: real-time algorithms for calibration, filtering, and alarms', *J Diabetes Sci Technol,* vol. 4, no. 2, pp. 404-18.

Bhat, MA, Kumar, P, Bhansali, A, Majumdar, S & Narang, A 2000, 'Hypoglycemia in small for gestational age babies', *Indian J Pediatr,* vol. 67, no. 6, pp. 423-7.

Bistrian, BR 2001, 'Hyperglycemia and infection: which is the chicken and which is the egg?', *JPEN J Parenter Enteral Nutr,* vol. 25, no. 4, pp. 180-1.

Bland, JM & Altman, DG 1999, 'Measuring agreement in method comparison studies', *Stat Methods Med Res,* vol. 8, no. 2, pp. 135-60.

Boyne, MS, Silver, DM, Kaplan, J & Saudek, CD 2003, 'Timing of changes in interstitial and venous blood glucose measured with a continuous subcutaneous glucose sensor', *Diabetes,* vol. 52, no. 11, pp. 2790-4.

Branco, RG, Garcia, PC, Piva, JP, Casartelli, CH, Seibel, V & Tasker, RC 2005, 'Glucose level and risk of mortality in pediatric septic shock', *Pediatr Crit Care Med,* vol. 6, no. 4, pp. 470-2.

Breton, M, Farret, A, Bruttomesso, D, Anderson, S, Magni, L, Patek, S, Dalla Man, C, Place, J, Demartini, S, Del Favero, S, Toffanin, C, Hughes-Karvetski, C, Dassau, E, Zisser, H, Doyle, FJ, 3rd, De Nicolao, G, Avogaro, A, Cobelli, C, Renard, E & Kovatchev, B 2012, 'Fully integrated artificial pancreas in type 1 diabetes: modular closed-loop glucose control maintains near normoglycemia', *Diabetes,* vol. 61, no. 9, pp. 2230-7.

Breton, M & Kovatchev, B 2008, 'Analysis, modeling, and simulation of the accuracy of continuous glucose sensors', *J Diabetes Sci Technol,* vol. 2, no. 5, pp. 853-62.

Bridges, BC, Preissig, CM, Maher, KO & Rigby, MR 2010, 'Continuous glucose monitors prove highly accurate in critically ill children', *Crit Care,* vol. 14, no. 5, p. R176.

Brunkhorst, FM, Engel, C, Bloos, F, Meier-Hellmann, A, Ragaller, M, Weiler, N, Moerer, O, Gruendling, M, Oppert, M, Grond, S, Olthoff, D, Jaschinski, U, John, S, Rossaint, R, Welte, T, Schaefer, M, Kern, P, Kuhnt, E, Kiehntopf, M, Hartog, C, Natanson, C, Loeffler, M & Reinhart, K 2008, 'Intensive insulin therapy and pentastarch resuscitation in severe sepsis', *N Engl J Med,* vol. 358, no. 2, pp. 125-39.

Brunner, R, Adelsmayr, G, Herkner, H, Madl, C & Holzinger, U 2012, 'Glycemic variability and glucose complexity in critically ill patients: a retrospective analysis of continuous glucose monitoring data', *Crit Care,* vol. 16, no. 5, p. R175.

Brunner, R, Kitzberger, R, Miehsler, W, Herkner, H, Madl, C & Holzinger, U 2011, 'Accuracy and reliability of a subcutaneous continuous glucose-monitoring system in critically ill patients', *Crit Care Med,* vol. 39, no. 4, pp. 659-64.

Buckingham, BA, Kollman, C, Beck, R, Kalajian, A, Fiallo-Scharer, R, Tansey, MJ, Fox, LA, Wilson, DM, Weinzimer, SA, Ruedy, KJ & Tamborlane, WV 2006, 'Evaluation of factors affecting CGMS calibration', *Diabetes Technol Ther,* vol. 8, no. 3, pp. 318-25.

Burns, CM, Rutherford, MA, Boardman, JP & Cowan, FM 2008, 'Patterns of cerebral injury and neurodevelopmental outcomes after symptomatic neonatal hypoglycemia', *Pediatrics,* vol. 122, no. 1, pp. 65-74.

Capes, SE, Hunt, D, Malmberg, K & Gerstein, HC 2000, 'Stress hyperglycaemia and increased risk of death after myocardial infarction in patients with and without diabetes: a systematic overview', *Lancet,* vol. 355, no. 9206, pp. 773-8.

Carayon, P & Gurses, AP 2005, 'A human factors engineering conceptual framework of nursing workload and patient safety in intensive care units', *Intensive Crit Care Nurs,* vol. 21, no. 5, pp. 284-301.

Castle, JR & Ward, WK 2010, 'Amperometric glucose sensors: sources of error and potential benefit of redundancy', *J Diabetes Sci Technol,* vol. 4, no. 1, pp. 221-5.

Cembrowski, GS, Tran, DV & Higgins, TN 2010, 'The use of serial patient blood gas, electrolyte and glucose results to derive biologic variation: a new tool to assess the acceptability of intensive care unit testing', *Clin Chem Lab Med,* vol. 48, no. 10, pp. 1447-54.

Chase, JG, Andreassen, S, Jensen, K & Shaw, GM 2008a, 'Impact of Human Factors on Clinical Protocol Performance: A Proposed Assessment Framework and Case Examples', *Journal of Diabetes Science and Technology,* vol. 2, no. 3, pp. 409-16.

Chase, JG, Hann, CE, Jackson, M, Lin, J, Lotz, T, Wong, XW & Shaw, GM 2006, 'Integral-based filtering of continuous glucose sensor measurements for glycaemic control in critical care', *Comput Methods Programs Biomed,* vol. 82, no. 3, pp. 238-47.

Chase, JG, Le Compte, AJ, Suhaimi, F, Shaw, GM, Lynn, A, Lin, J, Pretty, CG, Razak, N, Parente, JD, Hann, CE, Preiser, JC & Desaive, T 2011, 'Tight glycemic control in critical care--the leading role of insulin sensitivity and patient variability: a review and model-based analysis', *Comput Methods Programs Biomed,* vol. 102, no. 2, pp. 156-71.

