2015

# Automated Processing Of Continuous Glucose Monitor (Cgm) Data To Study Onset Of Diabetes

Aman Khatri
*University of Alabama at Birmingham*

# AUTOMATED PROCESSING OF CONTINUOUS GLUCOSE MONITOR (CGM) DATA TO STUDY ONSET OF DIABETES

by

AMAN KHATRI

KARTHIKEYAN LINGASUBRAMANIAN, COMMITTEE CHAIR
ARIE NAKHAMANI
MOHAMMAD RAFIQUL HAIDER
PAULA C CHANDLER-LANEY

A THESIS

Submitted to the graduate faculty of The University of Alabama at Birmingham
In partial fulfillment of the requirements for the degree of
Master of Science

Birmingham, Alabama

2015

AUTOMATED PROCESSING OF CONTINUOUS GLUCOSE MONITOR (CGM) DATA TO STUDY ONSET OF DIABETES

AMAN KHATRI

ELECTRICAL ENGINEERING

## ABSTRACT

Diabetes is a long term condition that causes high levels of blood glucose, and it is necessary to get a complete picture of glucose levels which can lead to better treatment decision and better glucose control. The advent of Continuous Glucose Monitoring (CGM) is helping researches to track the blood glucose levels continuously and to understand the effects of impaired glucose levels on human body which in turn can lead to better treatment of diabetes. Current CGM systems process enormous amounts of data and have limitations in regards to data accuracy, precision, and reliability of raw glucose data. The inaccuracy in data also produces larger relative error in the estimates of glycemic variability than in the estimates of mean glucose and other related multiple measure of variability and multiple clinical end points. One of the objectives of the study is to automate the cleaning process of raw CGM data so as to replace the manual approach which very methodical but time consuming. Additionally, this study presents an

automated procedure that predicts meal consumption and provides the intake time and glycemic load. The proposed study is performed on data generated from three categories of participants, normal weight, over weight and obese. The success of the automated model to inspect and clean data is based on comparison of the datasets resulting from automated processing to those resulting from manual inspection and cleaning. These results obtained from automated protocol are found to be in agreement with the results obtained from manual inspection. Additionally, the average percent of correctly detected meals for normal weight participants is 82.777, the average percent of correctly detected meals for overweight participants is 85.933, and the average percent of correctly detected meals for obese participants is 80.589. The overall success rate for determining the meal times is 83.099. We have also shown some success in determining more specific information about the nutrition values of meal like glycemic load. The ability of the proposed model to predict meal intake is essential to understand the onset of diabetes. This proposed study results in a flexible platform which can facilitate important clinical studies on diabetes and possibly on other biological issues related to blood glucose.

Keywords: Data Processing, Continuous Glucose Monitoring System, Diabetes, Meal Detection

## ACKNOWLEDGEMENTS

This thesis was possible because of a strong collaboration between Department of Electrical and Computer Engineering and Department of Nutrition Science.

I would like to express gratitude to my academic advisor and committee chair, Dr. Karthikeyan Lingasubramanian, for helping me with each step of my academic and professional growth. He continually conveyed a spirit of adventure in regards to research. Without his guidance and persistent help this dissertation would not have been possible.

I would also like to thank Dr. Paula C. Chandler-Laney for providing me with the opportunity to work on diabetes research and for useful discussions throughout my program of study. She helped me with getting started and was always available to answer my questions. I would also like to thank her for offering me her expertise on diabetes.

My sincere thanks to Dr. Arie Nkhamani and Dr. Mohammad R. Haider for serving in my committee.

TABLE OF CONTENTS

*Page*

LIST OF TABLES

LIST OF FIGURES

# CHAPTER 1:

# INTRODUCTION

## 1.1 BACKGROUND

Glucose is a major source of energy and therefore is an essential component of blood. In the fasting state, such as in morning, to maintain blood glucose levels, body releases glucose which is stored in liver in form of glycogen. This breakdown and release of glucose into the bloodstream is stimulated by the glucagon hormone. However, when body is not fasting, that is while consuming food, the enzymes in the mouth and the small intestines breakdown the nutrients of the consumed food. These nutrients are then absorbed across the small intestine and transferred into blood stream. Carbohydrates are one such type of nutrients which gets absorbed by the small intestine and gets converted into glucose. The change in blood glucose due to carbohydrates is determined by the rate of digestion and absorption into the bloodstream and the ability of body to release insulin to clear the glucose from the circulation. In a non-diabetic person the rise in glucose causes the body to release insulin to remove glucose in order to maintain normal levels of blood glucose; whereas, in a diabetic person, the body does not produce adequate

amount of insulin, produces no insulin, or has cells that do not respond properly to the insulin produced by pancreas. This metabolic impairment is referred to as diabetes. Diabetes is usually classified in three types:

1) Type 1 Diabetes: In this type of Diabetes, the body does not produce insulin.
2) Type 2 Diabetes: In this type of Diabetes, the body does not produce enough insulin for proper function, or makes the body cells unresponsive of insulin.
3) Gestational Diabetes: This type of Diabetes effects pregnant women, and impairs their ability to produce adequate amount of insulin to transport all of the glucose into different cells.

Type 2 Diabetes (T2D) is a public health crisis. Approximately 8.3% of adults in the United States are estimated to have diagnosed or undiagnosed T2D, and another 35% are pre-diabetic **[1]**. Alarmingly, if these trends continue, by the year 2050 an estimated 25-33% of the US population will have T2D **[2]**. Although improvements in diet and lifestyle have been shown to delay the onset of, or possibly even prevent, T2D **[3]**, research is still needed to identify those at risk for T2D prior to any deterioration in metabolic health. Consequently, it would be advantageous to identify patterns of free-living glycaemia among diabetic and non-diabetic individuals that will predict response to standardized clinical measures of glucose tolerance, with the long-term goal of developing models for the prediction of time to β-cell failure and thereby T2D onset, given specific diet and lifestyle conditions.

The overall assessment of glycemic pattern has typically been performed using hemoglobin A1c (HbA1c) in combination with fasting blood glucose and self-monitored capillary blood glucose profiles **[4]**. HbA1c is "gold standard" measure of glycemic exposure as it provides a biologically integrated indication of average glucose control during the 6-8 weeks prior to sampling **[5]**. However, this process does not provide a more specific understanding of lifestyle events that contribute to glucose excursions and introduces the problem of irregular and infrequent glucose sampling. The advent of Continuous Glucose Monitoring (CGM) has eliminated the problem of irregular sampling and is helping researches to track the blood glucose levels continuously and to understand the factors that contribute to high glucose under free living conditions.

Continuous Glucose Monitor (CGM) is a tool to help people manage their blood glucose levels [Fig.1.1]. CGM consists of a multilayered electro-enzymatic sensor which transmits the glucose levels of the interstitial fluid to a receiver via radio waves. The sensor is inserted into the fatty tissue below the skin where it can access the interstitial fluid which resides between the cells as shown in Fig.1.2. The sensor then transmits the glucose reading to the receiver in every one to five minutes. CGM is capable to detect sudden change in the glucose levels occurring in individuals with diabetes. Increased availability of CGM system enables the researchers to investigate glycemic patterns throughout a given period of time. Developed specifically for the clinical purposes of detecting hyper and hypo glycaemia excursions among individuals with Type 1 Diabetes, CGM can also be used in conjunction with an insulin pump as a type of artificial pancreas.

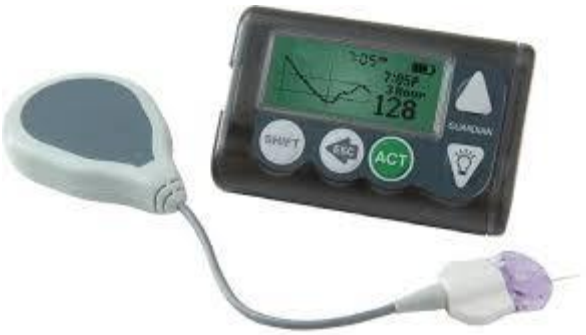


Fig.1.1 Continuous Glucose Monitor

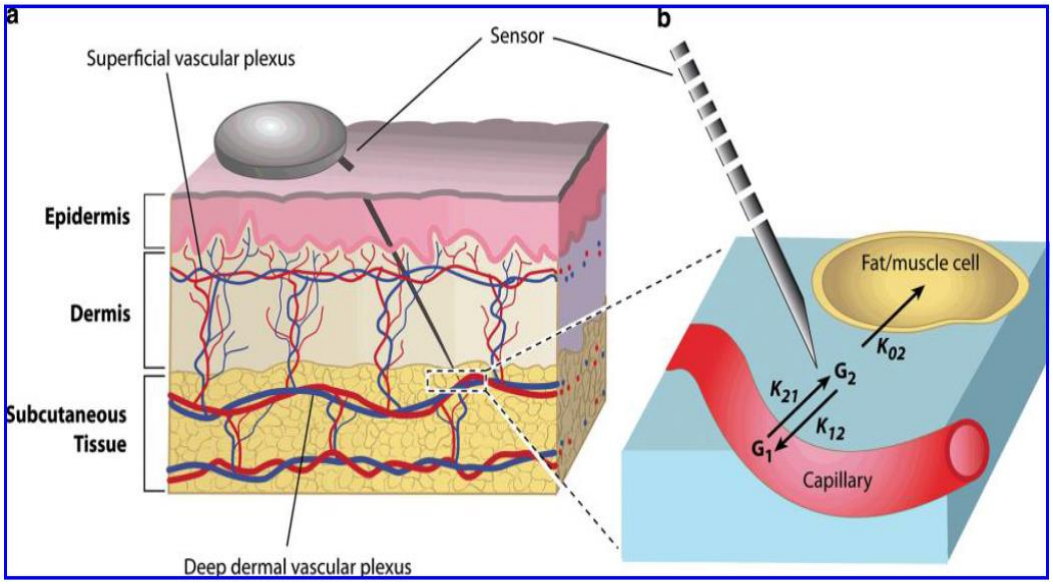


Fig.1.2 **[6]** a) Skin Layers with probe inserted into subcutaneous tissue b) Diffusion of glucose from blood to interstitial fluid in subcutaneous tissue.

## 1.2 MOTIVATION

The CGM monitoring has benefited the care of individuals with diabetes for over a decade. In clinical research, CGM technology can be used to understand dynamic and unpredictable phenomena like glycemic fluctuations in pregnant women and to characterize the role of free living glucose on fetal growth for analysis of early childhood onset of obesity. However, the lack of an automated

standardized approach to perform the labor intensive evaluation of glucose measures associated with maternal outcomes limits the use of CGM data for pregnancy related studies. Also current CGM system has limitations in regards to data accuracy, and precision of raw glucose data. The inaccuracy in enormous data produced by CGM also introduces larger relative error in the estimates of glycemic variability than in the estimates of mean glucose and other related multiple measure of variability and multiple clinical end points **[7]**. This lack of accuracy and absence of replicable standardized automated approach to detect incorrect readings from the raw data limits the development of useful autonomous systems like artificial pancreas. Additionally, the current algorithms and software which accompany CGM do not reveal all aspects of free-living glucose profile. In our study we propose a platform to characterize all aspects of free living glucose, including indices of glycemic control, variability, and the rate and acceleration of glucose.

