DALHOUSIE UNIVERSITY

Internship Report: Informatics Student Innovator with IWK Health Centre

Halifax, NS
Masters of Health Informatics
HINF 7000
Dalhousie University
August 15, 2016

Travis Stenerson
B00733087
tstenerson@gmail.ca

Supervisor: Dr. Brett Taylor

Acknowledgements and Endorsement

Certainly this internship would not have been possible without the endorsement and support of Dr. Brett Taylor, Pediatric Emergency Physician and Head of Informatics at the IWK Health Center. His support, feedback and encouragement has been integral to my project this summer and I can not express my gratitude enough. As well, I would like to thank Dr Arati Mokashi, Pediatric Endocrinologist, and the Diabetes Care team, including the nurses, dieticians, and administrators in the Peds Endocrinology department at IWK. Observing the clinical workflow and exactly how diabetic patient data is utilized in a meaningful way dramatically shaped my vision for the final product that is to be the result of this internship. Further thanks must go to the hospital administrators, including Tracey Kitch, President and CEO of IWK, who approved the funding for this position such that I might realize my vision. Of course, my gratitude must also be expressed for Dr Raza Abidi, as well as all the MHI professors at Dalhousie, whose feedback and support has helped me realize how to properly bring informatics solutions to health care, and how to present these solutions to increase the likelihood that they might see some use in a clinical workflow.

This report was written entirely by the author, Travis Stenerson, for the purposes of reporting on my internship work term, HINF 7000, during the summer semester of 2016, for the Internship portion of the Master of Health Informatics. No credit has been received for this report at any other institution.

Executive Summary

This summer I have worked with IWK Health Centre and their Informatics Department in a novel position titled 'Student Innovator in Health Informatics', seeking to bring a health informatics based solution of my design to reality. The position is the result of Dr Brett Taylor, a graduate of the Master of Health Informatics at Dalhousie, and hospital administrators, recognizing the growing importance of bringing technology to healthcare, hope to foster innovation and provide support to health informatics students with a vision. I was fortunate enough to be the first student in this position.

This position is unique in that, while there is ample clinical and technical support, the majority of the work is self-directed. I designed, planned, adapted, and taught myself that which was necessary for the project. With the help and feedback of my advisors, supervisor and clinical consultants, the project nears completion.

The goal is to create a system that would enable diabetic patients to record, display, analyze and understand data from their lives, including blood sugar, insulin doses, carbohydrate consumption and exercise sessions. As well, to prepare this data for transfer to their care team in a clinically meaningful way. At present, this data is addressed with their care team, but it is often presented incomplete or recorded with pen and paper. The project is ongoing and will continue after the internship comes to a close, but together we have built a near complete application for iOS mobile devices that accomplishes most of these goals.

Design decisions were deeply informed by clinical observation in hospital with Dr Arati Mokashi and the diabetes team at the Pediatric Endocrinology Department at IWK. Many aspects of health informatics are involved in bringing an information technology solution like this to a health care problem. Fitting seamlessly into existing clinical workflow, technical system design choices, statistical analysis, and creating a clinically meaningful tool that can also operate as a patient teaching tool are all aspects that I addressed in the design and implementation of this project. This report will detail many facets of these informatics associated topics, and will seek to justify the continuation of this substantial position for future students.

Discussed herein are two specific informatics topics that were among the more complex addressed this summer. The first is a visual data representation of blood glucose, carbohydrate consumption and insulin action relative to time of day. This is a chart adapted from observations of patient-caregiver discussions about their blood sugar readings at home. The second is a novel algorithm for calculating a specific measure of glycemic variability, the mean amplitude of glycemic excursions. This is a well studied and validated measure of blood glucose swings that has been noted to be difficult to compute. (Marics et al., 2015) Both these challenges were solved using personal research, support from my advisors, knowledge taken from my course work and experimentation. The chart will visually support and improve the conversation between caregiver and patient about blood sugar at specific times of days. The algorithm demonstrates taking a clinical concept and translating it to computer understandable code.

The project as a whole has been challenging as well as deeply rewarding. Entrusting a student that seeks to bring informatics solutions to healthcare, and having project success be contingent on their ability is extremely motivating. The volume of information and technical skills

I have acquired to see this project through are ample evidence of how important a position like this can be. Certainly, this experience should contribute greatly to future opportunities in the health informatics field. I also have a great deal of gratitude to the institution for giving me this opportunity and intend to contribute to the continuation of the Student Innovator position as well as the Informatics Department at IWK in the future in any way I can. Beyond what I personally have taken from this position, we also have a patient-centered data recording and analysis tool prototype that compares very favorably to existing products. I feel it will improve the experience of recording blood sugar for patients and of dissecting the data for physicians

  The unique elements of this position allowed me to see informatics from a both a top down and bottom up perspective. Mapping clinical context to technical choices and then implementing those technical choices allowed me to apply everything I have learned in health informatics to a real world project. The position required confidence, humility, curiosity and implicit motivation to accomplish what I set out to. I sincerely hope that future students have this opportunity and that my advisors and the institution see the value that this position has had for me. Health informatics is a field that is evolving quickly. That evolution and the extent to which we can transform healthcare with technology will depend on the sort of support to bring an idea from concept to reality that I have had this summer.

Table of Contents

1. Introduction

During the summer semester of 2016, I have worked at an internship with the IWK Health Centre in Halifax, NS, on a project of my design. This report will detail this internship, the project, the organization and departments with which I worked, and how this internship has contributed to my understanding of health informatics.

This internship represents the commencement of a program set up by the Informatics Department at IWK that hopes to encourage students who seek to bring technological innovation to healthcare. I am one such student, and I was fortunate enough to meet the individual physician and informatician, Dr Taylor, that made this program a reality. This position enabled me to acquire skills I have wanted to acquire since starting my post-secondary education, and learn more and work harder on a wide array of subjects in four months than any similar period in my life. I can not thank Dr Taylor and IWK enough for this opportunity, and I hope that the work I have done this summer will encourage their continuation of the program. They provided me with the clinical contacts, guidance and access to patients that are necessary in health informatics and enabled me to work from a conceptual design to a tangible healthcare application nearly ready for consumer release.

The internship is a self-directed personal project that I have been envisioning since being re-exposed to information technology early in the Master of Health Informatics program at Dalhousie. It draws on my clinical experience, my knowledge of health informatics and my technical skills in a way that a typical internship within an organization could not. I had to plan the project, outline the steps and tasks, learn how to perform those steps that I did not know prior and then genuinely implement them. The project began with lofty set of goals, many of which have been accomplished.