Chase, JG, LeCompte, A, Shaw, GM, Blakemore, A, Wong, J, Lin, J & Hann, CE 2008b, 'A benchmark data set for model-based glycemic control in critical care', *J Diabetes Sci Technol,* vol. 2, no. 4, pp. 584-94.

Chase, JG, Shaw, G, Le Compte, A, Lonergan, T, Willacy, M, Wong, X-W, Lin, J, Lotz, T, Lee, D & Hann, C 2008c, 'Implementation and evaluation of the SPRINT

protocol for tight glycaemic control in critically ill patients: a clinical practice change', *Critical Care,* vol. 12, no. 2, p. R49.

Chase, JG, Shaw, GM, Lotz, T, LeCompte, A, Wong, J, Lin, J, Lonergan, T, Willacy, M & Hann, CE 2007, 'Model-based insulin and nutrition administration for tight glycaemic control in critical care', *Curr Drug Deliv,* vol. 4, no. 4, pp. 283-96.

Chase, JG, Suhaimi, F, Penning, S, Preiser, JC, Le Compte, AJ, Lin, J, Pretty, CG, Shaw, GM, Moorhead, KT & Desaive, T 2010, 'Validation of a model-based virtual trials method for tight glycemic control in intensive care', *Biomed Eng Online,* vol. 9, p. 84.

Chee, F, Fernando, T & van Heerden, PV 2002, 'Closed-loop control of blood glucose levels in critically ill patients', *Anaesth Intensive Care,* vol. 30, no. 3, pp. 295-307.

Chee, F, Fernando, T & van Heerden, PV 2003, 'Closed-loop glucose control in critically ill patients using continuous glucose monitoring system (CGMS) in real time', *IEEE Trans Inf Technol Biomed,* vol. 7, no. 1, pp. 43-53.

Clarke, WL 2005, 'The original Clarke Error Grid Analysis (EGA)', *Diabetes Technol Ther,* vol. 7, no. 5, pp. 776-9.

Clarke, WL, Anderson, S, Breton, M, Patek, S, Kashmer, L & Kovatchev, B 2009, 'Closed-loop artificial pancreas using subcutaneous glucose sensing and insulin delivery and a model predictive control algorithm: the Virginia experience', *J Diabetes Sci Technol,* vol. 3, no. 5, pp. 1031-8.

Clarke, WL, Cox, D, Gonder-Frederick, LA, Carter, W & Pohl, SL 1987, 'Evaluating clinical accuracy of systems for self-monitoring of blood glucose', *Diabetes Care,* vol. 10, no. 5, pp. 622-8.

Clarke, WL & Kovatchev, B 2007, 'Continuous Glucose Sensors: Continuing Questions about Clinical Accuracy', *J Diabetes Sci Technol,* vol. 1, no. 5, pp. 669-75.

Cornblath, M, Hawdon, JM, Williams, AF, Aynsley-Green, A, Ward-Platt, MP, Schwartz, R & Kalhan, SC 2000, 'Controversies regarding definition of neonatal hypoglycemia: suggested operational thresholds', *Pediatrics,* vol. 105, no. 5, pp. 1141-5.

Corstjens, AM, Ligtenberg, JJ, van der Horst, IC, Spanjersberg, R, Lind, JS, Tulleken, JE, Meertens, JH & Zijlstra, JG 2006, 'Accuracy and feasibility of point-of-care and continuous blood glucose analysis in critically ill ICU patients', *Crit Care,* vol. 10, no. 5, p. R135.

Critchell, CD, Savarese, V, Callahan, A, Aboud, C, Jabbour, S & Marik, P 2007, 'Accuracy of bedside capillary blood glucose measurements in critically ill patients', *Intensive Care Med,* vol. 33, no. 12, pp. 2079-84.

D'Orazio, P, Burnett, RW, Fogh-Andersen, N, Jacobs, E, Kuwa, K, Kulpmann, WR, Larsson, L, Lewenstam, A, Maas, AH, Mager, G, Naskalski, JW & Okorodudu, AO 2006, 'Approved IFCC recommendation on reporting results for blood glucose: International Federation of Clinical Chemistry and Laboratory Medicine Scientific Division, Working Group on Selective Electrodes and Point-of-Care Testing (IFCC-SD-WG-SEPOCT)', *Clin Chem Lab Med,* vol. 44, no. 12, pp. 1486-90.

Das, UN 2003, 'Insulin in sepsis and septic shock', *J Assoc Physicians India,* vol. 51, pp. 695-700.

Dassau, E, Cameron, F, Lee, H, Bequette, BW, Zisser, H, Jovanovic, L, Chase, HP, Wilson, DM, Buckingham, BA & Doyle, FJ, 3rd 2010, 'Real-Time hypoglycemia prediction suite using continuous glucose monitoring: a safety net for the artificial pancreas', *Diabetes Care,* vol. 33, no. 6, pp. 1249-54.

De La Rosa Gdel, C, Donado, JH, Restrepo, AH, Quintero, AM, Gonzalez, LG, Saldarriaga, NE, Bedoya, M, Toro, JM, Velasquez, JB, Valencia, JC, Arango, CM, Aleman, PH, Vasquez, EM, Chavarriaga, JC, Yepes, A, Pulido, W & Cadavid, CA 2008, 'Strict glycaemic control in patients hospitalised in a mixed medical and surgical intensive care unit: a randomised clinical trial', *Crit Care,* vol. 12, no. 5, p. R120.

Djakoure-Platonoff, C, Radermercker, R, Reach, G, Slama, G & Selam, JI 2003, 'Accuracy of the continuous glucose monitoring system in inpatient and outpatient conditions', *Diabetes Metab,* vol. 29, no. 2 Pt 1, pp. 159-62.

Duvanel, CB, Fawer, CL, Cotting, J, Hohlfeld, P & Matthieu, JM 1999, 'Long-term effects of neonatal hypoglycemia on brain growth and psychomotor development in small-for-gestational-age preterm infants', *J Pediatr,* vol. 134, no. 4, pp. 492-8.

Egi, M, Bellomo, R, Stachowski, E, French, CJ & Hart, G 2006, 'Variability of blood glucose concentration and short-term mortality in critically ill patients', *Anesthesiology,* vol. 105, no. 2, pp. 244-52.