## 1.3 CONTRIBUTIONS

- This is the first and only automated methodical approach adapted from the manual method described in Hernandez & Barbour **[8]** for reliably inspecting and cleaning inaccurate raw CGM data.
- Our platform offers an automated method to define and analyze time-series glucose measurements to predict meal times and glycemic loads of meals based on the technique of pattern recognition.

- The developed platform can also identify and extract important glucose characteristics like glycemic variability, the rate and acceleration of glucose changes, along with pre meal glucose, post meal glucose peak, and the area under the curve associated with meal intake.

The proposed platform can facilitate diverse applications of CGM to support multiple research studies conducted by clinical investigators to diagnose glucose impairments as well as gives researchers the ability to understand the onset of diabetes in non-diabetic individuals. Our platform offers a standardized automated method to define and analyze time-series glucose measurements and to predict meal times and glycemic loads of meals to increase the understanding of glycemic profiles contributing to onset of diabetes. Using the versatility of the platform we can produce multiple diagnostic studies like estimating the characteristics of hyper and hypo glycemic excursions occurring due to lifestyle related factors.

## 1.4 ORGANIZATION

This work is divided into 5 chapters. We discuss relevant issues and research work in Chapter 2. This chapter revisits the previous approaches to process CGM data using mathematical models and signal processing techniques. Chapter 3 is divided into seven parts which gives us the background information about the participants, the process of data collection, and the methods adapted for processing of glucose data in order to achieve reliable outcomes. Chapter 3 also deals with the explanation of automated protocol to clean glucose data, and the

techniques used to detect meal times for a given dataset. Chapter 4 discusses the results associate with cleaning of glucose data and detection of meal times. In chapter 4 we also present the methodology and results related to prediction of the Glycemic Load based on characteristics of glucose profile like Area Under Curve (AUC) during meal consumption, the duration of time to reach the peak value based on the duration of the excursions produced due to meal, and the peak value of the glucose after the meal consumption. In chapter 5 we conclude the study.

# CHAPTER 2

# PRIOR WORK

Diabetes accounts for significant morbidity, mortality and can result in diminishing quality of life in individuals [10], [11]. Being a serious condition, diabetes over time may cause significant organ damage to heart, blood vessels, kidneys, eyes, and nerves. According to National diabetes statistics report, diabetes is one of the leading causes of heart disease and stroke by damage of blood vessels, kidney failures, limb amputations by diabetic foot syndrome and early onset of blindness **[12]**. To address the problems associated with diabetes, researchers throughout the world and the United States are working to understand and treat this disease **[13]**. The researchers and the medical organizations involved in diabetes research have established aggressive targets for controlling glucose levels in diabetic individuals **[14], [15]**. The traditional tools like finger-stick to measure blood glucose can only be used three to four times a day because of their invasive nature. These tools fail to provide continuous information regarding the hyper and hypo glycemic excursions of the blood glucose and thereby missing all the important details **[16]**. However, now with the advent of CGM, we have an opportunity to understand the full story about the hyper and

hypo glycemic excursions. CGM enables us to access updated glucose values every few minutes and can be used for understanding glucose trends in an individual. According to Klonoff **[17], [18]**, CGM is likely to become a routine part of diabetes management. To derive full potential benefit from CGM it is important to understand the information about slope, magnitude and duration of fluctuations occurring in interstitial glucose.

Although CGM sensors are useful in detection of increasing or decreasing trends of the blood glucose, there are ongoing concerns about the accuracy and the sources of error in the CGM data **[19]**. Despite the fact that sensor accuracy improves with each new system generation, CGM-generated glucose values still vary from simultaneous fingersticks by an average of 10-20% **[20]**. These accuracy issues are due to lag time caused in measuring the glucose from interstitial fluid, or due to lack of accurate and timely calibration by patients or because of wrong choice of insertion site with inadequate subcutaneous fat. The clinical researchers involved in these studies come across enormous amounts of inaccurate data and have tried to address the issue by conducting several controlled experiments. These challenges related to cleaning of CGM data have been handled before using various mathematical models. Kalman filter and Moving Average filer based smoothing technique was used to handle inter-individual and intra individual signal to noise ratio in CGM data **[21]**. The validity of the resultant data set as an output of Kalman filter was based on Monte Carlo simulations and the performance of the Kalman filter was compared to Moving Average filter in terms of glucose measurement delay. A similar filtering

technique that can limit higher rate of change in CGM data was proposed to aid smoothing and calibration **[22]**. In another study, an integral based fitting and filtering method was used to reduce the effect of large errors in CGM sensors **[23]**. Several studies with use of autoregressive moving average to handle consecutive CGM sensor errors in order to improve CGM accuracy have also been proposed **[24]**. Powerful statistical techniques like Support Vector Machines (SVMs) have also been used to detect therapeutically incorrect measurements made by CGMs **[24], [25]**. However the studies so far do not offer a methodical approach to identify and extract relevant data from the cleaned CGM data for clinicians that can yield several other characteristics like Area Under Curve, duration, slope, pre and post meal excursion values. These time series glucose characteristics derived from CGM data can facilitate comparison among different glucose studies based on CGM to increase our understanding of individual glycemic profiles.

In this work we are presenting a flexible platform which can be used to clean the raw CGM data and derive the above described characteristics. Additionally, the platform can be used to evaluate the glycemic response of food by combining meal diaries and glucose data from CGM. This is the only platform which offers a capability to combine the glycemic response from CGM to the meal diaries, and an ability to detect glycemic load. According to Freeman and Lynee evaluating the glycemic response to a meal is difficult not only because of the composition of the meal, but also because of the dependence of glucose values on meal intake and the insulin resistance of an individual **[26], [27]**. However,

our platform provides researchers with an ability to understand the excursion in the glucose values which are dependent on the characteristics like glycemic variability, the rate and acceleration of glucose changes, pre meal glucose, post meal glucose peak, and the area under associated with meal intake.

# CHAPTER 3

# METHODS

## 3.1 PARTICIPANTS

This study involved secondary analysis of CGM data obtained as part of a study on healthy African American women above the age of 16 years during their third trimester of pregnancy. The recruits had a varying risk for developing impaired glucose tolerance and were stratified by Body Mass Index (BMI) during pregnancy. The normal weight women had BMI less than 25.0 $Kg/m^2$, the over-weight women had a BMI in the range of 25.0 to 29.9 $Kg/m^2$ and the obese women had a BMI more than 30.0 $Kg/m^2$. The women were recruited from among those planning to deliver at a large urban University Hospital and had initiated prenatal care prior to 19 weeks' gestation, and were experiencing a healthy singleton pregnancy. Whereas, women with pre-existing Diabetes Type 1 and Diabetes Type 2, who had developed gestational diabetes in current pregnancy or had previously delivered prematurely, that is the birth of a baby of less than 37 weeks, were excluded. The women with growth restricted infants and/or with infants weighing less than 2.5 Kgs were also excluded from the study.

The approval from Institutional Review Board (IRB) and the University of Alabama at Birmingham (UAB) was obtained for this study and the women recruited for the study were provided informed consent before the data collection.

## 3.2 DATA COLLECTION PROCEDURE

In the study, the raw CGM data files, along with the concordant food diaries were made available by Dr. Paula C Chandler-Laney to automate the cleaning and processing of CGM data based on the adapted protocol from Hernandez & Barbour[8]. The women were enrolled in the study at 32.0 to 34.6 weeks of their gestation. Maternal weight for each individual was obtained from medical records, and height (cm) and weight (Kg) were measured during the placement of the CGM. These participants wore a CGM device manufactured by Medtronic for four consecutive days, starting from day 0 to day 3, and were provided with a glucometer for calibrating the CGM within every 12 hour period. An accelerometer to record the physical activity was also provided as a part of this study. These participants were also instructed to record their food intake with the portion size of their meal during the free-living period for day 1 and day 2 where each meal was defined as the duration between the start and end of a distinct intake. During the study, the participants were neither restricted to a specific diet nor were they asked to perform any physical activity. In fact, participants were encouraged to carry on their usual every day activities during the period of study with an exception of an overnight fast prior to returning to the clinic on day 3 on

which an Oral Glucose Tolerance Test (OGTT) was conducted. The data obtained from food diary was entered into web tool called automated self-administered 24-hour recalls (ASA24) **[28]** and dietary outcomes of interest over 24 hours and per meal, intake were extracted, along with the kilo-calories from proteins, carbohydrates and fat. We received 33 datasets of raw CGM data, as shown in Fig. 3.2.1, along with the concordant food diaries, as shown in Fig 3.2.2, to automate the cleaning of CGM data and to predict the time for meal consumption along with the detection of glycemic load from glucose variables extracted for day 2. Each raw CGM file contains several columns however for our study we will use the Sample#, Day, Date, Sensor Glucose and the ratio of ISIG (interstitial glucose) and Meter BG, which is the glucose value from finger stick used for calibrating CGM.

| Sample # | Day | Date | Time | ISIG | VCTR | Meter BG | Paired Meter BG | Slope | Offset | Valid ISIG | Sensor Glucose | Glucose Units |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 91 | Tue | 3-Jan-12 | 5:42 PM | 23.47 | -0.58 | | | 3.03 | -6 | 23.51 | 90 | mg/dL |
| 92 | Tue | 3-Jan-12 | 5:47 PM | 23.56 | -0.58 | | | 3.03 | -6 | 23.56 | 90 | mg/dL |
| 93 | Tue | 3-Jan-12 | 5:52 PM | 23.51 | -0.58 | | | 3.03 | -6 | 23.91 | 91 | mg/dL |
| 94 | Tue | 3-Jan-12 | 5:57 PM | 23.56 | -0.58 | | | 3.03 | -6 | 24.87 | 94 | mg/dL |
| 95 | Tue | 3-Jan-12 | 6:02 PM | 23.91 | -0.58 | | | 3.03 | -6 | 25.79 | 96 | mg/dL |
| 96 | Tue | 3-Jan-12 | 6:07 PM | 24.87 | -0.58 | | | 3.03 | -6 | 26.81 | 100 | mg/dL |
| 97 | Tue | 3-Jan-12 | 6:12 PM | 25.79 | -0.58 | | | 3.03 | -6 | 27.66 | 102 | mg/dL |
| 98 | Tue | 3-Jan-12 | 6:17 PM | 26.81 | -0.58 | | | 3.03 | -6 | 28.1 | 103 | mg/dL |
| 99 | Tue | 3-Jan-12 | 6:22 PM | 27.66 | -0.58 | | | 3.03 | -6 | 28.19 | 104 | mg/dL |
| 100 | Tue | 3-Jan-12 | 6:27 PM | 28.1 | -0.58 | | | 3.03 | -6 | 27.81 | 103 | mg/dL |
| 101 | Tue | 3-Jan-12 | 6:32 PM | 28.19 | -0.58 | | | 3.03 | -6 | 26.49 | 99 | mg/dL |
| 102 | Tue | 3-Jan-12 | 6:37 PM | 27.81 | -0.58 | | | 3.03 | -6 | 24.95 | 94 | mg/dL |
| 103 | Tue | 3-Jan-12 | 6:42 PM | 26.49 | -0.58 | | | 3.03 | -6 | 23.76 | 90 | mg/dL |