The project is a mobile healthcare application and analytical tool for diabetic patients and their caregivers. Initially, I had hoped to build both a patient-centric app for data collection, visualization and teaching, a server-side environment for data transfer, and a caregiver-centric app for customizable display of patient data. This would have involved learning two programming languages that I was only partially exposed to prior to this summer. As of August 15th, I have a patient-centric app that performs all the necessary analytics and data display, and I have begun learning how to build the remaining portions. The majority of my summer was spent coding this application, with regular meetings with my advisors and supervisor, as well as intermittent clinical shadowing to observe the clinical workflow in pediatric diabetes care.

The major component of this report will detail my progress on the project and how I see this project fitting within health informatics. I will elaborate on two of the more complex informatics issues I faced and overcame. I will conclude with recommendations for the position and for future students that are fortunate enough to work in the 'Student Innovator' position.

2. The Organization: The IWK Health Center

IWK Health Centre, in Halifax, is a 1252 bed women and children's hospital that provides care to women, children, families and youth in the Maritime provinces and beyond. ("IWK Health Centre | About Us", 2016) My supervisor is Dr. Brett Taylor, a pediatric emergency physician and the Lead Physician, Informatics for the IWK Health Centre. ("Dalhouse Division of Pediatric Emerency Medicine | Our People", 2016) Dr Taylor is a health informatician, in fact a graduate of the Dalhousie MHI program, and has made it a goal of his to help bring innovation in health information technology to IWK and to health care at large.

Dr Taylor and I met last November, at the mHealth conference in Halifax. I had the opportunity to speak briefly about my impressions of health informatics in a two minute pitch to the conference, and I touched on my desire to contribute to the field and to learn from this group involved in the industry. I had a vision for a specific product that I felt could be useful to primary care physicians and diabetes care teams that I was designing and building. Dr Taylor, always seeking to foster such ideas and to provide the relevant clinical experience and patient data, approached me. We discussed what I was hoping to build and he told me that his Informatics department would like to set up a program that enabled students to do exactly what I hoped to do in my spare time. He approached the new CEO and President of IWK, Tracy Kitch, who too sees the value in bringing health IT solutions to the hospital. They put aside a small amount of funding for a 'student innovator' position, and provided me with clinical contacts in the hospital that could guide me. I became the first 'Student Innovator', and I hope to contribute to future students like myself in this position. I have learned so much this semester in numerous fields directly related to Health Informatics, and feel that this position could provide future MHI students a similar experience.

The major clinical contact was Dr. Arati Mokashi, pediatric endocrinologist, and her diabetes care team at the IWK Pediatric Endocrinology department. I worked with her, the nurses, the dieticians, and the patients to observe clinical workflow and mold my vision into something that aligns more closely with clinical care for Type 1 Diabetics. Their contribution was invaluable to my work this summer. I am deeply gracious to all those who came together to help me bring my vision to fruition.

3. The Position: Student Innovator in Health Informatics

The position title is 'Student Innovator in Health Informatics', and as such the description of the position is deeply dependent on the particular innovation the student is hoping to accomplish. The position is unique in that it is not exactly an intern - observing and contributing in small increments where they can. It is more than that - a self learning opportunity to see health IT innovation from its infancy to a tangible product. This comes with some disadvantages relative to regular internships, but at the same time inherently has many more advantages over typical internships.

The disadvantage to a position involving as much auto-didactics as this is that you don't get to see the industry after implementation or development of products. Though I had many great teachers, mentors and ample guidance, there is something to be learnt from actual industry that I may have missed out on, such as the business side of the industry, challenges with funding and working within a large organization. However, this tiny disadvantage is overshadowed by the genuine experience of having a vision, learning how it could manifest in clinical workflow, and obtaining the relevant knowledge that is required to bring it to life. When looking retrospectively on the position, I feel that this knowledge would be infinitely more valuable to me in future positions in Health Informatics, in government, industry or academics.

The timeline for the project was entirely proposed by me, and I had some flexibility in what I needed to achieve. I set my goals, in retrospect, too high to accomplish in a single semester. I had hoped to have a working prototype of an entire system for data transfer and display in both a local environment on the patient's mobile, and a server side environment for caregivers. I vastly underestimated the amount of time it takes to code a project, and this is owing to my inexperience with programming. Table 1: Tasks and Progress outlines what I had hoped to achieve and estimations of what I have achieved thus far. Currently, I am very close to having a fully functional app for patients to record and analyze blood sugar. I expect to be ready to submit to Apple for approval and distribution in the app store by the end of the term. However, major changes are imminent to the Swift language used for the project, and this may delay my projected completion date as syntax changes will be necessary after the release of Swift 3. (Apple Inc., 2016)

Table 1: Tasks and Progress

| Main Task	Sub Task	Started	Projected End Date	Progress as of August 15th
				10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Patient App: SugarTracker	Database	Yes	Done	
	Data Entry System	Yes	Done	
	Analytics	Yes	Mid June	
	Achievements System	Yes	End of June	
	Data Visualization System	Yes	End of June	
	Data Output System	Yes	Early July	
	Data Transfer System	No	August	
	Data Encryption	No	August	
Server Side	Learn Ruby	Yes	End of June	
	Learn Rails	Yes	End of June	
	API List	Yes	July	
	API code	No	July	
	Security	No	August	
Physician App: Remote Monitoring System	Database	No	August	
	Login System	No	August	
	Data Transfer System	No	August	
	Data Visualization System	No	August	
	Data Printing Option	No	August	
	Data Decryption	No	August	

4. Relation to Health Informatics

My position related to Health Informatics(HI) in numerous ways. Particularly in the field of patient centric mobile health, however many aspects of health informatics were involved. The 'innovation' I aspire to bring to health care with this position is a data transfer and analysis platform for diabetic patients and their care team. The product would take self measured blood glucose(SMBG) data, as well as other data affecting blood glucose, from patients. It would analyze and graphically display the data, and provide the users with pre-formatted, clinically relevant data output that can be brought to their care team and would fit seamlessly into the clinical workflow. Certainly, similar products exist, and I reviewed the literature surrounding these products prior to and again during the term. I have been selectively choosing project topics surrounding this theme throughout my time in the MHI program to guide my design decisions for the product I envisioned.

As I see it, the product and the work I have done this summer relates to informatics as a mobile health product, as a technical system, as a statistical analysis platform, as a clinical tool, and as a patient centered learning tool.