Egi, M, Bellomo, R, Stachowski, E, French, CJ, Hart, GK, Taori, G, Hegarty, C & Bailey, M 2010, 'Hypoglycemia and outcome in critically ill patients', *Mayo Clin Proc,* vol. 85, no. 3, pp. 217-24.

Eke, A, Herman, P, Kocsis, L & Kozak, LR 2002, 'Fractal characterization of complexity in temporal physiological signals', *Physiol Meas,* vol. 23, no. 1, pp. R1-38.

Ellahham, S 2010, 'Molecular mechanisms of hyperglycemia and cardiovascular-related events in critically ill patients: rationale for the clinical benefits of insulin therapy', *Clin Epidemiol,* vol. 2, pp. 281-8.

Eren-Oruklu, M, Cinar, A & Quinn, L 2010, 'Hypoglycemia prediction with subject-specific recursive time-series models', *J Diabetes Sci Technol,* vol. 4, no. 1, pp. 25-33.

Evans, A, Le Compte, A, Tan, CS, Ward, L, Steel, J, Pretty, CG, Penning, S, Suhaimi, F, Shaw, GM, Desaive, T & Chase, JG 2012, 'Stochastic targeted (STAR) glycemic control: design, safety, and performance', *J Diabetes Sci Technol,* vol. 6, no. 1, pp. 102-15.

Finfer, S, Chittock, DR, Su, SY, Blair, D, Foster, D, Dhingra, V, Bellomo, R, Cook, D, Dodek, P, Henderson, WR, Hebert, PC, Heritier, S, Heyland, DK, McArthur, C, McDonald, E, Mitchell, I, Myburgh, JA, Norton, R, Potter, J, Robinson, BG & Ronco, JJ 2009, 'Intensive versus conventional glucose control in critically ill patients', *N Engl J Med,* vol. 360, no. 13, pp. 1283-97.

Finfer, S & Delaney, A 2008, 'Tight glycemic control in critically ill adults', *JAMA,* vol. 300, no. 8, pp. 963-5.

Fisk, LM, Le Compte, AJ, Shaw, GM, Penning, S, Desaive, T & Chase, JG 2012, 'STAR development and protocol comparison', *IEEE Trans Biomed Eng,* vol. 59, no. 12, pp. 3357-64.

Gandhi, GY, Kovalaske, M, Kudva, Y, Walsh, K, Elamin, MB, Beers, M, Coyle, C, Goalen, M, Murad, MS, Erwin, PJ, Corpus, J, Montori, VM & Murad, MH 2011, 'Efficacy of continuous glucose monitoring in improving glycemic control and reducing hypoglycemia: a systematic review and meta-analysis of randomized trials', *J Diabetes Sci Technol,* vol. 5, no. 4, pp. 952-65.

Garg, SK, Voelmle, M & Gottlieb, PA 2010, 'Time lag characterization of two continuous glucose monitoring systems', *Diabetes Res Clin Pract,* vol. 87, no. 3, pp. 348-53.

Gearhart, MM & Parbhoo, SK 2006, 'Hyperglycemia in the critically ill patient', *AACN Clin Issues,* vol. 17, no. 1, pp. 50-5.

Ginsberg, BH 2009, 'Factors affecting blood glucose monitoring: sources of errors in measurement', *J Diabetes Sci Technol,* vol. 3, no. 4, pp. 903-13.

Girardin, CM, Huot, C, Gonthier, M & Delvin, E 2009, 'Continuous glucose monitoring: a review of biochemical perspectives and clinical use in type 1 diabetes', *Clin Biochem,* vol. 42, no. 3, pp. 136-42.

Goldberg, PA, Siegel, MD, Russell, RR, Sherwin, RS, Halickman, JI, Cooper, DA, Dziura, JD & Inzucchi, SE 2004, 'Experience with the continuous glucose

monitoring system in a medical intensive care unit', *Diabetes Technol Ther,* vol. 6, no. 3, pp. 339-47.

Goldberger, AL, Amaral, LA, Hausdorff, JM, Ivanov, P, Peng, CK & Stanley, HE 2002, 'Fractal dynamics in physiology: alterations with disease and aging', *Proc Natl Acad Sci U S A,* vol. 99 Suppl 1, pp. 2466-72.

Gonzalez-Michaca, L, Ahumada, M & Ponce-de-Leon, S 2002, 'Insulin subcutaneous application vs. continuous infusion for postoperative blood glucose control in patients with non-insulin-dependent diabetes mellitus', *Arch Med Res,* vol. 33, no. 1, pp. 48-52.

Greisen, G & Pryds, O 1989, 'Neonatal hypoglycaemia', *Lancet,* vol. 1, no. 8650, pp. 1332-3.

Griesdale, DE, de Souza, RJ, van Dam, RM, Heyland, DK, Cook, DJ, Malhotra, A, Dhaliwal, R, Henderson, WR, Chittock, DR, Finfer, S & Talmor, D 2009, 'Intensive insulin therapy and mortality among critically ill patients: a meta-analysis including NICE-SUGAR study data', *CMAJ,* vol. 180, no. 8, pp. 821-7.

Gross, TM, Bode, BW, Einhorn, D, Kayne, DM, Reed, JH, White, NH & Mastrototaro, JJ 2000, 'Performance evaluation of the MiniMed continuous glucose monitoring system during patient home use', *Diabetes Technol Ther,* vol. 2, no. 1, pp. 49-56.

Gschwendtner, J 2007, *Continuous Glycemic Control in Type 1 Diabetes* Technical University of Munich,

Halamek, LP, Benaron, DA & Stevenson, DK 1997, 'Neonatal hypoglycemia, Part I: Background and definition', *Clin Pediatr (Phila),* vol. 36, no. 12, pp. 675-80.

Harris, DL, Battin, MR, Weston, PJ & Harding, JE 2010, 'Continuous glucose monitoring in newborn babies at risk of hypoglycemia', *J Pediatr,* vol. 157, no. 2, pp. 198-202 e1.

Hausdorff, JM, Purdon, PL, Peng, CK, Ladin, Z, Wei, JY & Goldberger, AL 1996, 'Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations', *J Appl Physiol,* vol. 80, no. 5, pp. 1448-57.