Fig. 3.2.1 Raw CGM Data File

| Date of Intake | Meal File ID | Meal Time | Energy (kcal) | Total Carbohydrate (g) | Total Sugars (g) | Added Sugars (by Available Carbohydrate) (g) | Glycemic Index (bread reference) | Glycemic Load (bread reference) |
|---|---|---|---|---|---|---|---|---|
| 5/20/2013 | 1 | 12:20 | 1115.599 | 50.009 | 111.841 | 52.752 | 14.866 | 12.632 |
| 5/20/2013 | 2 | 18:59 | 733.821 | 19.876 | 98.942 | 37.452 | 57.483 | 68.663 |
| 5/20/2013 | 3 | 19:25 | 16.329 | 0.044 | 4.341 | 0.092 | 3.436 | 0 |
| 5/21/2013 | 1 | 9:33 | 108.17 | 1.044 | 25.114 | 1.479 | 12.325 | 12.105 |
| 5/21/2013 | 2 | 13:33 | 366.454 | 19.755 | 31.086 | 16.031 | 3.48 | 0.438 |
| 5/21/2013 | 3 | 15:50 | 1082.903 | 43.966 | 134.227 | 43.506 | 34.579 | 20.829 |
| 5/21/2013 | 4 | 17:32 | 250.24 | 0 | 63.733 | 0.782 | 63.733 | 63.733 |
| 5/21/2013 | 5 | 22:29 | 434.286 | 27.365 | 19.253 | 26.088 | 0.068 | 0 |
| 5/22/2013 | 1 | 10:40 | 857.572 | 34.664 | 107.258 | 30.002 | 12.145 | 9.991 |
| 5/22/2013 | 2 | 17:27 | 928.072 | 27.044 | 145.139 | 26.924 | 86.505 | 88.081 |
| 5/20/2013 | 1 | 12:20 | 1115.599 | 50.009 | 111.841 | 52.752 | 14.866 | 12.632 |

Fig. 3.2.2 Meal Diary Data File

## 3.3 ADAPTED PROTOCOL FOR CLEANING RAW CGM DATA

According to manufacturer's instruction, an indwelling glucose sensor, iPro and a Medtronic CGM was placed on the abdomen of each qualified participant on the side reported to be less likely to be slept on. Participants of the study were also provided with an OneTouch glucometer for measuring their blood glucose, three times a day, to calibrate the Continuous Glucose Monitoring System. Since the number one reason for inaccuracy in CGM is the time of calibration, the participants were instructed to calibrate the sensor in the morning, while fasting, and before a meal intake, when the blood glucose is very stable.

After day 3, Dr. Chandler-Laney's team retrieved the CGM from the participants and data were uploaded from CGM recorder using Solutions Software

version 2.2 which was provided by the manufacturer. The meal diaries from the participants were also retrieved by Dr. Chandler-Laney's team. Data from the CGM was exported in the tabulated form, in excel sheets and was manually cleaned using a modified version of a published protocol by Hernandez & Barbour[8]. Data from days during which less than two calibrations were conducted was discarded. Also as a part of initial inspection, any glucose series of 40s and 50s with at least one 40 was discarded along the reading for which the ratio of metered blood glucose, from glucometer, and Interstitial Fluid Glucose (ISIG) reading was outside the range of 0.5 to 15. After initial inspection, the mean and the standard deviation of two preceding and two following glucose values for each five minute glucose value was calculated in Microsoft excel. If the glucose value in consideration was higher than two standard deviation from the mean of the surrounding glucose value, it was marked as an outlier and was replaced by the average of the two prior glucose values and two following glucose values of the glucose value in consideration with an exception of last two glucose values. As a part of the study to verify the correct manual inspection using the same protocol, a random selection of 20% of the files was given to a second investigator. If these files inspected by second investigator were found to be in agreement with the files inspected by first investigator they were considered to be as valid entries. This process of manual inspection and cleaning of the raw CGM data consumed about two hours for each file, which motivated the development of automated protocol.

## 3.4 AUTMOATED PROTOCOL FOR CLEANING RAW CGM DATA

This section covers the standardized automated protocol, Fig. 3.4.1, for reliably inspecting and cleaning raw CGM data. The raw CGM data files used in this section were made available to us along with the modified version of adapted protocol. The provided protocol was automated using MATLAB® which is a high level language and an interactive environment that enabled us to clean the data at a very fast pace. With the help of MATLAB®, not only were we able to develop a model to automate the cleaning and processing of raw CGM data, which was the goal for initial phase of study, but also we were able to efficiently detect meal consumptions and other parameters which characterize free living glucose.

In the automation of the protocol, the excel file from the CGM was imported into MATLAB® workspace. The data was imported using inbuilt method to read excel sheets which returned the unprocessed data, numbers and text from the raw CGM file. The dates and times of the glucose measurements were also imported from excel files in MATLAB® format which was later converted to the readable dates and times. Data imported in workspace was then inspected to find the day during which at least two calibrations were conducted. This day was marked as the first day for starting the analysis and the first glucose value occurring on this day was marked as the first reading for counting the number of glucose values occurring on each day. Furthermore, the validity of the data for each day was checked by counting the number of calibration on each day following the first day and the ratio of metered blood glucose (BG), from glucometer, and Interstitial Fluid Glucose (ISIG). If a day had less than two

calibrations and/or if the ratio of BG and ISIG was outside the range of 0.5-1.5 then the day was discarded.

The next part of the program for automation dealt with calculating the mean and the standard deviation of two preceding and two following glucose values for each five minute glucose value in raw CGM file. This was achieved by implementing equation 3.4.1 and equation 3.4.2 in the MATLAB® script. In the following equations $X^r$ denotes the mean and i denote the glucose value about which the mean is to be calculated, n is a constant with the value of 4, and $Y^r$ denotes the standard deviation to analyze how widely values of the glucose are dispersed from the mean.

$$X_i^r = \frac{X_{i-2}^r + X_{i-1}^r + X_{i+1}^r + X_{i+2}^r}{n} \quad (3.4.1)$$

$$Y_i^r = \frac{(X_i^r - X_{i-2}^r)^2 + (X_i^r - X_{i-2}^r)^2 + (X_i^r - X_{i-2}^r)^2 + (X_i^r - X_{i-2}^r)^2}{n} \quad (3.4.2)$$

Furthermore, to find if the glucose value in consideration was higher than two standard deviations from the surrounding glucose values, two separate functions were coded to implement equations 3.4.3 and 3.4.4. The output parameters of these functions were then used to flag and replace the outliers, $O_{i+}$ and $O_{i-}$, with the average of two prior and two following glucose values given by $X^c$ in equation 3.4.5. The decision of implementing separate functions for equations 3.4.1, 3.4.2, 3.4.3 and 3.4.4 was made to give user the ability to readjust the

readings as many times as they would like in order to produce all the physiologically feasible values.

$$O_{i+} = X_i + 2Y_i \tag{3.4.3}$$

$$O_{i-} = X_i - 2Y_i \tag{3.4.4}$$

$$X_i^c = \frac{X_{i-2}^r + X_{i-1}^r + X_{i+1}^r + X_{i+2}^r}{n} = X_i^r \tag{3.4.5}$$

The resultant of implementation was then exported into a new excel file. This was achieved using MATLAB® to handle windows object in its object oriented environment.

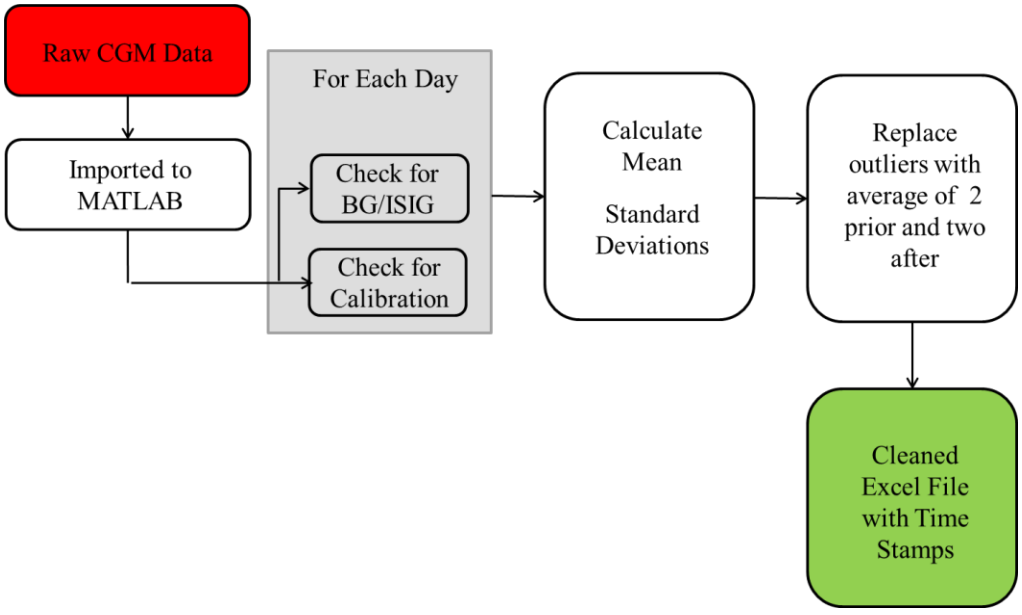


Fig 3.4.1 Automated Process for Cleaning and Processing Raw CGM Data

## 3.5 DETECTION OF MEAL INTAKE

This section covers the automated protocol **[8]** for efficiently processing the time series cleaned CGM data, to detect meal consumptions with the associated meal times which is important to depict hyper glycemic excursions in glucose. For the automated inspection of the glucose data for characterizing meal consumption, glucose values derived from CGM were used after they have been automatically cleaned and processed. As per the protocol, to detect meal consumption, the difference between each glucose concentration and the one preceding was calculated using MATLAB cleaned CGM data file. If the calculated slope of the glucose curve sustained an increase of at least 4mg/dl or greater over a period of time, it was considered to be suggestive of calorie intake. The times of suspected caloric intake were then flagged. The process of automation for meal detection is shown in Fig 3.5.1.