Mobile health is an aspect of HI that seeks to support health care with the ubiquity of mobile devices. (Canada Health Infoway, 2014) My product is one such platform on mobile devices, specifically for iOS devices. As such I had to be aware of diabetes specific apps, competitors and the current state of the mobile diabetes application industry. As of 2016, 952 diabetes targeted apps exist in Apple and Google's mobile stores. (Basilico, Marceglia, Bonacina & Pinciroli, 2016) My literature inquiry into this field primarily demonstrated the speed of evolution of the industry and, prominently, the value of publishing a framework mapping clinical workflow to design decisions and clinical justification with controlled outcome measurement. Furthermore, it is important to know the target population such that the design can enable them with an easy-to-use product. Compliance to use, like any health intervention, is integral to success. The application is intended for use by Type 1 and Type 2 diabetics of any age, as well as their parents (in the case of young Type 1 diabetics). Early design choices focused on this wide range of users. An intuitive interface, limited typing, large, readable fonts, entry error correction and simple explanations were goals from early development.

Technical aspects of the system fit within HI and reflect much of what I have learned from the MHI program. I had to design a simple database for local storage on the device. I suspect I will also need to design a similar database that includes user IDs and passwords for the final product, however I am not at that stage yet. The local database was implemented using SQLite, and the entity-relationship diagram is displayed in Figure 1: ER Diagram of Local Database. This database was instantiated as a

Figure 1: ER Diagram of Local Database
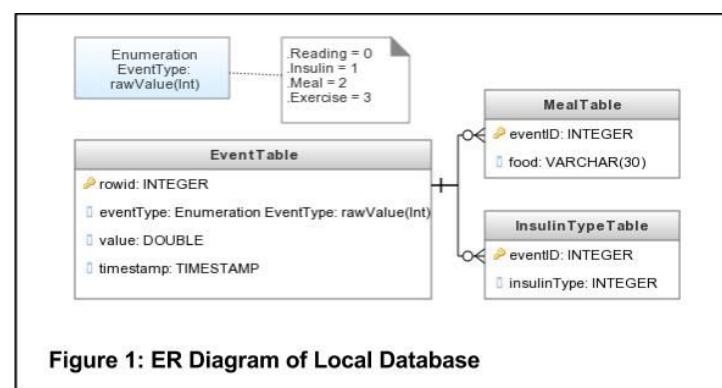
singleton, a shared instance across all classes, and is interacted with through a publicly available SQLite library for the language I coded the app in. Access to the database follows the model-view-controller paradigm that I learned of early in my self-taught coding experience.

Other technical aspects were, of course, the immense amount of coding I learned through books, documentation, websites, tutorials and experimentation. The project is coded in Swift, Apple's new open source language with a focus on readability. While this is more technical than informatics, I feel having an understanding of all aspects of information technology is helpful to bringing technological solutions to health care. Informaticians, in my opinion, often operate as 'translators' between those with technical skills and those with clinical skills. Much like teaching a language, understanding the native language of they whom you are teaching makes passing knowledge much easier. Though programming made up the vast majority of my time, this report will not focus on it. However, I am confident that being able to cite my programming ability will help with obtaining and succeeding in future positions.

Statistics and analytics are a portion of health informatics that were featured prominently in my project. This involved retrieving, collating, visually displaying data and mining that data for knowledge. I drew on much of what was taught during the MHI Statistics course with Dr Stewart, often referencing his lecture slides. Using this, I have written methods(class specific functions) for linear regression, for estimation of area under the curve, range, standard deviation, mean and many other necessary data manipulations.

The clinical end of the technical-clinical spectrum of HI was also notable in my position. Observing clinical workflow was important in shaping my vision, as product uptake requires a deep understanding of the existing workflow. I have a good amount of experience in observation and learning in a clinical context, and went into this observation prepared for what I was looking for. I kept my questions directed and did my best not to intervene or discuss what I needed, instead allowing the patient-caregiver interactions to flow as they normally do. Furthermore, I reviewed a great deal of literature on the specifics of using diabetes data for clinical decisions. My review revealed that many aspects of self measured blood glucose data are overlooked in clinical decisions, and that this data may become increasingly important in establishing strong glycemic control as well as controlling for hypo- and severe hyper-glycemic events. (Kohnert, Vogt, & Salzseider, 2016) With this in mind, my product provides the algorithms for calculating measures of glycemic variability that are inaccessible to clinicians with patients not yet on an insulin pump, and occasionally not available at all.

Finally, I had to be aware of the capability of mobile applications to act as patient-centered learning tools. My goal was to give full control of SMBG data and analysis to clinicians as well as to patients. Much of the analysis may appear inaccessible and confusing to those without proper training, so the product will present the data output in ways that are meaningful and simplified for non-professionals. For instance, clinical language is usually replaced with terms immediately understandable('postprandial' becomes 'after eating'), and wherever possible simple data visualizations are used. Any clinical terms used come with prominent "What is?" expanding buttons or "Explain on tap" checkboxes in the patient app. This can be difficult to implement when trying to avoid lengthy text that users shy away from. So these learning opportunity are always initially hidden but immediately accessible for those patients who do want to learn more.

5. Discussion of Informatics Specific Issues

As was outlined in the previous section, my position involved a varied set of health informatics specific issues that I overcame. I will touch on two specific issues that I feel exemplify a range of problems in HI. The first of these issues is a data visualization that is an extension of some patient-dietician conversations I observed in the hospital. The second is an algorithm for the calculation of a well studied measure of glycemic variability that is typically unavailable for many patients. The issues discussed are microcosms of larger health informatics issues, such as fitting a technological solution within a clinical context and translating clinical values into a computer algorithm. This differs slightly from the requested description of problems that *could* have an informatics based solution. The project as a whole is a solution of my design meant to address the clinical issue of curating and analyzing self measured blood glucose data. I feel that specific examples of informatics related issues I faced and how I came to their solution would be more valuable in describing my work this summer. Furthermore, a complete map of design choices to clinical relevance would be beyond the scope of this report.

5.1 The Average Day Chart

My clinical observation at the pediatric endocrinology department revealed many aspects of clinical workflow that I hoped to emulate and extend with the product. I will demonstrate this here with a data visualization that is the result of the frequent discussion between dietician and patient regarding blood glucose values relative to time of day. I feel my solution to this issue demonstrates design decisions that were made to fit this application into the already existing clinical workflow, while extending the workflow in a way that will make this patient-caregiver discussion easier for both parties.

Prior to this observation, I imagined a product that presented all blood glucose reading to the caregiver in a time versus blood glucose chart to demonstrate simple rates of hyperglycemia. This is shortsighted, particularly in the context of Type 1 diabetics. Type 1 diabetics are completely responsible for manipulating their blood glucose directly with exogenous insulin. As such, patterns of blood glucose within a day are important. Large portions of the day spent hyperglycemic are a detriment to long and short term goals of glycemic control. I noted that, while going over a table of time of day versus blood sugar, the dietician was looking for patterns specific to a day and time. For instance, the patient tended to have very high readings in the mid-afternoon, 3pm to 5pm, and particularly so on Thursdays. So the dieticians inquired what they were eating for lunch or afternoon snack at that time, and what was particular about Thursdays.