Hawdon, JM, Ward Platt, MP & Aynsley-Green, A 1992, 'Patterns of metabolic adaptation for preterm and term infants in the first neonatal week', *Arch Dis Child,* vol. 67, no. 4 Spec No, pp. 357-65.

Hay, WW, Jr. 2006, 'Placental-fetal glucose exchange and fetal glucose metabolism', *Trans Am Clin Climatol Assoc,* vol. 117, pp. 321-39; discussion 39-40.

Helton, KL, Ratner, BD & Wisniewski, NA 2011, 'Biomechanics of the sensor-tissue interface-effects of motion, pressure, and design on sensor performance and foreign body response-part II: examples and application', *J Diabetes Sci Technol,* vol. 5, no. 3, pp. 647-56.

Hermanides, J, Bosman, RJ, Vriesendorp, TM, Dotsch, R, Rosendaal, FR, Zandstra, DF, Hoekstra, JB & DeVries, JH 2010a, 'Hypoglycemia is associated with intensive care unit mortality', *Crit Care Med,* vol. 38, no. 6, pp. 1430-4.

Hermanides, J, Vriesendorp, TM, Bosman, RJ, Zandstra, DF, Hoekstra, JB & Devries, JH 2010b, 'Glucose variability is associated with intensive care unit mortality', *Crit Care Med,* vol. 38, no. 3, pp. 838-42.

Ho, HT, Yeung, WK & Young, BW 2004, 'Evaluation of "point of care" devices in the measurement of low blood glucose in neonatal practice', *Arch Dis Child Fetal Neonatal Ed,* vol. 89, no. 4, pp. F356-9.

Hoedemaekers, CW, Klein Gunnewiek, JM, Prinsen, MA, Willems, JL & Van der Hoeven, JG 2008, 'Accuracy of bedside glucose measurement from three glucometers in critically ill patients', *Crit Care Med,* vol. 36, no. 11, pp. 3062-6.

Hoeks, LB, Greven, WL & de Valk, HW 2011a, 'Real-time continuous glucose monitoring system for treatment of diabetes: a systematic review', *Diabet Med,* vol. 28, no. 4, pp. 386-94.

Hoeks, LBEA, Greven, WL & de Valk, HW 2011b, 'Real-time continuous glucose monitoring system for treatment of diabetes: a systematic review', *Diabetic Medicine,* vol. 28, no. 4, pp. 386-94.

Holzinger, U, Kitzberger, R, Fuhrmann, V, Schenk, P, Kramer, L, Funk, G, Zauner, C & Madl, C 2005, 'ICU-staff education and implementation of an insulin therapy algorithm improve blood glucose control', *18th ESICM Annual Congress*, Amsterdam, Netherlands

Holzinger, U, Warszawska, J, Kitzberger, R, Herkner, H, Metnitz, PG & Madl, C 2009, 'Impact of shock requiring norepinephrine on the accuracy and reliability of subcutaneous continuous glucose monitoring', *Intensive Care Med,* vol. 35, no. 8, pp. 1383-9.

Holzinger, U, Warszawska, J, Kitzberger, R, Wewalka, M, Miehsler, W, Herkner, H & Madl, C 2010, 'Real-time continuous glucose monitoring in critically ill patients: a prospective randomized trial', *Diabetes Care,* vol. 33, no. 3, pp. 467-72.

Iglesias Platas, I, Thio Lluch, M, Pociello Alminana, N, Morillo Palomo, A, Iriondo Sanz, M & Krauel Vidal, X 2009, 'Continuous glucose monitoring in infants of very low birth weight', *Neonatology,* vol. 95, no. 3, pp. 217-23.

Ihlen, EA 2012, 'Introduction to multifractal detrended fluctuation analysis in matlab', *Front Physiol,* vol. 3, p. 141.

Jopling, J, Henry, E, Wiedmeier, SE & Christensen, RD 2009, 'Reference ranges for hematocrit and blood hemoglobin concentration during the neonatal period: data from a multihospital health care system', *Pediatrics,* vol. 123, no. 2, pp. e333-7.

Kamath, A, Mahalingam, A & Brauker, J 2009, 'Analysis of time lags and other sources of error of the DexCom SEVEN continuous glucose monitor', *Diabetes Technol Ther,* vol. 11, no. 11, pp. 689-95.

Kanji, S, Buffie, J, Hutton, B, Bunting, PS, Singh, A, McDonald, K, Fergusson, D, McIntyre, LA & Hebert, PC 2005, 'Reliability of point-of-care testing for glucose measurement in critically ill adults', *Crit Care Med,* vol. 33, no. 12, pp. 2778-85.

Kantelhardt, JW, Zschiegner, SA, Koscielny-Bunde, E, Havlin, S, Bunde, A & Stanley, HE 2002, 'Multifractal detrended fluctuation analysis of nonstationary time series', *Physica a-Statistical Mechanics and Its Applications,* vol. 316, no. 1-4, pp. 87-114.

Karon, BS, Gandhi, GY, Nuttall, GA, Bryant, SC, Schaff, HV, McMahon, MM & Santrach, PJ 2007, 'Accuracy of roche accu-chek inform whole blood capillary, arterial, and venous glucose values in patients receiving intensive intravenous insulin therapy after cardiac surgery', *Am J Clin Pathol,* vol. 127, no. 6, pp. 919-26.

Keenan, DB, Mastrototaro, JJ, Voskanyan, G & Steil, GM 2009, 'Delays in minimally invasive continuous glucose monitoring devices: a review of current technology', *J Diabetes Sci Technol,* vol. 3, no. 5, pp. 1207-14.

Keenan, DB, Mastrototaro, JJ, Zisser, H, Cooper, KA, Raghavendhar, G, Lee, SW, Yusi, J, Bailey, TS, Brazg, RL & Shah, RV 2012, 'Accuracy of the Enlite 6-day glucose sensor with guardian and Veo calibration algorithms', *Diabetes Technol Ther,* vol. 14, no. 3, pp. 225-31.

King, C, Anderson, SM, Breton, M, Clarke, WL & Kovatchev, BP 2007, 'Modeling of Calibration Effectiveness and Blood-to-Interstitial Glucose Dynamics as Potential Confounders of the Accuracy of Continuous Glucose Sensors during Hyperinsulinemic Clamp', *J Diabetes Sci Technol,* vol. 1, no. 3, pp. 317-22.