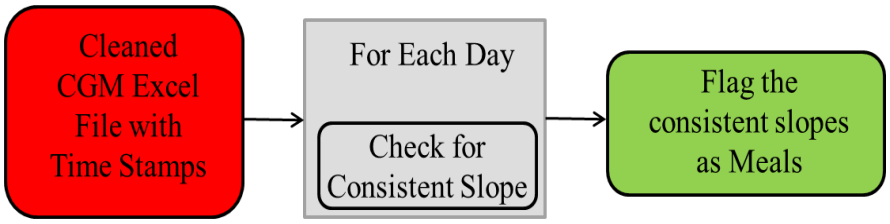


Fig 3.5.1 Automated Process for Meal Detection

The time of suspected caloric intake were then noted and compared with reported meal times from the meal diary. If the time of glucose increase reported by CGM was within an hour of either side of suspected caloric intake time, or if there was no other increase which was close in time to the participants reported

meal time, then the start time of the glucose increase was noted down as the meal initiation. Each meal for all reporting purposes was defined as the duration between the start and end of a distinct intake. However, if any snacks were reported on a given day during which fewer than three meals were reported in the diary, they were considered as a meal. In case where there were no meals reported in diary and the glucose values reported by CGM were consistently increasing with a slope of at least 4mg/dl, then those increases were noted down as missed meal from the participant's side. This process of manual inspection, detection of meal intake and the times associated with the meal intake required about 30 minutes for processing each file, and motivated the development of automated protocol which reduces the time for processing raw data and combining meal diaries to a few seconds.

## 3.6 GLYCEMIC LOAD DETECTION

To understand the hyper and the hypo glycemic excursions, it is important to relate the meal intake times with the nutrition facts of the meals. The most important nutrition value which determines the ability of meal to affect the glucose is glycemic load **[26]**. One unit of glycemic load approximates the effect of consuming one gram of glucose. Since the glucose response of every individual is unique, it is important to individualize the process of detecting the glycemic load using pattern recognition techniques. The proposed technique recognizes glycemic patterns in terms of glucose variables like Area Under Curve (AUC) during meal consumption, the duration of Time to reach the peak value, the

average of pre meal glucose, and the peak value of the excursion while evaluating the glycemic of a meal. These variables are then used to predict glycemic load of meals. A visual perspective of these glucose variables is given in Fig 3.6.1 where AUC is obtained by multiplying the duration, and height of the green line which represents a meal intake.

As the next step in processing the data and detecting glycemic load, the meal diaries, in excel format, were imported into MATLAB® workspace. The data for day 1 from the meal diary was imported using inbuilt method to read excel sheets which returned the unprocessed data, numbers and text from the meal diaries. The dates and times of the glucose measurements were also imported from excel files in MATLAB® format which were later converted to the readable dates and times. The meal diary imports were combined with the cleaned CGM glucose times and dates with the corresponding meal times and nutrition facts for day 1, this was done within a precision of 5 minutes of reported meals. The slopes of the flagged meal for day 1 were then associated with the imported nutrition facts of day 1 as shown in Fig. 3.6.2. This was done only if the time of glucose increase reported by CGM was within an hour of either side of reported meal intake time. These combined values for day 1 were used to derive glucose variables like Area Under Curve (AUC) during meal consumption, the duration of Time to reach the peak value based on the duration of the excursions produced due to meal, the average of pre meal glucose commonly referred to as Pre-prandial Blood Glucose, and the peak value of the excursion due to meal intake. The pattern for these glucose variables for day 1 were used to predict glycemic

load of the predicted meals on day 2. This was achieved by matching the glucose variable patterns occurring on day 1 to predict glycemic load for day 2. The process to obtain the above listed derived parameters was implemented in separate functions of the program. Each of these parameters was individually used to predict the glycemic load for the flagged meal on day 2 based on day 1. These parameters along with the predicted meals were at last exported to an excel sheet using the object oriented MATLAB code. At the completion of this aim, a resultant model was produced to process the cleaning of data, detection of meal times with the glycemic loads associated with them, and the parameters which define free-living glycemic profile.

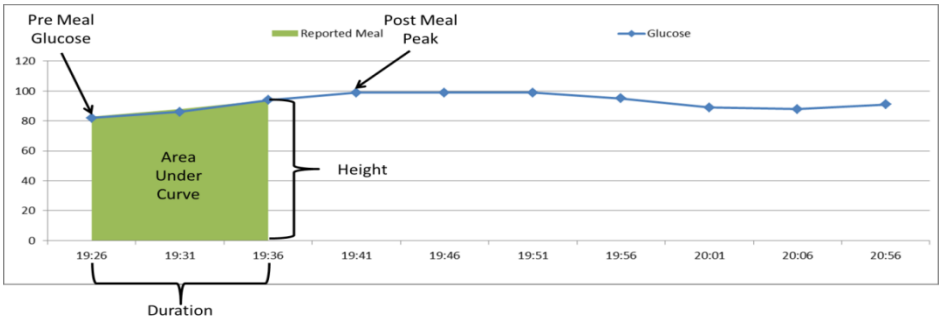


Fig. 3.6.1 Visual Perspective of Glucose Variables

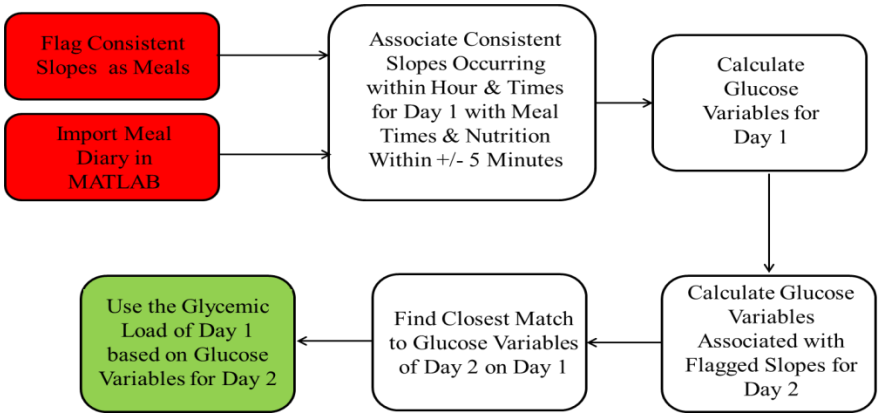


Fig 3.6.2 Automated Process for Glycemic Load Detection

# CHAPTER 4

# RESULTS

This study involves analysis of raw CGM data collected as part of one cohort studies in healthy, predominantly African American women with varying risk for the development of impaired glucose tolerance. Participants in the cohort wore a CGM device (Medtronic; Northridge, CA) to examine free-living glucose profile, with concurrent assessment of diet and underwent a standardized oral glucose or liquid mixed meal test. In the proposed study, first goal was addressed by automated identification of outliers and invalid or improbable data through data mining concepts and adaptable algorithms, described in section 3.4. To validate the results, the free-living glucose datasets processed for goal were compared to those manually derived by two independent coders, using intra-class correlation analyses. The second goal for the study was addressed by developing the algorithm to detect meal times and glycemic loads of meals based on the technique of pattern recognition. As a resultant product a comprehensive platform was developed to process the cleaning of data, detection of meal times with the glycemic loads associated with them, and the parameters which define free-living glycemic profile.

In this chapter we discuss the results of the automated protocols for cleaning raw CGM data in section 4.1, section 4.2 covers the results for prediction of meal times, and section 4.3 covers the results associated with the prediction of glycemic loads for day 2. Each section of the chapter is subdivided into three subsections for the Normal Weight (NW) women with BMI less than 25.0 $Kg/m^2$, the Over-Weight (OW) women with BMI in the range of 25.0 to 29.9 $Kg/m^2$ and the Obese (OB) women with a BMI more than 30.0 $Kg/m^2$.

## 4.1 RESULTS ASSOCIATED WITH CLEANING RAW CGM DATA

In this section of the chapter we describe the results obtained from automated standardized protocol developed for cleaning raw CGM data. For the study, we were provided with 33 raw files from the CGM with a glucose data of 4 days for each participant. The systems collected glucose data continuously, every 5 minutes, for a period of 72 hours. Of the given 33 datasets, three results for each category of NW, OW, and OB are described here. The blue line on the graphs in figures 4.1.1, 4.1.2, and 4.1.3 depict the glucose values from the raw CGM data file, the green line on the graphs depict the glucose values which were cleaned manually, and the red line on the graphs depict the glucose values which were cleaned by automated cleaning process.

In the figures for following subsections of NW, OW, and OB women, x axis shows the times in hours, y axis, marked as glucose values, shows the range of glucose values in mg/dl for an individual. In the figures, the overlapping blue,

green and red lines shows that the raw data, the manually cleaned data and the automatically cleaned data are closely in agreement and therefore no manual or automatic cleaning was performed. Whereas, the non over lapping blue and/or green/red lines demonstrate the use of cleaning protocol to smoothen the curve for that discreet value of the glucose. In the figures some instances occur where the manually cleaned data, shown in green, is not in complete agreement with automatically cleaned data, show in red and the difference between such values is shown by the y axis titled as "Difference" on the figures. The orange color diamond ♦ on the graph shows the difference between manually and automatically cleaned data, and the blue x shows the difference between raw and manually cleaned data.

### 4.1.1 RESULTS ASSOCIATED WITH NW WOMEN

In this section results associated with NW women, patient 103, with BMI less than 25.0 $Kg/m^2$ are described. The figures 4.1.1 shows times in hours on x-axis and the glucose associated at the given times on y-axis. The data on the graph also shows the excursions in the estimated circulating glucose in interstitial fluid for every five minutes. Since the orange color diamonds ♦ are aligned on x axis, the difference between manual and automated cleaning is zero and the manual cleaning and automated cleaning are in complete agreement.

### 4.1.2 RESULTS ASSOCIATED WITH OW WOMEN

In this section results associated with OW women, patient 201, with BMI in the range of 25.0 to 29.9 Kg/m$^2$ are described. The figures 4.1.2 shows times in hours on x-axis and the glucose associated at the given times on y-axis. The data on the graph also shows the magnified excursion in the estimated circulating glucose in interstitial fluid for every five minutes. Since not all the orange color diamonds ♦ are aligned on x axis, the difference between manual and automated cleaning is not zero and hence the automated cleaning was successful in cleaning the data points which were missed during manual process. Additionally, the over lapping orange color diamonds ♦ and blue x shows the minute disagreement between the manually cleaned and automatically cleaned data. This scenario arises because in manually cleaned data the outliers were replaced by visual inspection and hence they may not have been the exact value of the average of two prior and two following glucose values surrounding the outlier.

### 4.1.3 RESULTS ASSOCIATED WITH OB WOMEN

In this section results associated with OB women, patient 316, with a BMI more than 30.0 Kg/m$^2$ are described. The figures 4.1.3a shows times in hours on x-axis and the glucose associated at the given times on y-axis. The data on the graph also shows the magnified excursion in the estimated circulating glucose in interstitial fluid for every five minutes. Since not all the orange color diamonds ♦ are aligned on x axis, the difference between manual and automated cleaning is

not zero and hence the automated cleaning was successful in cleaning the data points which were missed during manual process. Additionally, the over lapping orange color diamonds ♦ and blue x shows the disagreement between the manually cleaned and automatically cleaned data. This scenario arises because in manually cleaned data the outliers were replaced by visual inspection and hence they may not have been the exact value of the average of two prior and two following glucose values surrounding the outlier.