I saw the opportunity for a chart that showed what an 'Average Day' looked like for the patient, showing mean blood glucose at all times of day, as well as meals, insulin doses and exercise sessions, averaged over a set time period. This dietician was basing her discussion on only the last two weeks of values, and was scanning a table of tiny numbers looking for patterns. This is likely a place where proper data visualization could be used to aid in locating

patterns. Making a chart with controls to filter data to day or time and easily accessible lists of food consumed at that time could reveal the source of this elevation.

I also noted that the data output sheets from insulin pumps provide a version of this chart, though they restrict it to two weeks of data. They provide charts showing all values recorded over those two weeks amalgamated into a single day. This chart is unavailable to those patients not on insulin pumps, which represent 75% of patients. ("Pediatric Endocrinology Department", 2016)

My implementation of this chart required a good deal of research on the statistics involved in amalgamating data in this way. I had to determine the mean blood glucose at each hour of the day, as well as the average time and average size of meals and insulin doses. Furthermore, to make the chart visually appealing, the final values were plotted using a smooth line that was the product of interpolation between the points, such that all points fell on the line and the line would represent smooth transitions from point to point.

When retrieved from the database, the glucose reading events contain a decimal value representing blood sugar, as well as an integer representing time-since-reference-point in seconds. This integer can be easily converted into a Date object which makes it possible to extract specific calendar units from the date like year, month, day, hour, minute, etc. For this chart, the data was converted into a point with a y value representing blood glucose and x value representing hour, minute and second of day in decimal value. So a reading of 8.5 mmol/L blood sugar performed on August 15th, at 5:15pm (17:15) would become a Point(x: 17.25, y: 8.5).

With this set of transformed points, it was possible to determine mean blood sugar at each time of day. This was accomplished using a 'weighted moving average', a technique prominent in financial statistics. Points were distributed into sets for each hour or half hour of day(depending on point density), and values were weighted inversely by time from the center x-point of the set. This differs slightly from most implementations of a weighted moving average, in that contributing points were taken both ahead of and behind the center x-point. Weighted moving average is calculated using the following equation:

$$Mean = (\sum_{i=1}^{n} \omega_i \times BG_i)/\sum_{i=1}^{n} \omega_i \qquad \text{("Weighted Moving Average Model", 2016)}$$

Where ω_i is the weight of reading i, ranging from a value of 0 if the reading is 30 minutes (or 60 minutes) away, or a value of 1 if the reading is directly on the center x-point, and BG_i is the blood glucose value for reading i.

This calculation was repeated at any interval containing more than one point(to avoid outliers pulling the line far away from what was the likely mean), or a single point if the full data set was sufficiently small. From there, the points were plotted on a chart with hour of day on the x axis and blood glucose on the y axis. Directly drawing a line between the points would, in my opinion, be less easily understood to patients than a smooth curve, so a smoothing operation was done between the points. This is called 'spline' or 'polynomial interpolation', specifically Catmull Rom spline interpolation. This type of interpolation will more closely follow the points and is less prone to pulling the curve far outside the values contained in the points. This is important as pulling the curve above or below the high or low values would misrepresent the time spent hyper or hypoglycemic. These interpolations are complex and involve calculus that

would have taken a long time to learn and translate into code. Fortunately they are common in computer graphics, and I implemented it by referencing existing public domain code and tailoring it to my needs. ("Swift-Catmullrom", 2016). Figure 2a: Average Day Chart demonstrates the result of these calculations. The data set used here is a 'random-but-realistic' set of points, the glucose values scale relative to time since last meal and number of carbohydrates in the last meal. This test data class was used for testing the charting and analysis. I intend to test real patient data obtained voluntarily as I come to the close of the project. Figure 2b: With Points demonstrates the points that went into the calculation. In these plots, the y axis represents blood glucose in mmol/L and the x axis represents the time of day of the reading, regardless of the date of the reading.

Figure 2a: Average Day Chart

To further the efficacy of the visualization, meal and insulin dose time were important. Insulin doses are recorded in units of insulin, as well as what type of insulin, and meals are recorded as grams of carbohydrates. Both of these values are stored in the database discussed above in the 'value' column. Demonstrating doses and meals on the above chart gives rise to two issues. First, insulin dose time and number of

Figure 2b: With Points

units dosed is valuable, but equally valuable is the duration of action and the peak time of insulin action, such that the user can visualize when certain insulin types are acting in their system. Thus, the program can be configured to the specific insulin type that the user is currently taking. Within the program, each type of insulin is associated with a start time, a time to peak, and a duration that is used when drawing an insulin dose to a chart. These values were obtained from the Clinical Practice Guidelines available on the Canadian Diabetes Association website. (McGibbon, Richardson, Hernandez, & Dornan, 2016) The shape of the insulin dose curves is another interpolation using the start, peak and end times, and is an imitation of the curves from literature on the pharmacokinetics of injectable insulin. (Heinemann et al., 2000) A similar curve could be drawn using carbohydrates following meals, but this would require a knowledge of the exact type of carbohydrates consumed and their effect on blood glucose, or their place on the glycemic and insulin indices - measures of impact to specific types of carbohydrates. Patients would not have this information and thus when recording a meal,
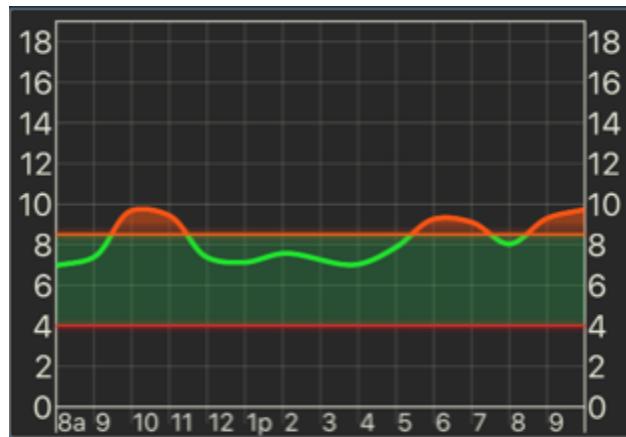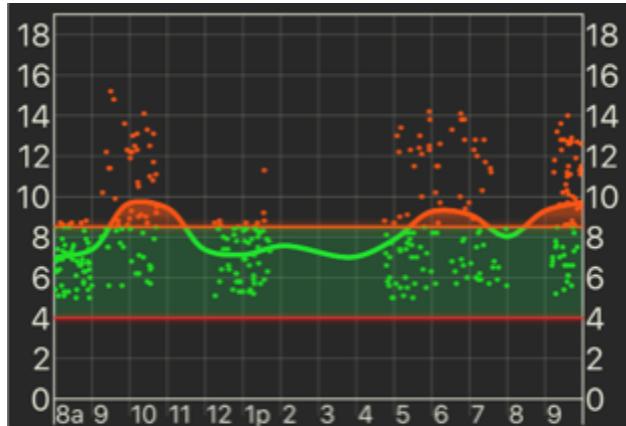
patients are asked only for number of carbohydrates or, alternatively, if it was a small, medium or large meal.