Klonoff, DC 2000, 'The Importance of Continuous Glucose Monitoring in Diabetes', *Diabetes Technology & Therapeutics,* vol. 2, no. Sup. 1, pp. S1-S3.

Klonoff, DC 2005a, 'Continuous glucose monitoring: roadmap for 21st century diabetes therapy', *Diabetes Care,* vol. 28, no. 5, pp. 1231-9.

Klonoff, DC 2005b, 'A review of continuous glucose monitoring technology', *Diabetes Technol Ther,* vol. 7, no. 5, pp. 770-5.

Koh, TH, Eyre, JA & Aynsley-Green, A 1988, 'Neonatal hypoglycaemia--the controversy regarding definition', *Arch Dis Child,* vol. 63, no. 11, pp. 1386-8.

Kosiborod, M, Gottlieb, R, Sekella, J, Peterman, D, Grodzinsky, A, Kennedy, P & Borkon, M 2013, 'Performance of the Medtronic Sentrino(R) continuous glucose management system in the cardiac ICU', *Critical Care,* vol. 17, no. Suppl 2, p. P462.

Kovatchev, B, Anderson, S, Heinemann, L & Clarke, W 2008, 'Comparison of the numerical and clinical accuracy of four continuous glucose monitors', *Diabetes Care,* vol. 31, no. 6, pp. 1160-4.

Kovatchev, BP, Gonder-Frederick, LA, Cox, DJ & Clarke, WL 2004, 'Evaluating the accuracy of continuous glucose-monitoring sensors: continuous glucose-error grid analysis illustrated by TheraSense Freestyle Navigator data', *Diabetes Care,* vol. 27, no. 8, pp. 1922-8.

Kovatchev, BP, Shields, D & Breton, M 2009, 'Graphical and numerical evaluation of continuous glucose sensing time lag', *Diabetes Technol Ther,* vol. 11, no. 3, pp. 139-43.

Krinsley, JS 2003, 'Association between hyperglycemia and increased hospital mortality in a heterogeneous population of critically ill patients', *Mayo Clin Proc,* vol. 78, no. 12, pp. 1471-8.

Krinsley, JS 2004, 'Effect of an intensive glucose management protocol on the mortality of critically ill adult patients', *Mayo Clin Proc,* vol. 79, no. 8, pp. 992-1000.

Krinsley, JS 2008, 'Glycemic variability: a strong independent predictor of mortality in critically ill patients', *Crit Care Med,* vol. 36, no. 11, pp. 3008-13.

Krinsley, JS & Jones, RL 2006, 'Cost analysis of intensive glycemic control in critically ill adult patients', *Chest,* vol. 129, no. 3, pp. 644-50.

Langouche, L, Vanhorebeek, I & Van den Berghe, G 2005, 'The role of insulin therapy in critically ill patients', *Treat Endocrinol,* vol. 4, no. 6, pp. 353-60.

Larson, NS & Pinsker, JE 2013, 'The role of continuous glucose monitoring in the care of children with type 1 diabetes', *Int J Pediatr Endocrinol,* vol. 2013, no. 1, p. 8.

Leal, Y, Gonzalez-Abril, L, Lorencio, C, Bondia, J & Vehi, J 2013, 'Detection of correct and incorrect measurements in real-time continuous glucose

monitoring systems by applying a post-processing support vector machine', *Biomedical Engineering, IEEE Transactions on,* vol. PP, no. 99, pp. 1-.

Lee, JM, Kim, DJ, Kim, IY, Suk Park, K & Kim, SI 2004, 'Nonlinear-analysis of human sleep EEG using detrended fluctuation analysis', *Med Eng Phys,* vol. 26, no. 9, pp. 773-6.

Ligtenberg, JJ, de Plaa, ME & Zijlstra, JG 2011, 'Continuous subcutaneous glucose monitoring: good enough to use in glucose regulation protocols?', *Crit Care,* vol. 15, no. 1, p. 403; author reply

Lin, J 2007, *Robust modelling and control of the glucose-insulin regulatory system for tight glycemic control of critical care patients* University of Canterbury, PhD Thesis.

Lin, J, Lee, D, Chase, JG, Shaw, GM, Le Compte, A, Lotz, T, Wong, J, Lonergan, T & Hann, CE 2008, 'Stochastic modelling of insulin sensitivity and adaptive glycemic control for critical care', *Comput Methods Programs Biomed,* vol. 89, no. 2, pp. 141-52.

Lonergan, T, Le Compte, A, Willacy, M, Chase, JG, Shaw, GM, Wong, XW, Lotz, T, Lin, J & Hann, CE 2006, 'A simple insulin-nutrition protocol for tight glycemic control in critical illness: development and protocol comparison', *Diabetes Technol Ther,* vol. 8, no. 2, pp. 191-206.

Lorencio, C, Leal, Y, Bonet, A, Bondia, J, Palerm, CC, Tache, A, Sirvent, JM & Vehi, J 2012, 'Real-time continuous glucose monitoring in an intensive care unit: better accuracy in patients with septic shock', *Diabetes Technol Ther,* vol. 14, no. 7, pp. 568-75.

Lucas, A, Morley, R & Cole, TJ 1988, 'Adverse neurodevelopmental outcome of moderate neonatal hypoglycaemia', *BMJ,* vol. 297, no. 6659, pp. 1304-8.

Lundelin, K, Vigil, L, Bua, S, Gomez-Mestre, I, Honrubia, T & Varela, M 2010, 'Differences in complexity of glycemic profile in survivors and nonsurvivors in an intensive care unit: a pilot study', *Crit Care Med,* vol. 38, no. 3, pp. 849-54.

Mahoney, JJ & Ellison, JM 2007, 'Assessing glucose monitor performance--a standardized approach', *Diabetes Technol Ther,* vol. 9, no. 6, pp. 545-52.

Makaya, T, Memmott, A & Bustani, P 2012, 'Point-of-care glucose monitoring on the neonatal unit', *J Paediatr Child Health,* vol. 48, no. 4, pp. 342-6.

Marieb, EN & Hoehn, K 2005, *Human anatomy and physiology*, Pearson Benjamin Cummings.