Fig. 4.1.1 Glucose Value Versus Time for NW Patient 103

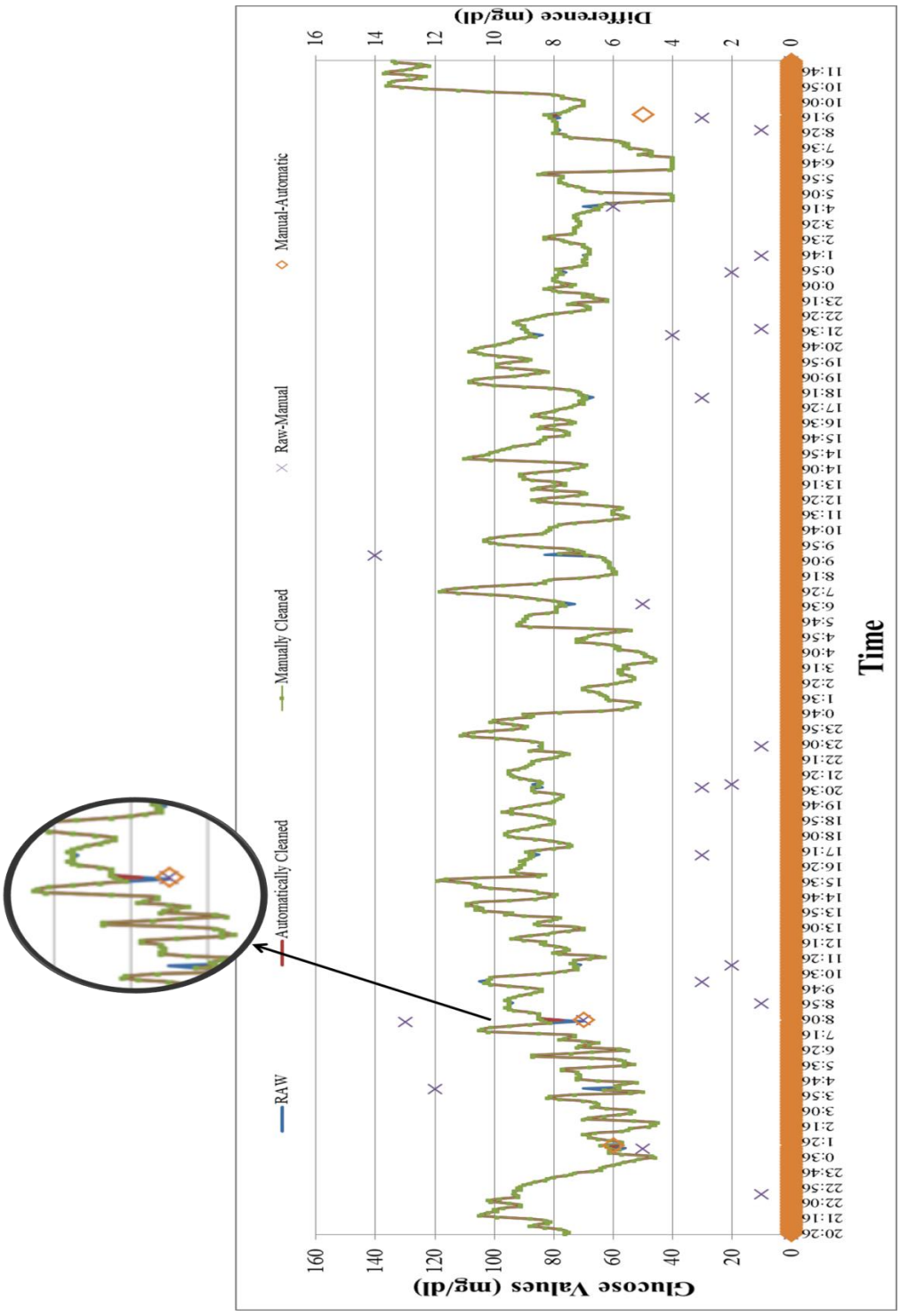


Fig. 4.1.2 Glucose Value Versus Time for OW Patient 201

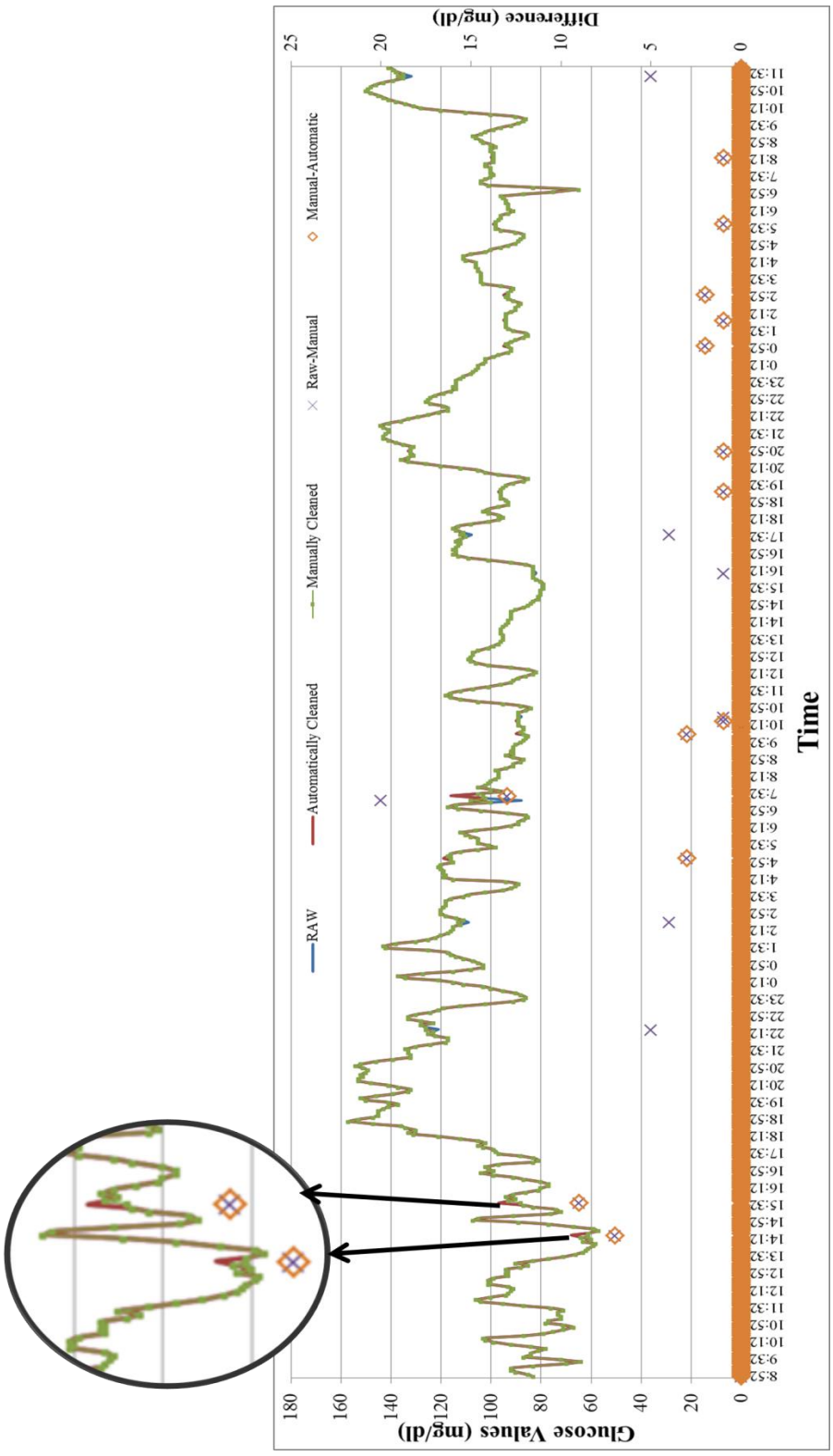


Fig. 4.1.3 Glucose Value Versus Time for OB Patient 316

## 4.2 RESULTS ASSOCIATED WITH DECTECTING MEAL INTAKE

In this section of the chapter we describe the results obtained from automated standardized protocol developed for detecting the meal times from processed CGM data files. For the study, we were provided with one Excel file with the meal diaries concordant to the 33 raw CGM files. The provided meal diaries had data entry for day 1 and day 2 with the dates, times, and nutrition facts corresponding to portion size of the meals consumed. Of the given 33 datasets, three results for each category of NW, OW, and OB with low, average and maximum accuracies are described here. Each table in this section is titled with the participant identification number, Patient ID, and has three rows which describe the Correctly Detected Meal Times from automated protocol, total reported meal times in meal diaries by participants, and the percent of Correctly Detected Meals. The percent of correctly detected meals is calculated by using equation 4.2.1

$$Percent\ of\ Correctly\ Detected\ Meals = \frac{Correctly\ Detected\ Meal\ Times\ from\ Automated\ Protocols}{Total\ Reported\ Meal\ Times\ in\ Meal\ Dairies} \times 100 \qquad (4.2.1)$$

Even though the glucose response of each participants is unique and reported times for the meal consumption in the meal diary were not always accurate, the overall success rate for determining the meal times was 83.099.

### 4.2.1 RESULTS ASSOCIATED WITH NW WOMEN

In this section results associated with NW women with BMI less than less than 25.0 $Kg/m^2$ are described. The tables 4.2.1, 4.2.2, and 4.2.3 shows the Correctly Detected Meal Times from automated protocol, total reported meal times in Meal Diaries by participants, and the percent of Correctly Detected Meals.

Table 4.2.1 Percentage of Correctly Detected Meals for NW Women

| **Patient ID: GU103** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 4 |
| **Total Reported Meals in Meal Diary** | 9 |
| **Percent of Correctly Detected Meals** | 44.4 |

Table 4.2.2 Percentage of Correctly Detected Meals for NW Women

| **Patient ID: GU114** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 4 |
| **Total Reported Meals in Meal Diary** | 5 |
| **Percent of Correctly Detected Meals** | 80 |

Table 4.2.3 Percentage of Correctly Detected Meals for NW Women

| **Patient ID: GU112** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 9 |
| **Total Reported Meals in Meal Diary** | 9 |
| **Percent of Correctly Detected Meals** | 100 |

### 4.2.2 RESULTS ASSOCIATED WITH OW WOMEN

In this section results associated with OW women with BMI in the range of 25.0 to 29.9 $Kg/m^2$ are described. The tables 4.2.4, 4.2.5, and 4.2.6 shows Correctly Detected Meal Times from Automated Protocol, Total Reported meal times in Meal Diaries by participants, and the percent of Correctly Detected Meals.