The second issue is that these doses or these meals must be 'clustered' into the average time and average value at that time. To accomplish this a similar transformation into time of day as the x value is applied to the insulin and meal events retrieved from the database. Clustering data was something I was unfamiliar with and had to learn prior to performing this data manipulation. While this is not exactly clustering - which places the data points into sets determined by their x and y values, minimizing within set variance and maximizing between set distance - it is clustering in one dimension, time. Clustering data is a computationally expensive process, classified as NP-Hard ("K-Means Clustering", 2016). Though this topic is above my understanding of computer process complexity, I knew that I had to limit the number of iterations through the array of data points to avoid lengthy load times within the application. My "one-dimensional clustering" algorithm was designed using k-means clustering as inspiration.

The first step is to determine the number of clusters to search for. This value is taken from the most frequently found number of meals or insulin doses(for that type) per day. So a patient eating between 3 and 6 meals a day, if they usually record four meals, the number of clusters is four. If four and five meals in a day are found with equal frequency, the number of clusters is five. Three, four and five meals per day translates to four clusters.

Next, a preliminary x values for the cluster centroids is determined. First, the initial array of points is copied so that points can be removed without disrupting their order. All point's x values are rounded to the nearest interval (of a half hour on first iteration). Using the point's index in the copied array, each point is placed into a set for it's rounded x value, as well as the set for x one and two intervals above and below that rounded value. A meal at 5:35 pm is rounded to x = 17.5, and placed into the sets for points near x = 16.5, 17.0, 17.5, 18.0 and 18.5.This gives an estimation of point density near each interval. The interval with the highest density is selected, and all points that mapped to that rounded x value are removed from the copied array. The process is repeated until the copied array is empty, or the number of clusters is reached. If the copied array is emptied before the number of clusters is reached, then the process restarts with a smaller interval.

The third and final step is to calculate the actual cluster x-centers, and move points between those clusters according to their absolute distance to the nearest cluster center. The mean x value for each previously determined cluster groups is calculated. Then, every point in the initial array is placed into new sets according to which cluster center it was closest too, and again the center-x of the clusters are recalculated. This process is repeated to a maximum of five times (to reduce complexity) or until the cluster points entering the loop are identical to those leaving the loop. Finally the cluster's mean Y value is determined. This gives an estimation of the average meal or insulin dose times each day
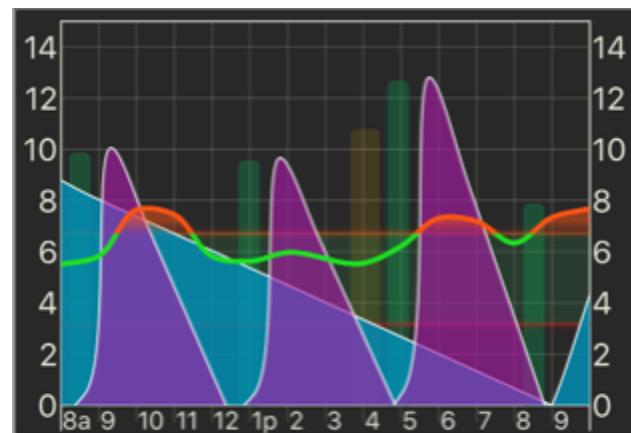

Figure 3a: Insulin Doses

and the average number of carbohydrates consumed or insulin units dosed at that time.

Using these values, it is now possible to see what the average meal time or insulin dose was slightly before the average blood sugar reading reached a peak at any time of day. Figures 3a: Insulin Doses and 3b: Meal Times demonstrate this. This is the same chart, however a control bar below the chart lets the user choose what layer is brought to the front, with the blood glucose curve always as the front-most layer. When the users changes the front layer, the y axis is redrawn, and in this case represents units of insulin and grams of carbohydrates, respectively. Meals are represented as vertical green bars, and insulin doses are represented by the action time curve and colored according to user specified colors for that insulin type. Here, short acting Humalog insulin is purple and long acting Lantus is light blue. If an insulin dose were to carry over to the next day, as the 24 hour duration Lantus insulin taken at just before 9 pm does in this case, the chart renderer detects this, shifts the drawing context left by 24 hours and redraws the curve to imply that the previous day's insulin is in effect throughout the day.

Figure 3b: Meal Times

This chart has the potential to improve the conversation I observed between the patient and dietician. They can view what time of days are spent in a hyperglycemic state and discuss what the patient typically eats in the meal preceding that time. The chart is interactive, and can be filtered to specific days. Tapping on a meal bar will bring up commonly recorded foods if the patient has chosen to record 'what' they eat as well as the number of carbs. Means for each day of the week are shown below this chart, and if one weekday or the weekend is problematic, the chart's data can be filtered to just those problem days and the weighted moving average glucose curve redrawn. The chart has other layers that describe variance at each time of day, range of readings, and a curve that shows what blood glucose looks like on days with exercise sessions recorded. Exercise here is displayed as the one hour wide yellow bar in the background, located at the mean time for exercise sessions.