Marik, PE & Raghavan, M 2004, 'Stress-hyperglycemia, insulin and immunomodulation in sepsis', *Intensive Care Medicine,* vol. 30, no. 5, pp. 748-56.

Mastrototaro, J, Gross, T & Shin, J 2002, *Glucose monitor calibration methods,* US 6,424,847 B1.

Mastrototaro, JJ 2000, 'The MiniMed continuous glucose monitoring system', *Diabetes Technol Ther,* vol. 2 Suppl 1, pp. S13-8.

McCowen, KC, Malhotra, A & Bistrian, BR 2001, 'Stress-induced hyperglycemia', *Crit Care Clin,* vol. 17, no. 1, pp. 107-24.

McGarraugh, G 2009, 'The chemistry of commercial continuous glucose monitors', *Diabetes Technol Ther,* vol. 11 Suppl 1, pp. S17-24.

McGarraugh, G 2010, 'Alarm characterization for a continuous glucose monitor that replaces traditional blood glucose monitoring', *J Diabetes Sci Technol,* vol. 4, no. 1, pp. 49-56.

Meijering, S, Corstjens, AM, Tulleken, JE, Meertens, JH, Zijlstra, JG & Ligtenberg, JJ 2006, 'Towards a feasible algorithm for tight glycaemic control in critically ill patients: a systematic review of the literature', *Crit Care,* vol. 10, no. 1, p. R19.

Minimed, M 2003, 'CGMS System Gold User Guide'.

Minimed, M 2006, 'Guardian Real-Time continuous glucose monitoring system user guide'.

Minimed, M 2010, 'iPro2 user guide'.

Mizock, BA 1995, 'Alterations in carbohydrate metabolism during stress: a review of the literature', *Am J Med,* vol. 98, no. 1, pp. 75-84.

Moser, E, Crew, L & Garg, S 2010, 'Role of continuous glucose monitoring in diabetes management', *Avances en Diabetología,* vol. 26, no. 2, pp. 73-8.

Nova 2011, 'Nova Stat Strip Glucose test strip packet insert'. Nova Biomedical Ltd, USA.

Nuntnarumit, P, Chittamma, A, Pongmee, P, Tangnoo, A & Goonthon, S 2011, 'Clinical performance of the new glucometer in the nursery and neonatal intensive care unit', *Pediatr Int,* vol. 53, no. 2, pp. 218-23.

Nylen, ES & Muller, B 2004, 'Endocrine changes in critical illness', *J Intensive Care Med,* vol. 19, no. 2, pp. 67-82.

Peng, CK, Havlin, S, Hausdorff, JM, Mietus, JE, Stanley, HE & Goldberger, AL 1995, 'Fractal mechanisms and heart rate dynamics. Long-range correlations and their breakdown with disease', *J Electrocardiol,* vol. 28 Suppl, pp. 59-65.

Peng, CK, Mietus, JE, Liu, Y, Lee, C, Hausdorff, JM, Stanley, HE, Goldberger, AL & Lipsitz, LA 2002, 'Quantifying fractal dynamics of human respiration: age and gender effects', *Ann Biomed Eng,* vol. 30, no. 5, pp. 683-92.

Penzel, T, Kantelhardt, JW, Grote, L, Peter, JH & Bunde, A 2003, 'Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea', *IEEE Trans Biomed Eng,* vol. 50, no. 10, pp. 1143-51.

Plank, J, Blaha, J, Cordingley, J, Wilinska, ME, Chassin, LJ, Morgan, C, Squire, S, Haluzik, M, Kremen, J, Svacina, S, Toller, W, Plasnik, A, Ellmerer, M, Hovorka, R & Pieber, TR 2006, 'Multicentric, randomized, controlled trial to evaluate blood glucose control by the model predictive control algorithm versus routine glucose management protocols in intensive care unit patients', *Diabetes Care,* vol. 29, no. 2, pp. 271-6.

Preiser, JC, Devos, P, Ruiz-Santana, S, Melot, C, Annane, D, Groeneveld, J, Iapichino, G, Leverve, X, Nitenberg, G, Singer, P, Wernerman, J, Joannidis, M, Stecher, A & Chiolero, R 2009, 'A prospective randomised multi-centre controlled trial on tight glucose control by intensive insulin therapy in adult intensive care units: the Glucontrol study', *Intensive Care Med,* vol. 35, no. 10, pp. 1738-48.

Pretty, C, Chase, JG, Lin, J, Shaw, GM, Le Compte, A, Razak, N & Parente, JD 2011, 'Impact of glucocorticoids on insulin resistance in the critically ill', *Comput Methods Programs Biomed,* vol. 102, no. 2, pp. 172-80.

Pretty, CG, Chase, JG, Le Compte, A, Shaw, GM & Signal, M 2010, 'Hypoglycemia detection in critical care using continuous glucose monitors: an in silico proof of concept analysis', *J Diabetes Sci Technol,* vol. 4, no. 1, pp. 15-24.

Quan, S, Qin, SJ & Doniger, KJ 2010, 'Online dropout detection in subcutaneously implanted continuous glucose monitoring', *American Control Conference (ACC), 2010*, pp. 4373-8

Rabiee, A, Andreasik, V, Abu-Hamdah, R, Galiatsatos, P, Khouri, Z, Gibson, BR, Andersen, DK & Elahi, D 2009, 'Numerical and clinical accuracy of a continuous glucose monitoring system during intravenous insulin therapy in the surgical and burn intensive care units', *J Diabetes Sci Technol,* vol. 3, no. 4, pp. 951-9.

Rao, LV, Jakubiak, F, Sidwell, JS, Winkelman, JW & Snyder, ML 2005, 'Accuracy evaluation of a new glucometer with automated hematocrit measurement and correction', *Clin Chim Acta,* vol. 356, no. 1-2, pp. 178-83.

Roberts, JR, Park, J, Helton, K, Wisniewski, N & McShane, MJ 2012, 'Biofouling of polymer hydrogel materials and its effect on diffusion and enzyme-based luminescent glucose sensor functional characteristics', *J Diabetes Sci Technol,* vol. 6, no. 6, pp. 1267-75.

Roche 2007, 'ACCU-CHEK Comfort Curve Test-strip packet insert'. Roche Diagnostics Ltd., USA.