Table 4.2.4 Percentage of Correctly Detected Meals for OW Women

| **Patient ID: GU203** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 4 |
| **Total Reported Meals in Meal Diary** | 5 |
| **Percent of Correctly Detected Meals** | 80 |

Table 4.2.5 Percentage of Correctly Detected Meals for OW Women

| **Patient ID: GU201** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 8 |
| **Total Reported Meals in Meal Diary** | 9 |
| **Percent of Correctly Detected Meals** | 88.9 |

Table 4.2.6 Percentage of Correctly Detected Meals for OW Women

| **Patient ID: GU207** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 3 |
| **Total Reported Meals in Meal Diary** | 3 |
| **Percent of Correctly Detected Meals** | 100 |

### 4.2.3 RESULTS ASSOCIATED WITH OB WOMEN

In this section results associated with OW women with a BMI more than 30.0 $Kg/m^2$ are described. The tables 4.2.7, 4.2.8, and 4.2.9 shows Correctly Detected Meal Times from Automated Protocol, Total Reported meal times in Meal Diaries by participants, and the percent of Correctly Detected Meals.

Table 4.2.7 Percentage of Correctly Detected Meals for OB Women

| **Patient ID: GU318** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 2 |
| **Total Reported Meals in Meal Diary** | 6 |
| **Percent of Correctly Detected Meals** | 33.33 |

Table 4.2.8 Percentage of Correctly Detected Meals for OB Women

| **Patient ID: GU314** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 3 |
| **Total Reported Meals in Meal Diary** | 4 |
| **Percent of Correctly Detected Meals** | 75 |

Table 4.2.9 Percentage of Correctly Detected Meals for OB Women

| **Patient ID: GU316** | |
|---|---|
| **Correctly Detected Meal from Automated Protocol** | 5 |
| **Total Reported Meals in Meal Diary** | 5 |
| **Percent of Correctly Detected Meals** | 100 |

## 4.3 RESULTS ASSOCIATED WITH GLYCEMIC LOAD PREDICTION

In this section of the chapter we describe the results obtained from automated protocol developed to predict glycemic load based on glucose parameters; namely Area Under Curve (AUC) during meal consumption, the duration of time to reach the peak value based on the duration of the excursions produced due to meal, the average of pre meal glucose commonly referred to as Pre-prandial glucose, and the peak value of the glucose after the meal consumption. Of the given 33 datasets, three results corresponding to the results of participants shown in section 4.2 for each category of NW, OW, and OB are described here. Tables with suffix "a" in the subsections of 4.3 for NW, OW, and OB shows the reported diary value of the glycemic load for day 2, and the predicted glycemic loads for day 2 based on glucose variable. These values were predicted based on the closest match of glucose variables occurring on day 1 in the meal diry of the given individual. Additional, each table in the subsections has two other tables with suffix "b" and "c" associated with it, which show the percent error in the predicted values of glycemic loads for each meal, and the average percent error associate with predicted glycemic loads based on each glucose variable for day 2 respectively. The results generated for detection of glycemic load have a very high percentage error. This is due to the limited knowledge available for building the training set for individuals for recognizing the patter of glycemic loads.

### 4.3.1 RESULTS ASSOCIATED WITH NW WOMEN

In this section results associated with NW women are described. Tables 4.3.1a, 4.3.2a, and 4.3.3a shows the predicted glycemic load, tables 4.3.1b, 4.3.2b, and 4.3.3b shows the errors associated with glycemic load for each meal and tables 4.3.1c, 4.3.2c, and 4.3.3c shows the errors associated with glycemic load in each parameters for entire day 2.

Table 4.3.1a Predicted Glycemic Load

| Patient ID: GU103 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 87.686 | 67.893 | 67.893 | 67.893 | 25.776 | 67.893 |
| 10.834 | 67.893 | 67.893 | 67.893 | 25.776 | 67.893 |
| 50.599 | 67.893 | 67.893 | 67.893 | 47.999 | 67.893 |
| 119.057 | 67.893 | 67.893 | 67.893 | 47.999 | 67.893 |

Table 4.3.1b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU103 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 22.573 | 22.573 | 22.573 | 70.604 | 22.573 | 32.179 |
| 526.666 | 526.666 | 526.666 | 137.918 | 526.666 | 448.916 |
| 34.179 | 34.179 | 34.179 | 5.138 | 34.179 | 26.316 |
| 42.974 | 42.974 | 42.974 | 59.684 | 42.974 | 46.316 |

Table 4.3.1c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU103 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 156.598 | 156.598 | 156.598 | 68.336 | 156.598 | 138.431 |

Table 4.3.2a Predicted Glycemic Load

| Patient ID: GU114 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 11.453 | 43.245 | 43.245 | 43.245 | 74.635 | 43.245 |
| 54.149 | 43.245 | 43.245 | 43.245 | 74.635 | 43.245 |
| 91.169 | 74.635 | 74.635 | 43.245 | 34.907 | 34.907 |
| 134.015 | 43.245 | 43.245 | 43.245 | 43.245 | 43.245 |

Table 4.3.2b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU114 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 277.587 | 277.587 | 277.587 | 551.663 | 277.587 | 332.402 |
| 20.137 | 20.137 | 20.137 | 37.833 | 20.137 | 8.543 |
| 18.136 | 18.136 | 52.566 | 61.712 | 61.712 | 42.452 |
| 67.731 | 67.731 | 67.731 | 67.731 | 67.731 | 67.731 |

Table 4.3.2c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU114 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 95.897 | 95.897 | 104.505 | 179.734 | 106.791 | 112.782 |

Table 4.3.3a Predicted Glycemic Load

| Patient ID: GU112 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 32.388 | 32.388 | 65.709 | 65.709 | 11.053 | 65.709 |
| 33.073 | 32.388 | 32.388 | 65.709 | 90.842 | 65.709 |
| 15.14 | 32.388 | 32.388 | 65.709 | 90.842 | 65.709 |
| 63.119 | 34.274 | 90.842 | 90.842 | 34.274 | 65.709 |
| 114.19 | 90.842 | 90.842 | 90.842 | 27.905 | 11.053 |
| 1.888 | 34.274 | 90.842 | 90.842 | 90.842 | 65.709 |
| 13.883 | 34.274 | 90.842 | 90.842 | 90.842 | 65.709 |
| 62.005 | 64.709 | 64.709 | 65.709 | 65.709 | 65.709 |
| 33.073 | 64.709 | 64.709 | 65.709 | 65.709 | 65.709 |

Table 4.3.3b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU112 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 0.000 | 102.881 | 102.881 | 65.873 | 102.881 | 48.554 |
| 2.071 | 2.071 | 98.679 | 174.671 | 98.679 | 73.577 |
| 113.923 | 113.923 | 334.009 | 500.013 | 334.009 | 279.175 |
| 45.699 | 43.922 | 43.922 | 45.699 | 4.103 | 0.110 |
| 20.447 | 20.447 | 20.447 | 75.563 | 90.321 | 45.445 |
| 1715.360 | 4711.547 | 4711.547 | 4711.547 | 3380.350 | 3846.070 |
| 146.877 | 554.340 | 554.340 | 554.340 | 373.305 | 436.640 |

Table 4.3.3c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU112 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 292.054 | 792.733 | 837.975 | 875.387 | 626.235 | 675.653 |

## 4.3.2 RESULTS ASSOCIATED WITH OW WOMEN

In this section results associated with OW women are described. Tables 4.3.4a, 4.3.5a, and 4.3.6a shows the predicted glycemic load, tables 4.3.4b, 4.3.5b,

and 4.3.6b shows the errors associated with glycemic load for each meal and tables 4.3.4c, 4.3.5c, and 4.3.6c shows the errors associated with glycemic load in each parameters for entire day 2.

Table 4.3.4a Predicted Glycemic Load

| Patient ID: GU203 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 54.386 | 59.152 | 59.152 | 59.152 | 59.152 | 59.152 |
| 20.091 | 185.557 | 185.557 | 185.557 | 59.152 | 59.152 |
| 32.801 | 59.152 | 59.152 | 59.152 | 185.557 | 59.152 |
| 17.714 | 185.557 | 185.557 | 59.152 | 185.557 | 59.152 |

Table 4.3.4b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU203 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 8.763 | 8.763 | 8.763 | 8.763 | 8.763 | 8.763 |
| 823.583 | 823.583 | 823.583 | 194.420 | 194.420 | 571.918 |
| 80.336 | 80.336 | 80.336 | 465.705 | 80.336 | 157.410 |
| 947.516 | 947.516 | 233.928 | 947.516 | 233.928 | 662.081 |

Table 4.3.4c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU203 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 465.050 | 465.050 | 286.652 | 404.101 | 129.362 | 350.043 |

Table 4.3.5a Predicted Glycemic Load

| Patient ID: GU201 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 21.215 | 0 | 71.623 | 33.09 | 142.22 | 0 |
| 34.201 | 71.623 | 71.623 | 56.591 | 142.22 | 33.09 |
| 11.627 | 29.316 | 71.623 | 0 | 0 | 33.09 |
| 14.146 | 29.316 | 71.623 | 0 | 0 | 33.09 |
| 0.871 | 29.316 | 71.623 | 0 | 0 | 33.09 |
| 61.665 | 142.22 | 33.09 | 29.316 | 33.09 | 29.316 |
| 28.562 | 142.22 | 33.09 | 0 | 142.22 | 56.591 |
| 187.573 | 142.22 | 33.09 | 33.09 | 56.591 | 29.316 |

Table 4.3.5b Percent Error Associated with Glycemic Load for Each Predicted Meals

| Patient ID: GU201 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 100.000 | 237.605 | 55.975 | 570.375 | 100.000 | 190.989 |
| 109.418 | 109.418 | 65.466 | 315.836 | 3.248 | 119.378 |
| 152.137 | 516.006 | 100.000 | 100.000 | 184.596 | 130.548 |
| 107.239 | 406.313 | 100.000 | 100.000 | 133.918 | 89.494 |
| 3265.786 | 8123.077 | 100.000 | 100.000 | 3699.082 | 2977.589 |
| 130.633 | 46.339 | 52.459 | 46.339 | 52.459 | 13.393 |
| 397.934 | 15.853 | 100.000 | 397.934 | 98.134 | 161.971 |
| 24.179 | 82.359 | 82.359 | 69.830 | 84.371 | 68.620 |

Table 4.3.5c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU201 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 535.916 | 1192.121 | 82.032 | 212.539 | 544.476 | 468.998 |

Table 4.3.6a Predicted Glycemic Load

| Patient ID: GU207 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 55.439 | 55.439 | 11.964 | 42.577 | 55.439 | 39.386 |
| 46.376 | 69.629 | 69.629 | 3.467 | 3.467 | 35.740 |
| 100.000 | 38.312 | 36.836 | 38.312 | 36.836 | 50.059 |

Table 4.3.6b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU207 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 55.439 | 55.439 | 11.964 | 42.577 | 55.439 | 39.386 |
| 46.376 | 69.629 | 69.629 | 3.467 | 3.467 | 35.740 |
| 100.000 | 38.312 | 36.836 | 38.312 | 36.836 | 50.059 |

Table 4.3.6c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU207 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 67.272 | 54.460 | 39.476 | 28.119 | 31.914 | 41.728 |

4.3.3 RESULTS ASSOCIATED WITH OB WOMEN

In this section results associated with OB women are described. Tables 4.3.7a, 4.3.8a, and 4.3.9a shows the predicted glycemic load, tables 4.3.7b, 4.3.8b, and 4.3.9b shows the errors associated with glycemic load for each meal and tables 4.3.7c, 4.3.8c, and 4.3.9c shows the errors associated with glycemic load in each parameters for entire day 2.