5.2 A Novel Algorithm for Mean Amplitude of Glycemic Excursions

The second informatics associated issue I will present is my implementation of one measure of glycemic variability that can be calculated from SMBG data, the mean amplitude of glycemic excursions(MAGE). I would like to highlight this issue as I strongly suspect that as diabetes data is more easily made available to physicians, glycemic variability will become more important in clinical decision making. As well, this particular measure of variability has been cited as the most challenging to compute. (Marics et al., 2015)

The design of two 'systems' for this calculation (Hill et al., 2011) (Czerwoniuk, Fendler, Walenciak, & Mlynarski, 2011) have been published and two separate algorithms (Fritzsche,
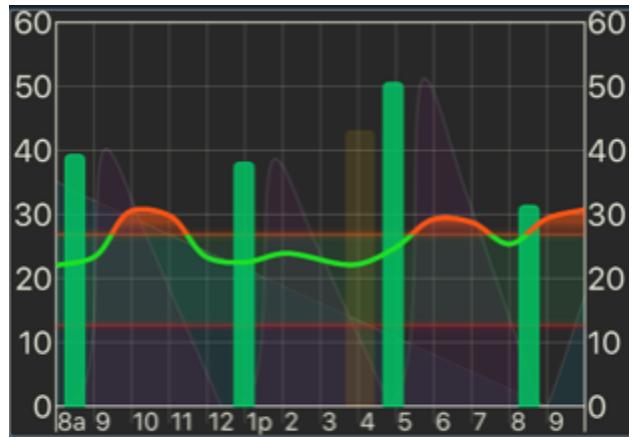
Kohnert, Heinke, Vogt, & Salzsieder, 2011)(Baghurst, 2011) have been published that demonstrate differing calculations of this measure. While I drew on these algorithms as inspiration, I opted to redesign the algorithm for these reasons: There is some disagreement between the algorithms when fed the same set of data.(Marics et al.) All of these publications focus on continuous glucose monitoring data only, as opposed to the SMBG data that my product analyzes. None of these publishings were contained in a patient-centric mobile application and presented in a way that a patient might understand this value. This issue demonstrates the technical aspects of designing an algorithm as well as how this value can be implemented as a teaching tool for patients and caregivers.

The mean amplitude of glycemic excursions(MAGE) was one of the first attempts to quantify glycemic variability, first published in 1970. (Service et al., 1970) It is well studied and validated. (Kohnert, 2015) In the design of my calculation, I drew on the description of MAGE from a 2013 review of glycemic variability from Service (Service, 2013), the primary author on the 1970 paper first describing MAGE. It is designed to quantify the number of 'swings' or excursions in blood sugar from a peak to a nadir(a low) or vice versa. The cut off for the determination of an excursion is an absolute value greater than one standard deviation for all values taken on that day. It has been applied to both self measured blood glucose and continuous glucose monitoring data. As the calculation is rooted in hand drawing excursions on a printout of blood sugar readings vs time, the resulting swings can be visualized, enabling me to verify the results of my algorithm.

The calculation involves finding all the alternating local peaks and nadirs that are separated by an excursion exceeding one standard deviation of the data that day. The iterative portion of the algorithm is available in a flow chart in the Appendix of this report.

First, the data set is spliced into sets for each day with data. Each set's standard deviation is calculated and placed in an indexed array(specifically in a hash or dictionary), where the index is the day for that set relative to the first day with data.

Next the dataset is sorted by x value and filtered, removing all points in which the incoming line and the outgoing line are moving in the same direction. This is accomplished by dividing the incoming line's slope by the outgoing line's slope. If the resulting ratio is positive (a positive value divided by a positive or a negative by a negative), the point is removed from the set.

The resulting array contains only points that represent a change in direction. The final value is calculated by counting the number of excursions either from peak to nadir or nadir to peak(not both), the direction being determined by the first excursion found in the data. A peak may only follow a nadir, and a nadir may only follow a peak, this means that even if a single point to point difference in blood sugar is greater than the cut off, the latter point does not count as a peak(or nadir) until the end point of that excursion is reached, ie the highest local peak before the next descending excursion.

Thus, when iterating over the array of turning points, the following points are needed for comparison: the last point, the low point this interval, the high point this interval, the last peak and the last nadir. To begin, an empty array of peaks and an empty array of nadirs are initialized. The last point, low this interval and high this interval are initialized to the first point of the array of turning points. Then the remaining points, starting at point two, are iterated through.

The slope from the last point to the current point is determined. If less than zero (down), then the point is compared to the lowest point in this interval. If the y value of the current point is lower than the lowest point this interval, the point becomes the low this interval and we check for an excursion. To do this, the absolute change in glucose relative to the high this interval is determined. Next, to determine the standard deviation for comparison, the midpoint in X between the two points is determined, specifying the day in which the majority of this excursion occurs. The integer of this value summons the standard deviation from the indexed standard deviation array. If the absolute change in glucose exceeds this standard deviation, we have determined our first nadir and this value is added to the nadir array. The last point, the low this interval and the high this interval are all reset to this point. If it is not, only the last point is set to the current point.

On the next loop, the slope is again determined and should be found to be greater than zero. The glucose value for this point is compared to the high this interval, and if found to be higher, it becomes the high this interval and we check for an excursion again. If the change exceeds the standard deviation, this peak is added to the array.

However, to account for potential excursions that do not end in a local peak or nadir, at each iteration a check is performed. If the slope from the last value is positive (and we are looking for a peak) the algorithm checks the boolean 'Was the last a peak?', if so, the y value for the current point is compared to the last peak and if higher, the last peak is removed and the current point is added to the peaks. This prevents two peaks being added in succession, and ensures that all excursions are bounded by a peak and a nadir.

Finally, if a peak was found first, the set of peaks and nadirs are zipped, and peak to nadir change is determined for each pair. This array of peak to nadir absolute changes (or nadir to peak if a nadir was found first) is averaged, and the resulting value is the mean amplitude of all glycemic excursions in this time period. The flowchart for the major portions of this algorithm is available in Appendix I: MAGE Algorithm Flow Chart.

When displayed graphically, the peaks and nadirs determined by this algorithm do in fact coincide with the proper peaks and nadirs on the readings chart suggesting that it has in fact calculated the MAGE for those readings. When testing the application with genuine user data, I will verify the results of all the glycemic variability calculations against previously published variability calculators. This value can be used to quantify patients with extreme swings in blood sugar data, and to alter treatment plans accordingly. Excursions tend to follow meals, and the application will map these swings to certain food types and amounts of carbohydrates, as well as explaining what this value implies so that it can be used as a teaching aid by the caregiver or a learning opportunity within the app by the patient.

These two examples of issues faced during my internship demonstrate the range of problems I contended with when trying to design a tool for both clinicians and patients. There were many others, as my position enabled me to work from a conceptual design of a health information technology solution to a tangible product, however these were two of the most complex both in concept and for coding.

6. Conclusion

This project has been challenging and deeply rewarding. The informatics specific issues I faced were as numerous as the technical issues in learning the language. This was my first opportunity to apply my enjoyment of programming to a genuine project. Likewise, it was my first opportunity to apply everything I have learned in medicine and health informatics to a product that I hope will improve the experience for any diabetic patient and their caregivers. While I cannot deny my bias, the product we have created offers a significant improvement on existing applications for collecting, curating and analyzing diabetes data. I have downloaded, experimented with and read reviews of a variety of existing products. When contrasted with these applications, the entry system is simpler and quicker, the visualizations more aesthetically pleasing and understandable, and the analysis significantly more robust. Furthermore, our app has a scale-to-the-user customizability that should appeal to a wider range of diabetic users. I am confident that there will be a market for patients that want to use this application for recording their blood sugar data, and to learn all that can be learned from this data when it is properly parsed. As well, the analytical and visualization classes coded for this product can be directly applied to a similar app for caregivers when that stage of the project is reached.