Roche 2008, *Evaluation Report of the ACCU-CHEK Comfort Curve Test Strip as a Plasma-like Test Strip*, Roche Diagnostics, Mannheim, Germany, 340-40436-0408.

Rossetti, P, Bondia, J, Vehi, J & Fanelli, CG 2010, 'Estimating plasma glucose from interstitial glucose: the issue of calibration algorithms in commercial continuous glucose monitoring devices', *Sensors (Basel),* vol. 10, no. 12, pp. 10936-52.

Scott, RJ, Deobald, G, Griesmann, l, Wockenfus, A & Karon, B 2008, 'Evaluation of Multiple Whole Blood Glucose Methods Compared With a Laboratory Plasma Hexokinase Reference Assay', *Point of Care,* vol. 7, no. 2, pp. 43-6.

Signal, M, Le Compte, A, Harris, DL, Weston, PJ, Harding, JE & Chase, JG 2012a, 'Using stochastic modelling to identify unusual continuous glucose monitor measurements and behaviour, in newborn infants', *Biomed Eng Online,* vol. 11, p. 45.

Signal, M, Le Compte, A, Shaw, GM & Chase, JG 2012b, 'Glycemic levels in critically ill patients: are normoglycemia and low variability associated with improved outcomes?', *J Diabetes Sci Technol,* vol. 6, no. 5, pp. 1030-7.

Signal, M, Pretty, CG, Chase, JG, Le Compte, A & Shaw, GM 2010, 'Continuous glucose monitors and the burden of tight glycemic control in critical care: can they cure the time cost?', *J Diabetes Sci Technol,* vol. 4, no. 3, pp. 625-35.

Singer, P, Berger, MM, Van den Berghe, G, Biolo, G, Calder, P, Forbes, A, Griffiths, R, Kreyman, G, Leverve, X & Pichard, C 2009, 'ESPEN Guidelines on Parenteral Nutrition: Intensive care', *Clinical Nutrition,* vol. 28, no. 4, pp. 387-400.

Solnica, B, Naskalski, JW & Sieradzki, J 2003, 'Analytical performance of glucometers used for routine glucose self-monitoring of diabetic patients', *Clin Chim Acta,* vol. 331, no. 1-2, pp. 29-35.

Stanley, CA & Baker, L 1999, 'The causes of neonatal hypoglycemia', *N Engl J Med,* vol. 340, no. 15, pp. 1200-1.

Suhaimi, F, Le Compte, A, Preiser, JC, Shaw, GM, Massion, P, Radermecker, R, Pretty, CG, Lin, J, Desaive, T & Chase, JG 2010, 'What makes tight glycemic control tight? The impact of variability and nutrition in two clinical studies', *J Diabetes Sci Technol,* vol. 4, no. 2, pp. 284-98.

Tarin, C, Traver, L, Bondia, J & Vehi, J 2010, 'A learning system for error detection in subcutaneous continuous glucose measurement using Support Vector Machines', *Control Applications (CCA), 2010 IEEE International Conference on*, pp. 1614-9

Treggiari, MM, Karir, V, Yanez, ND, Weiss, NS, Daniel, S & Deem, SA 2008, 'Intensive insulin therapy and mortality in critically ill patients', *Crit Care,* vol. 12, no. 1, p. R29.

Turina, M, Fry, DE & Polk, HC, Jr. 2005, 'Acute hyperglycemia and the innate immune system: clinical, cellular, and molecular aspects', *Crit Care Med,* vol. 33, no. 7, pp. 1624-33.

Umpierrez, GE, Hellman, R, Korytkowski, MT, Kosiborod, M, Maynard, GA, Montori, VM, Seley, JJ & Van den Berghe, G 2012, 'Management of hyperglycemia in hospitalized patients in non-critical care setting: an endocrine society clinical practice guideline', *J Clin Endocrinol Metab,* vol. 97, no. 1, pp. 16-38.

Valgimigli, F, Lucarelli, F, Scuffi, C, Morandi, S & Sposato, I 2010, 'Evaluating the clinical accuracy of GlucoMen(R)Day: a novel microdialysis-based continuous glucose monitor', *J Diabetes Sci Technol,* vol. 4, no. 5, pp. 1182-92.

van den Berghe, G, Wouters, P, Weekers, F, Verwaest, C, Bruyninckx, F, Schetz, M, Vlasselaers, D, Ferdinande, P, Lauwers, P & Bouillon, R 2001, 'Intensive insulin therapy in critically ill patients', *N Engl J Med,* vol. 345, no. 19, pp. 1359-67.

Van den Berghe, G, Wouters, PJ, Bouillon, R, Weekers, F, Verwaest, C, Schetz, M, Vlasselaers, D, Ferdinande, P & Lauwers, P 2003, 'Outcome benefit of intensive insulin therapy in the critically ill: Insulin dose versus glycemic control', *Crit Care Med,* vol. 31, no. 2, pp. 359-66.

Van den Berghe, G, Wouters, PJ, Kesteloot, K & Hilleman, DE 2006, 'Analysis of healthcare resource utilization with intensive insulin therapy in critically ill patients', *Crit Care Med,* vol. 34, no. 3, pp. 612-6.

Vashist, SK 2012, 'Non-invasive glucose monitoring technology in diabetes management: a review', *Anal Chim Acta,* vol. 750, pp. 16-27.

Vlkova, A, Dostal, P, Musil, F, Smahelova, A, Zadak, Z & Cerny, V 2009, 'Blood and tissue glucose level in critically ill patients: a comparison of different

methods of measuring interstitial glucose levels', *Intensive Care Med,* vol. 35, no. 7, p. 1318.

Vogelzang, M, van der Horst, IC & Nijsten, MW 2004, 'Hyperglycaemic index as a tool to assess glucose control: a retrospective study', *Crit Care,* vol. 8, no. 3, pp. R122-7.

Voskanyan, G, Barry Keenan, D, Mastrototaro, JJ & Steil, GM 2007, 'Putative delays in interstitial fluid (ISF) glucose kinetics can be attributed to the glucose sensing systems used to measure them rather than the delay in ISF glucose itself', *J Diabetes Sci Technol,* vol. 1, no. 5, pp. 639-44.