Table 4.3.7a Predicted Glycemic Load

| Patient ID: GU318 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 100.94 | 24.688 | 24.688 | 24.688 | 46.797 | 46.797 |
| 24.817 | 46.797 | 213.025 | 213.025 | 46.797 | 46.797 |

Table 4.3.7b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU318 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 75.542 | 75.542 | 75.542 | 53.639 | 53.639 | 66.781 |
| 88.568 | 758.383 | 758.383 | 88.568 | 88.568 | 356.494 |

Table 4.3.7c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU318 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 82.055 | 416.963 | 416.963 | 71.104 | 71.104 | 211.638 |

Table 4.3.8a Predicted Glycemic Load

| Patient ID: GU314 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 118.73 | 34.141 | 0.463 | 7.044 | 34.141 | 91.82 |
| 98.95 | 34.141 | 0 | 36.747 | 36.747 | 91.82 |
| 26.457 | 34.141 | 0 | 36.747 | 36.747 | 91.82 |

Table 4.3.8b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU314 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 71.247 | 99.610 | 94.068 | 71.247 | 22.671 | 71.769 |
| 65.497 | 100.000 | 62.863 | 62.863 | 7.206 | 59.686 |
| 29.043 | 100.000 | 38.893 | 38.893 | 247.054 | 50.777 |

Table 4.3.8c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU314 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 55.262 | 99.870 | 65.275 | 57.668 | 92.310 | 60.744 |

Table 4.3.9a Predicted Glycemic Load

| Patient ID: GU316 | | | | | |
|---|---|---|---|---|---|
| **Diary Value** | **Detected Value Based on Area** | **Detected Value Based on Duration** | **Detected Based on Peak** | **Detected Based on Pre Meal Glucose** | **Detected Based on Difference between Pre Meal and Peak** |
| 13.403 | 179.175 | 179.175 | 179.175 | 1.888 | 179.175 |
| 60.972 | 1.888 | 1.888 | 179.175 | 1.888 | 179.175 |
| 86.212 | 1.888 | 1.888 | 179.175 | 1.888 | 1.888 |
| 54.56 | 179.175 | 65.138 | 65.138 | 1.888 | 179.175 |
| 14.701 | 179.175 | 65.138 | 65.138 | 1.888 | 179.175 |

Table 4.3.9b Percent Error Associated with Glycemic Load for Each Predicted Meal

| Patient ID: GU316 | | | | | |
|---|---|---|---|---|---|
| **Percent Error in GL Predicted using area** | **Percent Error in GL Predicted Using Duration** | **Percent Error in GL Predicted Using Peak** | **Percent Error in GL Predicted Using Pre Meal Glucose** | **Percent Error in GL Predicted using Difference Between Pre Meal and Peak Glucose** | **Average Error for Meal** |
| 1236.828 | 1236.828 | 1236.828 | 85.914 | 1236.828 | 972.280 |
| 96.903 | 96.903 | 193.864 | 96.903 | 193.864 | 19.404 |
| 97.810 | 97.810 | 107.831 | 97.810 | 97.810 | 56.682 |
| 228.400 | 19.388 | 19.388 | 96.540 | 228.400 | 79.807 |
| 1118.79 | 343.086 | 343.086 | 87.157 | 1118.795 | 567.320 |

Table 4.3.9c Percent Error Associated with Glycemic Load for Day 2

| Patient ID: GU316 | | | | | |
|---|---|---|---|---|---|
| **Average Error Based on Area** | **Average Error Based on Duration** | **Average Error Based on Peak** | **Average Error Based on Pre Meal Glucose** | **Average Error Based on Difference between Pre Meal and Peak** | **Total Over All Average** |
| 555.746 | 358.803 | 380.199 | 92.865 | 575.139 | 339.099 |

# CHAPTER 5

# CONCLUSION

From the results presented in chapter 4, it can be concluded that the platform developed as a part of this study is successful in cleaning the data and predicting meal times. The average percent of correctly detected meals for NW participants was 82.777, the average percent of correctly detected meals for OW participants was 85.933, and the average percent of correctly detected meals for OB participants was 80.589. However, the automated platform has a high error percentage in detecting glycemic load due to the lack of information which can be eliminated by obtaining more data and conducting controlled experiments to understand the relationship of glycemic load with other glucose parameters. Our platform gives clinicians a basic playground to easily evaluate the glycemic response of the patients and the ability to see the effect of excursions due to meals on the glucose values. The proposed platform can facilitate diverse applications of CGM to support multiple research studies conducted by clinical investigators globally and at University of Alabama at Birmingham (UAB). For example, Dr. Chandler-Laney plans to investigate whether free-living glucose profile in early pregnancy can be used to predict the time-course for β-cell failure and the development of gestational diabetes in the absence of any intervention. This

investigation will also help in understanding childhood obesity. As technology underlying CGM improves, its use will expand in both research and clinical fields. The real strength of CGM, however, lies in the potential to characterize all aspects of free-living glucose profile, including indices of glycemic control, variability, and the rate and acceleration of glucose changes, along with examining associations with lifestyle factors and diurnal meal pattern. Our propose platform is a step forward in this direction.

# REFERENCES


1. PREVENTION CFDCA. National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. In: Services. USDoHaH, ed. Atlanta, GA, 2011.
2. BOYLE JP, THOMPSON TJ, GREGG EW, BARKER LE, WILLIAMSON DF. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. Popul Health Metr 2010;8:29.
3. ALBRIGHT AL, GREGG EW. Preventing type 2 diabetes in communities across the U.S.: the National Diabetes Prevention Program. Am J Prev Med 2013;44:S346-51.
4. Goldstein DE, Little RR, Wiedmeyer HM, England JD, McKenzie EM: Glycated hemoglobin: methodologies and clinical applications. Clin Chem 1986;32:B64-B70.
5. Hill Nathan R., Oliver Nick S., Chodhary Pratik, Levy Jonathan C., Hinmarsh Peter, and Matthew David R: Normal Regerence range for Mean Tissue Glucose and Glycemic Variablity Derived from Continuous Glucose Monitoring for Subjects Without Diabetes in Different Ethnic Groups.

6. Cengiz Eda, Tamborlane William V.: A Tale of Two Compartments: Interstitial Versus Blood Glucose Monitoring. Diabetes Technology & Therapeutics Volume 11, Supplement 1, 2009.
7. RODBARD D. The Challenges of measuring glycemic Variablity.J Diabetes Science and Technology 2012.
8. HERNANDEZ TL, BARBOUR LA. A standard approach to continuous glucose monitor data in pregnancy for the study of fetal growth and infant outcomes. Diabetes Technology and Therapeutics 2013;15:172-9.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
9. WAGNER J, TENNEN H, WOLPERT H. Continuous glucose monitoring: a review for behavioral researchers. Psychosom Med 2012;74:356-65.
10. Stratton IM, Adler AI, Neil HA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ. 2000; 321(7258):405–412.
11. United Kingdom Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet. 1998;352(9131):837–853.

12. National diabetes statistics report, 2014. Centers for Disease Control and Prevention website.www.cdc.gov/diabeteS/pubs/statsreport14.htm. Updated June 13, 2014. Accessed June 16, 2014.
13. http://diabetes.niddk.nih.gov/dm/pubs/causes/
14. Nathan DM, Buse JB, Davidson MB, et al. Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy. Diabetes Care.2009;32(1):193–203.
15. Rodbard HW, Jellinger PS, Davidson JA, et al. Statement by an American Association of Clinical Endocrinologists/American College of Endocrinology consensus panel on type 2 diabetes mellitus: an algorithm for glycemic control. Endocr Pract. 2009;15(6):540–559).
16. Scheiner Gary, Schwartz Stanley, The role of continuous Glucose Monitoring in Management of Type-1 and Type-2 Diabetes. Evidence Based Management of Diabetes, 2012.
17. Klonoff DC. Continuous glucose monitoring: roadmap for 21st century diabetes therapy. Diabetes Care. 2005;28(5):1231-9.
18. Klonoff DC. A review of continuous glucose monitoring technology. Diabetes Technol Ther. 2005;7(5):770-775
19. VADDIRAJU S, BURGESS DJ, TOMAZOS I, JAIN FC, PAPADIMITRAKOPOULOS F. Technologies for continuous glucose monitoring: current problems and future promises. J Diabetes Sci Technol 2010;4:1540-62.

20. STATEMENT BY THE AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS CONSENSUS PANEL ON CONTINUOUS GLUCOSE MONITORING Thomas C. Blevins, et al ENDOCRINE PRACTICE 2010.16 :440-56
21. FACCHINETTI A, SPARACINO G, COBELLI C. An online self-tunable method to denoise CGM sensor data. IEEE Transactions on Biomedical Engineering 2010;57:634-41
22. MAHMOUDI Z, JOHANSEN MD, CHRISTIANSEN JS, HEJLESEN OK. A multistep algorithm for processing and calibration of microdialysis continuous glucose monitoring data. Diabetes Technology and Therapeutics 2013;15:825-35.
23. CHASE JG, HANN CE, JACKSON M, LIN J, LOTZ T, WONG XW, SHAW GM. Integral-based filtering of continuous glucose sensor measurements for glycaemic control in critical care. Computer Methods and Programs in Biomedicine 2006;82:238-47.
24. BRETON M, KOVATCHEV B. Analysis, modeling, and simulation of the accuracy of continuous glucose sensors. J Diabetes Science and Technology 2008;2:853-62.
25. MAHMOUDI Z, JOHANSEN MD, CHRISTIANSEN JS, HEJLESEN OK. A multistep algorithm for processing and calibration of microdialysis continuous glucose monitoring data. Diabetes Technology and Therapeutics 2013;15:825-35.

26. Freeman Janine and Lynee Lyons. The Use of Continuous Glucose Monitoring to Evaluate the Glycemic Response to Food. Diabetes Spectrum 2008;21:2134-137.

27. Sheard NF, Clark NG, Brand-Miller JC, Franz MJ, Pi-Sunyer FX, Mayer-Davis E, Kulkarni K, Geil P: Dietary carbohydrate (amount and type) in the prevention and management of diabetes [American Diabetes Association Statement]. *Diabetes Care***27**:2266 –2271, 2004.