This is an ongoing project that will continue after my position at IWK comes to a close this summer. I am determined to release the patient-centric application in as little as a month. Between classes this fall, I will continue to work on the caregiver and server side of the system. Even accomplishing the patient-centric tool, I feel, represents a success. My business acumen certainly lags behind my clinical and technical abilities, but I hope there is a market for what I have created with my advisors' help. Regardless of financial success, I intend to support this position at IWK in the future in any way I can. This experience has taught me more than I could outline in a brief report and I what I have accomplished gives me a great deal of pride. I suspect there are aspects of my system, like the MAGE algorithm above or the framework for the application, that have potential for papers and conference posters and I do hope to have the opportunity to present it in the future.

The position was unique in the amount of self-directed learning and planning it involved. I was given as much support as necessary from the hospital and from my advisors and supervisor, for which I am deeply gracious. Together, we have taken an information technology solution to a health care problem from a conceptual stage to near complete product. Hopefully, in the process, we have justified the investment in a 'student innovator'. Health informatics is an evolving field, and those training in it will be a large part of determining how profoundly we can change healthcare for the better.

7. Recommendations

Foremost in my recommendations is that this program should continue. The investment is small and the returns in human capital are potentially immense. Putting the onus on the student to envision, plan, learn and implement a project of their own design is deeply motivational. To me, the project felt like it was 'mine', that success was contingent only on my ability to follow through. No position I have ever held has given me that sort of encouragement and in turn, I worked harder than any position held previously. Regardless of success of the product, the amount I have learned and the debt of gratitude I owe to the institution that supported me will result in returns on their investment. Selecting the right student is probably important, but I know I am not the first MHI student to dream up a technical solution to a health care problem, and I am certain there will be others.

To future students that might get this opportunity I have some recommendations. Be certain of yourself and your project. People will doubt you, and, more so, you will doubt yourself as I have many times. I will not be able to fully dispel those doubts until I have a product that patients and caregivers use and appreciate. But when I get deep into my work those doubts fade and all I do is learn and build, then I am able to see that endpoint that I am near certain is approaching. Have a flexible plan before you start, and do not presume to know how your product can fit into a clinical context without seeing the context. I had only seen half of the context - type 2 diabetes care. Clinical observership of type 1 care was eye-opening. In fact I feel I did not fully take advantage of that opportunity, however time was precious with the amount of coding I had, and the majority of time spent shadowing does not exactly contribute to informing your decisions - only brief moments do. My clinical experience allowed me to very precisely see what I needed to and I was ready to take advantage of those brief moments when they occurred. Reviewing my notes from observing, I have implemented nearly all of what I observed that would help my product fit into the existing clinical context. Make sure to constantly be humble. You will have many bad ideas and some good ideas. Being told an idea is bad is just as helpful as being told one is good when the feedback is in proper context. Thus, leave any feelings for or dedication to specific ideas aside. Finally, be confident. It turns out that if you can dream it up, there is almost definitely a way to do it and the resources are out there for you to figure out how.

To my advisors and supervisor, your support was integral to my project. Even having an individual aware of clinical and informatics issues to demonstrate to was helpful, and of course the feedback even more so. I appreciate feedback as harsh as "that's a bad idea", to that more carefully worded. When it is a bad idea (and I had many!) make sure to say so, and concise wording is just as helpful as elaborate wording. I was forging a path to my envisioned end point and I often felt half blind in doing so. You have the experience that I needed to draw on to give me just enough vision to see where I was going.

8. References

1. Apple Inc.,. (2016). *Swift 3.0 Release Process*. *Swift.org*. Retrieved 13 August 2016, from
    https://swift.org/blog/swift-3-0-release-process/

2. Baghurst, P. (2011). Calculating the Mean Amplitude of Glycemic Excursion from Continuous
    Glucose Monitoring Data: An Automated Algorithm. *Diabetes Technology & Therapeutics*, *13*(3),
    296-302. http://dx.doi.org/10.1089/dia.2010.0090

3. Basilico, A., Marceglia, S., Bonacina, S., & Pinciroli, F. (2016). Advising patients on selecting trustful
    apps for diabetes self-care. Computers In Biology And Medicine, 71, 86-96.
    http://dx.doi.org/10.1016/j.compbiomed.2016.02.005

4. Canada Health Infoway. (April 2014). Mobile Health Computing Between Clinicians and Patients
    White Paper. Emerging Technology Series. Retrieved from:
    https://www.infoway-inforoute.ca/en/component/edocman/resources/technical-documents/emerging-
    technology/1883-mobile-health-computing-between-clinicians-and-patients-white-paper-full-report

5. Czerwoniuk, D., Fendler, W., Walenciak, L., & Mlynarski, W. (2011). GlyCulator: A Glycemic
    Variability Calculation Tool for Continuous Glucose Monitoring Data. *Journal Of Diabetes Science
    And Technology*, *5*(2), 447-451. http://dx.doi.org/10.1177/193229681100500236

6. *Dalhouse Division of Pediatric Emerency Medicine | Our People*. (2016). *Dalhousie University*.
    Retrieved 12 August 2016, from
    http://medicine.dal.ca/departments/department-sites/emergency/divisions/pediatric-emergency-medi
    cine/our-people.html

7. Fritzsche, G., Kohnert, K., Heinke, P., Vogt, L., & Salzsieder, E. (2011). The Use of a Computer
    Program to Calculate the Mean Amplitude of Glycemic Excursions. *Diabetes Technology &
    Therapeutics*, *13*(3), 319-325. http://dx.doi.org/10.1089/dia.2010.0108

8. Heinemann, L., Linkeschova, R., Rave, K., Hompesch, B., Sedlak, M., & Heise, T. (2000).
    Time-action profile of the long-acting insulin analog insulin glargine (HOE901) in comparison with
    those of NPH insulin and placebo. *Diabetes Care*, *23*(5), 644-649.
    http://dx.doi.org/10.2337/diacare.23.5.644

9. Hill, N., Oliver, N., Choudhary, P., Levy, J., Hindmarsh, P., & Matthews, D. (2011). Normal
    Reference Range for Mean Tissue Glucose and Glycemic Variability Derived from Continuous
    Glucose Monitoring for Subjects Without Diabetes in Different Ethnic Groups. *Diabetes Technology
    & Therapeutics*, *13*(9), 921-928. http://dx.doi.org/10.1089/dia.2010.0247