Watkinson, PJ, Barber, VS, Amira, E, James, T, Taylor, R & Young, JD 2012, 'The effects of precision, haematocrit, pH and oxygen tension on point-of-care glucose measurement in critically ill patients: a prospective study', *Ann Clin Biochem,* vol. 49, no. Pt 2, pp. 144-51.

Weinzimer, SA, Steil, GM, Swan, KL, Dziura, J, Kurtz, N & Tamborlane, WV 2008, 'Fully automated closed-loop insulin delivery versus semiautomated hybrid control in pediatric patients with type 1 diabetes using an artificial pancreas', *Diabetes Care,* vol. 31, no. 5, pp. 934-9.

Weiss, R & Lazar, I 2007, 'The need for continuous blood glucose monitoring in the intensive care unit', *J Diabetes Sci Technol,* vol. 1, no. 3, pp. 412-4.

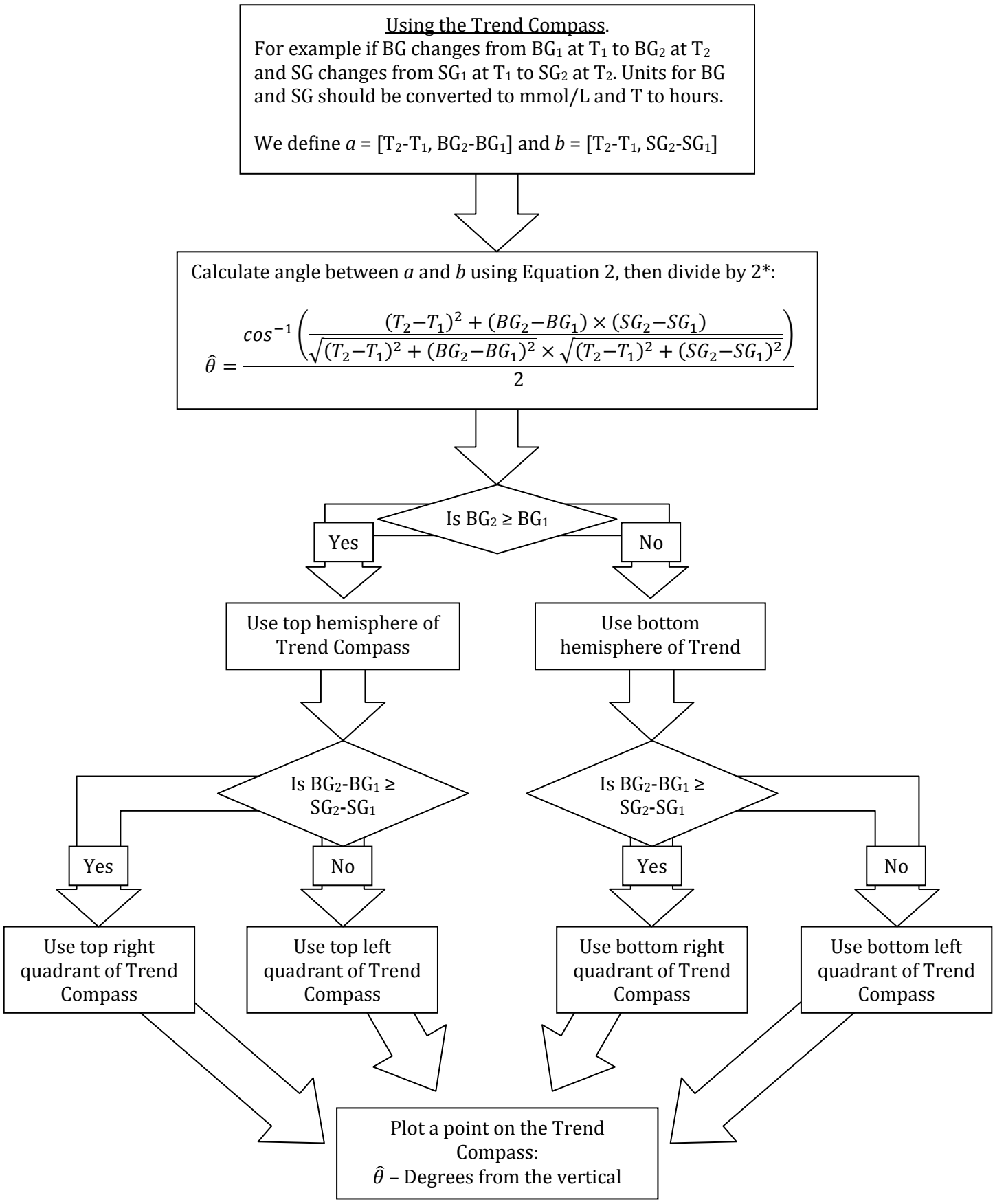
Wentholt, IM, Hoekstra, JB & Devries, JH 2006, 'A critical appraisal of the continuous glucose-error grid analysis', *Diabetes Care,* vol. 29, no. 8, pp. 1805-11.

Westhoff, D, Bosman, R, Oudemans-van Straaten, H, DeVries, J, Wester, J, van Stijn, I, Zandstra, D & van der Voort, P 2010, 'Validation and feasibility of two Continuous Glucose Monitoring Systems (CGMS) against point-of-care AccuChek® in critically ill patients; a pilot study', *Blood,* vol. 3, p. 17.6.

Wong, XW, Chase, JG, Shaw, GM, Hann, CE, Lotz, T, Lin, J, Singh-Levett, I, Hollingsworth, LJ, Wong, OS & Andreassen, S 2006, 'Model predictive glycaemic regulation in critical illness using insulin and nutrition input: a pilot study', *Med Eng Phys,* vol. 28, no. 7, pp. 665-81.

Zimmerman, CR, Mlynarek, ME, Jordan, JA, Rajda, CA & Horst, HM 2004, 'An insulin infusion protocol in critically ill cardiothoracic surgery patients', *Ann Pharmacother,* vol. 38, no. 7-8, pp. 1123-9.

Appendix A



*Note: The angle (θ) is divided by 2 to give $\hat{\theta}$ so the full range of theoretic angles can be displayed on each quadrant. The theoretical limit is BG rising vertically and SG falling vertically (or vice versa), which would result in 180° between the vectors (displayed as $\hat{\theta}$=90° on the Trend Compass).