28. National Cancer Institute (NCI) Automated Self-administered 24-hour Recall (ASA24). Applied Research Program, Risk Factor Monitoring and Methods Branch. 2008. [July 16, 2008]. http://riskfactor.cancer.gov/tools/instruments/asa24.html

# APPENDIX A

# INSTITUTIONAL REVIEW BOARD APPROVAL

# IRB Exemption Review Application

- **To complete the form**, click the underlined areas and type or paste in your text; double-click checkboxes to check/uncheck. For more tips, see www.uab.edu/irb/forms.
- **Mail or deliver all materials to AB 470**, 701 20th Street South, Birmingham, AL 35294-0104.

**1. Project Identification**

a. Title of Project: **Development of automated processing techniques for time-series data derived via continuous glucose monitoring (CGM): Working toward the identification of patterns in free-living glucose profile that may predict future glucose intolerance**

b. Principal Investigator (PI): **Karthikeyan Lingasubramanian** PI's BlazerID or E-Mail Address: **klinga@uab.edu**

If the PI is a student, fellow, or resident, provide the name, number, and email of the faculty advisor or course instructor as contact information and obtain the person's signature.
Advisor/Instructor's Name:_____ Telephone Number:_____ BlazerID:_____
Advisor/Instructor's Signature:_____

c. PI's Address (on-campus or home)
On-Campus: Department:**Electrical and Computer Engineering** Building:**BEC** Room:**255D** UAB Zip:**35294**
Phone:**2059753385** FAX:**2059753337**
-OR-
Home Address:_____ Street:_____ City:_____ State:_____ ZIP:_____
and Campus Affiliation:_____

d. List all staff who will be involved with the research, their degree(s) and job title, and any additional qualifications. Include individuals who will be involved in the consent process. *Repeat the table below for each individual.*
*Note. For studies involving investigational drugs, include all investigators who will be listed on FDA Form 1572 and attach a copy, if applicable. Send the IRB a copy of Form 1572 anytime you update the form with the FDA.*

| | |
|---|---|
| Role: | ☒Co- -OR- ☐Other |
| Full Name: | **Paula Chandler-Laney** |
| Primary UAB Dept.: (Employer if not UAB) | **Nutrition Sciences** |
| Degree(s) / Job Title: | **PhD** |
| Additional Qualifications pertinent to the study: | |

| | |
|---|---|
| Role: | ☐Co- -OR- ☒Other |
| Full Name: | **Camille Schneider** |
| Primary UAB Dept.: (Employer if not UAB) | **Nutrition Sciences** |
| Degree(s) / Job Title: | **BS, RD** |
| Additional Qualifications pertinent to the study: | |

| | |
|---|---|
| Role: | ☐Co- -OR- ☒Other |
| Full Name: | **Aman Khatri** |
| Primary UAB Dept.: (Employer if not UAB) | **Electrical and Computer Engineering** |
| Degree(s) / Job Title: | **BSEE** |
| Additional Qualifications pertinent to the study: | |

e. Is this activity funded in any way? ☒Yes ☐No
If yes, attach 1 copy of completed application and complete (i)-(iv):

i. Grant or Contract Title: **Development of automated processing techniques for time-series CGM data**

ii. PI of Grant or Contract: **Karthikeyan Lingasubramanian**

iii. OSP Proposal Number:

iv. Funding Source

☐Gov't Agency or Agencies:_____

☒UAB Departmental Funds: **DRC (Diabetes Research Center/Department of Nutrition Sciences)**

☐Other:_____

**2. Mark the category or categories below that describe the proposed research:**

☐1. Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods. The research is not FDA regulated and does not involve prisoners as participants.

☐2. Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless: (i) Information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation. Attach questionnaire(s) and/or surveys. If the research involves children as participants, the procedures are limited to educational tests and observation of public behavior where the investigators do not participate in the activities being observed. The research is not FDA regulated and does not involve prisoners as participants.

☐3. Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior that is not exempt under category (2), if: (i) the human subjects are elected or appointed public officials or candidates for public office; or (ii) federal statute(s) require(s) without exception that the confidentiality of the personally identifiable information will be maintained throughout the research and thereafter. Attach to this application a copy of any questionnaire or survey to be used. The research is not FDA regulated and does not involve prisoners as participants.

☒4. Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the Investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. Attach a specimen release form if applicable. (*Specimens must be preexisting.*) The research is not FDA regulated and does not involve prisoners as participants.

☐5. Research and demonstration projects which are conducted by or subject to the approval of department or agency heads, and which are designed to study, evaluate, or otherwise examine: (i) public benefit or service programs; (ii) procedures for obtaining benefits or services under those programs;(iii) possible changes in or alternatives to those programs or procedures; or (iv) possible changes in methods or levels of payment for benefits or services under those programs. The protocol will be conducted pursuant to specific federal statutory authority; has no statutory requirement for IRB review; does not involve significant physical invasions or intrusions upon the privacy interests of the participant; has authorization or concurrent by the funding agency and does not involve prisoners as participants.

☐6. Taste and food quality evaluation and consumer acceptance studies, (i) if wholesome foods without additives are consumed or (ii) if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the U.S. Department of Agriculture. The research does not involve prisoners as participants.

3. Briefly describe the proposed research: **Continuous glucose monitoring (CGM) is a rapidly advancing technology that has been used to facilitate glucose management in individuals with type 1 and 2 diabetes. There are numerous applications for CGM, however, beyond clinical management of diagnosed disease. Increasingly, CGM is being utilized in research studies to characterize glucose profiles of both diabetic and non-diabetic individuals. One limitation of CGM use in research is that it produces an enormous amount of data that must be inspected, cleaned, and reduced, in order for meaningful outcomes to be derived. The overall objective of the proposed study is to develop an automated method to process CGM data in order to derive outcomes that fully characterize free-living glucose profile and have the potential to predict future glucose intolerance and diabetes.**

4. Describe how subjects/data/specimens will be selected. If applicable, include the sex, race, and ethnicity of the subject population: **Data to be used for this study has previously been collected as part of protocols F110714003 (PI: Chandler-Laney), and F100310019 (PI: Biggio). In these protocols, pregnant women wore continuous glucose monitors for 3 days and concurrently recorded their food intake on diet records, and their activity was recorded via accelerometer. All participants were female and the majority (~90% or more) in each cohort was African American. The remainder were Caucasian or Hispanic.**

5. Does the research involve deception? ☐Yes ☒No

6. Describe why none of the research procedures would cause a subject either physical or psychological discomfort or be perceived as harassment above and beyond what the person would experience in daily life: **Only de-identified data will be provided to the PI and his graduate research assistant. No contact between investigators and the former participants of the parent study will take place as part of the proposed analysis.**

7. Describe the provisions to maintain confidentiality of data: **All data provided to the PI and his assistant will be de-identified by Dr. Chandler-Laney who is the PI of the F110714003 protocol and co-I of the F100310019 protocol. The data will be coded with a unique identifier that cannot be used to determine the participants' identities.**

8. Describe the provisions included in the research to protect the privacy interests of participants (e.g., others will not overhear your conversation with potential participants, individuals will not be publicly identified or embarrassed).: **No conversations or interactions with participants are planned. Further, during conversations among investigators, no names will be used.**

9. Will the research involve interacting with the subjects? ☐Yes ☒No

**If yes,** describe the consent process and information to be presented to subjects, including:

- That the activities involve research.
- The procedures to be performed.
- That participation is voluntary.
- Name and contact information for the investigator.

10. Additional Information

In the space below, provide any additional information that you believe may help the IRB review the proposed research, or enter "None."

**The variables that will be provided to the PI and his research assistant for the purposes of this study are listed below. These are the variables that are likely or possibly relevant to the development of the automated processing method.**

**All CGM data, participant height, weight, age, gestational age at time of CGM data collection, sex of the fetus, all dietary variables including meal times and composition of meals, all activity data including number of steps taken and intensity of activity each hour of the recording period, serum measures of glucose, insulin, and C-peptide.**

11. Findings? **(applicable for Continuing Review or Final Report only)**

State both the positive and negative results received to date:

**Since the last IRB review,** have any of the following occurred?

a. Have participants experienced any harms (expected or unexpected)? ☐Yes ☐No
If yes, attach Problem Summary Sheet, and briefly describe here the harms (serious and/or non-serious) experienced by participants:____

b. Have there been any unanticipated problems involving risks to participants or others? ☐Yes ☐No
If yes, attach Problem Report, and briefly describe here the unanticipated problems involving risks to participants or others:____

c. Have you have any problems obtaining informed consent? ☐Yes ☐No ☐N/A
If yes, briefly describe the problems here:____

d. Have any participants or others complained about the research? ☐Yes ☐No
If yes, briefly describe the number and nature of the complaints:____

e. Have any participants withdrawn from the research? ☐Yes ☐No
If yes, indicate the number of withdrawals and include the reason for each:____

f. Have any obvious, study-related benefits occurred for participants? ☐Yes ☐No
If yes, briefly describe the benefits here:____

g. Have the risks or potential benefits of this research changed? ☐Yes ☐No
If yes, briefly describe the changes here:____

h. Has there been any published literature? ☐Yes ☐No
If yes, attach a copy and summarize the published findings here:____

Principal Investigator's Signature:_____ Date: 5/13/14

UAB THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Institutional Review Board for Human Use

Form 4: IRB Approval Form
Identification and Certification of Research
Projects Involving Human Subjects

UAB's Institutional Review Boards for Human Use (IRBs) have an approved Federalwide Assurance with the Office for Human Research Protections (OHRP). The Assurance number is FWA00005960 and it expires on January 24, 2017. The UAB IRBs are also in compliance with 21 CFR Parts 50 and 56.

| | |
|---|---|
| Principal Investigator: | LINGASUBRAMANIAN, KARTHIKEYAN |
| Co-Investigator(s): | CHANDLER-LANEY, PAULA C |
| Protocol Number: | **E140417004** |
| Protocol Title: | *Development of Automated Processing Techniques for Time-Series CGM Data* |

The above project was reviewed on 5/16/14. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services. This project qualifies as an exemption as defined in 45CF46.101, paragraph 4.

This project received EXEMPT review.

IRB Approval Date: 5/16/14

Date IRB Approval Issued: 5/16/14

Cari Oliver
Assistant Director, Office of the
Institutional Review Board for Human
Use (IRB)

Investigators please note:

IRB approval is given for one year unless otherwise noted. For projects subject to annual review research activities may not continue past the one year anniversary of the IRB approval date.

Any modifications in the study methodology, protocol and/or consent form must be submitted for review and approval to the IRB prior to implementation.

Adverse Events and/or unanticipated risks to subjects or others at UAB or other participating institutions must be reported promptly to the IRB.

470 Administration Building
701 20th Street South
205.934.3789
Fax 205.934.1301
irb@uab.edu

The University of
Alabama at Birmingham
Mailing Address:
AB 470
1720 2ND AVE S
BIRMINGHAM AL 35294-0104