10. *IWK Health Centre | About Us*. (2016). *Iwk.nshealth.ca*. Retrieved 12 August 2016, from
    http://www.iwk.nshealth.ca/page/about-us

11. *K-means clustering*. (2016). *Wikipedia*. Retrieved 12 August 2016, from
    https://en.wikipedia.org/wiki/K-means_clustering#Complexity

12. Kohnert, K. (2015). Utility of different glycemic control metrics for optimizing management of diabetes. *World Journal Of Diabetes*, *6*(1), 17. http://dx.doi.org/10.4239/wjd.v6.i1.17

13. Kohnert, K., Vogt, L., & Salzseider, E. (2016). *Advances in Understanding Glucose Variability and the Role of Continuous Glucose Monitoring | Touch Endocrinology | Independent Insight for Medical Specialists*. *Doi.org*. Retrieved 12 August 2016, from http://doi.org/10.17925/EE.2010.06.00.53

14. Marics, G., Lendvai, Z., Lódi, C., Koncz, L., Zakariás, D., & Schuster, G. et al. (2015). Evaluation of an open access software for calculating glucose variability parameters of a continuous glucose monitoring system applied at pediatric intensive care unit. *Biomedical Engineering Online*, *14*(1). http://dx.doi.org/10.1186/s12938-015-0035-3

15. McGibbon, A., Richardson, C., Hernandez, C., & Dornan, J. (2016). *Pharmacotherapy in Type 1 Diabetes*. *Guidelines.diabetes.ca*. Retrieved 12 August 2016, from http://guidelines.diabetes.ca/browse/Chapter12

16. *Pediatric Endocrinology Department*. (2016). *Dalhousie University*. Retrieved 12 August 2016, from http://medicine.dal.ca/departments/department-sites/pediatrics/divisions/endocrinology.html

17. Service, F. (2013). Glucose Variability. *Diabetes*, *62*(5), 1398-1404. http://dx.doi.org/10.2337/db12-1396

18. Service, F., Molnar, G., Rosevear, J., Ackerman, E., Gatewood, L., & Taylor, W. (1970). Mean Amplitude of Glycemic Excursions, a Measure of Diabetic Instability. *Diabetes*, *19*(9), 644-655. http://dx.doi.org/10.2337/diab.19.9.644

19. *swift-catmullrom*. (2016). Author: Andre Lind. *GitHub*. Retrieved 12 August 2016, from https://github.com/andrelind/swift-catmullrom/blob/master/CatmullRom.swift

20. *Weighted Moving Average Model*. (2016). *Help.sap.com*. Retrieved 12 August 2016, from http://help.sap.com/saphelp_scm70/helpdata/en/5d/a2ca53f1a1214be10000000a174cb4/content.htm

9. Glossary

Clinical Terms	
Diabetes	A chronic disorder characterized by elevated plasma glucose levels.
Type 1 Diabetes	The elevated plasma glucose is due to an inability to produce insulin, patients must self administer insulin daily to reduce blood sugar. Patients must test their blood sugar multiple times a day.
Type 2 Diabetes	The elevation is due to a cellular resistance to the effect of insulin, typical patients take oral medication for sugar control and test their blood sugar less frequently than type 1 diabetics.
Blood Glucose	Circulating sugar in plasma, increased with carbohydrate consumption and reduced by insulin, measured in mmol/L (in Canada) and mg/dL (in the US).
Glucose Variability	Within day or between day swings in blood glucose. Quantified by various metrics including standard deviation and MAGE.
Self Measured Blood Glucose	The protocol that diabetics adhere to concerning daily testing of blood glucose with glucometers and finger sticks.
Continuous Glucose Monitoring	A device that measures interstitial (capillary) blood glucose at a set interval throughout the day, as low as five minutes. Use of CGM enables diabetics to avoid SMBG.
Insulin	A gatekeeper hormone that opens the door to a cell and lets glucose enter a cell from the space outside a cell. Released from the pancreas in response to elevation in blood glucose.
Insulin Pump	A device that can infuse insulin directly to a patient's system without the need for frequent injections. The device is calibrated with the number of carbohydrates consumed and the blood sugar glucometer readings.
Carbohydrates	A single unit or a set of chemically bound units of sugar. These are the nutrients most responsible for elevating blood glucose and, in turn, insulin release.
Postprandial	The period following a meal, typically beginning one hour after eating and ending 2.5 hours after eating, where blood glucose would peak following a meal.
Technical terms	
Swift	Apple's open source programming language which can be used to code applications for iOS, the operating system for Apple devices.
SQLite	A simplified relational database management language for interacting with a database.
Entity Relationship Diagram	A diagram that demonstrates the set of tables, columns and relationships between tables in a relational database.
Array	A collection type of variables, typically all of the same type, zero-indexed with integers.
Hash	A collection type representing pairings of keys (or indices) and values. Keys do not have to be integers. Called a dictionary in Swift.
Boolean	A variable that is only either true or false. If the variable is 'optional' it may also be nil.
Model-View-Controller	A paradigm in object oriented programming in which the model defines the database, the view defines what a user sees and interacts with, and the controller is responsible for interacting with both view and model, moving data between them and handling user input

10. Appendix I: MAGE Algorithm Flow Chart

Remove all points in which the preceding and ensuing lines are both positive slope or both negative slope

Initialize optional boolean FirstIsDescending to **nil**
Initialize optional boolean LastWasPeak to **nil**
Initialize empty arrays [peaks] and [nadirs]

Initialize last point, high point this interval, low point this interval to the first point

Move to next point

Loop through all points

Determine slope to last point

Slope > 0 (line goes up)

Slope < 0 (line goes down)

Was last a peak
AND
current point Y value >
last **peak** y value?

Same steps as otherwise replacing **peak** with **nadir** and **highThisInterval** with **lowThisInterval**

No last was not peak

Yes:
Remove last peak, replace with current point.
Set high this interval to current point

Is current point Y value >
high this interval Y?

No. This point is not possibly a peak and lower than the current high this interval

Yes:
Set **high this interval** to current point
Check for excursion

Obtain values between current point and **low this interval:**
- Slope
- Standard deviation on day of this potential excursion
- Absolute change in Y

Is absolute change >
standard deviation

No. This point is not a peak

Yes: local peak found

Is **firstWasDescending** still nil?

No: already determined first excursion direction

Yes: Set first was descending to false.
Add low point this interval to Nadirs

- Append current point to peaks
- Set last was peak to true
- Reset low this interval AND high this interval to current point

Set **lastPoint** to current point